

ESTADÍSTICA INFERENCIAL I

APUNTES

MARÍA CONCEPCIÓN FUENTES MORALES

AGOSTO 2021 - FEBRERO 2022



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

CARTA DE RECONOCIMIENTO DEL AUTOR DE LOS DERECHOS A FAVOR DEL
TECNM

Ciudad de México, 02/febrero /2022

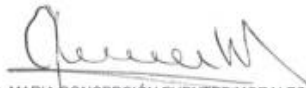
TECNOLÓGICO NACIONAL
DE MÉXICO.PRESENTE

Bajo protesta de decir verdad, **María Concepción Fuentes Morales**, personal docente adscrito a **Ciencias Básicas en el Instituto Tecnológico de Cd. Juárez** del Tecnológico Nacional de México, manifiesto que en cumplimiento de mis actividades relacionadas con el Año Sabático elaboré la obra titulada **"ESTADÍSTICA INFERENCIAL I"**.

Con base en lo anterior, y con fundamento en los artículos 83 de la Ley Federal del Derecho de Autor y 46 de su Reglamento, reconozco que el Tecnológico Nacional de México es titular de los derechos patrimoniales sobre la misma y le corresponden las facultades relativas a la divulgación, integridad de la obra y de colección, conservando el derecho a figurar como autor.

Asimismo, respondo por la autoría y originalidad de la citada obra; y relevo de toda responsabilidad al Tecnológico Nacional de México de cualquier demanda o reclamación que llegara a formular alguna persona física o moral que considere que con esta obra es afectado en alguno de los derechos protegidos por la Ley en cita, asumiendo todas las consecuencias legales y económicas.

ATENTAMENTE



M.C. MARIA CONCEPCIÓN FUENTES MORALES
AUTOR



Av. Universidad 1200, Col. Xoco, Alcaldía Benito Juárez,
C.P. 03330, Ciudad de México.
Tel. (55) 3601-7500, Ext. 65050, e-mail: direccion@tecnm.mx
tecnm.mx



Contribución académica. Análisis de la contribución del tema a los objetivos educativos del programa educativo al cual pertenece la asignatura:

El presente, apuntes de la materia de Estadística Inferencial I abarcan el programa de la asignatura, con ejercicios resueltos paso a paso en su desarrollo para proporcionarle una mejor comprensión al estudiante de esta materia; desafiándolo con ejercicios propuestos, así mismo se le presenta la solución en un anexo donde podrá comprobar su resultado, cuenta con el apoyo de la enseñanza en la aplicación del software estadístico minitab® para la solución de problemas. Aunado a esto se refuerza el conocimiento con una evaluación final de los conceptos importantes del contenido por unidad. Por otro lado, la contribución al docente del área de estadística ayuda a objetivar su conocimiento y aplicación en el área de ingeniería industrial y logística proporcionándole un banco de reactivos de aplicación práctica. Así mismo es un análisis como un desarrollo e innovación en el área de ciencias básicas para incrementar el prestigio académico del Instituto Nacional de México y del Instituto Tecnológico de Cd. Juárez, egresando profesionistas con altas competencias profesionales y puedan desarrollar en cualquier campo de trabajo del mercado global.

Índice/Contenido.

1.	Distribuciones Fundamentales para el Muestreo.....	1
1.1.	Introducción a la Estadística Inferencial.....	3
1.2.	Muestreo: Introducción al muestreo y tipos de muestreo.....	3
1.3.	Teorema del límite central.....	9
1.4.	Distribuciones fundamentales para el muestreo.....	10
1.4.1	Distribución muestral de la media.....	19
1.4.2	Distribución muestral de la diferencia de medias.....	21
1.4.3	Distribución muestral de la proporción.....	24
1.4.4	Distribución muestral de la diferencia de proporciones.....	31
1.4.5	Distribución t-student.....	37
1.4.6	Distribución muestral de la varianza.....	41
1.4.7	Distribución muestral de la relación de varianzas.....	44
2.	Estimación.....	50
2.1	Introducción.....	52
2.2	Características de un estimador.....	53
2.3	Estimación puntual.....	53
2.4	Estimación por intervalos.....	54
2.4.1	Intervalo de confianza para la media.....	56
2.4.2	Intervalo de confianza para la diferencia de medias.....	60
2.4.3	Intervalos de confianza para la proporción.....	62
2.4.4	Intervalos de confianza para la diferencia de proporciones.....	64
2.4.5	Intervalos de confianza para la varianza.....	66
2.4.6	Intervalos de confianza para la relación de varianzas.....	68
2.5	Determinación del tamaño de muestra.....	68
2.5.1	Basado en la media de la Población.....	69
2.5.2	Basado en la proporción de la Población.....	70
3.	Prueba de hipótesis.....	76
3.1	Introducción.....	78
3.2	Confiabilidad y significancia.....	79
3.3	Errores tipo I y tipo II.....	80
3.4	Potencia de la prueba.....	81
3.5	Formulación de Hipótesis estadísticas.....	82
3.6	Prueba de hipótesis para la media.....	86
3.7	Prueba de hipótesis para la diferencia de medias.....	87
3.8	Prueba de hipótesis para la proporción.....	89
3.9	Prueba de hipótesis para la diferencia de proporciones.....	91
3.10	Prueba de hipótesis para la varianza.....	93

3.11 Prueba de hipótesis para la relación de varianzas.....	95
4. Pruebas de bondad de ajuste y pruebas no paramétricas.....	103
4.1 Bondad de ajuste.....	103
4.1.1 Análisis Ji-Cuadrada.....	105
4.1.2 Prueba de independencia.....	107
4.1.3 Prueba de la bondad del ajuste.....	110
4.1.4 Tablas de contingencia.....	113
4.2 Pruebas no paramétricas.....	116
4.2.1 Escala de medición.....	116
4.2.2 Métodos estadísticos contra no paramétricos.....	118
4.2.3 Prueba de Kolmogorov – Smirnov.....	121
4.2.4 Prueba de Anderson – Darling.....	129
4.2.5 Prueba de Ryan – Joiner.....	134
4.2.6 Prueba de Shappiro – Wilk.....	139
5. Regresión Lineal Simple.....	145
5.1 Prueba de hipótesis en la regresión lineal simple.....	153
5.2 Calidad del ajuste en regresión lineal simple.....	155
5.3 Estimación y predicción por intervalo en regresión lineal simple.....	155
6. Bibliografía.....	163
ANEXOS.....	165
Anexo 1: Solución a Problemas propuestos mediante enlace virtual.....	166
Anexo 2: Exámenes Propuestos mediante enlace virtual.....	177
Anexo 3: Uso de software estadístico minitab® mediante enlace virtual.....	178
Anexo 4: Tablas de Distribuciones Estadísticas.....	179
Anexo 5: Instrumentación Didáctica completa por asignatura.....	196

Introducción.

Las “ciencias exactas, “duras” requieren del razonamiento y abstracción para su estudio y comprensión; puede indicar dificultad para unos, y facilidad para otros, lo cierto es que el alumno debe de constituirse en una aventura y lograr aprendizajes significativos en estas ciencias. En la actualidad, la enseñanza de la estadística adquiere cada vez mayor importancia en los currículos de los países. Es necesario que el estudio de la metodología sea una enseñanza activa, y representativa del significado de conceptos. La estadística inferencial posee aplicaciones en un gran número de áreas, en las cuales resulta ser una herramienta fundamental. El papel tan importante que desempeña esta rama de las matemáticas en numerosos campos de aplicación, donde podemos citar las ingenierías en su gran diversidad de especialidades, las ciencias actuariales, la salud pública, las ciencias de la vida y ambientales, la física, la investigación de operaciones, la informática, etc.

Los conceptos estadísticos son a menudo mal entendidos por los usuarios, las consecuencias de esta confusión pueden ser claramente significativas debido al uso incorrecto de los datos no hay manera de obtener resultados, a diferencia de usar cálculos estadísticos y representaciones gráficas de los mismos; ya que éstas ilustran tendencias y relaciones de forma rápida y sencilla de interpretar. La estadística inferencial es una herramienta fundamental para muchas investigaciones, pero debido al desconocimiento, también es un instrumento que se utiliza a veces de manera incorrecta o inapropiada. Para solucionar este problema es necesario mejorar la formación estadística de cualquier estudiante, por lo que de esta forma no solo dominará los conceptos básicos, sino también poseerá una visión crítica de los métodos y del uso de los datos. El docente debe tener en su preparación académica y pedagógica las suficientes herramientas de trabajo actualizadas como incluir la tecnología y poder estar al mismo ritmo que sus estudiantes.

Hoy en día un elemento importante en la enseñanza- aprendizaje es la tecnología en conocimientos de las matemáticas, pero sobre todo indispensable en la rama de la estadística. Provoca que el estudiante se enfrente retos y desafíos en el uso y manejo de ésta conlleva al desarrollo de los principios éticos del conocimiento y funcionamiento productivo del mismo. La herramienta tecnológica proporciona contribución a la comprensión y despertará el interés en particular en el ámbito

escolar. El software estadístico más utilizado en las empresas es el Minitab®, por tal motivo es indispensable los estudiantes egresen con un alto dominio en el mismo aportándole un alto valor a su perfil profesional; permitiéndole realizar el análisis en profundidad de las operaciones de fabricación y potenciar sus modelos predictivos e iniciativas en la mejora continua al obtener resultados rápidos y precisos para su interpretación e inferencia. Por tal motivo el uso de este software es considerado en esta propuesta de trabajo de Elaboración de Apuntes describiendo con ilustraciones detalladas la aplicación en temas del programa mediante un power point®.

Estos apuntes de la materia de Estadística Inferencial I abarcan el programa de la asignatura, con ejercicios resueltos y desafiando al estudiante con ejercicios propuestos con solución anexo, cuenta con el apoyo de la enseñanza del software estadístico. Aunado a esto se refuerza el conocimiento con una evaluación final de los conceptos importantes del contenido por unidad. Por otro lado, al docente ayudará a objetivar su conocimiento y aplicación en el área de ingeniería industrial y logística. Siendo un análisis como un desarrollo e innovación en el área de ciencias básicas para incrementar el prestigio académico del Instituto Nacional de México egresando profesionistas con altas competencias profesionales.

UNIDAD 1

DISTRIBUCIÓN DE MUESTREO



Instrumentación Didáctica

Competencia Específica	Competencia Genérica	Actividades de Aprendizaje
Escoge el tipo de muestreo al que corresponde un experimento según la selección de la variable de estudio para observar su comportamiento.	• Capacidad de abstracción, análisis y síntesis. • Capacidad de aplicar los conocimientos en la práctica. • Capacidad para identificar, plantear y resolver problemas.	• Buscar información sobre conceptos relacionados con el muestreo. • Discutir en grupo sobre los conceptos investigados. • Distinguir entre muestreo aleatorio probabilístico y no probabilístico. • Proporcionar al estudiante situaciones hipotéticas de procesos y/o poblaciones finitas para que obtengan un conjunto de datos para su análisis. • Obtener los valores de probabilidad de t, χ^2, F y Z de las diferentes distribuciones muestrales. • Interpretar los resultados obtenidos.

1. Distribuciones Fundamentales para el Muestreo.

1.1 Introducción a la Estadística Inferencial.

La estadística inferencial es la parte de la Estadística que se encarga de estudiar procedimientos para la obtención de conclusiones, referentes al total de la población, a partir de la información proporcionada por la muestra o muestras seleccionadas.

Es de gran utilidad en todos aquellos estudios de investigación llevados a cabo en poblaciones demasiados grandes como para poder realizar mediciones en todos y cada uno de los individuos de dichas poblaciones.

Para poder llevar a cabo el proceso de Inferencia Estadística es habitual representar a dicha variable como una variable aleatoria X se supone desconocida, en todo o en parte.

- Si suponemos que la distribución de probabilidad X es conocida en su totalidad salvo por uno o varios parámetros, diremos que estamos ante un problema de *Inferencia Paramétrica*.
- Si suponemos que la distribución de probabilidad es totalmente desconocida, diremos que estamos ante un problema de **Inferencia No Paramétrica**.

La inferencia No Paramétrica es más complicada, y no será objeto de estudio en esta asignatura.

En términos generales, hay dos grandes tipos de procedimientos inferenciales:

- Los procedimientos de estimación
- Los procedimientos de contraste de hipótesis.

Teniendo en cuenta que las técnicas a utilizar con cada uno de ellos serán diferentes, es importante saber distinguir claramente cuando estamos ante un problema de estimación y cuando estamos ante un problema de contraste de hipótesis.

1.2 Muestreo: Introducción al muestreo y tipos de muestreo.

El definir el tamaño muestral y el método de selección de las unidades de análisis, es tal vez el punto más crucial en el proceso investigativo, ya que dependiendo de éste, será posible generar inferencias o generalizaciones a toda la población, y por ende, definir políticas y tomar acciones que impacten el entorno del objeto de estudio.

Es evidente que el reducir costos en todas las actividades que se realizan en la cotidianidad es una prioridad, la investigación no es la excepción, por lo tanto, es menester en lo posible trabajar con una parte de la población objeto de estudio. Por lo que es necesario conocer los siguientes conceptos:

- *Población*: Conjunto de sujetos sobre los que queremos aplicar las conclusiones de una investigación.
- *Muestra*: Parte representativa de un conjunto mas amplio llamado población. Debe reproducir lo más fielmente las características de la población.
- *Unidad de Investigación/ Unidad de Muestra*: Es el objeto de estudio que reúne las características que me interesan para mi estudio.
- *Relación Población-Muestra*: puede ser cuantitativa o cualitativa.
 - *Cuantitativa*: tiene que ver con el error y el tamaño de la muestra, es

decir, la muestra debe tener un tamaño proporcional al de la población que le permita ser lo suficientemente representativa como para poder generalizar los resultados con ese grupo reducido que es la muestra, a toda la población.

- *Cualitativa*: significa que la muestra debe tener las mismas características que la población de la que procede y a la que representa. Si no las tuviese, no sería representativa.

- *Base de la muestra*: es un registro en el que las unidades de la muestra aparecen individualizadas. Cada una de esas unidades que forman la población es la unidad de muestra, y todas ellas son potencialmente elegibles, tienen la posibilidad de ser elegidas parte de la muestra. Hay que tener cuidado que la base de la muestra no puede estar ordenada atendiendo a un criterio. De ser así el muestreo nunca sería realmente aleatorio.

- *Error muestral*: Se refiere a la variación natural existente entre muestra

tomadas de la misma población. Aún si se ha tenido gran cuidado para asegurar que dos muestras del mismo tamaño sean representativas de una cierta población, no esperaríamos que las dos sean idénticas en todos los detalles.

- **Sesgo:** es un tipo de error no muestral. Se refiere a una tendencia sistemática

inherente a un método de muestreo que da estimaciones de un parámetro que son, en promedio, menores) sesgo negativo=, o mayores (sesgo positivo), que el parámetro real.

- **Aleatorización:** se refiere a cualquier proceso de selección de una muestra de

la población en el que la selección es imparcial o no está sesgada; una muestra elegida con procedimientos aleatorios se llama muestra aleatoria.

Tipos de Muestreo: pueden ser probabilísticos y No probabilísticos.

- **Probabilísticos:** Se basan en la selección aleatoria de los sujetos a partir de la

base de la muestra. Los elementos de la población son equiprobables, todos tienen la misma posibilidad de ser seleccionados para la muestra.

- **No Probabilísticos:** Son muestras cuyos individuos no se han escogido al

Azar y por lo tanto no permiten hacer inferencias de los resultados obtenidos con la muestra, a la población.

Los tipos más comunes de técnicas de muestreo aleatorios son el muestreo

1. *Muestreo aleatorio simple:* Es la técnica de muestreo en la que todos los elementos que forman el universo y que por lo tanto están descritos en el marco

muestral, tienen idéntica probabilidad de ser seleccionados para la muestra.

Un procedimiento para seleccionar una muestra aleatoria simple de una población finita es elegir los elementos para la muestra de uno en uno, de manera que, en cada paso, cada uno de los elementos que quedan en la población tenga la misma

probabilidad de ser seleccionado. Al seleccionar n elementos de esta manera, será satisfecha la definición de muestra aleatoria simple seleccionada de una población finita. Para seleccionar una muestra aleatoria simple de la población finita de administradores de EAI, primero se le asigna a cada administrador un número. Por ejemplo, se les asignan los números del 1 al 2500 en el orden en que aparecen sus nombres en el archivo de personal de EAI. A continuación se consulta la tabla de dígitos aleatorios que se muestran en la tabla 7.1. Al consultar el primer renglón de la tabla se da cuenta que cada dígito, 6, 3, 2, ... es un dígito aleatorio con la misma oportunidad de aparecer que cualquier otro. Como el número mayor en la lista de la población de administradores de EAI, 2500, tiene cuatro dígitos, se seleccionarán números aleatorios de la tabla en conjuntos o grupos de cuatro dígitos. Aun cuando para la selección de números aleatorios se puede empezar en cualquier lugar de la tabla y avanzar sistemáticamente en una de las cuatro direcciones, aquí se usará el primer renglón de la tabla de números aleatorios y se avanzará de izquierda a derecha. Los primeros 7 números aleatorios de cuatro dígitos son:

6327 1599 8671 7445 1102 1514 1807

63271	59986	71744	51102	15141	80714	58683	93108	13554	79945
88547	09896	95436	79115	08303	01041	20030	63754	08459	28364
55957	57243	83865	09911	19761	66535	40102	26646	60147	15702
46276	87453	44790	67122	45573	84358	21625	16999	13385	22782
55363	07449	34835	15290	76616	67191	12777	21861	68689	03263
69393	92785	49902	58447	42048	30378	87618	26933	40640	16281
13186	29431	88190	04588	38733	81290	89541	70290	40113	08243
17726	28652	56836	78351	47327	18518	92222	55201	27340	10493
36520	64465	05550	30157	82242	29520	69753	72602	23756	54935
81628	36100	39254	56835	37636	02421	98063	89641	64953	99337
84649	48968	75215	75498	49539	74240	03466	49292	36401	45525
63291	11618	12613	75055	43915	26488	41116	64531	56827	30825
70502	53225	03655	05915	37140	57051	48393	91322	25653	06543
06426	24771	59935	49801	11082	66762	94477	02494	88215	27191
20711	55609	29430	70165	45406	78484	31639	52009	18873	96927
41990	70538	77191	25860	55204	73417	83920	69468	74972	38712
72452	36618	76298	26678	89334	33938	95567	29380	75906	91807
37042	40318	57099	10528	09925	89773	41335	96244	29002	46453
53766	52875	15987	46962	67342	77592	57651	95508	80033	69828
90585	58955	53122	16025	84299	53310	67380	84249	25348	04332
32001	96293	37203	64516	51530	37069	40261	61374	05815	06714
62606	64324	46354	72157	67248	20135	49804	09226	64419	29457
10078	28073	85389	50324	14500	15562	64165	06125	71353	77669
91561	46145	24177	15294	10061	98124	75732	00815	83452	97355
13091	98112	53959	79607	52244	63303	10413	63839	74762	50289

Como los números de la tabla son aleatorios, estos números de cuatro dígitos son todos igualmente posibles. Ahora se pueden usar estos números aleatorios de cuatro dígitos para darle a cada uno de los administradores que constituyen la población la misma oportunidad de ser incluido en la muestra aleatoria. El primer número, 6327, es mayor que 2500. No corresponde a ninguno de los administradores numerados que forman la población y por tanto se descarta. El

segundo número, 1599, está entre 1 y 2500. Por tanto, el primer administrador seleccionado para la muestra aleatoria es el administrador que tiene el número

1599 en la lista de los administradores de EAI. Siguiendo este proceso, se ignoran los números 8671 y 7445 antes de identificar a los administradores con los números 1102, 1514 y 1807 e incluirlos en la muestra aleatoria. Este proceso sigue hasta que se tiene la muestra aleatoria de 30 administrativos de EAI

2. *Muestreo aleatorio sistematizado*: es un tipo de muestreo probabilístico donde se hace una selección aleatoria del primer elemento para la muestra, y luego se seleccionan los elementos posteriores utilizando intervalos o sistemáticos hasta alcanzar el tamaño de la muestra deseada.

es una implementación de la mismísima técnica de la probabilidad. Con este método de muestreo siempre hay una oportunidad igual para que cada miembro de una población sea seleccionado.

Por ejemplo, si una ONG busca formar un muestreo sistemático de 500 voluntarios de una población de 5000, puede seleccionar a cada décima persona de la población (esto es básicamente de lo que se trata el muestreo sistemático).

- Es necesario desarrollar una audiencia estructural definida para que el investigador comience a trabajar en el aspecto del muestreo.
- El investigador a cargo debe determinar el tamaño de la muestra ideal, es decir cuántas personas de la población completa se van a elegir.
- La clave para obtener resultados precisos razonables y prácticos es tener una audiencia grande.
- Una vez que se decida el número de tamaño de la muestra, se debe asignar un número a cada miembro de la muestra.
- Decide cuál será el intervalo de esta muestra. Este es básicamente la distancia estándar entre los elementos.
- El ejemplo mencionado anteriormente sugiere que el intervalo de muestra debe

ser 10, que es el resultado de la división de 5000 (N= tamaño de la población) y 500 (N = tamaño de la muestra) $(i) = N/n = 5000/500 = 10$

- El investigador debe seleccionar miembros que cumplan con este criterio.

En

este caso sería una persona de cada 10.

- Se debe elegir un número de manera aleatoria como miembro inicial (r) de la muestra, y este intervalo se agregará al número aleatorio para seguir agregando miembros r, r+i, r+2i etc. que serán elementos de la muestra.

3. *Muestreo aleatorio estratificado*: es un tipo de muestreo probabilístico que divide la población en grupos, y cada grupo tienen características similares, con una baja variación en los datos dentro del grupo, pero de un grupo a otro tiene una alta variación en los datos.

Ejemplo resuelto:

En una fábrica que consta de 600 trabajadores queremos tomar una muestra de 20. Sabemos que hay 200 trabajadores en la sección A, 150 en la B, 150 en la C y 100 en la D.

$$\frac{20}{600} = \frac{x_1}{200} \quad x_1 = 6.6 \approx 7 \text{ trabajadores de A}$$

$$\frac{20}{600} = \frac{x_2}{150} \quad x_2 = 5 \quad 5 \text{ trabajadores de B}$$

$$\frac{20}{600} = \frac{x_3}{150} \quad x_3 = 5 \quad 5 \text{ trabajadores de C}$$

$$\frac{20}{600} = \frac{x_4}{100} \quad x_4 = 3.3 \approx 3 \text{ trabajadores de D}$$

4. *Muestreo aleatorio conglomerado*: Es una técnica utilizada cuando hay agrupamientos "naturales" relativamente homogéneos en una población estadística. Existe una alta variación dentro del grupo, pero una baja variación de un grupo a otro en sus datos

El primer paso para aplicar esta técnica es definir los conglomerados: Se trata de identificar una característica que permita dividir la población en grupos disjuntos (sin superposición) y exhaustiva (todos los individuos deben estar en un grupo),

de tal manera que los grupos no difieran entre sí en relación con aquello que queremos medir.

Una vez hemos definido estos conglomerados, seleccionaremos al azar algunos de ellos para estudiarlos.

Por último, una vez hemos seleccionado los conglomerados a estudiar, podemos investigar a todos los sujetos que forman parte de estos, o bien aplicar un nuevo proceso de muestreo dentro del conglomerado.

Por ejemplo, obteniendo una muestra mediante muestreo aleatorio simple o sistemático. Si optamos por esta posibilidad, estaremos hablando de un muestreo en dos etapas o bietápico: la primera etapa es la selección del conglomerado, la segunda la de individuos dentro del conglomerado. Si por el contrario estudiamos todos los individuos del conglomerado, hablaremos de muestreo por conglomerados unietápico

1.3 Teorema del límite central.

Teorema del límite central: Si $\bar{X}$ es la media de una muestra aleatoria de tamaño n , tomada de una población con media μ y varianza finita σ^2 , entonces la forma límite de la distribución

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

a medida que $n \rightarrow \infty$, es la distribución normal estándar $n(z; 0, 1)$

La aproximación normal para $\bar{x}$ por lo general será buena si $n \geq 30$, siempre y cuando la distribución de la población no sea muy asimétrica. Si $n < 30$, la aproximación será buena sólo si la población no es muy diferente de una distribución normal y, como x antes se estableció, si se sabe que la población es normal, la distribución muestral de $\bar{x}$ seguirá siendo una distribución normal exacta, sin importar qué tan pequeño sea el tamaño de las muestras. El tamaño de la muestra $n = 30$ es un lineamiento para el teorema del límite central. Sin embargo, como indica el planteamiento del teorema, la suposición de normalidad

en la distribución de $\bar{x}$ se vuelve más precisa a medida que n se hace más grande. La figura 8.1 ilustra cómo funciona el teorema.

La figura indica cómo la distribución de $\bar{x}$ se acerca más a la normalidad a medida que aumenta n , empezando con la distribución claramente asimétrica de una observación individual ($n = 1$). También ilustra que la media de $\bar{x}$ sigue siendo μ para cualquier tamaño de la muestra y que la varianza de $\bar{x}$ se vuelve más pequeña a medida que aumenta n .

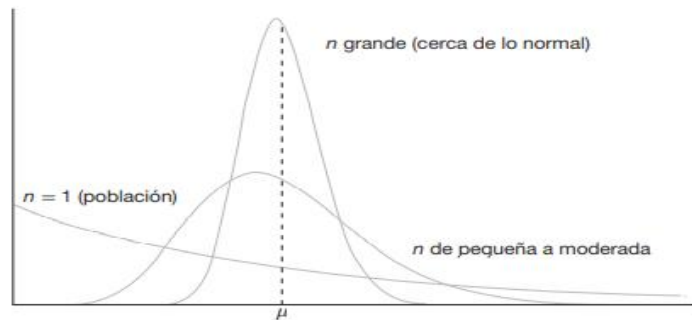


Figura 1.3 Ejemplo del teorema del límite central (distribución de $\bar{x}$ para $n = 1$, n moderada y n grande).

1.4 Distribuciones fundamentales para el muestreo.

El campo de la inferencia estadística trata básicamente con generalizaciones y predicciones. Por ejemplo, con base en las opiniones de varias personas entrevistadas en la calle, los estadounidenses podrían afirmar que en una próxima elección 60% de los votantes de la ciudad de Detroit favorecerían a cierto candidato. En este caso tratamos con una muestra aleatoria de opiniones de una población finita muy grande. Por otro lado, con base en las estimaciones de 3 contratistas seleccionados al azar, de los 30 que laboran actualmente en esta ciudad, podríamos afirmar que el costo promedio de construir una residencia en Charleston, Carolina del Sur, está entre \$330,000 y \$335,000. La población que se va a muestrear aquí también es finita, pero muy pequeña. Finalmente, consideremos una máquina despachadora de bebida gaseosa que está diseñada para servir en promedio 240 mililitros de bebida. Un ejecutivo de la empresa calcula la media de 40 bebidas servidas y obtiene $\bar{x} = 236$ mililitros y, con base en este valor, decide que la máquina está sirviendo bebidas con un contenido promedio de $\mu = 240$ mililitros. Las 40 bebidas servidas representan una muestra de la población infinita de posibles bebidas que despachará esta máquina.

Inferencias sobre la población a partir de información de la muestra En cada uno de los ejemplos anteriores calculamos un estadístico de una muestra que se selecciona de la población, y con base en tales estadísticos hicimos varias afirmaciones respecto a los valores de los parámetros de la población, que pueden ser o no ciertas. El ejecutivo de la empresa decide que la máquina despachadora está sirviendo bebidas con un contenido promedio de 240 mililitros, aunque la media de la muestra fue de 236 mililitros, porque conoce la teoría del muestreo según la cual, si $\mu = 240$ mililitros, tal valor de la muestra podría ocurrir fácilmente. De hecho, si realiza pruebas similares, cada hora, por ejemplo, esperaría que los valores del estadístico $\bar{x}$ fluctuaran por arriba y por abajo de $\mu = 240$ mililitros. Sólo cuando el valor de $\bar{x}$ difiera considerablemente de 240 mililitros el ejecutivo de la empresa tomará medidas para ajustar la máquina. Como un estadístico es una variable aleatoria que depende sólo de la muestra observada, debe tener una distribución de probabilidad.

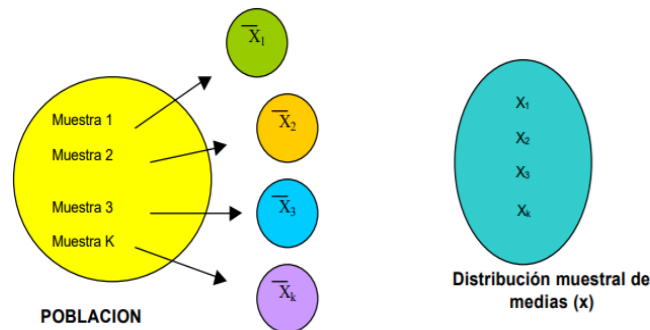
La distribución muestral de un estadístico depende de la distribución de la población, del tamaño de las muestras y del método de selección de las muestras. En lo que resta de este capítulo estudiaremos varias de las distribuciones muestrales más importantes de los estadísticos que se utilizan con frecuencia. Las aplicaciones de tales distribuciones muestrales a problemas de inferencia estadística se consideran en la mayoría de los capítulos posteriores. La distribución de probabilidad de $\bar{x}$ se llama distribución muestral de la media

La distribución de probabilidad de un estadístico se denomina **distribución muestral**

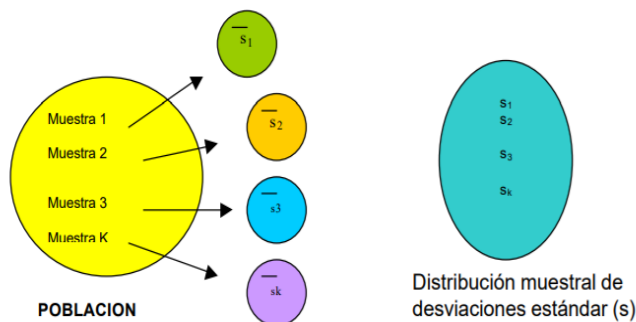
Distribuciones Muestrales Las muestras aleatorias obtenidas de una población son, por naturaleza propia, impredecibles. No se esperaría que dos muestras aleatorias del mismo tamaño y tomadas de la misma población tenga la misma media muestral o que sean completamente parecidas; puede esperarse que cualquier estadístico, como la media muestral, calculado a partir de las medias en una muestra aleatoria, cambie su valor de una muestra a otra, por ello, se quiere estudiar la distribución de todos los valores posibles de un estadístico. Tales distribuciones serán muy importantes en el estudio de la estadística inferencial, porque las inferencias sobre las poblaciones se harán usando estadísticas muestrales. Como el análisis de las distribuciones asociadas con los estadísticos muestrales, podremos juzgar la confiabilidad de un estadístico muestral como un instrumento para hacer inferencias sobre un parámetro poblacional desconocido. Como los valores de un estadístico, tal como x , varían de una muestra aleatoria a otra, se le puede considerar como una variable aleatoria con su correspondiente

distribución de frecuencias. La distribución de frecuencia de un estadístico muestral se denomina distribución muestral. En general, la distribución muestral de un estadístico es la de todos sus valores posibles calculados a partir de muestras del mismo tamaño.

Suponga que se han seleccionado muestras aleatorias de tamaño 20 en una población grande. Se calcula la media muestral $\bar{x}$ para cada muestra; la colección de todas estas medias muestrales recibe el nombre de distribución muestral de medias, lo que se puede ilustrar en la siguiente figura:



Suponga que se eligen muestras aleatorias de tamaño 20, de una población grande, y se calcula la desviación estándar de cada una. La colección de todas estas desviaciones estándar muestrales se llama distribución muestral de la desviación estándar, y lo podemos ver en la siguiente figura:

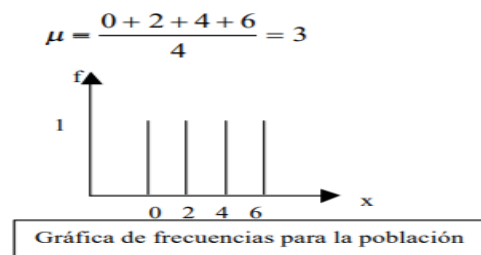


Ejemplo resuelto:

Se eligen muestras ordenadas de tamaño 2, con reemplazo, de la población de valores 0, 2, 4 y 6. Encuentre:

- a) μ , la media poblacional.
- b) σ , la desviación estándar poblacional.
- c) μ_x , la media de la distribución muestral de medias.
- d) σ_x , la desviación estándar de la distribución muestral de medias.
- e) Además, grafique las frecuencias para la población y para la distribución muestral de medias.

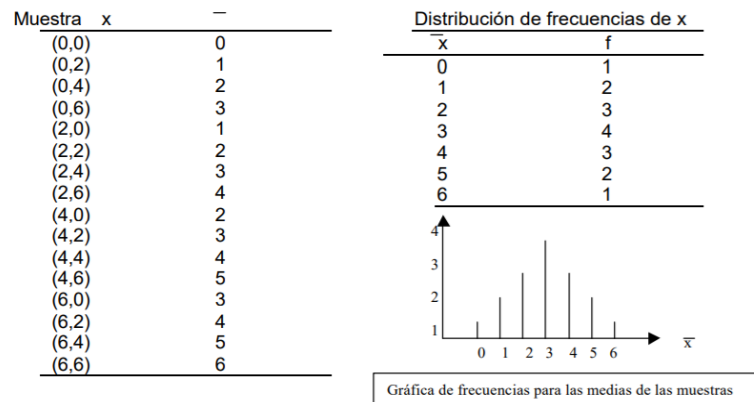
- a) La media poblacional es:



- b) La desviación estándar de la población es:

$$\sigma = \sqrt{\frac{(0-3)^2 + (2-3)^2 + (4-3)^2 + (6-3)^2}{4}} = 2.236$$

- c) A continuación, se listan los elementos de la distribución muestral de la media y la correspondiente distribución de frecuencias.



La media de la distribución muestral de medias es:

$$\mu_{\bar{x}} = \frac{\sum(fx)}{\sum f} = \frac{(0)(1)+(1)(2)+(2)(3)+(3)(4)+(4)(3)+(5)(2)+(6)(1)}{16} = \frac{48}{3} = 3$$

d) La desviación estándar de la distribución muestral de medias es:

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum(\bar{x} - \mu_{\bar{x}})^2 f}{\sum f}}$$

$$\sigma_{\bar{x}} = \sqrt{\frac{(0-3)^2 1 + (1-3)^2 2 + (2-3)^2 3 + (3-3)^2 4 + (4-3)^2 3 + (5-3)^2 2 + (6-3)^2 1}{16}} = 1.58$$

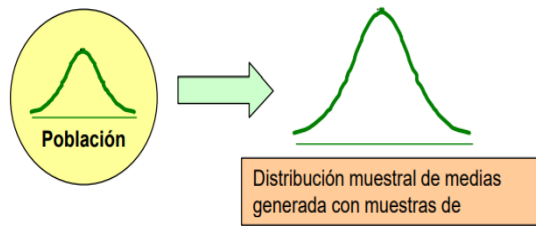
Se deduce:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.236}{2} = 1.58$$

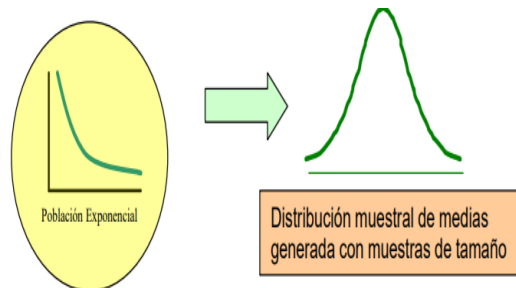
Como para cualquier variable aleatoria, la distribución muestral de medias tiene una media o valor esperado, una varianza y una desviación estándar, se puede demostrar que la distribución muestral de medias tiene una media igual a la media poblacional. Esto es:

$$\mu_{\bar{x}} = E(\bar{x}) = \mu = 3$$

Distribuciones muestrales Después de haber realizado el ejercicio anterior se puede ver que una distribución muestral se genera extrayendo todas las posibles muestras del mismo tamaño de la población y calculándoles a éstas su estadístico. Si la población de la que se extraen las muestras es normal, la distribución muestral de medias será normal sin importar el tamaño de la muestra.



Si la población de donde se extraen las muestras no es normal, entonces el tamaño de la muestra debe ser mayor o igual a 30, para que la distribución muestral tenga una forma acampanada. Mientras mayor sea el tamaño de la muestra, más cerca estará la distribución muestral de ser normal. Para muchos propósitos, la aproximación normal se considera buena si se cumple $n=30$. La forma de la distribución muestral de medias sea aproximadamente normal, aún en casos donde la población original es bimodal, es realmente notable.



La desviación estándar de la distribución muestral de un estadístico se conoce como error estándar del estadístico. Para el ejercicio anterior el error estándar de la media denotado por σ_x , es 1.58. Con esto se puede demostrar que, si de una población se eligen muestras de tamaño n con reemplazo, entonces el error estándar de la media es igual a la desviación estándar de la distribución de los errores muestrales.

En general se tiene: $\sigma_x = \sigma_e$

Cuando las muestras se toman de una población pequeña y sin reemplazo, se puede usar la fórmula siguiente para encontrar σ_x :

$$\sigma_x = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

donde σ es la desviación estándar de la población de donde se toman las muestras, n es el tamaño de la muestra y N el de la población.

Como regla de cálculo, si el muestreo se hace sin reemplazo y el tamaño de la población es al menos 20 veces el tamaño de la muestra ($N \geq 20$), entonces se puede usar la fórmula.

$$\sqrt{\frac{N-n}{N-1}}$$

Este factor se denomina factor de corrección para una población finita.

Ejemplo resuelto:

Suponga que la tabla siguiente muestra la antigüedad en años en el trabajo de tres maestros universitarios de matemáticas:

Maestro de matemáticas	Antigüedad
A	6
B	4
C	2

Suponga además que se seleccionan muestras aleatorias de tamaño 2 sin reemplazo. Calcule la antigüedad media para cada muestra, la media de la distribución muestral y el error estándar, o la desviación estándar de la distribución muestral.

Solución:

Se pueden tener ${}_3C_2 = 3$ muestras posibles. La tabla lista todas las muestras posibles de tamaño 2, con sus respectivas medias muestrales.

Muestras	Antigüedad	Media muestral
A, B	(6,4)	5
A, C	(6,2)	4
B, C	(4,2)	3

La media poblacional es:

$$\mu = \frac{2+4+6}{3} = 4$$

La media de la distribución muestral es:

$$\mu_{\bar{x}} = \frac{5+4+3}{3} = 4$$

La desviación estándar de la población es:

$$\sigma = \sqrt{\frac{(6-4)^2 + (4-4)^2 + (2-4)^2}{3}} = 1.63$$

El error estándar o la desviación estándar de la distribución muestral es:

$$\sigma_{\bar{x}} = \sqrt{\frac{(5-4)^2 + (4-4)^2 + (3-4)^2}{3}} = 0.816$$

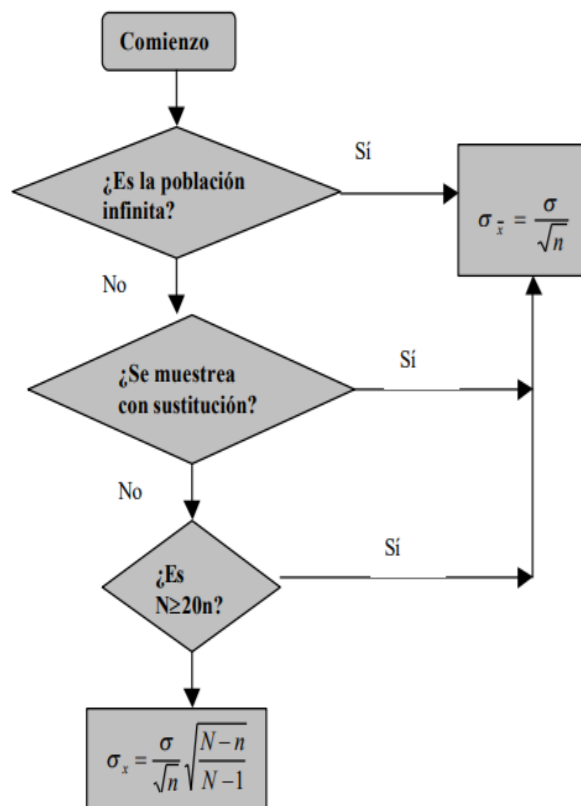
Si utilizamos la fórmula del error estándar sin el factor de corrección tendríamos que:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.63}{\sqrt{2}} = 1.152$$

Por lo que observamos que este valor no es el verdadero. Agregando el factor de corrección obtendremos el valor correcto:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{1.63}{\sqrt{2}} \sqrt{\frac{3-2}{3-1}} = 0.816$$

El diagrama de flujo resume las decisiones que deben tomarse cuando se calcula el valor del error estándar:



1.4.1 Distribución muestral de la media.

Si recordamos a la distribución normal, esta es una distribución continua, en forma de campana en donde la media, la mediana y la moda tienen un mismo valor y es simétrica. Con esta distribución podíamos calcular la probabilidad de algún evento relacionado con la variable aleatoria, mediante la siguiente fórmula:

$$z = \frac{x - \mu}{\sigma}$$

En donde z es una variable estandarizada con media igual a cero y varianza igual a uno. Con esta fórmula se pueden hacer los cálculos de probabilidad para cualquier ejercicio, utilizando la tabla de la distribución z .

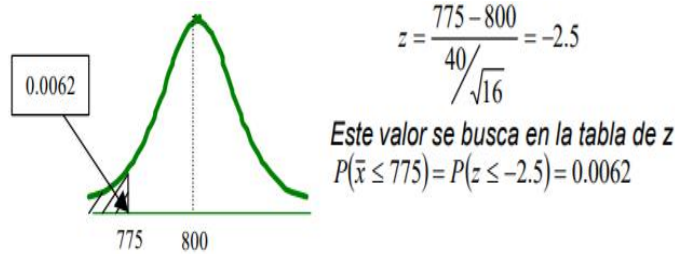
Sabemos que cuando se extraen muestras de tamaño mayor a 30 o bien de cualquier tamaño de una población normal, la distribución muestral de medias tiene un comportamiento aproximadamente normal, por lo que se puede utilizar la fórmula de la distribución normal con μ y μ_x , entonces la fórmula para calcular la probabilidad del comportamiento del estadístico, en este caso la media de la muestra, quedaría de la siguiente manera:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Ejemplo resuelto:

Una empresa eléctrica fabrica focos que tienen una duración que se distribuye aproximadamente en forma normal, con media de 800 horas y desviación estándar de 40 horas. Encuentre la probabilidad de que una muestra aleatoria de 16 focos tenga una vida promedio de menos de 775 horas.

Solución:



La interpretación sería que la probabilidad de que la media de la muestra de 16 focos sea menor a 775 horas es de 0.0062.

Ejemplo resuelto:

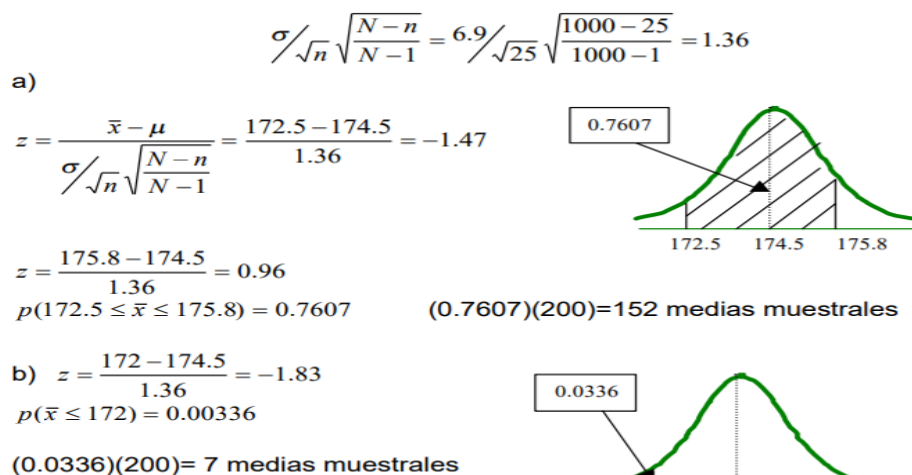
Las estaturas de 1000 estudiantes están distribuidas aproximadamente en forma normal con una media de 174.5 centímetros y una desviación estándar de 6.9 centímetros. Si se extraen 200 muestras aleatorias de tamaño 25 sin reemplazo de esta población, determine:

a) El número de las medias muestrales que caen entre 172.5 y 175.8 centímetros.

b) El número de medias muestrales que caen por debajo de 172 centímetros.

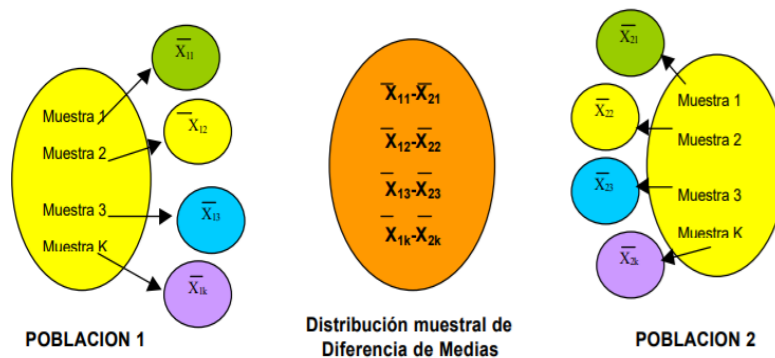
Solución: Como se puede observar en este ejercicio se cuenta con una población finita y un muestreo sin reemplazo, por lo que se tendrá que agregar el factor de corrección. Se procederá a calcular el denominador de Z para sólo sustituirlo en cada inciso.

Solución:



1.4.2 Distribución muestral de la diferencia de medias.

Suponga que se tienen dos poblaciones distintas, la primera con media μ_1 y desviación estándar σ_1 , y la segunda con media μ_2 y desviación estándar σ_2 . Más aún, se elige una muestra aleatoria de tamaño n_1 de la primera población y una muestra independiente aleatoria de tamaño n_2 de la segunda población; se calcula la media muestral para cada muestra y la diferencia entre dichas medias. La colección de todas esas diferencias se llama distribución muestral de las diferencias entre medias o la distribución muestral del estadístico $\bar{x}_1 - \bar{x}_2$.



La distribución es aproximadamente normal para $n_1 \geq 30$ y $n_2 \geq 30$. Si las poblaciones son normales, entonces la distribución muestral de medias es normal sin importar los tamaños de las muestras.

En ejercicios anteriores se había demostrado que $\mu = \mu_{\bar{x}}$ y que $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, por lo que no es difícil deducir que:

$$\mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2 \text{ y que } \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

La fórmula que se utilizará para el cálculo de probabilidad del estadístico de diferencia de medias es:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Ejemplo resuelto:

En un estudio para comparar los pesos promedio de niños y niñas de sexto grado en una escuela primaria se usará una muestra aleatoria de 20 niños y otra de 25 niñas. Se sabe que tanto para niños como para niñas los pesos siguen una distribución normal. El promedio de los pesos de todos los niños de sexto grado de esa escuela es de 100 libras y su desviación estándar es de 14.142, mientras que el promedio de los pesos de todas las niñas del sexto grado de esa escuela es de 85 libras y su desviación estándar es de 12.247 libras. Si $\bar{x}_1$ representa el promedio de los pesos de 20 niños y $\bar{x}_2$ es el promedio de los pesos de una muestra de 25 niñas, encuentre la probabilidad de que el promedio de los pesos de los 20 niños sea al menos 20 libras más grande que el de las 25 niñas.

Solución:

Datos:

$$\mu_1 = 100 \text{ libras}$$

$$\mu_2 = 85 \text{ libras}$$

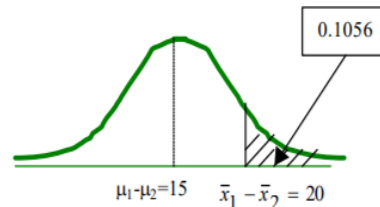
$$\sigma_1 = 14.142 \text{ libras}$$

$$\sigma_2 = 12.247 \text{ libras}$$

$$n_1 = 20 \text{ niños}$$

$$n_2 = 25 \text{ niñas}$$

$$p(\bar{x}_1 - \bar{x}_2 > 20) = ?$$



$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{20 - (100 - 85)}{\sqrt{\frac{(14.142)^2}{20} + \frac{(12.247)^2}{25}}} = 1.25$$

Por lo tanto, la probabilidad de que el promedio de los pesos de la muestra de niños sea al menos 20 libras más grande que el de la muestra de las niñas es 0.1056.

Ejemplo resuelto:

Uno de los principales fabricantes de televisores compra los tubos de rayos catódicos a dos compañías. Los tubos de la compañía A tienen una vida media de 7.2 años con una desviación estándar de 0.8 años, mientras que los de la B tienen una vida media de 6.7 años con una desviación estándar de 0.7. Determine la probabilidad de que una muestra aleatoria de 34 tubos de la compañía A tenga una vida promedio de al menos un año más que la de una muestra aleatoria de 40 tubos de la compañía B.

Solución:

Datos:

$$\mu_A = 7.2 \text{ años}$$

$$\mu_B = 6.7 \text{ años}$$

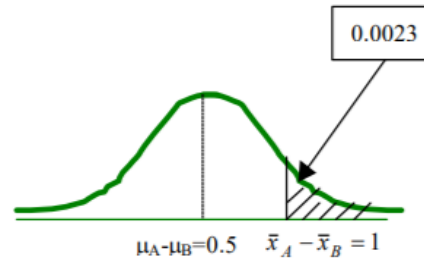
$$\sigma_A = 0.8 \text{ años}$$

$$\sigma_B = 0.7 \text{ años}$$

$$n_A = 34 \text{ tubos}$$

$$n_B = 40 \text{ tubos}$$

$$p(\bar{x}_A - \bar{x}_B > 1) = ?$$



$$z = \frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} = \frac{1 - (7.2 - 6.7)}{\sqrt{\frac{(0.8)^2}{34} + \frac{(0.7)^2}{40}}} = 2.84$$

Ejemplo resuelto:

Se prueba el rendimiento en km/L de 2 tipos de gasolina, encontrándose una desviación estándar de 1.23km/L para la primera gasolina y una desviación estándar de 1.37km/L para la segunda gasolina; se prueba la primera gasolina en 35 autos y la segunda en 42 autos.

a) ¿Cuál es la probabilidad de que la primera gasolina de un rendimiento promedio mayor de 0?45km/L que la segunda gasolina?

b) ¿Cuál es la probabilidad de que la diferencia en rendimientos promedio se encuentre entre 0?65 y 0.83km/L a favor de la gasolina 1?

Solución: En este ejercicio no se cuenta con los parámetros de las medias en ninguna de las dos poblaciones, por lo que se supondrán que son iguales.

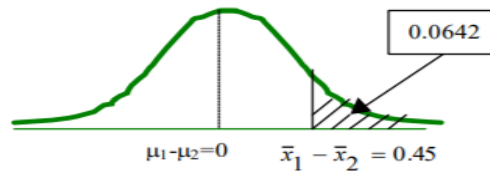
Datos:

$$\sigma_1 = 1.23 \text{ Km/Lto}$$

$$\sigma_2 = 1.37 \text{ Km/Lto}$$

$$n_1 = 35 \text{ autos}$$

$$n_2 = 42 \text{ autos}$$

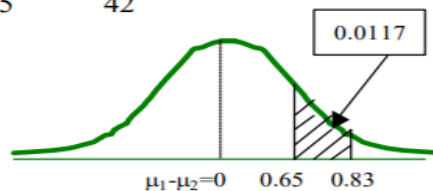


$$\text{a) } p(\bar{x}_1 - \bar{x}_2 > 0.45) = ? \quad z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.45 - 0}{\sqrt{\frac{(1.23)^2}{35} + \frac{(1.37)^2}{42}}} = 1.52$$

$$\text{b) } p(0.65 \leq \bar{x}_1 - \bar{x}_2 \leq 0.83) = ?$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.65 - 0}{\sqrt{\frac{(1.23)^2}{35} + \frac{(1.37)^2}{42}}} = 2.19$$

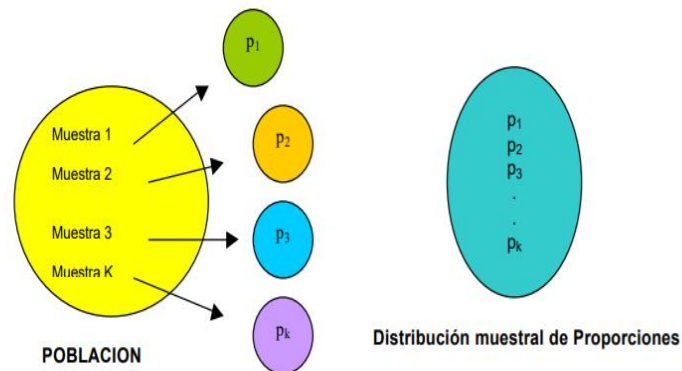
$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.83 - 0}{\sqrt{\frac{(1.23)^2}{35} + \frac{(1.37)^2}{42}}} = 2.80$$



La probabilidad de que la diferencia en rendimientos promedio en las muestras se encuentre entre 0.65 y 0.83 Km/Lto. a favor de la gasolina 1 es de 0.0117.

1.4.3 Distribución muestral de la proporción.

Distribución muestral de Proporciones Existen ocasiones en las cuales no estamos interesados en la media de la muestra, sino que queremos investigar la proporción de artículos defectuosos o la proporción de alumnos reprobados en la muestra. La distribución muestral de proporciones es la adecuada para dar respuesta a estas situaciones. Esta distribución se genera de igual manera que la distribución muestral de medias, a excepción de que al extraer las muestras de la población se calcula el estadístico proporción ($p=x/n$ en donde “x” es el número de éxitos u observaciones de interés y “n” el tamaño de la muestra) en lugar del estadístico media.



Una población binomial está estrechamente relacionada con la distribución muestral de proporciones; una población binomial es una colección de éxitos y fracasos, mientras que una distribución muestral de proporciones contiene las posibilidades o proporciones de todos los números posibles de éxitos en un experimento binomial, y como consecuencia de esta relación, las afirmaciones probabilísticas referentes a la proporción muestral pueden evaluarse usando la aproximación normal a la binomial, siempre que $np \geq 5$ y $n(1-p) \geq 5$. Cualquier evento se puede convertir en una proporción si se divide el número obtenido entre el número de intentos.

Ejemplo resuelto:

Suponga que se cuenta con un lote de 12 piezas, el cual tiene 4 artículos defectuosos. Se van a seleccionar 5 artículos al azar de ese lote sin reemplazo. Genere la distribución muestral de proporciones para el número de piezas defectuosas. Como se puede observar en este ejercicio la Proporción de artículos defectuosos de esta población es $4/12=1/3$. Por lo que podemos decir que el 33% de las piezas de este lote están defectuosas. El número posible de muestras de tamaño 5 a extraer de una población de 12 elementos es $12C5=792$, las cuales se pueden desglosar de la siguiente manera:

ARTICULOS BUENOS	ARTICULOS MALOS	PROPORCION DE ARTICULOS DEFECTUOSOS	NUMERO DE MANERAS EN LAS QUE SE PUEDE OBTENER LA MUESTRA
1	4	4/5= 0.8	8C1*4C4=8
2	3	3/5= 0.6	8C2*4C3=112
3	2	2/5= 0.4	8C3*4C2=336
4	1	1/5= 0.2	8C4*4C1=280
5	0	0/5= 0	8C5*4C0=56

TOTAL= 792

Para calcular la media de la distribución muestral de proporciones se tendría que hacer la sumatoria de la frecuencia por el valor de la proporción muestral y dividirla entre el número total de muestras. Esto es:

$$\mu_p = \frac{(.8)(8) + (.6)(112) + (.4)(336) + (.2)(280) + (0)(56)}{792} = \frac{1}{3} = 0.3333$$

Como podemos observar la media de la distribución muestral de proporciones es igual a la Proporción de la población.

$$\mu_p = P$$

También se puede calcular la desviación estándar de la distribución muestral de proporciones:

$$\sigma_p = \sqrt{\frac{(0.8-1/3)^2 * 8 + (0.6-1/3)^2 * 112 + (0.4-1/3)^2 * 336 + (0.2-1/3)^2 * 280 + (0-1/3)^2 * 56}{792}} = 0.1681$$

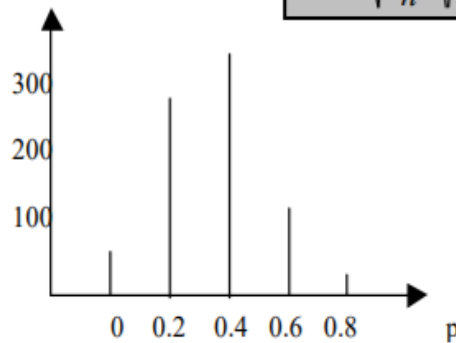
La varianza de la distribución binomial es $\sigma^2 = npq$, por lo que la varianza de la distribución muestral de proporciones es $\sigma^2_p = (Pq)/n$. Si se sustituyen los valores en esta fórmula tenemos que:

$$\sigma_p = \sqrt{\frac{(1/3)(2/3)}{5}} = 0.2108$$

este valor no coincide con el de 0.1681, ya que nos falta agregar el factor de corrección para una población finita y un muestreo sin reemplazo:

$$\sigma_p = \sqrt{\frac{(1/3)(2/3)}{5}} \sqrt{\frac{12-5}{12-1}} = 0.1681$$

$$\sigma_p = \sqrt{\frac{Pq}{n}} \sqrt{\frac{N-n}{N-1}}$$



Gráfica de frecuencias para las proporciones de las muestras

La fórmula que se utilizará para el cálculo de probabilidad en una distribución muestral de proporciones está basada en la aproximación de la distribución normal a la binomial. Esta fórmula nos servirá para calcular la probabilidad del comportamiento de la proporción en la muestra.

$$z = \frac{p - P}{\sqrt{\frac{Pq}{n}}}$$

A esta fórmula se le puede agregar el factor de corrección de $\sqrt{N - n/N - 1}$ si se cumple con las condiciones necesarias.

Ejemplo resuelto:

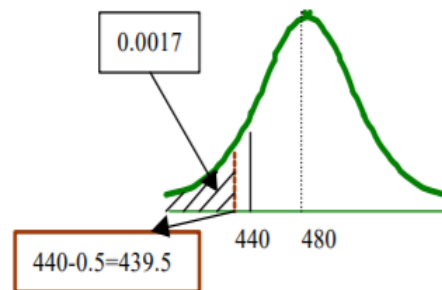
Se ha determinado que 60% de los estudiantes de una universidad grande fuman cigarrillos. Se toma una muestra aleatoria de 800 estudiantes. Calcule la

probabilidad de que la proporción de la muestra de la gente que fuma cigarrillos sea menor que 0.55. Solución: Este ejercicio se puede solucionar por dos métodos. El primero puede ser con la aproximación de la distribución normal a la binomial y el segundo utilizando la fórmula de la distribución muestral de proporciones.

Aproximación de la distribución normal a la binomial:

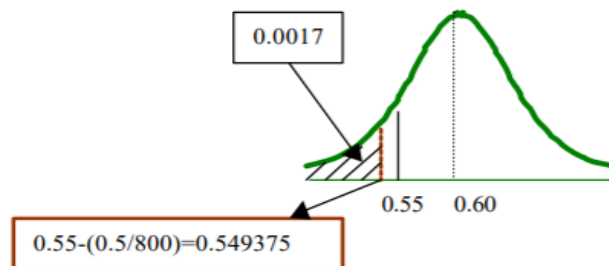
Datos:
n=800 estudiantes
p=0.60
x= (.55)(800) = 440 estudiantes
p(x<440) = ?
Media= np= (800)(0.60)= 480

$$z = \frac{x - np}{\sqrt{npq}} = \frac{439.5 - 480}{\sqrt{800(0.60)(0.40)}} = -2.92$$



P (<440)=0.0017. Este valor significa que existe una probabilidad de 0.17% de que, al extraer una muestra de 800 estudiantes, menos de 440 fuman cigarrillos.

Datos:
n=800 estudiantes
P=0.60
p= 0.55
p(p<0.55) = ?



$$z = \frac{p - P}{\sqrt{\frac{Pq}{n}}} = \frac{0.549375 - 0.60}{\sqrt{\frac{(0.60)(0.40)}{800}}} = -2.92 \quad \text{Observe que este valor es igual al obtenido}$$

En el método de la aproximación de la distribución de la normal a la binomial, por lo que si lo buscamos en la tabla Z nos da la misma probabilidad de 0.0017.

También se debe tomar en cuenta que el factor de corrección de 0.5 se está dividiendo entre el tamaño de la muestra, ya que estamos hablando de una proporción. La interpretación en esta solución estaría enfocada a la proporción de la muestra, por lo que diríamos que la probabilidad de que, al extraer una muestra de 800 estudiantes de esa universidad, la proporción de estudiantes que fuman cigarrillos sea menor al 55% es del 0.17%.

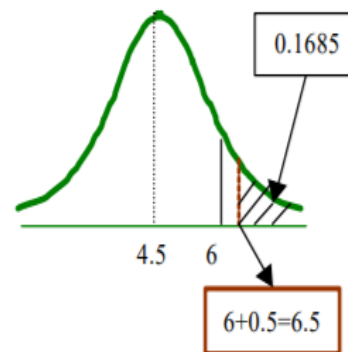
Ejemplo resuelto:

Un medicamento para malestar estomacal tiene la advertencia de que algunos usuarios pueden presentar una reacción adversa a él, más aún, se piensa que alrededor del 3% de los usuarios tienen tal reacción. Si una muestra aleatoria de 150 personas con malestar estomacal usa el medicamento, encuentre la probabilidad de que la proporción de la muestra de los usuarios que realmente presentan una reacción adversa exceda el 4%. a) Resolverlo mediante la aproximación de la normal a la binomial b) Resolverlo con la distribución muestral de proporciones.

a) Aproximación de la distribución normal a la binomial:

Datos:
 $n=150$ personas
 $p=0.03$
 $x = (0.04)(150) = 6$ personas
 $p(x>6) = ?$
 $\text{Media} = np = (150)(0.03) = 4.5$

$$z = \frac{x - np}{\sqrt{npq}} = \frac{6.5 - 4.5}{\sqrt{150(0.03)(0.97)}} = 0.96$$



$p(x>6) = 0.1685$. Este valor significa que existe una probabilidad del 17% de que al extraer una muestra de 150 personas, más de 6 presentarán una reacción adversa.

b) Distribución Muestral de Proporciones

b) Distribución Muestral de Proporciones

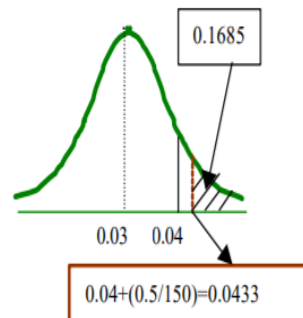
Datos:

$n=150$ personas

$P=0.03$

$p=0.04$

$p(p>0.04) = ?$



$$z = \frac{p - P}{\sqrt{\frac{Pq}{n}}} = \frac{0.0433 - 0.03}{\sqrt{\frac{(0.03)(0.97)}{150}}} = 0.96$$

Obsérvese que el valor es igual al obtenido y la interpretación es: existe una probabilidad del 17% de que al tomar una muestra de 150 personas se tenga una proporción mayor de 0.04 presentando una reacción adversa.

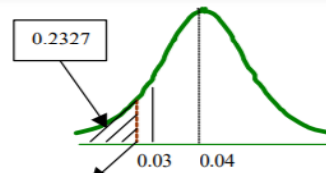
Ejemplo resuelto:

Se sabe que la verdadera proporción de los componentes defectuosos fabricadas por una firma es de 4%, y encuentre la probabilidad de que una muestra aleatoria de tamaño 60 tenga:

a) Menos del 3% de los componentes defectuosos.

b) Más del 1% pero menos del 5% de partes defectuosas.

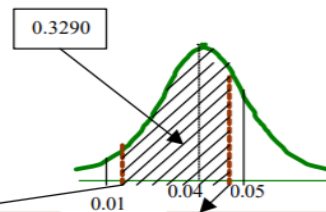
a) Datos:
n= 60 artículos
P=0.04
p= 0.03
p(p<0.03) = ?



$$0.03 - (0.5/60) = 0.0216$$

$z = \frac{p - P}{\sqrt{\frac{Pq}{n}}} = \frac{0.0216 - 0.04}{\sqrt{\frac{(0.04)(0.96)}{60}}} = -0.73$ La probabilidad de que en una muestra de 60 artículos exista una proporción menor de 0.03 artículos defectuosos es de 0.2327.

b) Datos:
n= 60 artículos
P=0.04
p= 0.01 y 0.05
p(0.01<p<0.05) = ?



$$0.01 - (0.5/60) = 0.0183$$

$$0.05 - (0.5/60) = 0.0416$$

$$z = \frac{p - P}{\sqrt{\frac{Pq}{n}}} = \frac{0.0183 - 0.04}{\sqrt{\frac{(0.04)(0.96)}{60}}} = -0.86 \quad z = \frac{p - P}{\sqrt{\frac{Pq}{n}}} = \frac{0.0416 - 0.04}{\sqrt{\frac{(0.04)(0.96)}{60}}} = 0.06$$

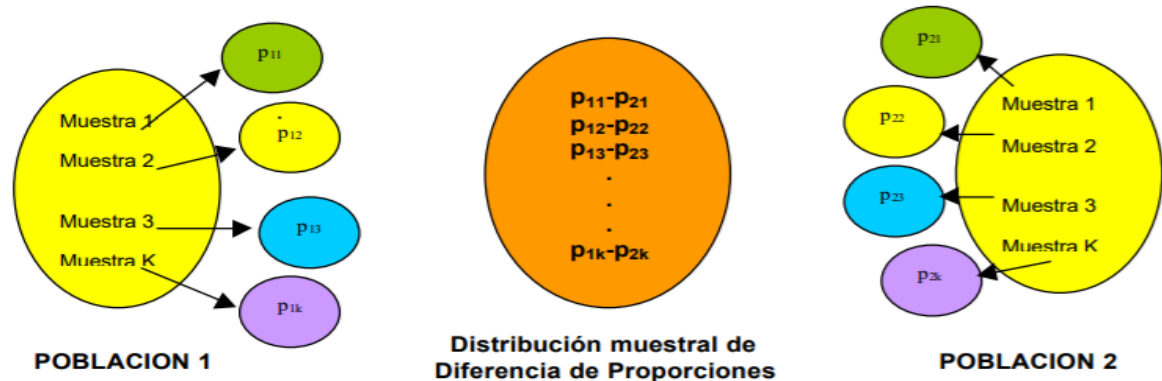
1.4.4 Distribución muestral de la diferencia de proporciones.

Muchas aplicaciones involucran poblaciones de datos cualitativos que deben compararse utilizando proporciones o porcentajes. A continuación, se citan algunos ejemplos:

- Educación. - ¿Es mayor la proporción de los estudiantes que aprueban matemáticas que las de los que aprueban inglés?
- Medicina. - ¿Es menor el porcentaje de los usuarios del medicamento A que presentan una reacción adversa que el de los usuarios del fármaco B que también presentan una reacción de ese tipo?
- Administración. - ¿Hay diferencia entre los porcentajes de hombres mujeres en posiciones gerenciales?

- Ingeniería. - ¿Existe diferencia entre la proporción de artículos defectuosos que genera la máquina A a los que genera la máquina B?

Cuando el muestreo procede de dos poblaciones binomiales y se trabaja con dos proporciones muestrales, la distribución muestral de diferencia de proporciones es aproximadamente normal para tamaños de muestra grande ($n_1p_1 \geq 5$, $n_1q_1 \geq 5$, $n_2p_2 \geq 5$ y $n_2q_2 \geq 5$). Entonces p_1 y p_2 tienen distribuciones muestrales aproximadamente normales, así que su diferencia $p_1 - p_2$ también tiene una distribución muestral aproximadamente normal.



Cuando se estudió a la distribución muestral de proporciones se comprobó que $P = \mu_p$ y que $\sigma_p = \sqrt{\frac{Pq}{n}}$, por lo que no es difícil deducir que $\mu_{p_1} - \mu_{p_2} = P_1 - P_2$ y

que $\sigma_{p_1 - p_2} = \sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}$.

La fórmula que se utilizará para el cálculo de probabilidad del estadístico de diferencia de proporciones es:

$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}}$$

Ejemplo resuelto:

Los hombres y mujeres adultos radicados en una ciudad grande del norte difieren en sus opiniones sobre la promulgación de la pena de muerte para personas culpables de asesinato. Se cree que el 12% de los hombres adultos están a favor de la pena de muerte, mientras que sólo 10% de las mujeres adultas lo están. Si se pregunta a dos muestras aleatorias de 100 hombres y 100 mujeres su opinión sobre la promulgación de la pena de muerte, determine la probabilidad de que el porcentaje de hombres a favor sea al menos 3% mayor que el de las mujeres.

Solución:

Datos:

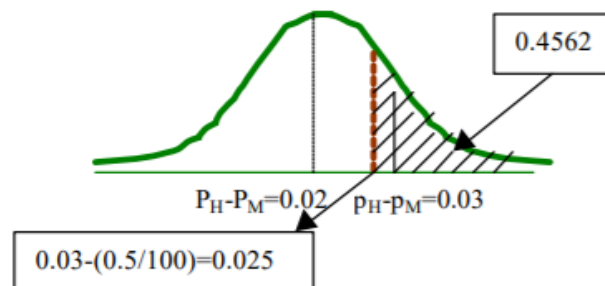
$$P_H = 0.12$$

$$P_M = 0.10$$

$$n_H = 100$$

$$n_M = 100$$

$$p(p_H - p_M \geq 0.03) = ?$$



Se recuerda que se está incluyendo el factor de corrección de 0.5 por ser una distribución binomial y se está utilizando la distribución normal.

$$z = \frac{(p_H - p_M) - (P_H - P_M)}{\sqrt{\frac{P_H q_H}{n_H} + \frac{P_M q_M}{n_M}}} = \frac{0.025 - (0.12 - 0.10)}{\sqrt{\frac{(0.12)(0.88)}{100} + \frac{(0.10)(0.90)}{100}}} = 0.11$$

Se concluye que la probabilidad de que el porcentaje de hombres a favor de la pena de muerte, al menos 3% mayor que el de mujeres es de 0.4562.

Ejemplo resuelto:

Una encuesta del Boston College constó de 320 trabajadores de Michigan que fueron despedidos entre 1979 y 1984, encontró que 20% habían estado sin trabajo durante por lo menos dos años. Supóngase que tuviera que seleccionar otra muestra aleatoria de 320 trabajadores de entre todos los empleados despedidos

entre 1979 y 1984. ¿Cuál sería la probabilidad de que su porcentaje muestral de trabajadores sin empleo durante por lo menos dos años difiera del porcentaje obtenido en la encuesta de Boston College, en 5% o más?

Solución:

En este ejercicio se cuenta únicamente con una población, de la cual se están extrayendo dos muestras y se quiere saber la probabilidad de la diferencia de los porcentajes en esas dos muestras, por lo que se debe de utilizar la distribución muestral de proporciones con $P_1 = P_2$, ya que es una misma población. Otra de las situaciones con la cual nos topamos es que desconocemos la proporción de trabajadores despedidos entre 1979 y 1984 que estuvieron desempleados por un período de por lo menos dos años, sólo se conoce la $p_1 = 0.20$ ya que al tomar una muestra de 320 trabajadores se observó esa proporción. En la fórmula de la distribución muestral de proporciones para el cálculo de probabilidad se necesita saber las proporciones de las poblaciones, las cuales en este ejercicio las desconocemos, por lo que se utilizará el valor de 0.20 como una estimación puntual de P .

También debe de comprenderse la pregunta que nos hace este problema, ¿cuál sería la probabilidad de que su porcentaje muestral de trabajadores sin empleo durante por lo menos dos años **difiera** del porcentaje obtenido en la encuesta de Boston College, en 5% o más?, la palabra **difiera** quiere decir que puede existir una diferencia a favor de la muestra uno, o a favor de la muestra dos, por lo que se tendrán que calcular dos áreas en la distribución y al final sumarlas.

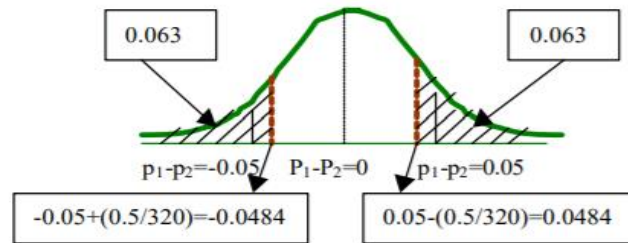
Datos:

$$p_1 = 0.20$$

$$n_1 = 320 \text{ trabajadores}$$

$$n_2 = 320 \text{ trabajadores}$$

$$P_1 = P_2$$



$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{0.0484 - 0}{\sqrt{\frac{(0.20)(0.80)}{320} + \frac{(0.20)(0.80)}{320}}} = 1.53$$

$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{-0.0484 - 0}{\sqrt{\frac{(0.20)(0.80)}{320} + \frac{(0.20)(0.80)}{320}}} = -1.53$$

$$p(-0.05 \geq p_1 - p_2 \geq 0.05) = 0.063 + 0.063 = 0.1260$$

La probabilidad de que su proporción muestral de trabajadores sin empleo durante por lo menos dos años difiera del porcentaje obtenido en la encuesta de Boston College, en 0.05 o más es de 0.1260.

Ejemplo resuelto:

Se sabe que 3 de cada 6 productos fabricados por la máquina 1 son defectuosos y que 2 de cada 5 objetos fabricados por la máquina 2 son defectuosos; se toman muestras de 120 objetos de cada máquina: a) ¿cuál es la probabilidad de que la proporción de artículos defectuosos de la máquina 2 rebase a la máquina 1 en por lo menos 0.10? b) ¿cuál es la probabilidad de que la proporción de artículos defectuosos de la máquina 1 rebase a la máquina 2 en por lo menos 0.15?

Solución:

Datos:

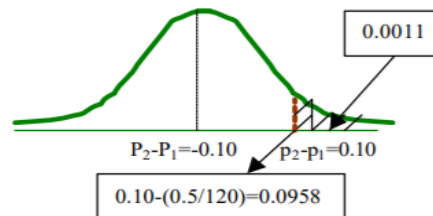
$$P_1 = 3/6 = 0.5$$

$$P_2 = 2/5 = 0.4$$

$$n_1 = 120 \text{ objetos}$$

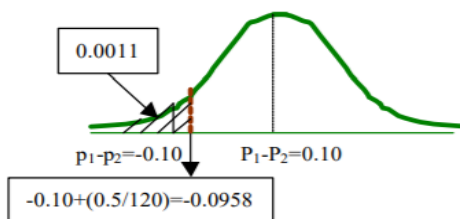
$$n_2 = 120 \text{ objetos}$$

a) $p(p_2 - p_1 \geq 0.10) = ?$



$$z = \frac{(p_2 - p_1) - (P_2 - P_1)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{0.0958 - (-0.10)}{\sqrt{\frac{(0.50)(0.50)}{120} + \frac{(0.40)(0.60)}{120}}} = 3.06$$

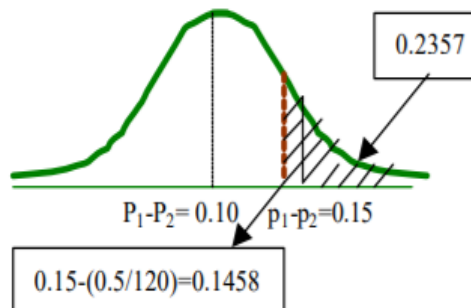
Otra manera de hacer este ejercicio es poner $P_1 - P_2$:



$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{-0.0958 - 0.10}{\sqrt{\frac{(0.50)(0.50)}{120} + \frac{(0.40)(0.60)}{120}}} = -3.06$$

La probabilidad de que exista una diferencia de proporciones de artículos defectuosos de por lo menos 10% a favor de la máquina 2 es de 0.0011.

b) $p(p_1 - p_2 \geq 0.15) = ?$



$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{0.1458 - 0.10}{\sqrt{\frac{(0.50)(0.50)}{120} + \frac{(0.40)(0.60)}{120}}} = 0.72$$

La probabilidad de que exista una diferencia de proporciones de artículos defectuosos de por lo menos 15% a favor de la máquina 1 es de 0.2357.

1.4.5 Distribución t-student.

La distribución t de Student o distribución t es un modelo teórico utilizado para aproximar el momento de primer orden de una población normalmente distribuida cuando el tamaño de la muestra es pequeño y se desconoce la desviación típica.

En otras palabras, la distribución t es una distribución de probabilidad que estima el valor de la media de una muestra pequeña extraída de una población que sigue una distribución normal y de la cual no conocemos su desviación típica.

Fórmula de la distribución t de Student

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Supóngase que se toma una muestra de una población normal con media μ y varianza σ^2 . Si $\bar{x}$ es el promedio de las n observaciones que contiene la muestra aleatoria, entonces la distribución es una distribución normal estándar. Supóngase que la varianza de la población σ^2 es desconocida. ¿Qué sucede con la distribución de esta estadística si se reemplaza σ por s ? La distribución t proporciona la respuesta a esta pregunta. La media y la varianza de la distribución t son $\mu = 0$ y $\sigma^2 = \text{gl}-2$ para $\text{gl} > 2$, respectivamente. La apariencia general de la distribución t es similar a la de la distribución normal estándar: ambas son simétricas y unimodales, y el valor máximo de la ordenada se alcanza en la media $\mu = 0$. Sin embargo, la distribución t tiene colas más amplias que la normal; esto es, la probabilidad de las colas es mayor que en la distribución normal. A medida que el número de grados de libertad tiende a infinito, la forma límite de la distribución t es la distribución normal estándar.

Dada una variable aleatoria continua L , decimos que la frecuencia de sus observaciones puede aproximarse satisfactoriamente a una distribución t con g grados de libertad tal que:

$$L \approx t_g$$

La variable aleatoria L sigue una distribución t con g grados de libertad.

Representación de la distribución t de Student

Función de densidad de una distribución t con 3 grados de libertad (df).



Función de densidad de la distribución t con 3 grados de libertad.

Como podemos ver, la representación de la distribución t se parece mucho a la distribución normal salvo que la distribución normal tiene las colas más anchas y es más apuntalada. En otras palabras, deberíamos añadir más grados de libertad a la distribución t para que la distribución “crezca” y se parezca más a la distribución normal.

Aplicación de la t de Student

La distribución t se utiliza cuando:

- Queremos estimar la media de una población normalmente distribuida a partir de una muestra pequeña.
- Tamaño de la muestra es inferior a 30 elementos, es decir, $n < 30$.

A partir de 30 observaciones, la distribución t se parece mucho a la distribución normal y, por tanto, utilizaremos la distribución normal.

- No se conoce la desviación típica o estándar de una población y tiene que ser estimada a partir de las observaciones de la muestra.
- Por cada curva t tiene forma de campana con centro en 0.
 - Cada curva t, está más dispersa que la curva normal estándar z.
 - A medida que v aumenta, la dispersión de la curva t correspondiente disminuye.
 - A medida que $gl \rightarrow \infty$, la secuencia de curvas t se aproxima a la curva normal estándar, por lo que la curva z recibe a veces el nombre de curva t con $gl = \infty$

propiedades de la distribución t

Ejemplo resuelto:

Suponemos que tenemos 28 observaciones de una variable aleatoria G que sigue una distribución t de Student con 27 grados de libertad (gl).

```
-0.66109321  0.87426449 -0.21001010  0.85784467 -2.48692810  0.17831511
-0.34501437 -0.95639484  0.35129492  2.27694001  2.36461405 -0.59664360
 0.31923232 -0.15809322 -0.86069232 -0.47580231  0.65589350 -0.20863280
 0.18857879  0.98738525 -0.08536776  0.85393439 -1.38541556 -0.80056519
 0.66046656 -0.56274230 -0.45095378  0.13224412
```

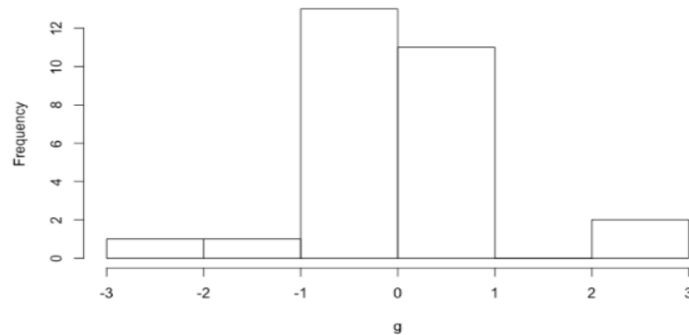
observaciones de la variable aleatoria G que sigue una distribución t con 27 grados de libertad.

Matemáticamente: $G(g_1, g_2, \dots, g_{28}) \approx t_{27}$

Variable aleatoria G que sigue una distribución t con 27 grados de libertad.

Dado que estamos trabajando con datos reales, siempre habrá un error de aproximación entre los datos y la distribución. En otras palabras, la media, mediana y moda no siempre serán cero (0) o exactamente iguales.

Representamos la frecuencia de cada observación de la variable G mediante un histograma.



Histograma de frecuencias de la variable aleatoria G.

Razones para considerar que la variable G sigue una distribución t:

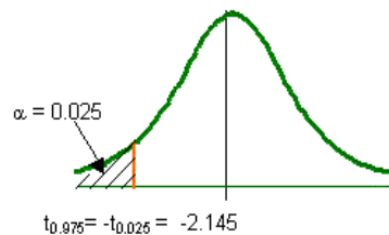
- La distribución es simétrica. Es decir, existe el mismo número de observaciones tanto a la derecha como a la izquierda del valor central. También, que la media y la mediana tienden a aproximarse al mismo valor. La media es aproximadamente cero, media = 0,016.
- Las observaciones con más frecuencia o probabilidad están alrededor del valor central. Las observaciones con menos frecuencia o probabilidad se encuentran lejos del valor central.

La diferencia es que la distribución normal es la distribución límite de numerosas variables aleatorias, discretas y continuas, como se demuestra a través de los teoremas centrales del límite. La distribución t de Student se construye como un cociente entre una normal y la raíz de una Chi-cuadrado independiente.

Ejemplo resuelto:

El valor t con $gl = 14$ grados de libertad que deja un área de 0.025 a la izquierda, y por tanto un área de 0.975 a la derecha, es

$$t_{0.975} = -t_{0.025} = -2.145$$



Si se observa la tabla, el área sombreada de la curva es de la cola derecha, es por esto que se tiene que hacer la resta de $1 - \alpha$. La manera de encontrar el valor de t es buscar el valor de α en el primer renglón de la tabla y luego buscar los grados de libertad en la primer columna y donde se intercepten α y gl se obtendrá el valor de t .

1.4.6 Distribución muestral de la varianza.

Definición: A la distribución muestral (s^2) se le conoce también como distribución Chi-cuadrada (χ^2). Es decir, que, si se extraen todas las muestras posibles de una población normal y a cada muestra se le calcula su varianza, se obtendrá la distribución muestral de las varianzas.

Fórmula:
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Propiedades:

- 1.- Los valores de χ^2 son mayores o iguales que 0
- 2.- La forma de una distribución de χ^2 depende de los grados de libertad $gl = n-1$ (grados de libertad).
- 3.- el área bajo una curva de ji cuadrada y sobre el eje horizontal es 1
- 4.- Las distribuciones χ^2 no son simétricas. (pendiente positiva).

Consiste:

El cálculo de probabilidad en una distribución muestral de varianzas nos sirve para saber cómo se va a comportar la varianza o desviación estándar en una muestra que proviene de una distribución normal. También es utilizada para obtener una estimación por intervalos y realizar pruebas de hipótesis para una varianza poblacional.

Usos:

Para estimar riesgos en inversiones. (el riesgo depende de la varianza para aplicar el teorema central del límite. Así como para estimar desigualdades en otros ingresos, rentas, entre otros.

Ejemplo resuelto.

Encuentre la probabilidad de que una muestra aleatoria de 25 observaciones, de una población normal con varianza igual a 6, tenga una varianza muestral:

- a) Mayor que 9.1
- b) Entre 3.462 y 10.745

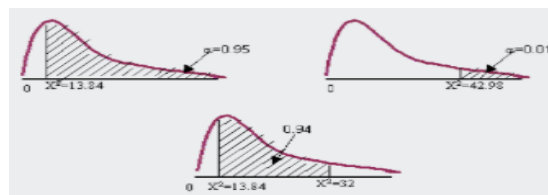
a)

$$X^2 = \frac{(25-1)(9.1)}{6} = 36.4$$

Al buscar este número en el renglón de 24 grados de libertad nos da un área a la derecha de 0.05 por lo tanto: $P(s^2 > 9.1) = 0.05$

$$X^2 = \frac{(25-1)(3.462)}{6} = 13.847 \quad \text{y} \quad X^2 = \frac{(25-1)(10.745)}{6} = 42.98$$

Al buscar estos valores en el renglón de 24 grados de libertad, nos dan unas áreas de 0.95 y 0.01 respectivamente. Como se está pidiendo la probabilidad entre dos valores se resta $(0.95 - 0.01)$ quedando 0.94



Sean las variables aleatorias X e Y tales que

Sean las variables aleatorias X e Y tales que

$$\left. \begin{array}{l} X \longrightarrow N(\mu_X, \sigma_X) \\ Y \longrightarrow N(\mu_Y, \sigma_Y) \end{array} \right\} \text{Independientes}$$

Consideramos:

$$\left. \begin{array}{l} \text{m.a.s. de tamaño } n_X \text{ de } X \\ X_1, X_2, \dots, X_{n_X} \end{array} \right\} \bar{X}, S_X^2$$

$$\left. \begin{array}{l} \text{m.a.s. de tamaño } n_Y \text{ de } Y \\ Y_1, Y_2, \dots, Y_{n_Y} \end{array} \right\} \bar{Y}, S_Y^2$$

$$\bar{X} = \frac{1}{n_X} \sum_{i=1}^{n_X} X_i \quad S_X^2 = \frac{1}{n_X - 1} \sum_{i=1}^{n_X} (X_i - \bar{X})^2$$

$$\bar{Y} = \frac{1}{n_Y} \sum_{j=1}^{n_Y} Y_j \quad S_Y^2 = \frac{1}{n_Y - 1} \sum_{j=1}^{n_Y} (Y_j - \bar{Y})^2$$

El estadístico definido como:

$$F = \frac{S_X^2 \times \sigma_Y^2}{S_Y^2 \times \sigma_X^2} \rightarrow F_{n_X - 1, n_Y - 1}$$

Tiene una distribución F con $n_X - 1$, y $n_Y - 1$ grados de libertad:

$$F = \frac{S_X^2 \times \sigma_Y^2}{S_Y^2 \times \sigma_X^2} \rightarrow F_{n_X - 1, n_Y - 1}$$

Ejemplo resuelto:

Se está comparando la variabilidad de los I.B-D de dos ríos A y B, que suponemos siguen distribuciones Normales. Se realizan 16 mediciones en el río A y se obtiene una cuasi varianza de 9.52, y 18 mediciones en el río B y se obtiene una cuasi varianza de 7. Obtener la probabilidad de que la varianza en el río B sea como mínimo el doble de la varianza en el río A.

$$X: \text{"I.B-D en el río A"} \rightarrow N(\mu_X; \sigma_X)$$

$$Y: \text{"I.B-D en el río B"} \rightarrow N(\mu_Y; \sigma_Y)$$

$$F = \frac{S_X^2 \times \sigma_Y^2}{S_Y^2 \times \sigma_X^2} \rightarrow F_{n_X - 1, n_Y - 1}$$

$$P(\sigma_Y^2 \geq 2\sigma_X^2) = P\left(\frac{\sigma_Y^2}{\sigma_X^2} \geq 2\right) = P\left(\frac{S_X^2 \sigma_Y^2}{S_Y^2 \sigma_X^2} \geq 2 \frac{S_X^2}{S_Y^2}\right) =$$

$$= P\left(F_{15,17} \geq 2 \times \frac{9.52}{7}\right) = P(F_{15,17} \geq 2.72) = 0.025$$

1.4.7 Distribución muestral de la relación de varianzas.

Es lo que resulta de considerar todas las muestras posibles que pueden ser tomadas de una población. Su estudio permite calcular la probabilidad que se tiene, dada una sola muestra, de acercarse al parámetro de la población. Mediante la distribución muestral se puede estimar el error para un tamaño de muestra dado. Sea $X_1, X_2, \dots, X_{n-1}$ una muestra aleatoria extraída de una población normal de parámetros σ^2 .

br > Del mismo modo, sea $Y_1, Y_2, \dots, Y_{n2}$ una muestra aleatoria extraída de una población normal de parámetros σ^2_2 .

Supondremos también que ambas poblaciones son independientes.

Nota 1:

Toda vez que se necesite resolver probabilidades de la forma $P(s^2_1 < s^2_2)$ o algunas de sus formas, deberemos realizar una transformación de variables hasta conseguir la forma cómo se define a T para luego utilizar la distribución F de Fisher a fin de encontrar la probabilidad buscada.

Nota 2:

Si las varianzas poblacionales son iguales u homogéneas entonces la variable muestral: cociente de varianzas muestrales debe tomar la forma $T = (s^2_1 / s^2_2)$ para tener una distribución F de Fisher con $n_1 - 1$ y $n_2 - 1$ grados de libertad en el denominador y denominador, respectivamente.

b) Dado $P(s^2_1 \leq a^2_2) = 0.88$, primero dividiremos entre s^2_2 y luego debemos transformarla en una variable que se distribuya como F ($n_1 - 1, n_2 - 1$) y finalmente lo igualamos a 0.88.

En efecto; Puesto que no se conoce el valor para el cual se tiene $P(F < 1.2143k) = 0.88$,

a) Supuestos

Las poblaciones siguen la distribución normal.

Los datos son independientes

b) Hipótesis.

$$\begin{array}{l} H_0 : \sigma_1^2 - \sigma_2^2 = 0 \\ H_0 : \sigma_1^2 - \sigma_2^2 \neq 0 \end{array}$$

c) Estadístico de contraste

$$F = \frac{\tilde{s}_1^2}{\tilde{s}_2^2}$$

d) Distribución del estadístico de contraste: F con n_1-1 y n_2-1 grados de libertad

e) Significación del estadístico de contraste.

Ejemplo resuelto:

Quieren someter a contraste la hipótesis de igualdad de las varianzas de las puntuaciones obtenidas por 2 gpos. De individuos en una prueba de extraversión.

Cuasivarianza del grupo1: 2.8

Cuasivarianza del gpo. 2: 2.5

Tamaño del gpo. 1: 31

Tamaño del gpo. 2: 41

$\alpha=0.01$

a) Supuestos: Dado que las muestras han sido generadas aleatoriamente, se supone que los datos son independientes

b) Hipótesis

$$\begin{aligned} H_0 : \sigma_1^2 - \sigma_2^2 &= 0 \\ H_a : \sigma_1^2 - \sigma_2^2 &\neq 0 \end{aligned}$$

c) Estadístico de contraste:

$$F = \frac{\tilde{s}_1^2}{\tilde{s}_2^2} = \frac{2.8}{2.5} = 1.12$$

d) Distribución del estadístico de contraste F con 31-1 y 41-1 grados de libertad

e) Significación del estadístico de contraste: 0.63

f) Decisión: la significación del estadístico de contraste es superior a 0.01, y en consecuencia se acepta la Hipótesis nula.

Ejemplos propuestos unidad 1

1. Se sabe que la resistencia a la ruptura de cierto tipo de cuerda se distribuye normalmente con media de 2000 libras y una varianza de 25,000 lbs. Si se selecciona una muestra aleatoria de 100 cuerdas; determine la probabilidad de que en esa muestra:
 - a) La resistencia media encontrada sea de por lo menos 1958 libras.
 - b) La resistencia media sea mayor de 2080 libras.
2. Como parte de un proyecto general de mejoramiento de la calidad, un fabricante textil decide controlar el número de imperfecciones encontradas en cada pieza de tela. Se estima que el número promedio de imperfecciones por cada pieza de tela es de 12, determine la probabilidad de que en la próxima pieza de tela fabricada se encuentren:
 - a) Entre 10 y 12 imperfecciones.
 - b) Menos de 9 y más de 15 imperfecciones.
3. En una prueba de aptitud la puntuación media de los estudiantes es de 72 puntos y la desviación estándar es de 8 puntos. ¿Cuál es la probabilidad de que dos grupos de estudiantes, formados de 28 y 36 estudiantes, respectivamente, difieran en su puntuación media en:
 - a) 3 ó más puntos?
 - b) 6 o más puntos.
 - c) Entre 2 y 5 puntos.
4. Un especialista en genética ha detectado que el 26% de los hombres y el 24% de las mujeres de cierta región del país tiene un leve desorden sanguíneo; si se toman muestras de 150 hombres y 150 mujeres, determine la probabilidad de que la diferencia muestral de proporciones que tienen ese leve desorden sanguíneo sea de:
 - a) Menos de 0.035 a favor de los hombres.

- b) Entre 0.01 y 0.04 a favor de los hombres.
5. Una urna contiene 80 bolas de las que 60% son rojas y 40% blancas. De un total de 50 muestras de 20 bolas cada una, sacadas de la urna con remplazamiento, ¿en cuántas cabe esperar
- Igual número de bolas rojas y blancas?
 - 12 bolas rojas y 8 blancas?
 - 8 bolas rojas y 12 blancas?
 - 10 ó mas bolas blancas?
6. Los pesos de 1500 cojinetes de bolas se distribuyen normalmente con media de 2.40 onzas y desviación estándar de 0.048 onzas. Si se extraen 300 muestras de tamaño 36 de esta población, determinar la media esperada y la desviación estándar de la distribución muestral de medias si el muestreo se hace:
- Con remplazamiento
 - Sin remplazamiento
7. La vida media de una máquina para hacer pasta es de siete años, con una desviación estándar de un año. Suponga que las vidas de estas máquinas siguen aproximadamente una distribución normal, encuentre:
- La probabilidad de que la vida media de una muestra aleatoria de 9 de estas máquinas caiga entre 6.4 y 7.2 años.
8. Se llevan a cabo dos experimentos independientes en lo que se comparan dos tipos diferentes de pintura. Se pintan 18 especímenes con el tipo A y en cada uno se registra el tiempo de secado en horas. Lo mismo se hace con el tipo B. Se sabe que las desviaciones estándar de la población son ambas 1.0. Suponga que el tiempo medio de secado es igual para los dos tipo de pintura. Encuentre la probabilidad de que la diferencia de medias en el tiempo de secado sea mayor a uno a favor de la pintura A.
9. Un ingeniero químico afirma que el rendimiento medio de la población de cierto proceso en lotes es 500 gramos por milímetro de materia prima. Para verificar esta afirmación toma una muestra de 25 lotes cada mes. Si el valor de t calculado cae entre $-t_{0.05}$ y $t_{0.05}$, queda satisfecho con su afirmación. ¿Qué conclusión extraería de una muestra que tiene una media de 518 gramos por milímetro y una desviación estándar de 40 gramos? Suponga que la distribución de rendimientos es aproximadamente normal. Encuentre

$P(k < t < -1.761) = 0.045$, para una muestra aleatoria de tamaño 15 que se selecciona de una distribución normal.

10. Encuentre la probabilidad de $-t_{0.025} < t < t_{0.05}$.
11. Suponga que los tiempos requeridos por un cierto autobús para alcanzar un de sus destinos en una ciudad grande forman una distribución normal con una desviación estándar $\sigma = 1$ minuto. Si se elige al azar una muestra de 17 tiempos, encuentre la probabilidad de que la varianza muestral sea mayor que 2.
12. Encuentre la probabilidad de que una muestra aleatoria de 25 observaciones, de una población normal con varianza $\sigma^2 = 6$, tenga una varianza muestral:
 - a. Mayor que 9.1
 - b. Entre 3.462 y 10.745

Unidad 2. Estimación.



Instrumentación Didáctica.

Competencia Específica	Competencia Genérica	Actividades de Aprendizaje
Aplica los fundamentos de la teoría de estimación en problemas que requieran el cálculo del tamaño de la muestra para determinar los diferentes intervalos de confianza según la variable que se está analizando en procesos industriales y logísticos.	Capacidad de abstracción, análisis y síntesis. • Capacidad de aplicar los conocimientos en la práctica. • Capacidad para identificar, plantear y resolver problemas.	• Buscar información sobre conceptos relacionados con estimación. • Discutir en grupo sobre los conceptos investigados. • Dado un conjunto de datos diferenciar la importancia de utilizar estimadores puntuales y estimadores por intervalos. • Proporcionar al estudiante situaciones hipotéticas de procesos y/o poblaciones finitas para que obtengan un conjunto de datos para establecer una estimación por intervalo dependiendo la variable que se está midiendo con el fin de obtener la muestra definitiva. • Interpretar el significado de los intervalos de confianza para: la media, diferencia de medias, la proporción, diferencia de proporciones, varianza y relación de varianzas. • Utilizar TIC's para obtener intervalos de confianza y tamaños de muestra.

2. Estimación

2.1 Introducción:

El objetivo principal de la estadística inferencial es la estimación, esto es que mediante el estudio de una muestra de una población se quiere generalizar las conclusiones al total de la misma. Los estadísticos varían mucho dentro de sus distribuciones muestrales, y mientras menor sea el error estándar de un estadístico, más cercanos serán unos de otros sus valores.

Si queremos estimar la media μ de una población, debemos utilizar la media $\bar{X}$ de una muestra representativa extraída de la población en estudio. Rara vez la media de la muestra seleccionada al azar será tal que $\bar{X} = \mu$. Sin duda es mucho más interesante realizar la inferencia con un intervalo de posible valores del parámetro. Al que denominaremos intervalo de confianza

La estimación puntual consiste en atribuir a un valor (la estimación) al parámetro poblacional. Si la muestra es representativa de la población, podemos esperar que los estadísticos calculados en las muestras tengan valores semejantes a los parámetros poblacionales, y la estimación consiste en asignar los valores de los estadísticos muestrales a los parámetros poblacionales. Los estadísticos con que obtenemos las estimaciones se denominan **estimadores**.

Ejemplo:

Se desea estimar la media de las puntuaciones del curso 2003/4, pero solo se dispone de 50 puntuaciones seleccionadas aleatoriamente. La media de la muestra (el estimador), es igual a 5.6 y atribuimos este valor (la estimación) a la media del curso completo.

Media de la población (curso 2003/4) μ = desconocida.

Estimador: media de la muestra $\bar{x}$

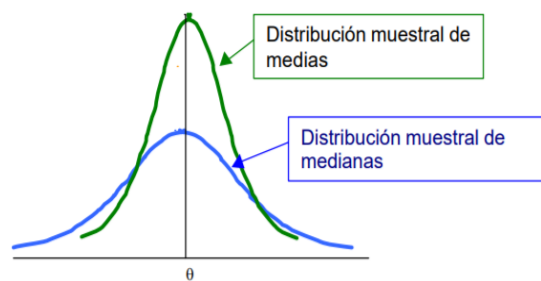
Estimación de $\mu = 5.6$

Podemos utilizar como estimadores de la media de la población, otros estadísticos de tendencia central como la mediana o la moda, pero no todos los estimadores son apropiados. Los estimadores deben satisfacer ciertos requisitos, y por esta razón, interesa conocer sus propiedades a fin de utilizar los que sean adecuados según las circunstancias de la estimación.

2.2 Características de un estimador.

Insesgado. - Se dice que un estimador puntual $\hat{\theta}$ es un estimador insesgado de θ si $E(\hat{\theta}) = \theta$, para todo valor posible de θ . En otras palabras, un estimador insesgado es aquel para el cual la media de la distribución muestral es el parámetro estimado. Si se usa la media muestral $\bar{x}$ para estimar la media poblacional μ , se sabe que $E(\bar{x}) = \mu$, por lo tanto, la media es un estimador insesgado.

Eficiente o con varianza mínima. - Suponga que $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores insesgados de θ . Entonces, aun cuando la distribución de cada estimador esté centrada en el valor verdadero de θ , las dispersiones de las distribuciones alrededor del valor verdadero pueden ser diferentes. Entre todos los estimadores de θ que son insesgados, seleccione al que tenga varianza mínima. El $\hat{\theta}$ resultante recibe el nombre de estimador insesgado con varianza mínima (MVUE, minimum variance unbiased estimator) de θ . En otras palabras, la eficiencia se refiere al tamaño de error estándar de la estadística. Si comparamos dos estadísticas de una muestra del mismo tamaño y tratamos de decidir cual de ellas es un estimador más eficiente, escogeríamos la estadística que tuviera el menor error estándar, o la menor desviación estándar de la distribución de muestreo. Tiene sentido pensar que un estimador con un error estándar menor tendrá una mayor oportunidad de producir una estimación más cercana al parámetro de población que se está considerando.



Como se puede observar las dos distribuciones tienen un mismo valor en el parámetro sólo que la distribución muestral de medias tiene una menor varianza, por lo que la media se convierte en un estimador eficiente e insesgado.

Coherencia. - Una estadística es un estimador coherente de un parámetro de población, si al aumentar el tamaño de la muestra se tiene casi la certeza de que

el valor de la estadística se aproxima bastante al valor del parámetro de la población. Si un estimador es coherente se vuelve más confiable si tenemos tamaños de muestras más grandes.

Suficiencia. - Un estimador es suficiente si utiliza una cantidad de la información contenida de la muestra que ningún otro estimador podría extraer información adicional de la muestra sobre el parámetro de la población que se está estimando. Es decir, se pretende que al extraer la muestra el estadístico calculado contenga toda la información de esa muestra. Por ejemplo, cuando se calcula la media de la muestra, se necesitan todos los datos. Cuando se calcula la mediana de una muestra sólo se utiliza a un dato o a dos. Esto es solo el dato o los datos del centro son los que van a representar la muestra. Con esto se deduce que si utilizamos a todos los datos de la muestra como es en el caso de la media, la varianza, desviación estándar, etc; se tendrá un estimador suficiente

2.3 Estimación puntual.

Es un solo número que se utiliza para estimar un parámetro de población desconocido.

Por ejemplo, al observar a un integrante de futbol americano y decimos que el peso promedio de los jugadores de es 125 kg. Eso es una estimación puntual. El jefe de departamento de una universidad estaría haciendo una estimación puntual si afirmara Nuestros datos actuales indican que en esta materia tendremos 350 estudiantes en el siguiente semestre.

A menudo, una estimación puntual es insuficiente debido a que solo tienen dos opciones: es correcta o equivocada. Si le dicen solamente que la afirmación sobre la inscripción está equivocada, no sabe que tanto está mal y no puede tener la certeza de que la estimación es confiable. Si se entera de que solo esta errada por 10 estudiantes, entonces 350 estudiantes es una buena estimación de la inscripción futura. Entonces, una estimación puntual es mucho más útil si viene acompañada por una estimación del error que podría estar implicado.

2.4 Estimación por intervalos.

Si queremos estimar la media μ de una población, debemos utilizar la media $\bar{X}$ de una muestra representativa extraída de la población en estudio. Rara vez la media

de la muestra seleccionada al azar será tal que $\bar{X} = \mu$. Sin duda es mucho más interesante realizar la inferencia con un intervalo de posibles valores del parámetro. Al que denominaremos intervalo de confianza.

Es un rango de valores que se utiliza para estimar un parámetro de la población e indica el error de dos maneras: por la extensión del intervalo y por la probabilidad de que el verdadero parámetro poblacional se encuentre dentro del intervalo. En este caso como ejemplo el jefe de departamento diría: Estimo que la inscripción real de este curso para el próximo semestre estará entre 330 y 380. Y es muy probable que la inscripción exacta caiga dentro de este intervalo. Con esto tiene una mejor idea de la confiabilidad de su estimación.

El objetivo de la estimación por intervalo es aportar información de qué tan cerca se encuentra la estimación puntual, obtenida de la muestra, del valor del parámetro poblacional. La fórmula general para obtener una estimación por intervalo para la media poblacional es

$$\bar{x} \pm \text{margen de error}$$

De manera similar, la fórmula general para obtener una estimación por intervalo para la proporción poblacional es

$$\bar{p} \pm \text{margen de error}$$

Las distribuciones muestrales de $\bar{x}$ y $\bar{p}$ de son clave para calcular estas estimaciones por intervalo.

La distribución muestral de es una distribución normal con error estándar de

$$\sigma_{\bar{x}} = \sigma/\sqrt{n}$$

los datos históricos muestran que la población constituida por las cantidades gastadas está distribuida normalmente y que su desviación estándar es $\sigma = 20$ y existe una muestra de 100. Entonces el error estándar de la media es:

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 20/\sqrt{100} = 2$$

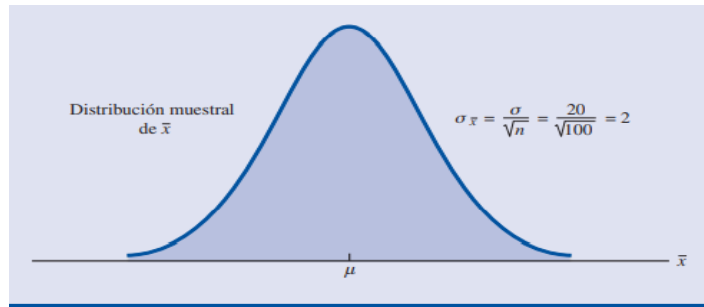
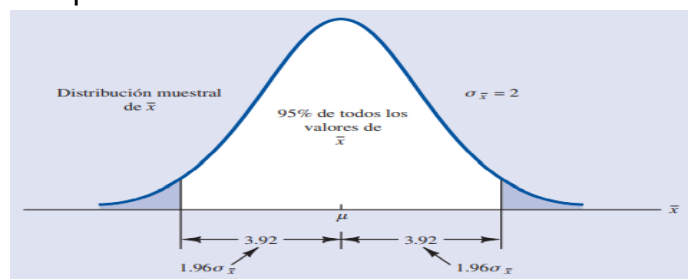


Fig. 2.1 Distribución muestral de la media muestral en muestras aleatorias simples de 100 clientes.

2.4.1 Intervalo de confianza para la media σ conocida

Puesto que la distribución muestral indica cómo están distribuidos los valores de en torno a la media poblacional μ , la distribución muestral de $\bar{x}$ proporciona información acerca de la posible diferencia entre $\bar{x}$ y μ .

En la tabla de probabilidad normal estándar se encuentra que 95% de los valores de cualquier variable aleatoria distribuida normalmente aparecen dentro de ± 1.96 desviaciones estándar de la media. Por tanto, si la distribución muestral se encuentra distribuida normalmente, 95% de los valores de deben estar dentro de $\bar{x} \pm 1.96$ de la media μ . Por el ejemplo se sabe que la distribución muestral de $\bar{x}$ está distribuida normalmente y que el error estándar es $\sigma_{\bar{x}} = 2$. se puede concluir que 95% de los valores de $\bar{x}$ obtenidos usando muestras de $n = 100$ estarán dentro de ± 3.92 de la media poblacional.



Como el margen de error está dado por $z_{\alpha/2}(\sigma/\sqrt{n})$, la fórmula general de una estimación por intervalo de la media poblacional con σ conocida, es la siguiente.

Estimación por intervalo de una media poblacional: σ conocida

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

donde $(1 - \alpha)$ es el coeficiente de confianza y $z_{\alpha/2}$ es el valor de z que proporciona un área $\alpha/2$ en la cola superior de la distribución de probabilidad normal estándar.

Se construye un intervalo de confianza de 95% con un coeficiente de confianza $(1-\alpha)= 0.95$ y, por tanto, $\alpha = 0.05$. En la tabla de la distribución normal estándar aparece que un área de $\alpha/2 = 0.05/2 = 0.025$ en la cola superior corresponde a $z_{0.025} = 1.96$. Como la media muestral es $\bar{x} = 82$, $\sigma = 20$ y el tamaño de la muestra es $n = 100$, se obtiene

$$82 \pm 1.96 \frac{20}{\sqrt{100}}$$

$$82 \pm 3.92$$

Por tanto, el margen de error es 3.92 y el intervalo de confianza de 95% va de $82 - 3.92 = 78.08$ a $82 + 3.92 = 85.92$

Por tanto, el margen de error es 3.92 y el intervalo de confianza de 95% va de $82 - 3.92 = 78.08$ a $82 + 3.92 = 85.92$.

Aunque a menudo se usa un nivel de confianza de 95%, también suelen usarse otros niveles de confianza como 90% y 99%.

Y con el intervalo de confianza de 90% en el ejemplo es

$$82 \pm 1.64 \frac{20}{\sqrt{100}}$$

$$82 \pm 3.29$$

A continuación se presentan los valores de Z más comúnmente usados:

Nivel de confianza	α	$\alpha/2$	$z_{\alpha/2}$
90%	.10	.05	1.64
95%	.05	.025	1.960
99%	.01	.005	2.57

σ desconocida

Cuando se calcula un intervalo de confianza para la media poblacional, suele no contarse con una buena estimación de la desviación estándar poblacional. En tales casos se usa la misma muestra para estimar μ y σ . Esta situación es el caso que se conoce como σ desconocida. Cuando se usa s para estimar σ , el margen de error y la estimación por intervalo de la media poblacional se basan en una distribución de probabilidad conocida como distribución t . Aunque la elaboración matemática de la distribución t parte de la suposición de que la población de la que se muestrea tiene una distribución normal, las investigaciones han demostrado que la distribución t se aplica en muchas situaciones en que la población se desvía significativamente de una población normal. La distribución t es una familia de distribuciones de probabilidad similares; cada distribución t depende de un parámetro conocido como grados de libertad. La distribución t para un grado de libertad es única, como lo es la distribución t para dos grados de libertad, para tres grados de libertad, etc. A medida que el número de grados de libertad aumenta, la diferencia entre la distribución t y la distribución normal estándar se va reduciendo.

Observe en la fig. 2.3 que una distribución t para más grados de libertad exhibe menos variabilidad y un mayor parecido con la distribución normal estándar, también que la media de toda distribución es cero.

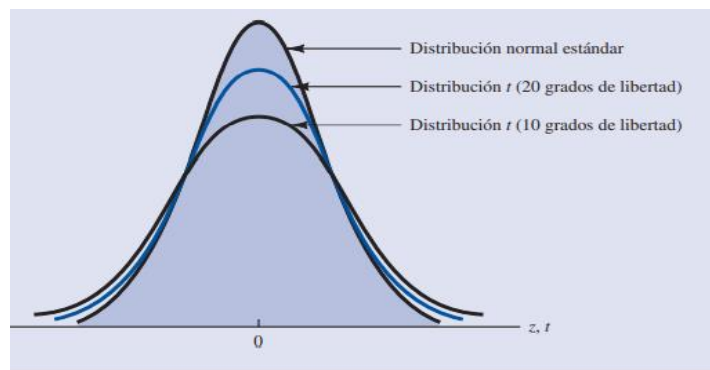
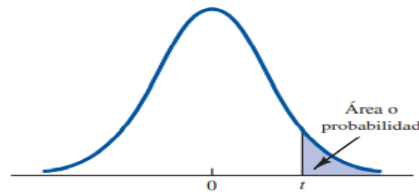


Fig.2.3 Comparación de la distribución normal estándar con las distribuciones t .

Para calcular una estimación por intervalo para μ en el caso en que no se conoce σ se usa la desviación estándar muestral s para estimar σ , y $z_{\alpha/2}$ se sustituye por el valor $t_{\alpha/2}$ de la distribución t .



Grados de libertad	Área en la cola superior					
	0.20	0.10	0.05	0.025	0.01	0.005
1	1.376	3.078	6.314	12.706	31.821	63.656
2	1.061	1.886	2.920	4.303	6.965	9.925
3	0.978	1.638	2.353	3.182	4.541	5.841
4	0.941	1.533	2.132	2.776	3.747	4.604
5	0.920	1.476	2.015	2.571	3.365	4.032
6	0.906	1.440	1.943	2.447	3.143	3.707
7	0.896	1.415	1.895	2.365	2.998	3.499
8	0.889	1.397	1.860	2.306	2.896	3.355
9	0.883	1.383	1.833	2.262	2.821	3.250
10	0.878	1.371	1.812	2.228	2.764	3.169
11	0.874	1.360	1.796	2.199	2.715	3.098
12	0.870	1.350	1.782	2.176	2.675	3.037
13	0.867	1.341	1.770	2.158	2.641	2.982
14	0.864	1.333	1.760	2.144	2.612	2.932
15	0.861	1.326	1.751	2.131	2.588	2.887
16	0.859	1.319	1.743	2.120	2.567	2.845
17	0.857	1.313	1.736	2.110	2.548	2.805
18	0.855	1.307	1.729	2.101	2.531	2.767
19	0.853	1.302	1.723	2.093	2.515	2.731
20	0.851	1.297	1.718	2.086	2.500	2.697
25	0.848	1.286	1.706	2.069	2.479	2.656
30	0.846	1.277	1.695	2.056	2.462	2.618
40	0.844	1.269	1.684	2.045	2.447	2.583
50	0.843	1.264	1.678	2.037	2.435	2.557
60	0.842	1.260	1.674	2.031	2.426	2.537
70	0.842	1.257	1.671	2.027	2.419	2.522
80	0.842	1.255	1.669	2.024	2.415	2.515
90	0.842	1.254	1.668	2.022	2.413	2.511
100	0.842	1.253	1.667	2.021	2.412	2.509
∞	0.842	1.252	1.667	2.020	2.411	2.508

Fig. 2.4 Algunos valores de la tabla de la distribución t

El margen de error está dado, entonces, por $t_{\alpha/2} \frac{s}{\sqrt{n}}$. Con este margen de error, la expresión general para una estimación por intervalo de la media poblacional cuando σ no se conoce es la siguiente.

Estimación por intervalo de la media poblacional: σ desconocida

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Donde s es la desviación estándar muestral, $(1 - \alpha)$ es el coeficiente de confianza y $t_{\alpha/2}$ es el valor de t que proporciona un área de $\alpha/2$ en la cola superior de la distribución t para $n-1$ grados de libertad.

2.4.2. Intervalo de confianza para la diferencia de medias.

Si se tienen dos poblaciones con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 , respectivamente, un estimador puntual de la diferencia entre μ_1 y μ_2 está dado por la estadística $\bar{x}_1 - \bar{x}_2$. Por tanto, para obtener una estimación puntual de $\mu_1 - \mu_2$, se seleccionan dos muestras aleatorias independientes, una de cada población, de tamaño n_1 y n_2 , se calcula la diferencia $\bar{x}_1 - \bar{x}_2$, de las medias muestrales. Recordando a la distribución muestral de diferencia de media. Recordando a la distribución muestral de diferencia de medias:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Al despejar de esta ecuación $\mu_1 - \mu_2$ se tiene:

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm z \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

En el caso en que se desconozcan las varianzas de la población y los tamaños de muestra sean mayores a 30 se podrá utilizar la varianza de la muestra como una estimación puntual.

Ejemplo resuelto:

Se lleva a cabo un experimento en que se comparan dos tipos de motores, A y B. Se mide el rendimiento en millas por galón de gasolina. Se realizan 50 experimentos con el motor tipo A y 75 con el motor tipo B. La gasolina que se utiliza y las demás condiciones se mantienen constantes. El rendimiento promedio de gasolina para el motor A es de 36 millas por galón y el promedio para el motor B es 24 millas por galón. Encuentre un intervalo de confianza de 96% sobre la diferencia promedio real para los motores A y B. Suponga que las desviaciones estándar poblacionales son 6 y 8 para los motores A y B respectivamente. Solución: Es deseable que la diferencia de medias sea positiva por lo que se recomienda restar la media mayor menos la media menor. En este caso será la media del motor B menos la media del motor A. El valor de z para un nivel de confianza del 96% es de 2.05.

$$\mu_B - \mu_A = (\bar{x}_B - \bar{x}_A) \pm z \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = (42 - 36) \pm 2.05 \sqrt{\frac{36}{50} + \frac{64}{75}}$$

$$3.43 < \mu_B - \mu_A < 8.57$$

La interpretación de este ejemplo sería que con un nivel de confianza del 96% la diferencia del rendimiento promedio esta entre 3.43 y 8.57 millas por galón a favor del motor B. Esto quiere decir que el motor B da más rendimiento promedio que el motor A, ya que los dos valores del intervalo son positivos.

Ejemplo resuelto:

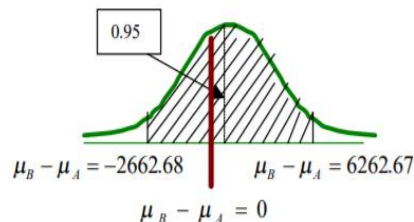
Compañía de taxis trata de decidir si comprar neumáticos de la marca A o de la B para su flotilla de taxis. Para estimar la diferencia de las dos marcas, se lleva a cabo un experimento utilizando 12 de cada marca. Los neumáticos se utilizan hasta que se desgastan, dando como resultado promedio para la marca A 36,300 kilómetros y para la marca B 38,100 kilómetros. Calcule un intervalo de confianza de 95% para la diferencia promedio de las dos marcas, si se sabe que las poblaciones se distribuyen de forma aproximadamente normal con desviación estándar de 5000 kilómetros para la marca A y 6100 kilómetros para la marca B.

Solución:

$$\mu_B - \mu_A = (\bar{x}_B - \bar{x}_A) \pm z \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = (38100 - 36300) \pm 1.96 \sqrt{\frac{5000^2}{12} + \frac{6100^2}{12}}$$

$$-2662.68 < \mu_B - \mu_A < 6262.67$$

Gráficamente:



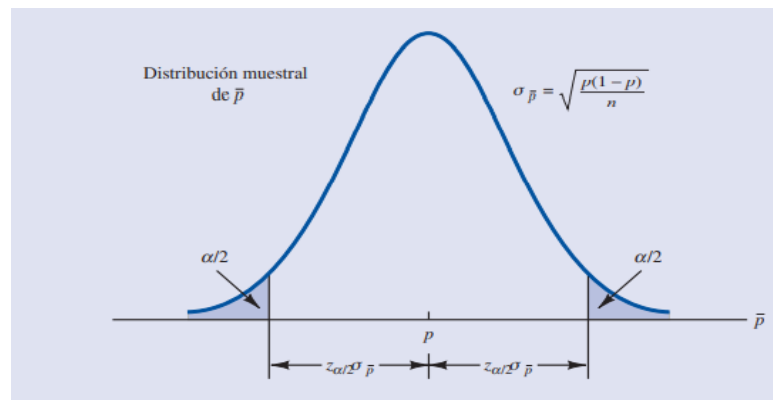
Como el intervalo contiene el valor “cero”, no hay razón para creer que el promedio de duración del neumático de la marca B es mayor al de la marca A, pues el cero nos está indicando que pueden tener la misma duración promedio.

2.4.3 Intervalos de confianza para la proporción.

En la introducción a la estimación se dijo que para obtener una estimación por intervalos de la porción poblacional p , la formula general era:

$$\bar{p} \pm \text{margen de error.}$$

En el cálculo de esta estimación por intervalo la distribución muestral es importante. la distribución muestral de se aproxima mediante una distribución normal siempre que $np \geq 5$ y $n(1 - p) \geq 5$ se muestra en la siguiente figura se muestra una aproximación normal a la distribución muestral de $\bar{p}$.



Aproximación Normal a la distribución Muestral de $\bar{p}$

La media de la distribución muestral de $\bar{p}$ es la proporción poblacional p y el error estándar $\sigma_{\bar{p}}$ es

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Como la distribución muestral de $\bar{p}$ es una distribución normal, si en la estimación por intervalo de la proporción poblacional se elige como margen de error $z_{\alpha/2} \sigma_{\bar{p}}$, entonces $100(1 - \alpha)\%$ de los intervalos que se obtengan contendrán la verdadera proporción poblacional. Pero para calcular el margen de error no se puede usar $\sigma_{\bar{p}}$, ya que no se conoce p ; p es lo que se está tratando de estimar. Lo que se hace, es que p se sustituye por $\bar{p}$ y de esta manera se calcula el margen de error y la estimación por intervalo de la proporción poblacional queda dada por

$$\text{Margen de error} = z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Con este margen de error la expresión general para la estimación por intervalo de la proporción poblacional es la siguiente.

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Donde $1 - \alpha$ es el coeficiente de confianza de $Z_{\alpha/2}$ es el valor de Z que deja un área $\alpha/2$ en la cola superior de la distribución normal estandar.

En el siguiente ejemplo se ilustra el cálculo del margen de error y de la estimación por intervalo para una proporción muestral. Un estudio en Estados Unidos encuestó a 900 golfistas para conocer su opinión acerca de cómo se les trataba en los cursos de golf. En el estudio se encontró que 396 golfistas estaban satisfechas con la disponibilidad de horarios de salida. Por tanto, la estimación puntual de la proporción poblacional de golfistas satisfechas con la disponibilidad de horarios de salida es $396/900 = 0.44$

$$\begin{aligned}\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} \\ 0.44 \pm 1.96 \sqrt{\frac{0.44(1 - 0.44)}{900}} \\ 0.44 \pm 0.0324\end{aligned}$$

En consecuencia, el margen de error es 0.0324 y la estimación por intervalo de confianza de 95% de la proporción poblacional es 0.4076 a 0.4724. Empleando porcentajes, los resultados de la investigación permiten decir con 95% de confianza que entre 40.76 y 47.24% de las golfistas están satisfechas con la disponibilidad de horarios de salida.

2.4.4 Intervalos de confianza para la diferencia de proporciones

En la sección anterior se vio el tema de la generación de las distribuciones muestrales, en donde se tenía el valor de los parámetros, se seleccionaban dos muestras y podíamos calcular la probabilidad del comportamiento de los estadísticos. Para este caso en particular se utilizará la distribución muestral de diferencia de proporciones para la estimación de las misma.

Recordando la fórmula

$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$$

Despejando $P_1 - P_2$ de esta ecuación:

$$P_1 - P_2 = (p_1 - p_2) \pm z \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}$$

Aquí se tiene el mismo caso que en la estimación de una proporción, ya que al hacer el despeje nos queda las dos proporciones poblacionales y es precisamente lo que queremos estimar, por lo que se utilizarán las proporciones de la muestra como estimadores puntuales:

$$P_1 - P_2 = (p_1 - p_2) \pm z \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Ejemplo resuelto:

Se considera cierto cambio en un proceso de fabricación de partes componentes. Se toman muestras del procedimiento existente y del nuevo para determinar si éste tiene como resultado una mejoría. Si se encuentra que 75 de 1500 artículos del procedimiento actual son defectuosos y 80 de 2000 artículos del procedimiento nuevo también lo son, encuentre un intervalo de confianza de 90% para la diferencia real en la fracción de defectuosos entre el proceso actual y el nuevo.

Solución:

Sean P_1 y P_2 las proporciones reales de defectuosos para los procesos actual y nuevo, respectivamente. De aquí, $p_1 = 75/1500 = 0.05$ y $p_2 = 80/2000 = 0.04$. con el uso de la tabla encontramos que z para un nivel de confianza del 90% es de 1.645.

$$P_1 - P_2 = (p_1 - p_2) \pm z \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} = (0.05 - 0.04) \pm 1.645 \sqrt{\frac{(0.05)(0.95)}{1500} + \frac{(0.04)(0.96)}{2000}}$$

$$-0.0017 < P_1 - P_2 < 0.0217$$

Como el intervalo contiene el valor de cero, no hay razón para creer que el nuevo procedimiento producirá una disminución significativa en la proporción de artículos defectuosos comparado con el método existente. 2. Un artículo relacionado con la salud, reporta los siguientes datos sobre la incidencia de disfunciones importantes entre recién nacidos con madres fumadoras de marihuana y de madres que no la fumaban:

	usuaria	No usuaria
Tamaño de la muestra	1246	11178
Número de disfunciones	42	294

Proporción muestral	0.0337	0.0263
----------------------------	--------	--------

Ejemplo resuelto:

Encuentre el intervalo de confianza del 99% para la diferencia de proporciones.

Solución:

Representemos P1 la proporción de nacimientos donde aparecen disfunciones entre todas las madres que fuman marihuana y definamos P2, de manera similar, para las no fumadoras. El valor de z para un 99% de confianza es de 2.58

$$P_1 - P_2 = (p_1 - p_2) \pm z \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} = (0.0337 - 0.0263) \pm 2.58 \sqrt{\frac{(0.0337)(0.9663)}{1246} + \frac{(0.0263)(0.9737)}{11178}}$$

$$\mathbf{-0.0064 < P_1 - P_2 < 0.0212}$$

Este intervalo es bastante angosto, lo cual sugiere que P1-P2 ha sido estimado de manera precisa.

2.4.5 Intervalos de confianza para la varianza.

Dada una muestra aleatoria simple $X_1, \dots, X_n$ de una población $N(\mu, \sigma)$, vamos a determinar el intervalo de confianza de σ^2 distinguiendo dos casos según sea desconocida o no la media de la población μ .

μ desconocida su fórmula es:

$$I = \left[\frac{(n-1)S^2}{\chi_{n-1; \alpha/2}^2}, \frac{(n-1)S^2}{\chi_{n-1; 1-\alpha/2}^2} \right]$$

Con S^2 la cuasivarianza muestral

μ conocida su fórmula es:

$$I = \left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n;\alpha/2}^2}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n;1-\alpha/2}^2} \right].$$

Ejemplo resuelto:

Un terapeuta desea estimar, con una confianza del 99%, la fuerza media de un músculo determinado en los individuos de una población. Admitiendo que las unidades de fuerza siguen una distribución normal de varianza 144, seleccionó una muestra aleatoria de 25 individuos de la población, para la que obtuvo una media muestral de 85. Si el terapeuta desea determinar un intervalo de confianza para la varianza de la variable en estudio será:

$$I = \left[\frac{(n-1)S^2}{\chi_{n-1;\alpha/2}^2}, \frac{(n-1)S^2}{\chi_{n-1;1-\alpha/2}^2} \right]$$

$$I = \frac{24.139}{45.54}, \frac{24.139}{9.886} = 7322.337$$

Obsérvese que para un tamaño muestral tan pequeño como el que tenemos, el intervalo de confianza al 99% determinado resulta poco informativo, al tener éste una longitud muy grande.

Correspondiente al 90%

$$I = \frac{24.139}{45.54}, \frac{24.139}{13.85} = 916.240$$

Tampoco resulta mucho mas informativo, perdiendo este, además parte del grado de confianza que el primero poseía. Una de las causas es que habitualmente, estaremos interesados en estimar la desviación y no la varianza, puesto que ésta viene expresada en unidades al cuadrado lo que distorsiona en parte el resultado. El intervalo de confianza para la desviación será el de extremos la raíz cuadrada del correspondiente de la varianza. Así, por ejemplo, el intervalo correspondiente al 90% será:

$$I = [\sqrt{916}, \sqrt{2409}] = [957, 15.52]$$

2.4.6 Intervalos de confianza para la relación de varianzas.

Supondremos que $X_1, \dots, X_{n1}$ e $Y_1, \dots, Y_{n2}$ son dos muestras de tamaños n_1 y n_2 extraídas respectivamente de dos poblaciones independientes $N(\mu_1, \sigma_1)$ y $N(\mu_2, \sigma_2)$.

Su fórmula es la siguiente:

$$\left(\frac{s^2}{S^2}\right)\left(\frac{1}{f_{\alpha/2}(gl_1, gl_2)}\right) \leq \frac{\sigma^2}{\sigma^2} \leq \left(\frac{s^2}{S^2}\right)(f_{\alpha/2}(gl_1, gl_2))$$

Donde $f_{\alpha/2}(gl_1, gl_2)$ es un valor f con $gl_1 = n_1 - 1$ y $gl_2 = n_2 - 1$ que deja un área de $\alpha/2$ a la derecha $f_{\alpha/2}(gl_2, gl_1)$ es un valor similar con $gl_2 = n_2 - 1$ y $gl_1 = n_1 - 1$.

Ejemplo resuelto:

El depto. De zoología del estado de virginia estimó la diferencia en el contenido medio de ortofosforo que se mide en miligramos/lit. En 2 estaciones sobre un rio. Mediante la suposición de que las varianzas normales de la población son diferentes. Justifique esta suposición mediante la construcción de un intervalo de confianza del 98% para la razón de varianzas.

Los datos son los siguientes:

Estación 1: $n_1 = 15$, $s_1 = 3.07$, $gl = 14$

Estación 2: $n_2 = 12$, $s_2 = 0.80$, $gl = 11$

$$\begin{aligned} \left(\frac{s^2}{S^2}\right)\left(\frac{1}{f_{\alpha/2}(gl_1, gl_2)}\right) &\leq \frac{\sigma^2}{\sigma^2} \leq (f_{\alpha/2}(gl_1, gl_2)) \\ (3.07^2/0.80^2)(1/4.30) &\leq \frac{\sigma^2}{\sigma^2} \leq (3.07^2/0.80^2)(3.87) \\ 3.425 &\leq \frac{\sigma^2}{\sigma^2} \leq 56.99 \end{aligned}$$

2.5 Determinación del tamaño de muestra.

El papel del tamaño de la muestra para obtener una buena aproximación a los intervalos de confianza en los casos en que la población no tiene una distribución normal, ahora se enfoca la atención a otro aspecto relacionado con el tamaño de

la muestra. Se describe cómo elegir un tamaño de muestra suficientemente grande para obtener un margen de error deseado.

2.5.1 Basado en la media de la Población.

Este tamaño de muestra proporciona el margen de error deseado al nivel de confianza elegido. En la ecuación E es el margen de error que el usuario está dispuesto a aceptar, y el valor $z_{\alpha/2}$ es consecuencia directa del nivel de confianza que se va a usar para calcular la estimación por intervalo. A reserva de la decisión del usuario, 95% de confianza es el valor más usado ($z_{0.025} = 1.96$). Y para usar la ecuación es necesario contar con el valor de la desviación estándar poblacional σ . Sin embargo, aun cuando este valor no se conozca, puede usarse la ecuación siempre que se tenga un valor preliminar o un valor planeado de σ . En la práctica, se suele usar alguno de los procedimientos siguientes para obtener este valor planeado de σ .

1. Usar como valor planeado de σ una estimación de la desviación estándar poblacional calculada a partir de datos de estudios anteriores.
2. Emplear un estudio piloto seleccionando una muestra preliminar. La desviación estándar muestral obtenida de la muestra preliminar puede usarse como valor planeado de σ .
3. Use su juicio para el valor de σ .

Por ejemplo, se puede empezar por estimar el mayor y el menor valor en los datos de la población. Esta diferencia entre el mayor y el menor valor proporciona una estimación del rango de los datos. Por último, este valor dividido entre 4 suele considerarse como una aproximación burda a la desviación estándar y tomarse como un valor planeado aceptable de σ .

La fórmula para determinar el tamaño de muestra para una estimación por intervalo de la media poblacional es la siguiente.

$$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2}$$

Ejemplo resuelto:

En un estudio previo para investigar el costo de la renta de automóviles en Estados Unidos se encontró que el costo medio de la renta de un automóvil mediano era aproximadamente \$55 por día. Suponga que la organización que

realizó dicho estudio quiere realizar un nuevo estudio para estimar la media poblacional de las rentas por día de automóviles medianos en Estados Unidos. Antes de iniciar, especificó que la media poblacional de las rentas por día debe estimarse con un margen de error de \$2 y que se desea un nivel de 95% de confianza. El margen de error especificado es $E = 2$, el nivel 95% de confianza indica que $z_{0.025} = 1.96$. Por tanto, sólo falta un valor planeado de la desviación estándar poblacional σ para calcular el tamaño de muestra deseado. El analista revisó los datos muestrales del estudio anterior y encontró que la desviación estándar poblacional del costo de la renta diaria era \$9.65. usando \$9.65 como valor planeado de σ , se tiene

$$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2} = \frac{(1.96)^2 (9.65)^2}{2^2} = 89.43$$

De esta manera, el tamaño de la muestra necesario para obtener un margen de error de \$2 debe ser de por lo menos 89.43 rentas de automóviles medianos. En casos como éste, en los que el valor de n no es un número entero, se redondea al siguiente valor entero; así que el tamaño de muestra que se aconseja es 90 rentas de automóviles medianos.

2.5.2 Basado en la proporción de la Población

Cuál debe ser el tamaño de la muestra para obtener una estimación de la proporción poblacional con una precisión determinada.

La función que tiene el tamaño de la muestra en la determinación de la estimación por intervalo de p es semejante a la que tiene en la estimación de la media poblacional. El margen de error en la estimación por intervalo de la proporción poblacional es $Z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$. Con p^* como valor planeado de la fórmula para calcular

el tamaño de la muestra con el que se obtendrá el error E queda como se presenta a continuación.

$$n = \frac{(z_{\alpha/2})^2 p^*(1 - p^*)}{E^2}$$

En la práctica, el valor planeado p^* se determina mediante alguno de los métodos siguientes.

1. Utilizar la proporción poblacional de una muestra previa de las mismas unidades o de unidades similares.
2. Utilizar un estudio piloto y elegir una muestra preliminar. La proporción muestral de esta muestra se usa como valor planeado, p^* .
3. Proponer una “mejor aproximación” para el valor de p^* .
4. Si no aplica ninguna de las alternativas anteriores, emplear como valor planeado $p^* = 0.50$.

Ejemplo resuelto:

De regreso al estudio con mujeres golfistas, suponga que la empresa desea llevar a cabo otro estudio para determinar la proporción en la población de golfistas que está satisfecha con la disponibilidad de horarios de salida. ¿De qué tamaño deberá ser la muestra si se desea que en la estimación de la proporción poblacional el margen de error sea 0.025 a 95% de confianza? Como $E = .025$ y $z_{\alpha/2} = 1.96$, se necesita un valor planeado p^* para responder esta pregunta sobre el tamaño de la muestra. Utilizando como valor planeado p^* , el resultado del estudio anterior, $\bar{p} = 0.44$.

$$n = \frac{(z_{\alpha/2})^2 p^*(1 - p^*)}{E^2} = \frac{(1.96)^2 (0.44)(1 - 0.44)}{(0.025)^2} = 1514.5$$

Ejercicios Propuestos unidad 2:

- 1.- En una muestra aleatoria simple de 50 artículos de una población en la que $\sigma = 6$ la media muestral fue 32.
 - a. Proporcione un intervalo de confianza de 90% para la media poblacional.
 - b. Establezca un intervalo de confianza de 95% para la media poblacional.
 - c. Proporcione un intervalo de confianza de 99% para la media poblacional.

- 2.- Con objeto de estimar la cantidad media que gasta un cliente en una comida en un importante restaurante, se recogieron los datos de una muestra de 49 clientes. Suponga que la desviación estándar de la población es \$5.
 - a. ¿Cuál es el margen de error para 95% de confianza?
 - b. Si la media poblacional es \$24.80, ¿cuál es el intervalo de confianza de 95% para la media poblacional?

- 3.- Nielsen Media Research llevó a cabo un estudio para saber cuánto tiempo se veía televisión en los hogares, en el horario de 8:00 a 11:00 de la noche. Los datos que se encuentran en el archivo Nielsen del disco compacto son consistentes con los hallazgos reportados (The World Almanac, 2003). Con base en estudios anteriores, la desviación estándar poblacional se considera conocida y es $\sigma = 3.5$ horas. Dé una estimación mediante un intervalo de confianza de 95% para la media del tiempo que se ve televisión por semana en el horario de 8:00 a 11:00 de la noche.

- 4.- The National Quality Research Center de la University of Michigan proporciona medidas trimestrales de las opiniones de los consumidores acerca de productos y servicios (The Wall Street Journal, 18 de febrero de 2003). En una encuesta sobre 10 restaurantes de comida rápida y pizza la media del índice de satisfacción de los consumidores fue 71. Datos anteriores indican que la desviación estándar ha sido relativamente estable y es $\sigma = 5$.
 - a. ¿Qué debe estar dispuesto a suponer el investigador si desea un margen de error?
 - b. Con 95% de confianza, ¿cuál es el margen de error?
 - c. ¿Cuál es el margen de error si se desea 99% de confianza?

5.- La durabilidad de 2 tipos de neumáticos para automóviles se comparó mediante muestras de prueba en carretera de 100 neumáticos de cada tipo. El número de millas hasta el desgaste final se definió como una cantidad específica de uso del neumático. Los resultados son:

Neumáticos

I

Media = 26400

Varianza (s^2) = 1,440,000

II

Media = 25,100

Varianza (s^2) = 1,960,000

Estime la diferencia en millas promedio hasta el desgaste final con un 99% de nivel de confianza. ¿Hay una diferencia en la durabilidad promedio para los 2 tipos de neumáticos?

6.- Se considera cierto cambio en un proceso de fabricación de partes de componentes. Se toman muestras del procedimiento existente y del nuevo para determinar si este tiene como resultado una mejoría. Si se encuentra que 75 de 1500 artículos del procedimiento actual son defectuosos, y 80 de 2000 artículos del procedimiento nuevo también lo son. Encuentre un I.C. del 90% para la diferencia real en la fracción de defectuosos

7.- Una muestra aleatoria de 20 estudiantes, obtiene una media de 72 y una varianza de 16 en un examen de colocación en matemáticas. Suponga que las calificaciones se distribuyen normalmente. Construya un I.C. del 98% para la varianza poblacional.

8.- Una máquina produce piezas metálicas de forma cilíndricas. Se toma una muestra de las piezas y los diámetros son:

1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01, 1.03 cms.

Construya un I.C. del 99% para la σ^2

9.- Los siguientes datos representan los tiempos de duración de películas que producen 2 compañías cinematográficos:

Compañía 1	103	94	110	87	98		
Compañía 2	97	82	123	92	175	88	118

Calcule el I.C. del 90% para σ_1^2 / σ_2^2

10.- En un conjunto de datos se estima que el rango es 36.

- a. ¿Cuál es el valor planeado para la desviación estándar poblacional?
- b. ¿De qué tamaño deberá ser la muestra para que el margen de error en un intervalo de confianza de 95% sea 3?
- c. ¿De qué tamaño deberá ser la muestra para que el margen de error en un intervalo de confianza de 95% sea 2?

11.- El costo promedio de la gasolina sin plomo en Grater Cincinnati es \$2.41 (The Cincinnati Enquirer, 3 de febrero de 2006). En una época de cambios en los precios, un periódico muestrea las gasolineras y presenta un informe sobre los precios de la gasolina. Suponga que en los precios del galón de la gasolina sin plomo la desviación estándar es \$0.15; dé el tamaño de muestra n que debe usar este periódico para tener 95% de confianza con cada uno de los márgenes de error siguientes.

- a. Un margen de error de \$0.07
- b. Un margen de error de \$0.05
- c. Un margen de error de \$0.03

12.- Según Thomson Financial, hasta el 25 de enero de 2006, la mayor parte de las empresas que informaban tener ganancias habían superado las estimaciones (BusinessWeek, 6 de febrero de 2006). En una muestra de 162 empresas, 104 superaron las estimaciones, 29 coincidieron y 29 se quedaron cortas.

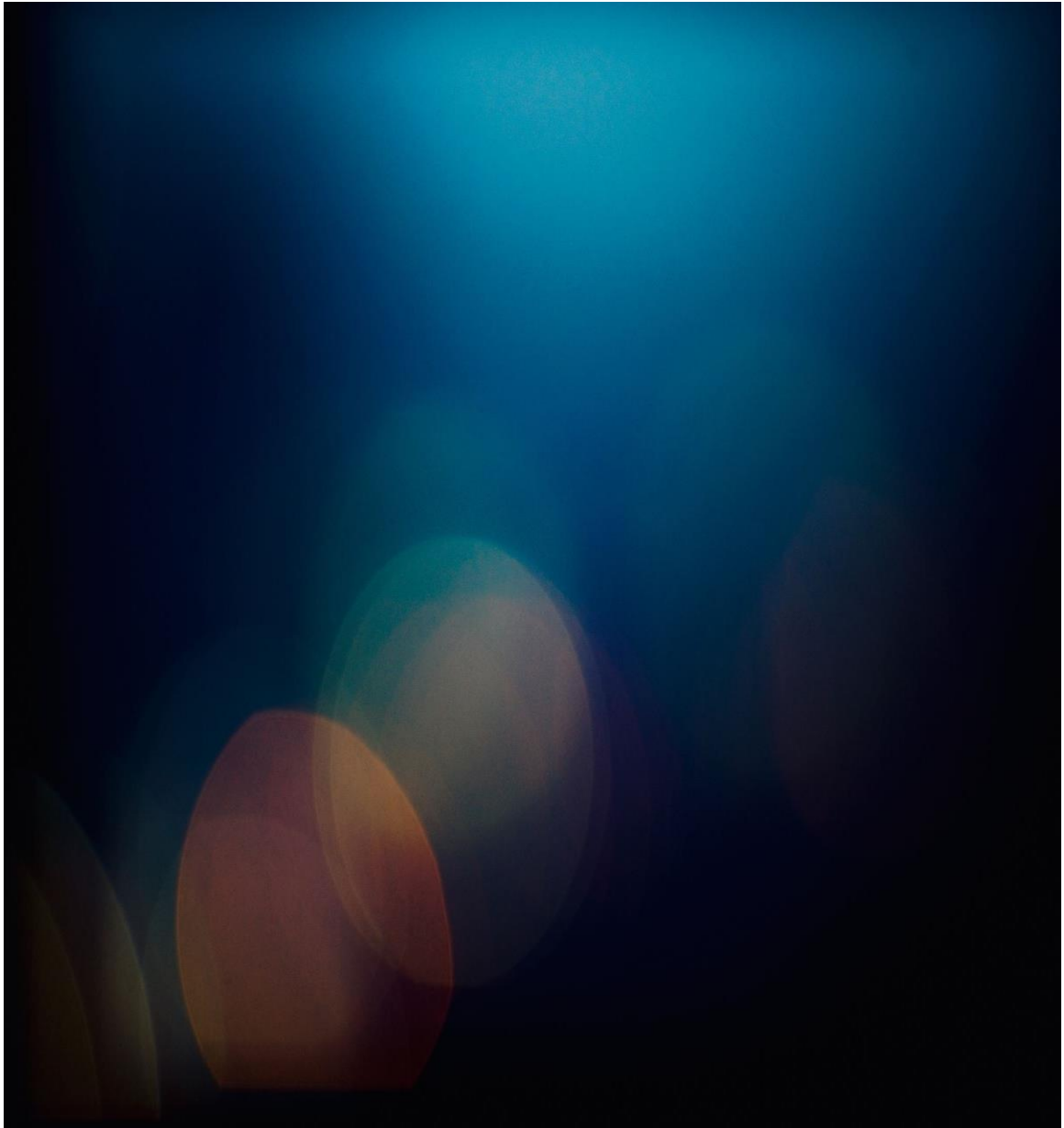
- a. ¿Cuál es la estimación puntual de la proporción de empresas que se quedaron cortas?
- b. Determine el margen de error y dé un intervalo de confianza de 95% para la proporción que superó las estimaciones.
- c. ¿De qué tamaño debe de ser la muestra si el margen de error es 0.05?

13.- En 2003 el porcentaje de personas que no tenía un seguro médico (en Estados Unidos) era 15.6% (Statistical Abstract of the United States, 2006). Se le pide a un comité del Congreso realizar un estudio para obtener información actualizada.

- a. ¿Qué tamaño de muestra le recomienda usted al comité, si el objetivo es que en la estimación de la proporción actual de individuos que no tienen seguro médico el margen de error sea 0.03? Use 95% de confianza.
- b. Repita el inciso a usando 99% de confianza.

Unidad 3.

Prueba de hipótesis.



Instrumentación Didáctica

Competencia Específica	Competencia Genérica	Actividades de Aprendizaje
Realiza pruebas de hipótesis para comparar si los valores de los estadísticos obtenidos de una muestra tienen una diferencia significativa con un valor supuesto asumiendo cierto nivel de confianza y tomando en cuenta los criterios de aceptación o rechazo en problemas de la industria y la logística que involucren errores tipo I o tipo II.	Capacidad de abstracción, análisis y síntesis. • Capacidad de aplicar los conocimientos en la práctica. • Capacidad para identificar, plantear y resolver problemas.	Buscar información sobre conceptos relacionados con pruebas de hipótesis. • Discutir en grupo sobre los conceptos investigados. • Formular y resolver ejercicios aplicando la metodología de prueba de hipótesis para la variable que se está midiendo y obtener tamaño de muestra para diferentes situaciones error tipo I, error tipo II y potencia de la prueba. • Simular un caso en donde se genere una hipótesis para una situación en donde el interés pueda ser, la media, diferencia de medias, proporción, diferencia de proporciones, varianza y relación de varianzas: Generar datos del caso. o Probar la hipótesis del caso. o Obtener conclusiones. o Cambiar el tamaño de muestra y mostrar su impacto. • Utilizar TIC's para realizar la prueba de hipótesis.

3.Pruebas de Hipótesis.

3.1 Introducción.

Las secciones anteriores han mostrado cómo puede estimarse un parámetro a partir de los datos contenidos en una muestra. Puede encontrarse ya sea un sólo número (estimador puntual) o un intervalo de valores posibles (intervalo de confianza). Sin embargo, muchos problemas de ingeniería, ciencia, y administración, requieren que se tome una decisión entre aceptar o rechazar una proposición sobre algún parámetro. Esta proposición recibe el nombre de hipótesis. Este es uno de los aspectos más útiles de la inferencia estadística, puesto que muchos tipos de problemas de toma de decisiones, pruebas o experimentos en el mundo de la ingeniería pueden formularse como problemas de prueba de hipótesis.

Una hipótesis estadística es una proposición o supuesto sobre los parámetros de una o más poblaciones.

Suponga que se tiene interés en la rapidez de combustión de un agente propulsor sólido utilizado en los sistemas de salida de emergencia para la tripulación de aeronaves. El interés se centra sobre la rapidez de combustión promedio. De manera específica, el interés recae en decir si la rapidez de combustión promedio es o no 50 cm/s. Esto puede expresarse de manera formal como

$$H_0; \mu = 50 \text{ cm/s} \quad H_1; \mu \neq 50 \text{ cm/s}$$

La proposición $H_0; \mu = 50 \text{ cm/s}$, se conoce como hipótesis nula, mientras que la proposición $H_1; \mu \neq 50 \text{ cm/s}$, recibe el nombre de hipótesis alternativa. Puesto que la hipótesis alternativa especifica valores de μ que pueden ser mayores o menores que 50 cm/s, también se conoce como hipótesis alternativa bilateral. En algunas situaciones, lo que se desea es formular una hipótesis alternativa unilateral, como en

$$H_0; \mu = 50 \text{ cm/s}$$

$$H_0; \mu = 50 \text{ cm/s}$$

ó

$$H_1; \mu < 50 \text{ cm/s}$$

$$H_1; \mu > 50 \text{ cm/s}$$

Es importante recordar que las hipótesis siempre son proposiciones sobre la población o distribución bajo estudio, no proposiciones sobre la muestra. Por lo general, el valor del parámetro de la población especificado en la hipótesis nula se determina en una de tres maneras diferentes:

1. Puede ser resultado de la experiencia pasada o del conocimiento del proceso, entonces el objetivo de la prueba de hipótesis usualmente es determinar si ha cambiado el valor del parámetro.
2. Puede obtenerse a partir de alguna teoría o modelo que se relaciona con el proceso bajo estudio. En este caso, el objetivo de la prueba de hipótesis es verificar la teoría o modelo.
3. Cuando el valor del parámetro proviene de consideraciones externas, tales como las especificaciones de diseño o ingeniería, o de obligaciones contractuales. En esta situación, el objetivo usual de la prueba de hipótesis es probar el cumplimiento de las especificaciones.

Un procedimiento que conduce a una decisión sobre una hipótesis en particular recibe el nombre de prueba de hipótesis. Los procedimientos de prueba de hipótesis dependen del empleo de la información contenida en la muestra aleatoria de la población de interés. Si esta información es consistente con la hipótesis, se concluye que ésta es verdadera; sin embargo, si esta información es inconsistente con la hipótesis, se concluye que esta es falsa. Debe hacerse hincapié en que la verdad o falsedad de una hipótesis en particular nunca puede conocerse con certidumbre, a menos que pueda examinarse a toda la población. Usualmente esto es imposible en muchas situaciones prácticas. Por tanto, es necesario desarrollar un procedimiento de prueba de hipótesis teniendo en cuenta la probabilidad de llegar a una conclusión equivocada.

La hipótesis nula, representada por H_0 , es la afirmación sobre una o más características de poblaciones que al inicio se supone cierta (es decir, la “creencia a priori”).

La hipótesis alternativa, representada por H_1 , es la afirmación contradictoria a H_0 , y ésta es la hipótesis del investigador. La hipótesis nula se rechaza en favor de la hipótesis alternativa, sólo si la evidencia muestral sugiere que H_0 es falsa. Si la muestra no contradice decididamente a H_0 , se continúa creyendo en la validez de la hipótesis nula. Entonces, las dos conclusiones posibles de un análisis por prueba de hipótesis son **rechazar H_0 o no rechazar H_0** .

3.2 Confiabilidad y significancia.

la confiabilidad de un instrumento se refiere a la constitución interna de las personas, a la mayor o menor acescencia de errores de medida. Un instrumento confiable significa que si lo aplicamos por más de una vez a un mismo elemento entonces obtendríamos iguales resultados.

3.3 Errores tipo I y tipo II.

La media muestral puede tomar muchos valores diferentes. Supóngase que si $48.5 \leq \bar{x} \leq 51.5$, entonces no se rechaza la hipótesis nula H_0 ; $m = 50$ cm/s, y que si $\bar{x} < 48.5$ o $\bar{x} > 51.5$, entonces se acepta la hipótesis alternativa H_1 ; $m \neq 50$ cm/s. Los valores de $\bar{x}$ que son menores que 48.5 o mayores que 51.5 constituyen la región crítica de la prueba, mientras que todos los valores que están en el intervalo $48.5 \leq \bar{x} \leq 51.5$ forman la región de aceptación. Las fronteras entre las regiones crítica y de aceptación reciben el nombre de valores críticos. La costumbre es establecer conclusiones con respecto a la hipótesis nula H_0 . Por tanto, se rechaza H_0 en favor de H_1 si el estadístico de prueba cae en la región crítica, de lo contrario, no se rechaza H_0 . Este procedimiento de decisión puede conducir a una de dos conclusiones erróneas. Por ejemplo, es posible que el valor verdadero de la rapidez promedio de combustión del agente propulsor sea igual a 50 cm/s. Sin embargo, para todos los especímenes bajo prueba, bien puede observarse un valor del estadístico de prueba $\bar{x}$ que cae en la región crítica. En este caso, la hipótesis nula H_0 será rechazada en favor de la alternativa H_1 cuando, de hecho, H_0 en realidad es verdadera. Este tipo de conclusión equivocada se conoce como error tipo I.

El error tipo I se define como el rechazo de la hipótesis nula H_0 cuando ésta es verdadera. También es conocido como α o nivel de significancia.

Si tuviéramos un nivel de confianza del 95% entonces el nivel de significancia sería del 5%. Análogamente si se tiene un nivel de confianza del 90% entonces el nivel de significancia sería del 10%. Ahora supóngase que la verdadera rapidez promedio de combustión es diferente de 50 cm/s, aunque la media muestral $\bar{x}$ caiga dentro de la región de aceptación. En este caso se acepta H_0 cuando ésta es falsa. Este tipo de conclusión recibe el nombre de **error tipo II**.

El error tipo II ó error β se define como la aceptación de la hipótesis nula cuando ésta es falsa.

Por tanto, al probar cualquier hipótesis estadística, existen cuatro situaciones diferentes que determinan si la decisión final es correcta o errónea.

Decisión	<i>Ho es verdadera</i>	<i>Ho es falsa</i>
Aceptar H_0	<i>No hay error</i>	<i>Error tipo II ó β</i>
Rechazar H_0	<i>Error tipo I ó α</i>	<i>No hay error</i>

1. Los errores tipo I y tipo II están relacionados. Una disminución en la probabilidad de uno por lo general tiene como resultado un aumento en la probabilidad del otro.
2. El tamaño de la región crítica, y por tanto la probabilidad de cometer un error tipo I, siempre se puede reducir al ajustar el o los valores críticos.
3. Un aumento en el tamaño muestral n reducirá a y b de forma simultánea.
4. Si la hipótesis nula es falsa, b es un máximo cuando el valor real del parámetro se aproxima al hipotético. Entre más grande sea la distancia entre el valor real y el valor hipotético, será menor β .

3.4 Potencia de la prueba.

La potencia de una prueba o el poder estadísticos es la probabilidad de que la hipótesis alternativa sea aceptada cuando la hipótesis alternativa es verdadera (es decir, la probabilidad de no cometer un error del tipo II). La potencia es en general una función de las distribuciones posibles, a menudo determinada por un parámetro, bajo la hipótesis alternativa. A medida que aumenta la potencia, las posibilidades de que se presente un error del tipo II disminuyen. La probabilidad de que ocurra un error de tipo II se conoce como la tasa de falsos negativos (β). Por lo tanto, la potencia es igual a $1 - \beta$, que también se conoce como la sensibilidad

El *poder estadístico* de un estudio depende de diferentes factores, como:

- *El tamaño del efecto a detectar*, es decir, la magnitud mínima de la diferencia o asociación entre los grupos que se considera clínicamente relevante. Cuanto mayor sea el tamaño del efecto que se desea detectar,

mayor será la probabilidad de obtener hallazgos significativos y, por lo tanto, mayor será el poder estadístico.

- La *variabilidad* de la respuesta estudiada. Así, cuanto mayor sea la variabilidad en la respuesta, más difícil será detectar diferencias entre los grupos que se comparan y menor será el poder estadístico de la investigación. De ahí que sea recomendable estudiar grupos lo más homogéneos posibles.

- El *tamaño de la muestra* a estudiar. Cuanto mayor sea el tamaño muestral, mayor será la potencia estadística de un estudio. Es por ello que en los estudios con muestras muy grandes se detectan como significativas diferencias poco relevantes, y en los estudios con muestras menores es más fácil obtener resultados falsamente negativos.

- El *nivel de significación estadística*. Si se disminuye el valor de α también se disminuye el poder de la prueba. Es decir, si disminuimos la probabilidad de cometer un error de tipo I aumentamos simultáneamente la probabilidad de un error de tipo II, por lo que se trata de encontrar un punto de “equilibrio” entre ambas. Habitualmente se trabaja con un nivel de significación del 95% ($\alpha = 0.05$), por lo que el equilibrio hay que encontrarlo finalmente entre el tamaño de la muestra que es posible estudiar y el poder que se quiere para el estudio.

Los cuatro factores anteriores, junto con el poder estadístico, forman un sistema cerrado. De este modo, una vez fijados tres de ellos, el cuarto queda completamente determinado.

3.5 Formulación de Hipótesis estadísticas.

Se pueden presentar tres tipos de ensayo de hipótesis que son:

1. Unilateral Derecho
2. Unilateral Izquierdo
3. Bilateral

Dependiendo de la evaluación que se quiera hacer se seleccionará el tipo de formulación.

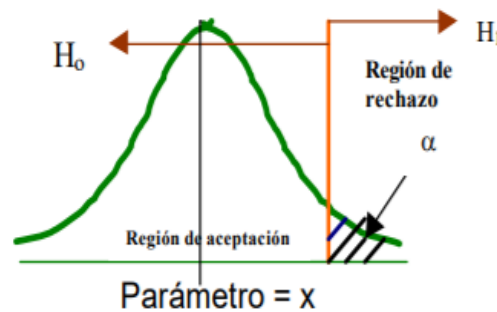
1. Unilateral Derecho. El investigador desea comprobar la hipótesis de un

aumento en el parámetro, en este caso el nivel de significancia se carga todo hacia el lado derecho, para definir las regiones de aceptación y de rechazo.

Ensayo de hipótesis:

H_0 : Parámetro $\leq x$

H_1 : Parámetro $> x$

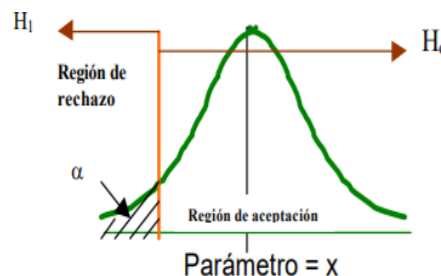


2. Unilateral Izquierdo: El investigador desea comprobar la hipótesis de una disminución en el parámetro, en este caso el nivel de significancia se carga todo hacia el lado izquierdo, para definir las regiones de aceptación y de rechazo.

Formulación de hipótesis:

H_0 : Parámetro $\geq x$

H_1 : Parámetro $< x$

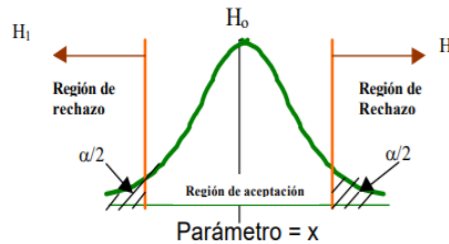


3. Bilateral: El investigador desea comprobar la hipótesis de un cambio en el parámetro. El nivel de significancia se divide en dos y existen dos regiones de rechazo.

Formulación de hipótesis:

H_0 : Parámetro = x

H_1 : Parámetro $\neq x$



Uso de valores P para la toma de decisiones

Al probar hipótesis en las que la estadística de prueba es discreta, la región crítica se puede elegir de forma arbitraria y determinar su tamaño. Si α es demasiado grande, se puede reducir al hacer un ajuste en el valor crítico. Puede ser necesario aumentar el tamaño de la muestra para compensar la disminución que ocurre de manera automática en la potencia de la prueba (probabilidad de rechazar H_0 dado que una alternativa específica es verdadera).

Por generaciones enteras de análisis estadístico, se ha hecho costumbre elegir un nivel de significancia de 0.05 ó 0.01 y seleccionar la región crítica en consecuencia. Entonces, por supuesto, el rechazo o no rechazo estricto de H_0 dependerá de esa región crítica. En la estadística aplicada los usuarios han adoptado de forma extensa la aproximación del valor P. La aproximación se diseña para dar al usuario una alternativa a la simple conclusión de “rechazo” o “no rechazo”.

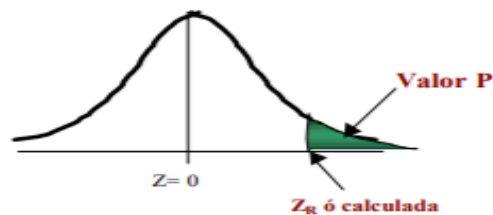
La aproximación del valor P como ayuda en la toma de decisiones es bastante natural pues casi todos los paquetes de computadora que proporcionan el cálculo de prueba de hipótesis entregan valores de P junto con valores de la estadística de la prueba apropiada.

- Un valor P es el nivel (de significancia) más bajo en el que el valor observado de la estadística de prueba es significativo.
- El valor P es el nivel de significancia más pequeño que conduce al rechazo de la hipótesis nula H_0 .
- El valor P es el mínimo nivel de significancia en el cual H_0 sería rechazada cuando se utiliza un procedimiento de prueba especificado con un conjunto dado de información. Una vez que el valor de P se haya determinado, la conclusión en cualquier nivel α particular resulta de comparar el valor P con α :

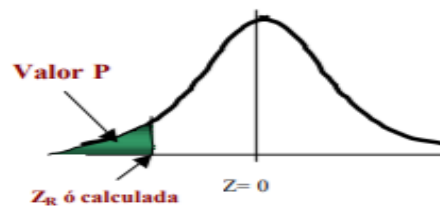
1. Valor $P \leq \alpha$ rechazar H_0 al nivel α .
2. Valor $P > \alpha$ No rechazar H_0 al nivel α .

Para realizar los ejemplos y ejercicios de ensayo de hipótesis se recomienda seguir los pasos mencionados anteriormente. Los ejemplos siguientes se solucionarán por los pasos recomendados, teniéndose una variedad de problemas en donde se incluirán a todas las distribuciones muestrales que se han visto hasta aquí.

Ensayo Unilateral Derecho:

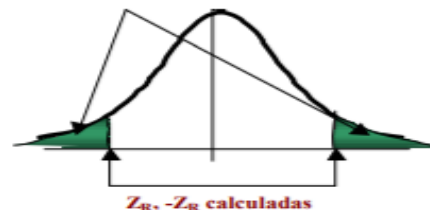


Ensayo Unilateral Izquierdo:



Ensayo Bilateral:

Valor P = Suma de las dos áreas



3.6 Prueba de hipótesis para la media.

Se tiene una población de la cual se selecciona una muestra de tamaño n , a la cual se le obtienen sus estadísticos para realizar la prueba de hipótesis donde se supondrá el parámetro de la población μ , el cual se le asigna el valor supuesto. Esto se realiza mediante la formulación del establecimiento de la hipótesis Nula y la Hipótesis alternativa.

Ejemplo resuelto:

1.- Una muestra aleatoria de 100 muertes registradas en Estados Unidos el año pasado muestra una vida promedio de 71.8 años. Suponga una desviación estándar poblacional de 8.9 años, ¿esto parece indicar que la vida media hoy en día es mayor que 70 años? Utilice un nivel de significancia de 0.05.

Solución:

1. Se trata de una distribución muestral de medias con desviación estándar conocida.

2. Datos:

$$\mu = 70 \text{ años}$$

$$\sigma = 8.9 \text{ años}$$

$$\bar{x} = 71.8 \text{ años}$$

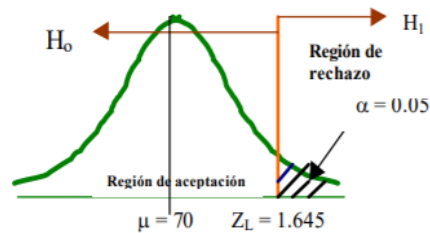
$$n = 100$$

$$\alpha = 0.05$$

3. Ensayo de hipótesis

$$H_0: \mu = 70 \text{ años.}$$

$$H_1: \mu > 70 \text{ años.}$$



4.Regla de decisión:

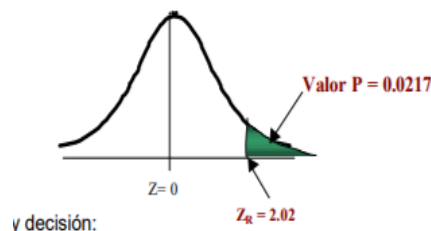
Si $P \leq .05$ se rechaza H_0 .

Si $P > .05$ No se rechaza H_0 .

5.Cálculos:

$$Z_R = \frac{\bar{x}_R - \mu}{\sigma / \sqrt{n}} = \frac{71.8 - 70}{8.9 / \sqrt{100}} = 2.02$$

Esta es el valor de Z que se utilizará para calcular el valor P, como es un ensayo unilateral derecho se calculará el área a la derecha de este valor



6.Justificación y decisión.

Como el valor P es 0.217 y es menor al valor del nivel de significancia de 0.05 por lo tanto se rechaza H_0 y se concluye que la edad media de los habitantes es mayor a 70 años.

3.7 Prueba de hipótesis para la diferencia de medias.

Existe 2 poblaciones donde se extraer 1 muestra a cada una y se analiza las medias a las muestras.

Un diseñador de productos está interesado en reducir el tiempo de secado de una pintura tapaporos. Se prueban dos fórmulas de pintura; la fórmula 1 tiene el contenido químico estándar, y la fórmula 2 tiene un nuevo ingrediente secante que debe reducir el tiempo de secado. De la experiencia se sabe que la desviación estándar del tiempo de secado es ocho minutos, y esta variabilidad inherente no debe verse afectada por la adición del nuevo ingrediente.

Ejemplo resuelto:

Se pintan diez especímenes con la fórmula 1, y otros diez con la fórmula 2. Los dos tiempos promedio de secado muestrales son 121 min y 112 min respectivamente. ¿A qué conclusiones puede llegar el diseñador del producto sobre la eficacia del nuevo ingrediente, utilizando $\alpha = 0.05$?

Solución:

1. Se trata de una distribución muestral de diferencia de medias con desviación estándar conocida.

2. Datos:

$$\sigma_1 = \sigma_2 = 8$$

$$\bar{x}_1 = 121 \text{ min}$$

$$\bar{x}_2 = 112 \text{ min}$$

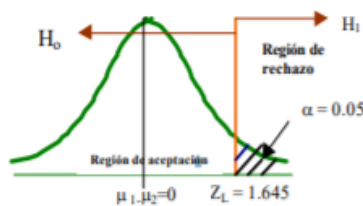
$$n_1 = n_2 = 10$$

$$\alpha = 0.05$$

3. Ensayo de hipótesis

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 > 0$$



Se desea rechazar H_0 si el nuevo ingrediente disminuye el tiempo promedio de secado, por eso se pone la diferencia mayor a cero o sea positiva para poder probar que μ_2 es menor que μ_1 .

4. Regla de decisión:

Si $P \leq 1.645$ se rechaza H_0 .

Si $P > 1.645$ No se rechaza H_0 .

5. Cálculos:

$$Z_k = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(121 - 112) - 0}{\sqrt{\frac{8^2}{10} + \frac{8^2}{10}}} = 2.52$$

Valor $P = .0059$, por lo tanto, $P < .05$ y H_0 se rechaza.

5. Justificación y decisión: y se concluye con un nivel de significancia de 0.05 que la adición del nuevo ingrediente a la pintura si disminuye de manera significativa el tiempo promedio de secado.

3.8 Prueba de hipótesis para la proporción.

Se tiene una población de la cual se obtiene una muestra, a la cual se le obtienen sus estadísticos para realizar la prueba de hipótesis donde se supondrá el valor de la Proporción de la población.

Ejemplo resuelto:

Un constructor afirma que se instalan bombas de calor en 70% de todas las casas que se construyen hoy en día en la ciudad de Richmond. ¿Estaría de acuerdo con esta afirmación si una investigación de casas nuevas en esta ciudad muestra que 8 de 15 tienen instaladas bombas de calor? Utilice un nivel de significancia de 0.10.

Solución:

1. Se trata de una distribución muestral de proporciones.

2. Datos:

$$P = 0.70$$

$$p = 8/15 = 0.5333$$

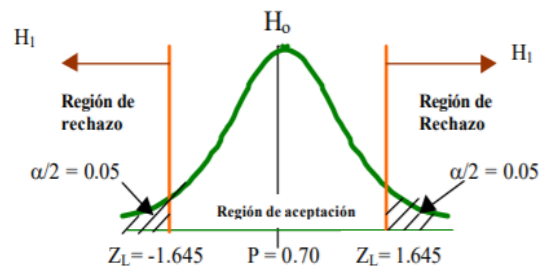
$$n = 15$$

$$\alpha = 0.10$$

3. Ensayo de hipótesis

$$H_0; P = 0.70$$

$$H_1; P \neq 0.70$$



4. Regla de Decisión:

P es mayor a alpha, No se rechaza H_0

Si P es menor a Alpha Se rechaza H_0

5. Cálculos:

$$Z_R = \frac{p - P}{\sqrt{\frac{Pq}{n}}} = \frac{0.533 - 0.70}{\sqrt{\frac{(0.70)(0.30)}{15}}} = -1.41$$

P de un solo lado es = .1586, pero como es una prueba bilateral, existen 2 P por lo tanto se multiplica por 2 para obtener un solo valor de P. $(.0793)(2) = .1586 = P$

6. Justificación y decisión:

Como P es mayor a α , No se rechaza H_0 y se concluye con un nivel de significancia de 0.10 que la afirmación del constructor es cierta.

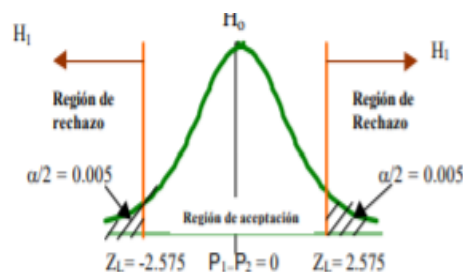
3.9 Prueba de hipótesis para la diferencia de proporciones.

Existe 2 poblaciones donde se extraer 1 muestra a cada una y se analiza las proporciones de las muestras.

Se evalúan dos tipos diferentes de soluciones para pulir, para su posible uso en una operación de pulido en la fabricación de lentes intraoculares utilizados en el ojo humano después de una cirugía de cataratas. Se pulen 300 lentes con la primera solución y, de éstos, 253 no presentaron defectos inducidos por el pulido. Después se pulen otros 300 lentes con la segunda solución, de los cuales 196 resultan satisfactorios. ¿Existe alguna razón para creer que las dos soluciones para pulir son diferentes? Utilice $\alpha = 0.01$

Solución:

1. Se trata de una distribución muestral de diferencia de proporciones.
2. Datos:
 $p_1 = 253/300 = 0.8433$
 $p_2 = 196/300 = 0.6533$
 $n_1 = n_2 = 300$
3. Ensayo de hipótesis:
 $H_0: P_1 - P_2 = 0$
 $H_1: P_1 - P_2 \neq 0$



4. Regla de Decisión:
 Si P es mayor a Alfa, No se rechaza H_0
 Si P es menor a Alfa Se rechaza H_0
5. Cálculos:

$$Z_R = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$$

En esta fórmula se puede observar que en el denominador se tienen a las proporciones poblacionales o sea los parámetros, los cuales no se conocen, por lo que en el ensayo de hipótesis la fórmula para poder calcular la Z_R cambia, estimando a el parámetro común P de la siguiente forma:

$$P = \frac{X_1 + X_2}{n_1 + n_2} \quad \text{o bien} \quad P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

Entonces la fórmula de Z_R quedaría de la siguiente manera: Entonces la fórmula de Z_R quedaría de la siguiente manera:

$$Z_R = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Se calculará el valor de P

$$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{253 + 196}{300 + 300} = 0.7483$$

$$Z_R = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{Pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.8433 - 0.6533) - 0}{\sqrt{(0.7483)(0.2517)\left(\frac{1}{300} + \frac{1}{300}\right)}} = 5.36$$

6. Justificación y decisión:

Puesto que es una prueba bilateral, se obtiene el valor P de un solo lado y se multiplica por 2 para obtener nuevamente el valor P total. Entonces el valor $P = .0013(2) = .0026 = P$ se rechaza la hipótesis nula y se concluye con un nivel de significancia de 0.01 que los dos fluidos para pulir son diferentes. 10. Se tomará el

voto entre los residentes de una ciudad y el condado circundante para determinar si se debe construir una planta química propuesta. El lugar de construcción está dentro de los límites de la ciudad y por esta razón muchos votantes del condado consideran que la propuesta pasará debido a la gran proporción de votantes que favorecen la construcción. Para determinar si hay una diferencia significativa en la proporción de votantes de la ciudad y votantes del condado que favorecen la propuesta, se realiza una encuesta. Si 120 de 200 votantes de la ciudad favorecen la propuesta y 240 de 500 residentes del condado también lo hacen, ¿estaría de acuerdo en que la proporción de votantes de la ciudad que favorecen la propuesta es más alto que la proporción de votantes del condado? Utilice un nivel de significancia de 0.025.

3.10 Prueba de hipótesis para la varianza.

Con σ_0^2 para denotar el valor hipotético de la varianza poblacional, las tres formas de una prueba de hipótesis para la varianza poblacional son las siguientes:

$$\begin{array}{lll} H_0: \sigma^2 \geq \sigma_0^2 & H_0: \sigma^2 \leq \sigma_0^2 & H_0: \sigma^2 = \sigma_0^2 \\ H_a: \sigma^2 < \sigma_0^2 & H_a: \sigma^2 > \sigma_0^2 & H_a: \sigma^2 \neq \sigma_0^2 \end{array}$$

En una prueba de hipótesis para la varianza poblacional se emplea el valor hipotético de la varianza poblacional σ_0^2 y la varianza muestral s^2 para calcular el valor del estadístico de prueba X^2 . Si la población tiene una distribución normal, el estadístico de prueba es el siguiente:

$$\chi^2 = \frac{(n - 1)s^2}{\sigma_0^2}$$

Donde X^2 tiene una distribución chi-cuadrada con $n-1$ grados de libertad.

Una vez calculado el valor del estadístico de prueba x^2 , para determinar si se rechaza la hipótesis nula se emplea el método del valor P o el método del valor crítico.

Prueba de Hipótesis para la Varianza Poblacional

	Prueba de la cola inferior	Prueba de la cola superior	Prueba de dos colas
Hipótesis	$H_0: \sigma^2 \geq \sigma_0^2$ $H_a: \sigma^2 < \sigma_0^2$	$H_0: \sigma^2 \leq \sigma_0^2$ $H_a: \sigma^2 > \sigma_0^2$	$H_0: \sigma^2 = \sigma_0^2$ $H_a: \sigma^2 \neq \sigma_0^2$
Estadístico de prueba	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$
Regla de rechazo: método del valor-p	Rechazar H_0 si $\text{valor-p} \leq \alpha$	Rechazar H_0 si $\text{valor-p} \leq \alpha$	Rechazar H_0 si $\text{valor-p} \leq \alpha$
Regla de rechazo: método del valor crítico	Rechazar H_0 si $\chi^2 \leq \chi_{(1-\alpha)}^2$	Rechazar H_0 si $\chi^2 \geq \chi_{\alpha}^2$	Rechazar H_0 si $\chi^2 \leq \chi_{(1-\alpha/2)}^2$ o si $\chi^2 \geq \chi_{\alpha/2}^2$

Ejemplo resuelto:

La St. Louis Metro Bus Company de Estados Unidos, desea dar una imagen de confiabilidad haciendo que sus conductores sean puntuales en los horarios de llegada de paradas. La empresa desea que haya poca variabilidad en dichos tiempos. En términos de la varianza de los tiempos de llegada a las paradas, la empresa desea que la varianza sea de 4 minutos o menos. Se formula la siguiente prueba de hipótesis para que la empresa pueda determinar si la varianza poblacional en los tiempos de llegada a las paradas es demasiado grande.

$$H_0 = \sigma^2 \leq 4$$

$$H_a = \sigma^2 > 4$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(24-1)(4.9)}{4} = 28.18$$

De acuerdo con el estadístico de prueba chi- cuadrada con $n-1=24-1=23$ grados de libertad. Como es una prueba de la cola superior, el área bajo la curva a la derecha del valor del estadístico de prueba $X^2= 28.18$ es el valor -p de la prueba.

Las tablas de la distribución chi-cuadrada no son suficientes detalladas para permitir determinar con exactitud el valor-p. Sin embargo, con las tablas de la distribución chi-cuadrada se obtiene un intervalo en el que se encuentra el valor-p. por ejemplo:

Área en la cola superior	0.10	0.05	0.025	0.01
Valor de X^2 (23 gl)	32.007	35.172	38.076	41.638

$$X^2=28.18$$

Como $X^2=28.18$ es menor que 32.007, el área en la cola superior (el valor-p) es mayor que 0.10. Como el valor-p es $>\alpha=0.05$, no se puede rechazar la hipótesis nula. La maestra no lleva a la conclusión de que la varianza en los tiempos de llegada a las paradas sea demasiado grande.

3.11 Prueba de hipótesis para la relación de varianzas.

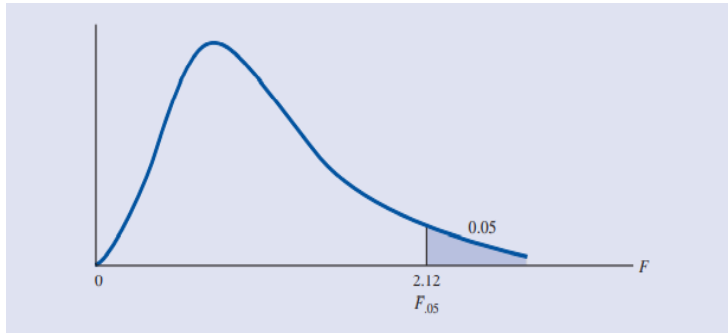
En algunas aplicaciones estadísticas interesa comparar las varianzas de las calidades de producto obtenido mediante dos métodos de producción diferentes, o las varianzas de tiempos de fabricación empleando dos métodos diferentes, o las varianzas de las temperaturas que se tienen con dos dispositivos distintos de calentamiento. Para comparar dos varianzas poblacionales, se emplean datos obtenidos de dos muestras aleatorias independientes, una de la población 1 y otra de la población 2. Para hacer las inferencias acerca de las dos varianzas poblacionales σ_1^2 y σ_2^2 se usan las dos varianzas muestrales S_1^2 y S_2^2 . Cuando las varianzas de dos poblaciones normales son iguales ($\sigma_1^2 = \sigma_2^2$), la distribución muestral de la proporción entre las dos varianzas muestrales S_1^2 / S_2^2 es la siguiente:

Cuando se toman muestras aleatorias simples independientes de tamaños n_1 y n_2 de dos poblaciones normales con varianzas iguales, la distribución muestral de

$$S_1^2 / S_2^2$$

es una distribución F con $n_1 - 1$ grados de libertad en el numerador y $n_2 - 1$ grados de libertad en el denominador; S_1^2 es la varianza muestral de la muestra aleatoria de n_1 elementos tomados de la población 1, y S_2^2 es la varianza muestral de la muestra aleatoria de n_2 elementos tomados de la población 2.

El estadístico de Prueba es la Distribución F, como se puede ver en la siguiente gráfica, la distribución F no es simétrica y los valores de F no pueden ser negativos. La forma de cada distribución F depende de los grados de libertad en el numerador y de los grados de libertad del denominador. La siguiente grafica muestra una distribución F con 20 grados de libertad en el numerador como en el denominador.



Ahora se verá cómo usar la distribución F para realizar una prueba de hipótesis para las varianzas de dos poblaciones. Se empieza con una prueba sobre la igualdad de las dos varianzas poblacionales. Las hipótesis son las siguientes:

$$H_0 = \sigma_1^2 = \sigma_2^2$$

$$H_a = \sigma_1^2 \neq \sigma_2^2$$

De manera tentativa se supone que las varianzas poblacionales son iguales. Si se rechaza H_0 se llega a la conclusión de que las varianzas poblacionales no son iguales.

Para realizar esta prueba de hipótesis se requieren dos muestras aleatorias independientes, una de cada población. Se calculan las dos varianzas muestrales. A la población en la que se encuentre la mayor varianza muestral se le considera la población 1. De manera que el tamaño de muestra n_1 y la varianza muestral s_1^2

corresponden a la población 1 y el tamaño de muestra n_2 y la varianza muestral s_2^2 corresponden a la población 2. Con base en la suposición de que las dos poblaciones tengan una distribución normal, la relación entre las varianzas muestrales proporciona el siguiente estadístico de prueba F.

estadístico de prueba para pruebas de hipótesis acerca de varianzas poblacionales con $\sigma_1^2 \neq \sigma_2^2$

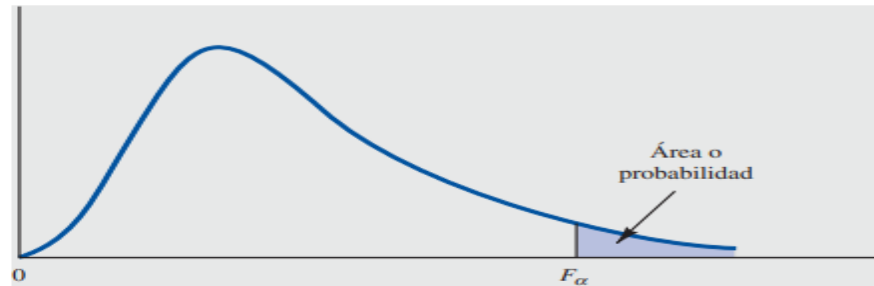
$$F = S_1^2 / S_2^2$$

Al denotar como población 1 a la población que tiene mayor varianza muestral, el estadístico de prueba tiene una distribución F con $n_1 - 1$ grados de libertad en el numerador y con $n_2 - 1$ grados de libertad en el denominador.

Resumen de las pruebas de Hipótesis a cerca de dos varianzas poblacionales

Hipótesis	Prueba de la cola superior	Prueba de la cola inferior
	$H_0: \sigma_1^2 \leq \sigma_2^2$ $H_a: \sigma_1^2 > \sigma_2^2$	$H_0: \sigma_1^2 = \sigma_2^2$ $H_a: \sigma_1^2 \neq \sigma_2^2$ Nota: La población 1 tiene la varianza muestral más grande
Estadístico de prueba	$F = \frac{s_1^2}{s_2^2}$	$F = \frac{s_1^2}{s_2^2}$
Regla de rechazo: método del valor-p	Rechazar H_0 si valor-p $\leq \alpha$	Rechazar H_0 si valor-p $\leq \alpha$
Regla de rechazo: método del valor crítico	Rechazar H_0 si $F \geq F_{\alpha}$	Rechazar H_0 si $F \geq F_{\alpha/2}$

Tabla de valores de la distribución F



Grados de libertad en el denominador	Área en la cola superior	Grados de libertad en el numerador				
		10	15	20	25	30
10	0.10	2.32	2.24	2.20	2.17	2.16
	0.05	2.98	2.85	2.77	2.73	2.70
	0.025	3.72	3.52	3.42	3.35	3.31
	0.01	4.85	4.56	4.41	4.31	4.25
15	0.10	2.06	1.97	1.92	1.89	1.87
	0.05	2.54	2.40	2.33	2.28	2.25
	0.025	3.06	2.86	2.76	2.69	2.64
	0.01	3.80	3.52	3.37	3.28	3.21
20	0.10	1.94	1.84	1.79	1.76	1.74
	0.05	2.35	2.20	2.12	2.07	2.04
	0.025	2.77	2.57	2.46	2.40	2.35
	0.01	3.37	3.09	2.94	2.84	2.78
25	0.10	1.87	1.77	1.72	1.68	1.66
	0.05	2.24	2.09	2.01	1.96	1.92
	0.025	2.61	2.41	2.30	2.23	2.18
	0.01	3.13	2.85	2.70	2.60	2.54
30	0.10	1.82	1.72	1.67	1.63	1.61
	0.05	2.16	2.01	1.93	1.88	1.84
	0.025	2.51	2.31	2.20	2.12	2.07
	0.01	2.98	2.70	2.55	2.45	2.39

**Nota: La tabla 4 del apéndice B es una tabla más completa.*

Ejemplo resuelto:

Dullus County Schools está por renovar el contrato del servicio de autobús para el año entrante y debe decidirse entre dos empresas que prestan el servicio, la empresa Milbank y la empresa Gulf Park. Como medida de la calidad del servicio se emplea la varianza en los tiempos en que llega a recoger/dejar a las personas. Poca varianza indica un mejor servicio, un servicio de mayor calidad. Si las varianzas de las dos empresas son iguales, la escuela elegirá la empresa que ofrezca mejores condiciones financieras. Pero si hay una diferencia significativa

en las varianzas, la escuela preferirá la empresa con la menor varianza o mejor servicio. Las hipótesis en este caso son las siguientes:

$$H_0 = \sigma_1^2 = \sigma_2^2$$

$$H_a = \sigma_1^2 \neq \sigma_2^2$$

Si se rechaza H_0 , se concluirá que los servicios no tienen la misma calidad. Para realizar esta prueba de hipótesis se usa como nivel de significancia $\alpha = 0.10$

En una muestra de 26 tiempos de llegada de la empresa Milbank la varianza muestral es 48 y en una muestra de 16 tiempos de llegada de la empresa Gulf Park la varianza muestral es 20. Como la varianza en la muestra de Milbank es la mayor, Milbank será la población 1. encuentre el valor del estadístico de prueba:

$$F = S_1^2 / S_2^2 = 48/20 = 2.40$$

La distribución F es la que tiene $n_1 - 1 = 26 - 1 = 25$ grados de libertad en el numerador y $n_2 - 1 = 16 - 1 = 15$ grados de libertad en el denominador.

Igual que en las otras pruebas de hipótesis, para llegar a una conclusión en esta prueba de hipótesis se puede emplear el método del valor-p o el método del valor crítico. En la tabla se encuentran las siguientes áreas en la cola superior correspondientes a los valores F de una distribución F con 25 grados de libertad en el numerador y 15 grados de libertad en el denominador.

Área en la cola superior	0.10	0.05	0.025	0.01
Valor F ($gl_1 = 25, gl_2 = 15$)	1.89	2.28	2.69	3.28

$F = 2.40$

Como $F = 2.40$ está entre 2.28 y 2.69, el área en la cola superior de la distribución está entre 0.05 y 0.025. Como se trata de una prueba de dos colas, se duplica el área de la cola superior, y se obtiene un valor-p $< \alpha = 0.10$. Por tanto, se rechaza la hipótesis nula. Esto lleva a la conclusión de que los dos servicios de autobús difieren en términos de la varianza de los tiempos en que llegan a recoger/dejar a las personas. Se le recomienda a la escuela el servicio de menor varianza o el servicio mejor que es el ofrecido por la empresa Gulf Park.

Ejemplos propuestos Unidad 3

1.- En una muestra de 16 elementos la desviación estándar muestral es 9.5. Pruebe la hipótesis siguiente usando $\alpha = 0.05$. ¿A que conclusión llega? Use el método del valor-p

$$H_0: \sigma^2 \leq 50$$

$$H_a: \sigma^2 > 50$$

2.- Una pieza para automóviles debe fabricarse con medidas de tolerancia muy estrechas para que sea aceptada por el cliente. Las especificaciones de producción indican que la varianza máxima en la longitud de la pieza debe ser 0.0004. Suponga que en 30 piezas la varianza muestral encontrada es $s^2 = 0.0005$. Use $\alpha = 0.05$ para probar si se está violando la especificación para la varianza poblacional.

3.- La desviación estándar promedio del rendimiento anual de fondos mutualistas de acciones de capital grande es 18.2% (The Top Mutual Funds, AAIL, 2004). La desviación estándar muestral en una muestra de 36 fondos mutualistas Vanguard PRIMECAP es 22%. Realice una prueba de hipótesis para determinar si la desviación estándar de los fondos Vanguard es mayor que la desviación estándar promedio de los fondos mutualistas de capital grande. Con nivel de significancia 0.05, ¿cuál es la conclusión?

4. En un estudio de Fortune se encontró que la varianza en la cantidad de vehículos que poseen o rentan los suscriptores de la revista Fortune es 0.94. Suponga que en una muestra de 12 suscriptores de otra revista se encuentran los datos siguientes sobre la cantidad de vehículos que poseen o rentan dichos suscriptores: 2, 1, 2, 0, 3, 2, 2, 1, 2, 1, 0 y 1.

a) Calcule la varianza muestral de la cantidad de vehículos que poseen o rentan estos 12 suscriptores.

b) Pruebe la hipótesis $H_0: \sigma^2 = 0.94$ para determinar si la varianza del número de vehículos que poseen o rentan los suscriptores de la otra revista difiere de la propia de Fortune, que es $\sigma^2 = 0.94$. Con un nivel de significancia 0.05, ¿cuál es la conclusión?

5. En un estudio para estimar la proporción de residentes de cierta ciudad y sus suburbios que están a favor de la construcción de una planta de energía nuclear, se encuentra que 63 de 100 residentes urbanos están a favor de la construcción mientras que sólo 59 de 125 residentes suburbanos la favorecen. ¿Hay una diferencia significativa entre la proporción de residentes urbanos y suburbanos que favorecen la construcción de la planta nuclear? Use un valor de P para su conclusión.

6. Una compañía petrolera afirma que un quinto de las casas en cierta ciudad se calienta con petróleo. ¿Tenemos razón en dudar de esta afirmación si, en una muestra aleatoria de 1000 casas en esta ciudad, se encuentra que 136 se calientan con petróleo? Utilice un nivel de significancia de 0.01. $(\bar{x}_1 - \bar{x}_2) = -0.019$. $\mu_1 - \mu_2 = 0$

$\alpha/2 = 0.025$ Región de rechazo Región de aceptación H1 Región de Rechazo
 $\alpha/2 = 0.025$ $(\bar{x}_1 - \bar{x}_2) = 0.019$ $\mu_1 - \mu_2 = 0.03$ $b = 0.1387$ 89

7. Se sabe que la duración, en horas, de un foco de 75 watts tiene una distribución aproximadamente normal, con una desviación estándar de 25 horas. Se toma una muestra aleatoria de 20 focos, la cual resulta tener una duración promedio de 1014 horas.

a) ¿Existe evidencia que apoye la afirmación de que la duración promedio del foco es mayor que 1000 horas? Utilice un $\alpha = 0.05$.

b) ¿Cuál es el valor P para la prueba?

c) ¿Cuál es el valor de b para la prueba del inciso a) si la verdadera duración promedio del foco es de 1050 horas?

8. Se estudia la tasa de combustión de dos propelentes sólidos utilizados en los sistemas de escape de emergencia de aeroplanos. Se sabe que la tasa de combustión de los dos propelentes tiene aproximadamente la misma desviación estándar de 3 cm/s. Se prueban dos muestras aleatorias de 20 especímenes cada una, obteniéndose medias de 18 y 24 cm/s respectivamente.

a) Pruebe la hipótesis de que los dos combustibles sólidos tienen la misma rapidez promedio de combustión. Utilice un $\alpha = 0.05$.

b) ¿Cuál es el valor de P de la prueba?

c) ¿Cuál es el valor de b para la prueba del inciso a) si la verdadera diferencia en la rapidez promedio de combustión es 2.5 cm/s?

9. Se supone que los neumáticos para automóvil de cierto tipo recién comprados deben llenarse a una presión de 30 lb/pulg². Se representa con m el verdadero promedio de presión. Encuentre el valor P asociado con cada valor del estadístico z dado para probar H_0 ; $m = 30$ contra H_1 ; $m \neq 30$.

- a) 210 b) -1.75 c) -0.55 d) 1.41 e) -5.3

10. Se realizó un experimento para comparar la resistencia a la fractura del acero con níquel maragizado, con el acero de pureza comercial del mismo tipo. Para 32 especímenes, la resistencia promedio muestral fue de 65.6 para el acero de alta pureza, mientras que se obtuvo una media muestral de 59.8 en 38 especímenes del acero comercial. Debido que el acero de alta pureza es más costoso, su uso para cierta aplicación puede justificarse sólo si su resistencia a la fractura excede la del acero de pureza comercial en más de 5. Suponga que ambas distribuciones de resistencias son normales.

a) Si se supone que $s_1 = 1.2$ y $s_2 = 1.1$, pruebe las hipótesis pertinentes usando $\alpha = 0.001$.

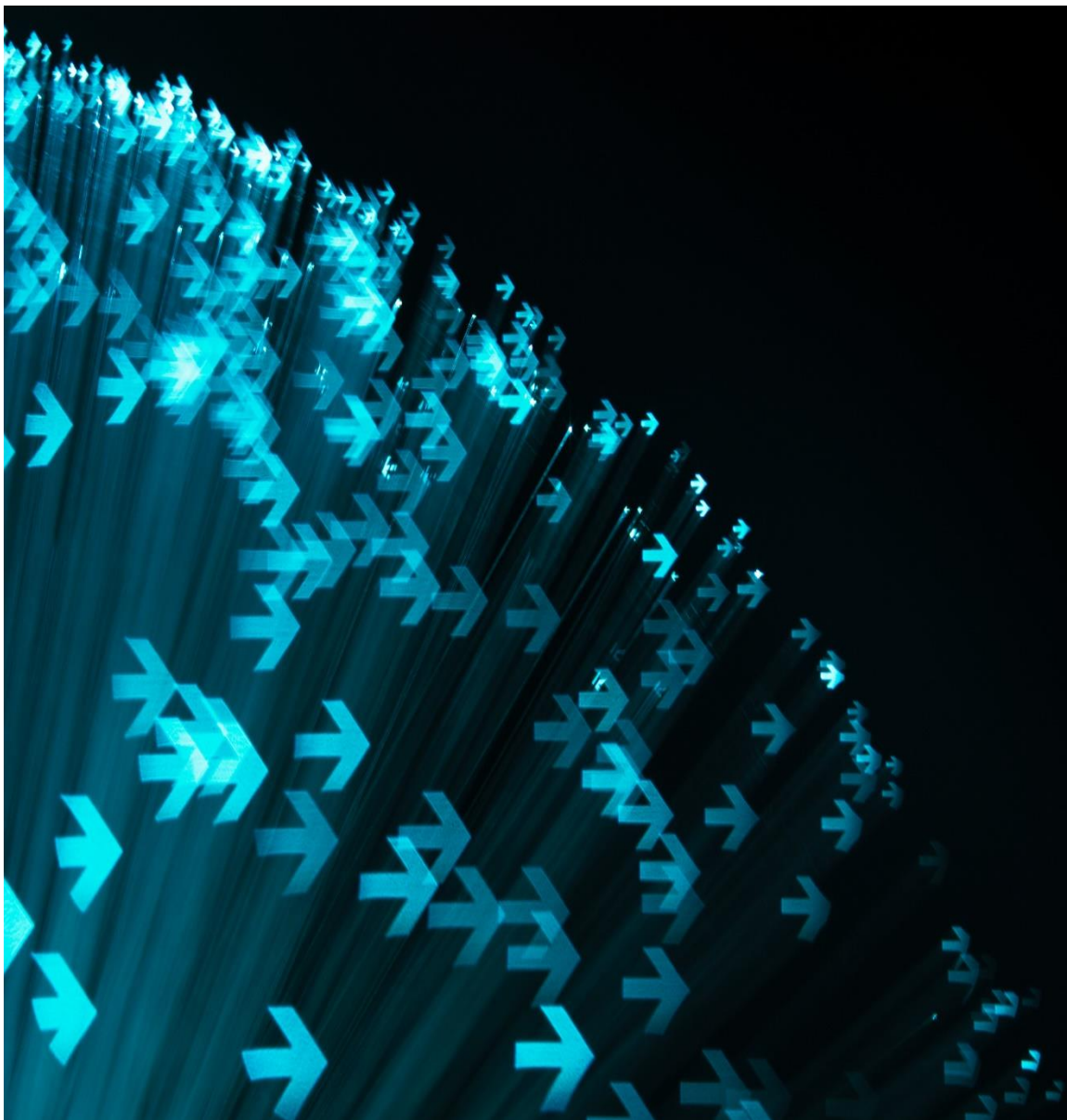
b) Calcule b) para la prueba del inciso anterior cuando $\mu_1 - \mu_2 = 6$.

11. Se cree que la portada y la naturaleza de la primera pregunta de encuestas por correo influyen en la tasa de respuesta. Un artículo probó esta teoría al experimentar con diferentes diseños de portadas. Una portada sencilla, y la otra utilizó la figura de un paracaidista. Los investigadores especularon que la tasa de devolución sería menor para la portada sencilla. Portada Número de envíos Número de devoluciones Sencilla 207 104 Paracaidista 213 109 ¿Esta información apoya la hipótesis de los investigadores? Haga la prueba con un nivel de significancia de 0.10, calculando primero un valor P .

Portada	Número de envíos	Número de devoluciones
Sencilla	207	104
Paracaidista	213	109

Unidad 4.

Pruebas de Bondad de Ajuste y Pruebas no Paramétricas



INSTRUMENTACIÓN DIDÁCTICA

Competencia Específica	Competencia Genérica	Actividades de Aprendizaje
Realiza pruebas de bondad de ajuste y no paramétricas para determinar si el comportamiento de un experimento se adecua a una distribución determinada en procesos de la industria y la logística.	• Capacidad de abstracción, análisis y síntesis. • Capacidad de aplicar los conocimientos en la práctica. • Capacidad para identificar, plantear y resolver problemas.	• Buscar información sobre conceptos relacionados con pruebas de bondad de ajuste y pruebas no paramétricas. • Discutir en grupo sobre los conceptos investigados. • Resolver ejercicios aplicando: o Prueba χ^2. o Prueba de Kolmogorov-Smirnov. o prueba de Anderson Darling. o Prueba de Ryan – Joiner. o Prueba de Shappiro – Wilk. • Dado un conjunto de datos: aplicar las pruebas, analizar los resultados y contrastar las pruebas. • Utilizar TIC's para realizar las pruebas de bondad y ajuste y pruebas no paramétricas.

4. Pruebas de Bondad de Ajuste y Pruebas no Paramétricas

4.1 Bondad de ajuste.

La bondad de ajuste de un modelo estadístico describe lo bien que se ajusta un conjunto de observaciones. Las medidas de bondad en general resumen la discrepancia entre los valores observados y los valores esperados en el modelo de estudio. Tales medidas se pueden emplear en el contraste de hipótesis, e.g. el test de normalidad de los residuos, comprobar si dos muestras se obtienen a partir de dos distribuciones idénticas (test de Kolmogorov-Smirnov), o si las frecuencias siguen una distribución específica (ji cuadrada).

Las pruebas de bondad de ajuste tratan de verificar si el conjunto de datos se puede ajustar o afirmar que proviene de una determinada distribución (Binomial, o Normal o Poisson).

Las pruebas básicas que pueden aplicarse son:

- Ji cuadrada.
- Smirnov-Kolmogorov.

Como su nombre lo indica miden el grado de ajuste que existe entre la distribución obtenida a partir de la muestra y la distribución teórica que se supone debe seguir esa muestra. Comúnmente se aplica con Ji cuadrada.

Ambas pruebas están basadas en la:

- *Hipótesis nula*: de que siguen una distribución sugerida.
- *Hipótesis alternativa*: siempre se enuncia de que los datos No siguen la distribución supuesta.

4.1.1 Análisis Ji-Cuadrada.

El estadístico **ji-cuadrado** (o **chi cuadrado**), que tiene distribución de probabilidad del mismo nombre, sirve para someter a prueba hipótesis referidas a distribuciones de frecuencias. En términos generales, esta prueba contrasta frecuencias observadas con las frecuencias esperadas de acuerdo con la hipótesis nula.

Sí, χ es la letra griega Ji.

- Si tiene una sola variable de medida, use una prueba ji cuadrado de bondad de ajuste.
- Si tiene dos variables de medida, use la prueba de ji cuadrado de independencia.
- Hay otras pruebas de ji cuadrado, pero estas dos son las más frecuentes.

Número de variables	uno	dos
Objetivo de la prueba	Determinar si es probable que una variable venga o no de una distribución dada	Decidir si dos pueden o no estar relacionadas
Ejemplo	Decidir si varias bolsas de caramelos tienen o no el mismo número de unidades de cada sabor.	Decidir si el consumo de snack de un espectador tiene una relación con el tipo de película que va a ver al cine
Hipótesis del ejemplo	Ho: La proporción de sabores es la misma Ha: la proporción de sabores es diferente	Ho: La proporción de gente que compra snacks es independiente del tipo de película. Ha: la proporción de gente que compra snacks es distinta para distintos tipos de película.
Prueba de Distribución	Ji cuadrada	Ji cuadrada
Grados de libertad	Número de categorías menos 1	Número de categorías de la primera variable menos 1, multiplicado por el número de categorías

		de la segunda variable menos 1
--	--	--------------------------------

4.1.2 Prueba de independencia.

Trata de la comparación de dos situaciones en las cuales podemos esperar que sean dependientes o independientes, esto quiere decir que, pueden o no estar relacionados sus datos debido a muchos factores que pueden influir en ellos, o bien, un problema no tenga relación con otro.

Su objetivo es determinar si alguna situación es afectada por otra, basándose en datos estadísticos y valores probabilísticos obtenidos de la fabulación de datos o de pronósticos por medio de fórmulas y tablas, se basa en un nivel de significancia en un caso y en el otro comparar, valiéndose de tablas de contingencia para obtener frecuencias esperadas y poder aplicarlas, para así obtener datos comparativos que son determinantes en la decisión de independencia.

Tabla de contingencia:

Son tablas que constan de renglones y columnas donde se presentan y organizan los datos recolectados y se calculan los esperados. Para todas las pruebas de Independencia, las hipótesis son:

H_0 : las dos variables de clasificación son INDEPENDIENTES

H_1 : las dos variables de clasificación son DEPENDIENTES.

De nuevo se compararán las frecuencias basándose en la prueba de ji cuadrada (χ^2). Si el valor del estadístico de prueba $\chi^2_{\text{calculado}} \geq \chi^2_{\text{tablas}}$ significará que H_0 se rechaza es decir que entre las dos variables de clasificación son dependientes.

Su fórmula:

$$\chi^2 = \sum_{i=1}^k \frac{[f_{o_i} - f_{e_i}]^2}{f_{e_i}}$$

grados de libertad: $gl = (\# r-1) * (\# c-1)$

$$\text{frecuencia esperada} = \frac{(\text{Total de la columna})(\text{Total del renglón})}{\text{Gran total.}}$$

Características:

- χ^2 toma valores no negativos; es decir puede ser cero o positiva.
- Se presenta la información en una tabla de contingencia.

Ejemplo Resuelto:

Supóngase que desea determinar si las opiniones de los residentes votantes estado de Illinois respecto a una nueva reforma impositiva son independientes de sus niveles de ingreso. Una muestra aleatoria de 1000 votantes registrados del estado de Illinois se clasifica de acuerdo a sus ingresos como bajo, medio y alto y si van a estar a favor o en contra de la nueva reforma impositiva. Las frecuencias observadas se presentan en la siguiente tabla de contingencia.

Tabla de contingencia 2 x 3

Reforma impositiva	Nivel de ingresos			Total
	Bajo	Medio	Alto	
A favor	182	213	203	598
En contra	154	138	110	402
Total	336	351	313	1000

$n=1000$

1. Establecer las hipótesis:

H_0 : Las opiniones de los residentes votantes del estado de Illinois respecto a una reforma impositiva son independientes de sus niveles de ingreso.

H_1 : Las opiniones de los residentes votantes del estado de Illinois respecto a una reforma impositiva son dependientes de sus niveles de ingreso.

Nota: todos los datos presentados en la tabla de contingencia es la frecuencia observada.

2. Calcular Frecuencia Esperada:

Con la siguiente formula realizar la operación para cada celda de la tabla de contingencia.

$$\text{frecuencia esperada} = \frac{(\text{Total de la columna})(\text{Total del renglón})}{\text{Gran total.}}$$

$$Fe_{1,1} = \frac{(336)(598)}{1000} = 200.92 \quad \text{etc. etc.}$$

Reforma Impositiva	Bajo		Medio		Alto		Total
	Fo	Fe	Fo	Fe	Fo	Fe	
A favor	182	200.92	213	209.89	203	187.17	598
En contra	154	135.07	138	141.10	110	125.82	402
total	336		351		313		1000

3. Calcular Chi cuadrada con la fórmula:

$$X^2 = \frac{(182-200.92)^2}{200.92} + \frac{(154-135.07)^2}{135.07} + \frac{(213-209.89)^2}{209.89} + \frac{(138-141.10)^2}{141.10} + \frac{(203-187.17)^2}{187.17} + \frac{(110-125.82)^2}{125.82}$$

$$X^2 = 1.78 + 2.65 + .046 + .068 + 1.33 + 1.98$$

$$X^2 = 7.854$$

4. obtenemos grados de Libertad:

$$gl = (\# \text{ renglones} - 1) * (\# \text{ columnas} - 1)$$

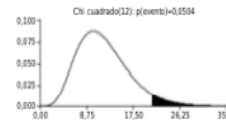
$$gl = (2 - 1) * (3 - 1) = 2$$

y un nivel de significancia de 5%. $\alpha = .05$

5. Se busca el valor en la tabla de Chi cuadrada con $gl=2$ y $\alpha=.05$

Tabla de distribución Chi-Cuadrado

En el margen superior se lee $P(\chi^2 \leq x)$ para los valores de x que figuran en el cuerpo de la tabla y en el margen izquierdo los grados de libertad (δ).



δ	0,99	0,975	0,95	0,9	0,05	0,025	0,01
1	0,000	0,001	0,004	0,016	3,842	5,024	6,635
2	0,020	0,051	0,103	0,211	5,992	7,378	9,210
3	0,115	0,216	0,352	0,584	7,815	9,348	11,345
4	0,297	0,484	0,711	1,064	9,488	11,143	13,277
5	0,554	0,831	1,146	1,610	11,071	12,833	15,086
6	0,872	1,237	1,635	2,204	12,592	14,449	16,812
7	1,239	1,690	2,167	2,833	14,067	16,013	18,475
8	1,647	2,180	2,733	3,440	15,507	17,535	20,090

6.Decisión:

$X^2_{calculada} > X^2_{tablas}$ por lo tanto H_0 se rechaza

7.Conclusión: Las opiniones de los residentes votantes del estado de Illinois respecto a una reforma impositiva son dependientes de sus niveles de ingreso.

4.1.3 Prueba de la bondad del ajuste.

Una prueba de bondad de ajuste permite docimar la hipótesis de que una variable aleatoria sigue cierta distribución de probabilidad y se utiliza en situaciones donde se requiere comparar una distribución observada con una teórica o hipotética, compararla con datos históricos o con la distribución conocida de otra.

Hay dos pruebas de ji cuadrado que se suelen usar: la prueba de bondad de ajuste de ji cuadrado y la prueba de independencia de ji cuadrado. Ambas pruebas implican variables que dividen los datos en categorías. Por ello, puede resultar confuso decidir qué prueba usar. La siguiente compara ambas pruebas.

El estadístico ji-cuadrado (o chi cuadrado), tiene distribución de probabilidad del mismo nombre, sirve para someter a prueba hipótesis referidas a distribuciones de frecuencias. Probar la bondad de un ajuste es ver en qué medida se ajustan los datos observados a una distribución teórica o esperada.

Su fórmula es la siguiente:

$$\chi^2 = \sum_{i=1}^k \frac{[f_{o_i} - f_{e_i}]^2}{f_{e_i}}$$

χ^2 = ji cuadrada

f_o = frecuencia obtenida u observada

f_e = frecuencia teórica o esperada

Σ = sumatoria

$i=1$ cada dato de uno en uno

k = en toda la celda

grados de libertad: $gl = V = K - 1$

Criterio de decisión:

Se rechaza H_0 si $\chi^2_{calculada} \geq \chi^2_{tabla}$

Cuanto más se aproxima a cero el valor de ji cuadrada, más ajustadas están ambas distribuciones.

Ejemplo Resuelto:

Durante 300 días de trabajo consecutivo, un pastelero prepara 3 pasteles de chocolate grandes, y los que no se venden el mismo día se donan a un banco de caridad de alimentos. Con respecto a los datos que aparecen en la tabla siguiente. Pruebe con un nivel de significancia de .05, si se puede considerar como valores de una muestra aleatoria que **tiene una distribución binomial**

NUM. DE PASTELES VENDIDOS (núm. de clases)	NUM. DE DIAS (frec. Observada)
0	2
1	14
2	46
3	238
total	300

Solución:

1.- Establecer las hipótesis de bondad de ajuste a la distribución sugerida (que mencionan en el ejercicio).

H₀: El número de pasteles vendidos tienen una distribución Binomial

H₁: El número de pasteles vendidos NO tienen una distribución Binomial

f_o ya lo tenemos, son los datos que observamos al muestrear

f_e como es la frecuencia que se espera, habrá que calcularla con la probabilidad sugerida

$$f_e = n p$$

n= núm. de datos.

p= probabilidad calculada con la distribución sugerida.

la fórmula de probabilidad de la distribución Binomial es:

$$P(x) = {}^nC_x p^x q^{n-x}$$

2.- Calcular el valor de ji cuadrada (χ^2)

$$P(0) = {}_3C_0 (.5)^0 (.5)^{3-0} = .125 \quad P=.5 \text{ porque la probabilidad de que venda es } .5 \text{ ya que no}$$

tenemos un dato específico

$$P(1) = {}_3C_1 (.5)^1 (.5)^{3-1} = .375$$

$$P(2) = {}_3C_2 (.5)^2 (.5)^{3-2} = .375$$

$$P(3) = {}_3C_3 (.5)^3 (.5)^{3-3} = .125$$

$$F_e = n p$$

$$(300) (.125) = 37.5$$

$$(300) (.375) = 112.5$$

$$(300) (.375) = 112.5$$

$$(300) (.125) = 37.5$$

3.- Calcular ji cuadrada (χ^2) con la fórmula

$$\chi^2 = \frac{(2-37.5)^2}{37.5} + \frac{(14-112.5)^2}{112.5} + \frac{(46-112.5)^2}{112.5} + \frac{(238-37.5)^2}{37.5} = 1231.16$$

4.- χ^2 de tablas se obtiene con la tabla de ji cuadrada con:

$V = \text{grados de libertad} = V = K-1$ y nivel de significancia ($\alpha = .05$)

$$V = 4 - 1 = 3$$

$K = \text{número de clases}$

χ^2 de tablas = 7.81 entonces:

$$\chi^2_{\text{calculada}} > \chi^2_{\text{tablas}}$$

$$1231.16 > 7.81$$

5. CONCLUSION: El número de pasteles vendidos no tiene una distribución binomial

Nota: Si la distribución que se sugiere es Poisson, entonces para calcular la probabilidad se hace mediante la formula Poisson en vez de la fórmula Binomial. TODO LO DEMAS ES EXACTAMENTE IGUAL

$$\text{Fórmula Poisson: } p(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \lambda = \text{promedio}$$

Ejemplo resuelto:

si el promedio $\lambda = .75$

$$P(0) = \frac{.75^0 (e^{-.75})}{0!} = .4723$$

Y todo de demás exactamente igual.

4.1.4 Tablas de contingencia.

Tablas de contingencia: son tablas que constan de renglones y columnas donde se presentan y organizan los datos recolectados y se calculan los esperados para cada una de las celdas de la tabla generalmente se usan las de 2X3; es decir 2 renglones por 3 columnas. Se emplean para resolver ejercicios de independencia. Se muestra la siguiente del ejercicio que ya se vió en pruebas de independencia anteriormente

Tabla de contingencia 2 x 3

Reforma impositiva	Nivel de ingresos			Total
	Bajo	Medio	Alto	
A favor	182	213	203	598
En contra	154	138	110	402
Total	336	351	313	1000

Ejercicios Propuestos unidad 4:

1.-Para analizar el número de artículos defectuosos en una fábrica en la cd. De Medellín, tomamos una muestra aleatoria de $n=60$ artículos y se observó el número de defectuosos y se obtuvieron los siguientes resultados:

Número de artículos defectuosos	Frecuencia observada
0	32
1	15
2	9
3	4
total	60

La distribución de los defectos de tarjeta de circuitos tiene una distribución poisson con un promedio de 0.75? Realice la prueba de bondad de ajuste mediante ji cuadrada con un nivel de significancia de 0.05.

2. -Una compañía requiere que los recién graduados de la universidad que buscan una colocación en la empresa, sean entrevistados por 3 ejecutivos diferentes. Cada uno de los ejecutivos califica al candidato en forma positiva o negativa.

Posibles calificaciones positivas	Núm. candidatos que obtienen calificación positiva
0	18
1	47
2	24
3	11
total	100

El director de la compañía piensa que el proceso de entrevistas puede ser aproximado por una distribución binomial con una probabilidad de 0.40, es decir una probabilidad del 40% de que cualquier candidato obtenga una calificación positiva en cualquiera de las entrevistas. El director desea probar esta hipótesis con un nivel de significancia de 0.10. de qué manera deberá proceder?

3.- Suponga que en 4 regiones NORTE, SUR, ESTE Y OESTE de cuidado de la salud muestrea las actitudes de trabajo de los empleados de sus hospitales con respecto al examen de desempeño en el trabajo. A los trabajadores se les da a escoger entre el método actual (2 exámenes al año) y un nuevo método propuesto (exámenes trimestrales). Queremos ver la respuesta de los empleados si la región estudiada depende de la preferencia de un método específico de evaluación. Utilice un nivel de significancia de .05. los datos muestreados son clasificados en la siguiente tabla de contingencia.

REGIONES DE APLICACIÓN

Núm. Empleados por preferencia de método.	NORTE	SUR	ESTE	OESTE	TOTAL
METODO ACTUAL (2 ex. al año)	68	75	57	79	
METODO NUEVO. (exámenes trimestrales)	32D	45	33	31	
TOTAL					

4.- Se registró un total de 309 muebles defectuosos con los tipos de defectos clasificados como A,B,C,D . al mismo tiempo se identificó cada mueble con el turno en el que fue producido. Los valores que se obtuvieron se presentan en la siguiente tabla de contingencia: utilice un nivel de significancia de .05.

TIPO DE DEFECTO

TURNO	A	B	C	D
1	15	21	45	13
2	26	31	34	5
3	33	17	49	20
TOTAL				

5.- Se desea estudiar la dependencia entre la práctica de un deporte y la depresión, se seleccionó una muestra aleatoria simple de 100 jóvenes con los siguientes resultados.

JÓVENES	SIN DEPRESIÓN	CON DEPRESIÓN	TOTAL
DEPORTISTA	38	9	
NO DEPORTISTA	31	22	
TOTAL			

4.2 Pruebas no paramétricas.

Engloban una serie de pruebas estadísticas que tienen como denominador común la ausencia de asunciones acerca de la ley de probabilidad que sigue la población de la que ha sido extraída la muestra. Por esa razón es común referirse a ellas como distribución libre.

4.2.1 Escala de medición.

Pueden ser de dos tipos:

1. Nominal
2. Ordinal

3. Intervalo
4. Proporción

1. Una escala nominal: es una escala de medición en la cual los números sirven como “*etiquetas*” solamente para identificar o clasificar un objeto. Una escala de medición nominal normalmente trata sólo con variables no numéricas (no cuantitativas).

Por ejemplo: supongamos que se realiza esta pregunta: “¿Podrías seleccionar el grado de incomodidad de tu enfermedad?”

Las opciones de respuesta utilizadas para este tipo de escala serían:

- 1.- Leve
- 2.- Moderado
- 3.- Severo

En este ejemplo en particular, 1 = leve, 2 = Moderado y 3 = Severo. Aquí los números simplemente son utilizados como etiquetas y no tienen ni un solo valor.

- La escala ordinal: es uno de los niveles de medición que nos otorga clasificación y el orden de los datos sin que realmente se establezca el grado de variación entre ellos. La escala de medición ordinal es la segunda de las 4 escalas de medición. Los datos ordinales son básicamente datos estadísticos que tienen la misma naturalidad, pero existe una diferencia entre ellos que es desconocida. Estos datos pueden ser agrupados o clasificados.

Ejemplo:

Una escala ordinal puede responder preguntas como: *¿Qué tan satisfecho estás con nuestros productos?*

- 1.- Totalmente satisfecho
- 2.- Satisfecho
- 3.- Neutral
- 4.- Insatisfecho
- 5.- Totalmente insatisfecho

Por lo tanto, se utiliza una escala ordinal como parámetro para comprender si las variables son mayores o menores. La tendencia central de la escala ordinal es mediana.

La escala de Likert es un ejemplo de porque la diferencia de intervalo entre las variables ordinales no se puede concluir. En esta escala de hecho, las opciones de respuesta suelen ser polares, como, por ejemplo, algo como “totalmente satisfecho” o “totalmente insatisfecho”.

- De intervalo: Una escala en intervalo considera el orden relativo de las mediciones, pero también considera el hecho de que una diferencia en escala entre dos medidas, corresponde a la misma diferencia en la cantidad del atributo medido. Por ejemplo, si se hablara del IQ (Coeficiente Intelectual), la diferencia entre una persona que tiene un IQ de 100 y otra de 101 debería ser igual a una que tiene un IQ de 130 y otra de 131.
- Proporción: El concepto de nivel de medición tipo proporción es igual al de intervalo, pero además se caracteriza por tener un punto cero. Debido a lo anterior, es posible hacer aseveraciones significativas acerca del atributo o variable medido. Por ejemplo, la mayoría de las medidas físicas (estatura, peso) tienen este tipo de nivel: tienen un punto cero y se pueden hacer afirmaciones sobre el peso de alguien en forma de proporción (una persona pesa el doble que otra).

4.2.2 Métodos estadísticos contra no paramétricos.

VARIABLE	ESTADÍSTICA PARAMETRICA	ESTADISTICA NO PARAMETRICA
DEFINICIÓN	Es una rama de la estadística inferencial que comprende los procedimientos estadísticos y de decisión que están basadas en distribuciones conocidas. Estas son	Engloban una serie de pruebas estadísticas que tienen como denominador común la ausencia de asunciones acerca de la ley de probabilidad que sigue la población de la que ha sido extraída la muestra.

	determinadas usado un número finito de parámetros. μ , σ^2 , σ , p etc.	Por esa razón es común referirse a ellas como distribución libre.
CARACTERÍSTICAS	La mayoría de los procedimientos paramétricos requiere conocer la forma de distribución para las mediciones resultantes de la población estudiada. Para la inferencia paramétrica es requerida como mínimo una escala de intervalo. Esto quiere decir nuestros datos pueden ser categorizados en: menores de 20 años, De 40 a 60 Mas de 60 etc. Ya que hay números con los cuales realizar cálculos estadísticos.	1. Son más fáciles de aplicar. 2. Son aplicables a los datos jerarquizados. 5. Se pueden usar cuando dos series de observaciones provienen de distintas poblaciones. 6. Son la única alternativa cuando el tamaño de la muestra es pequeño. 7. Son útiles a un nivel de significancia previamente especificado
SUPUESTOS	Se conoce el modelo de distribución de la población, objeto de estudio. Y se desconoce un número finito de parámetros de dicha distribución de la muestra para poder realizar inferencias sobre la población.	• Son métodos de distribución libre, no requieren conocer la distribución de la muestra. • Se utilizan estadísticos cuya distribución se denomina con independencia del cual sea la

		distribución de la población. • Son una alternativa para las pruebas paramétricas cuando los datos no cumplen los requisitos de las pruebas paramétricas.
VENTAJAS	1. Mas poder de eficiencia 2. Mas sensibles a los rasgos de los datos recolectados. 3. Menos posibilidad de errores. 4. Robustas (dan estimaciones probabilísticas muy exactas).	Más poder de eficacia, más sensibles a los rasgos de los datos recolectados. Son robustas (dan estimaciones probabilísticas muy exactas).
DESVENTAJAS	Más complicadas de calcular limitaciones en los tipos de datos que se pueden evaluar.	no son sistemáticas, se les relaciona con la convivencia, esto se debe a que no hay una distribución fija. Las tablas que se requieren están muy difundidas y aparecen en diferentes formatos, y puede ocasionar alguna confusión al investigador.

4.2.3 Prueba de Kolmogorov – Smirnov.

Fue desarrollada por Kolmogorov en 1933, pero recibió este nombre debido a la similitud entre esta prueba y la bondad de ajuste para 2 muestras independientes de Smirnov(1939). Ésta es usada para determinar si la distribución de los datos en una muestra se ajusta a una distribución poblacional específica teórica o empírica, es decir, para demostrar si una muestra se deriva de una distribución de tipo específico. Esta prueba está diseñada para trabajar con variables aleatorias continuas.

- Suposiciones

La muestra de n sujetos ha sido aleatoriamente elegida de la población que representa.

El nivel de medición de los datos es ordinal.

La variable bajo estudio es continua.

- Establecimiento de hipótesis

H_0 = Los datos analizados siguen una distribución M

H_1 = Los datos analizados siguen una distribución M

- Estadístico de contraste

$$D = \sup_{1 \leq i \leq n} |F_n(X_i) - F_0(X_i)|$$

Donde:

X_i es el i-ésimo valor observado en la muestra (cuyos valores se han ordenado previamente de menor a mayor).

$F_n(X_i)$ es un estimador de la probabilidad de observar valores menores o iguales que X_i .

$F_0(X)$ es la probabilidad de observar valores menores o iguales que X_i cuando H_0 es cierta.

D es la mayor diferencia absoluta observada entre la frecuencia acumulada observada $F_n(x)$ y la frecuencia acumulada teórica $F_0(x)$, obtenida a partir de la distribución de probabilidad que se especifica como hipótesis nula.

Si los valores observados $F_n(x)$ son similares a los esperados $F_0(x)$, el valor de D será pequeño. Cuanto mayor sea la discrepancia entre la distribución empírica $F_n(x)$ y la distribución teórica, mayor será el valor D .

Por tanto, el criterio para la toma de la decisión entre las dos hipótesis será la forma:

$$\begin{aligned} \text{Si } D \leq D_\alpha &\rightarrow \text{Aceptar } H_0 \\ \text{Si } D > D_\alpha &\rightarrow \text{Rechazar } H_0 \end{aligned}$$

Donde el valor D_α se elige de tal manera que:

$$P(\text{rechazar } H_0 / H_0 \text{ es cierta}) = P(D > D_\alpha / \text{los datos siguen la distribución } M) = \alpha$$

Siendo α el nivel mas bajo de significación del contraste.

Para el cálculo práctico del estadístico D deben obtenerse:

$$D^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F_0(x_i) \right\}, \quad D^- = \max_{1 \leq i \leq n} \left\{ F_0(x_i) - \frac{i-1}{n} \right\},$$

A partir de estos valores:

$$D = \max \{D^+, D^-\}$$

A su vez, el valor de $D_\alpha = \frac{C_\alpha}{k(n)}$

Donde C_α y $k(n)$ se encuentran en las siguientes tablas:

C_α	α		
	0.1	0.05	0.01
Modelo			
General	1.224	1.358	1.628
Normal	0.819	0.895	1.035
Exponencial	0.990	1.094	1.308
Weibull $n=10$	0.760	0.819	0.944
Weibull $n=20$	0.779	0.843	0.973
Weibull $n=50$	0.790	0.856	0.988
Weibull $n=\infty$	0.803	0.874	1.007

DISTRIBUCIÓN QUE SE CONTRASTA	$k(n)$
General. Parámetros conocidos.	$k(n) = \sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}$
Normal	$k(n) = \sqrt{n} - 0.01 + \frac{0.85}{\sqrt{n}}$
Exponencial	$k(n) = \sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}$
Weibull	$k(n) = \sqrt{n}$

Ejemplo resuelto:

Determinar los valores de la primera columna se conforman a una distribución normal.

Y	Y-ordenados	Orden	F	Z	Fo	D+	D-
6.0	1.9	1	0.1	-1.628	0.051	0.049	0.051
2.3	2.3	2	0.2	-1.332	0.091	0.109	-0.009
4.8	3.3	3	0.3	-0.592	0.276	0.024	0.076
5.6	3.4	4	0.4	-0.518	0.302	0.098	0.002
4.5	4.5	5	0.5	0.296	0.616	-0.116*	0.216*
3.4	4.5	6	0.6	0.296	0.616	-0.016	0.116
3.3	4.8	7	0.7	0.518	0.698	0.002	0.098
1.9	4.8	8	0.8	0.518	0.698	0.102	-0.002
4.8	5.6	9	0.9	1.11	0.867	0.033	0.067
4.5	6.0	10	1.0	1.406	0.920	0.080	0.020

Media =4.1, varianza=1.822

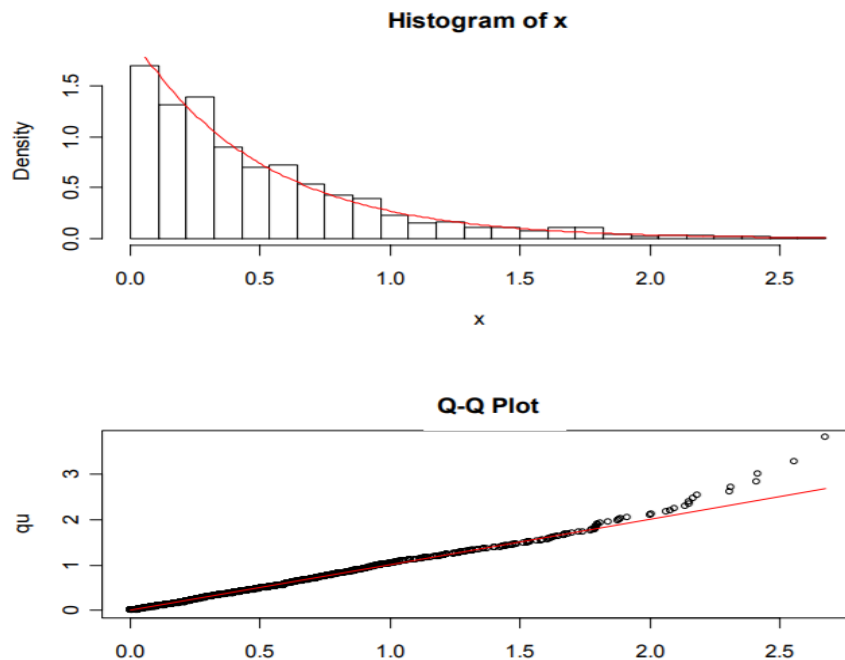
$$D_{\alpha} = \frac{0.895}{\sqrt{10} - 0.01 + \frac{0.85}{\sqrt{10}}} = \frac{0.895}{3.42} = 0.262$$

Como el valor $D=0.216 < 0.262$, no se rechaza H_0 y se acepta que los datos se distribuyen normalmente.

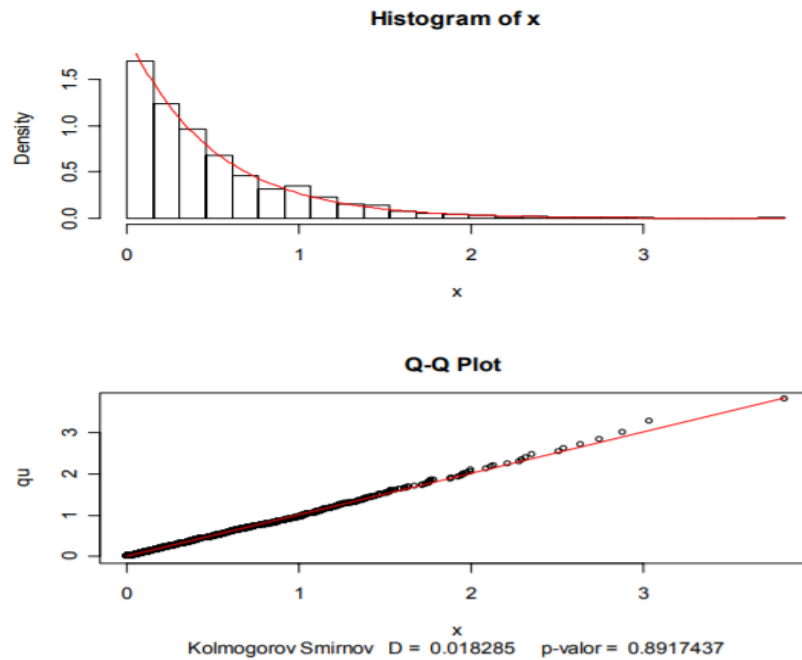
Ejemplo resuelto

En los siguientes ejemplos se han simulado datos con distribución exponencial o normal contrastándolos en todos los casos si puede aceptarse que los datos siguen una distribución exponencial. Se ha acompañado al contraste con el histograma de los datos y el grafico Q-Q plot (grafico cuantil-cuantil se representan los cuantiles de la distribución teórica supuesta frente a los cuantiles de la distribución empírica. En un buen ajuste, la mayoría de estos puntos deberían situarse sobre la recta $y=x$).

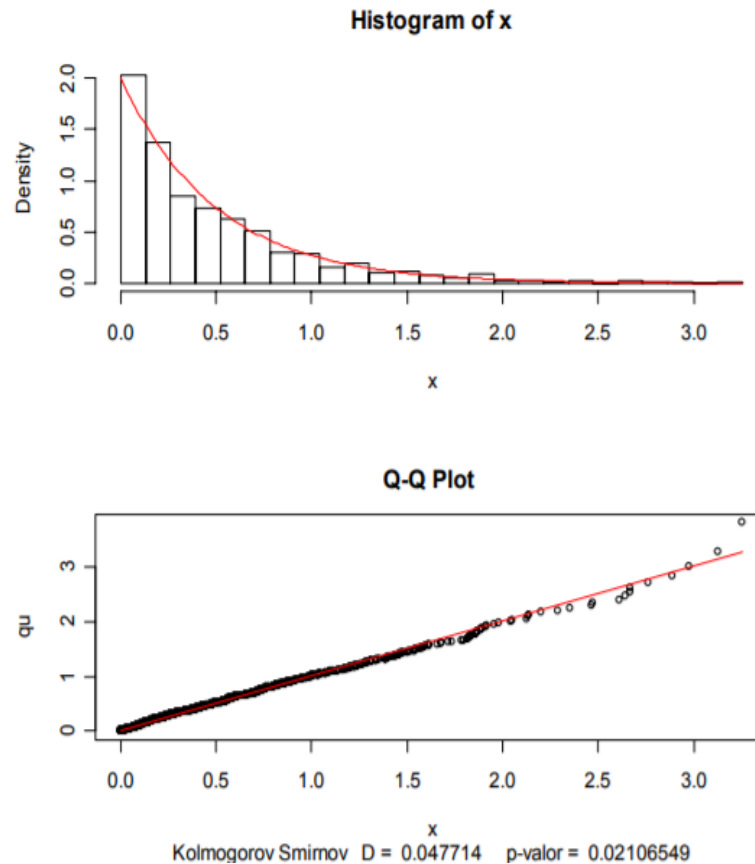
Simulación de datos con distribución exponencial **n=1000**



Simulación de datos con distribución exponencial:
n=1000

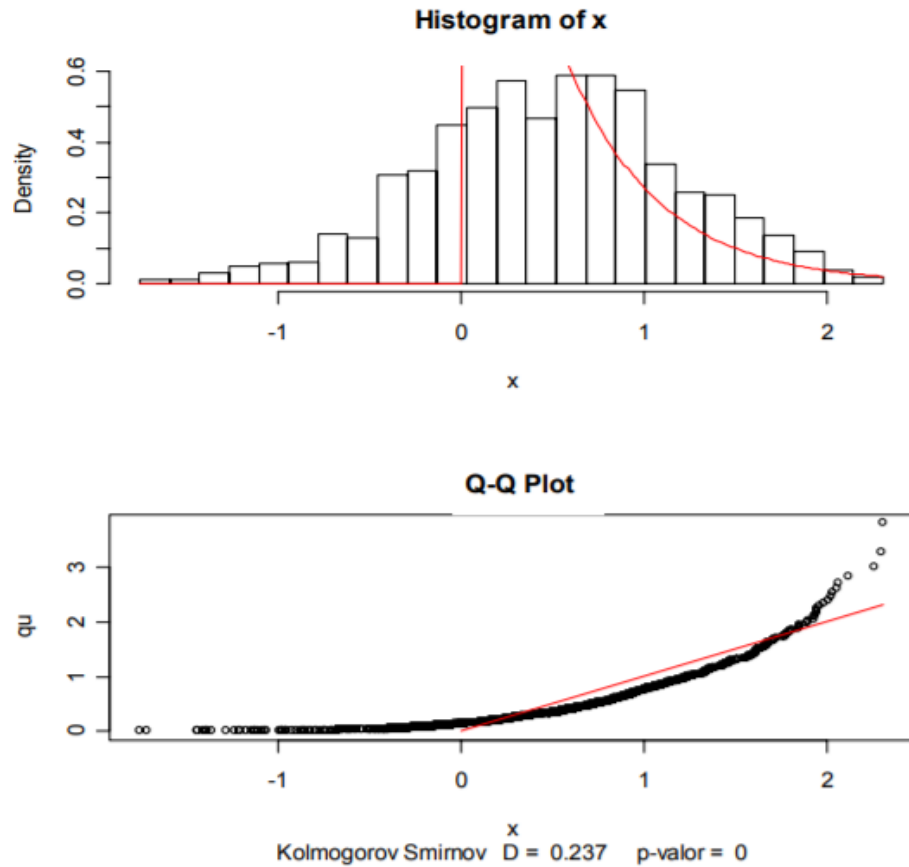


Simulación de datos con distribución exponencial
n=1000



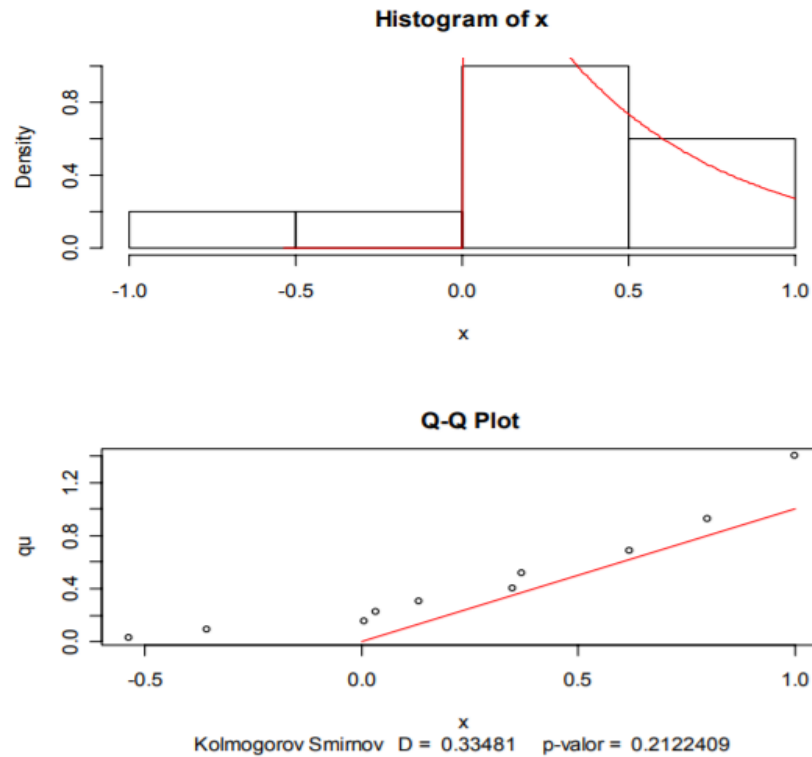
Nótese que, en este caso, aunque los datos se han generado realmente con distribución exponencial, el p-valor conduce que esta sea la distribución de los datos.

Simulación de datos con distribución normal
n=1000



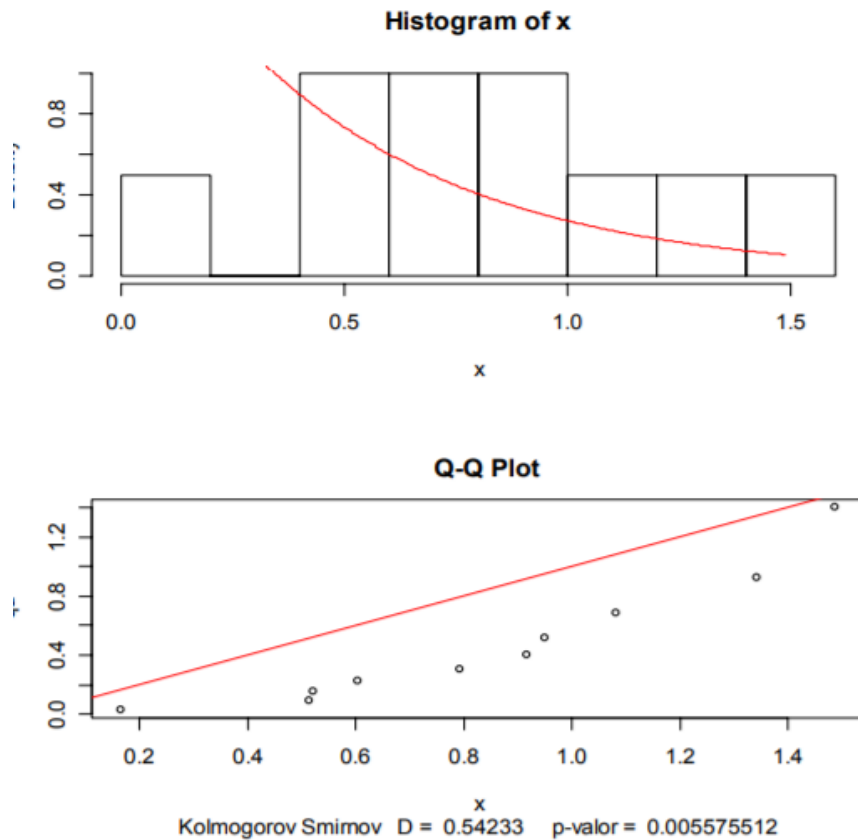
En este caso, obviamente se rechaza que la distribución sea exponencial, cosa que además se ve claramente en los gráficos.

Simulación de datos con distribución normal
n=10



En este caso, aunque los datos se han generado con distribución normal, el contraste conduce a aceptar que siguen la distribución normal. Ello se debe a que en general cuando hay poca información (en este caso sólo diez datos), la hipótesis nula tiende a no ser rechazada, salvo que haya una evidencia abrumadora en su contra.

Simulación de datos con distribución exponencial
n=10



Aquí ha ocurrido lo contrario al caso anterior; a pesar de que los datos son originalmente exponenciales, el contraste rechaza que lo sean.

4.2.4 Prueba de Anderson – Darling.f.d

El estadístico Anderson-Darling mide qué tan bien siguen los datos una distribución específica. Para un conjunto de datos y distribución en particular, mientras mejor se ajuste la distribución a los datos, menor será este estadístico. Por ejemplo, usted puede utilizar el estadístico de Anderson-Darling para determinar si los datos cumplen el supuesto de normalidad para una prueba t.

Las hipótesis para la prueba de Anderson-Darling son:

- H_0 : Los datos siguen una distribución especificada
- H_1 : Los datos no siguen una distribución especificada

Fórmula del estadístico: $A^2 = -N \cdot S$

Utilice el valor p correspondiente (si está disponible) para probar si los datos provienen de la distribución elegida. Si el valor p es menor que un nivel de significancia elegido (por lo general 0.05 o 0.10), entonces rechace la hipótesis nula de que los datos provienen de esa distribución. Minitab no siempre muestra un valor p para la prueba de Anderson-Darling, porque este no existe matemáticamente para ciertos casos.

También puede utilizar el estadístico de Anderson-Darling para comparar el ajuste de varias distribuciones con el fin de determinar cuál es la mejor. Sin embargo, para concluir que una distribución es la mejor, el estadístico de Anderson-Darling debe ser sustancialmente menor que los demás. Cuando los estadísticos están cercanos entre sí, se deben usar criterios adicionales, como las gráficas de probabilidad, para elegir entre ellos.

Ejemplo resuelto:

1.

20 números al azar:

19,55,30,79,97,75,65,90,77,22,45,16,57,66,30,91,88,58,29,86

$\mu = 58.75$

$\sigma = 26.83$

$\alpha = 0.05$

1. Crear la columna 1, 2 y en la 3 ordenar de menor a mayor, y en la 4 de mayor a menor

1	2	3	4
i	(2i-1)	Y_i	Y_{n+1-i}
1	1	16	97
2	3	19	91
3	5	22	90
4	7	29	88
5	9	30	86
6	11	30	79
7	13	45	77
8	15	55	75
9	17	57	66
10	19	58	65
11	21	65	58
12	23	66	57
13	25	75	55
14	27	77	45
15	29	79	30
16	31	86	30
17	33	88	29
18	35	90	22
19	37	91	19
20	39	97	16

Determinar Z de las columnas 3 y 4 $Z = \frac{x - \mu}{\sigma}$ y crear la columna 5. Y la columna 6 es la misma solo que ordenados inversamente

1	2	3	4	5	6
i	(2i-1)	Y_i	Y_{n+1-i}	Z_i	Z_{n+1-i}
1	1	16	97	-1.5117	1.3934
2	3	19	91	-1.4041	1.1782
3	5	22	90	-1.2965	1.1423
4	7	29	88	-1.0455	1.0706
5	9	30	86	-1.0096	0.9989
6	11	30	79	-1.0096	0.7478
7	13	45	77	-0.4716	0.6761
8	15	55	75	-0.1130	0.6043
9	17	57	66	-0.0412	0.2815
10	19	58	65	-0.0054	0.2457
11	21	65	58	0.2457	-0.0054
12	23	66	57	0.2815	-0.0412
13	25	75	55	0.6043	-0.1130
14	27	77	45	0.6761	-0.4716
15	29	79	30	0.7478	-1.0096
16	31	86	30	0.9989	-1.0096
17	33	88	29	1.0706	-1.0455
18	35	90	22	1.1423	-1.2965
19	37	91	19	1.1782	-1.4041
20	39	97	16	1.3934	-1.5117

Los valores de las columnas 7 y 8 son obtenidos de la tabla de distribución normal acumulada.

En Excel utiliza la función

=DISTR.NORM (valor, media, desviación estándar, acum)

7	8
F(Yi)	F(Y_{n+1-i})
0.0653	0.9182
0.0801	0.8806
0.0974	0.8733
0.1479	0.8578
0.1563	0.8411
0.1563	0.7727
0.3186	0.7505
0.4550	0.7272
0.4836	0.6109
0.4979	0.5970
0.5970	0.4979
0.6109	0.4836
0.7272	0.4550
0.7505	0.3186
0.7727	0.1563
0.8411	0.1563
0.8578	0.1479
0.8733	0.0974
0.8806	0.0801
0.9182	0.0653

Las columnas 9 y 10 se determina con logaritmos neperiano, para columna 9 se determina directo ($\text{LN}(< \text{valor columna7}>)$) y columna 10 se determina $\text{LN}((1- (< \text{valor columna8}>)))$ posteriormente se determina el resultado del logaritmo neperiano.

9	10
LN(F(Y _i))	LN(1-F(Y _{n+1-i}))
-2.7288	-2.5041
-2.5240	-2.1256
-2.3290	-2.0662
-1.9112	-1.9507
-1.8557	-1.8393
-1.8557	-1.4815
-1.1438	-1.3883
-0.7874	-1.2990
-0.7266	-0.9438
-0.6974	-0.9089
-0.5158	-0.6889
-0.4929	-0.6608
-0.3186	-0.6070
-0.2870	-0.3836
-0.2579	-0.1700
-0.1731	-0.1700
-0.1534	-0.1601
-0.1354	-0.1025
-0.1271	-0.0835
-0.0853	-0.0675

La última columna núm. 11 se determina con la fórmula

$$S_i = \frac{(2i-1)}{n} [\ln F(Y_i) + \ln(1 - F(Y_{n+1-i}))]$$

1	2	3	4	5	6	7	8	9	10	11
i	(2i-1)	Y _i	Y _{n+1-i}	Z _i	Z _{n+1-i}	F(Y _i)	F(Y _{n+1-i})	LN(F(Y _i))	LN(1-F(Y _{n+1-i}))	S _i
1	1	16	97	-1.5117	1.3934	0.0653	0.9182	-2.7288	-2.5041	-0.2616
2	3	19	91	-1.4041	1.1782	0.0801	0.8806	-2.5240	-2.1256	-0.6974
3	5	22	90	-1.2965	1.1423	0.0974	0.8733	-2.3290	-2.0662	-1.0988
4	7	29	88	-1.0455	1.0706	0.1479	0.8578	-1.9112	-1.9507	-1.3517
5	9	30	86	-1.0096	0.9989	0.1563	0.8411	-1.8557	-1.8393	-1.6628
6	11	30	79	-1.0096	0.7478	0.1563	0.7727	-1.8557	-1.4815	-1.8355
7	13	45	77	-0.4716	0.6761	0.3186	0.7505	-1.1438	-1.3883	-1.6459
8	15	55	75	-0.1130	0.6043	0.4550	0.7272	-0.7874	-1.2990	-1.5648
9	17	57	66	-0.0412	0.2815	0.4836	0.6109	-0.7266	-0.9438	-1.4198
10	19	58	65	-0.0054	0.2457	0.4979	0.5970	-0.6974	-0.9089	-1.5260
11	21	65	58	0.2457	-0.0054	0.5970	0.4979	-0.5158	-0.6889	-1.2649
12	23	66	57	0.2815	-0.0412	0.6109	0.4836	-0.4929	-0.6608	-1.3267
13	25	75	55	0.6043	-0.1130	0.7272	0.4550	-0.3186	-0.6070	-1.1570
14	27	77	45	0.6761	-0.4716	0.7505	0.3186	-0.2870	-0.3836	-0.9053
15	29	79	30	0.7478	-1.0096	0.7727	0.1563	-0.2579	-0.1700	-0.6204
16	31	86	30	0.9989	-1.0096	0.8411	0.1563	-0.1731	-0.1700	-0.5318
17	33	88	29	1.0706	-1.0455	0.8578	0.1479	-0.1534	-0.1601	-0.5171
18	35	90	22	1.1423	-1.2965	0.8733	0.0974	-0.1354	-0.1025	-0.4163
19	37	91	19	1.1782	-1.4041	0.8806	0.0801	-0.1271	-0.0835	-0.3897
20	39	97	16	1.3934	-1.5117	0.9182	0.0653	-0.0853	-0.0675	-0.2980

Se suman los valores de S_i(columna 11)

$$S = \sum_{i=1}^n S_i = -20.4916$$

Se aplicará el estadístico Anderson Darling:

$$A^2 = -N - S$$

$$A^2 = -20 - (-20.4916) = 0.4915$$

el valor del estadístico 0.4916 es menor al valor crítico (A²crítico=0.752), por lo tanto No se rechaza la hipótesis nula. Concluyendo que los datos observados tienen una naturaleza de Distribución Normal

4.2.5 Prueba de Ryan – Joiner.

Esta prueba evalúa la normalidad, calculando la correlación entre sus datos, y las puntuaciones normales de sus datos. Si el coeficiente de correlación se encuentra cerca de 1, es probable que la población sea normal. La estadística de Ryan-Joiner evalúa la solidez de esta correlación: si se encuentra del valor apropiado, usted rechazaría la hipótesis nula con normalidad en la población. Es una prueba

no paramétrica sobre si, los datos de una muestra provienen de una distribución específica.

La prueba de Ryan Joiner (es una extensión a la Kolmogorov y similar a la de Shapiro) está basado en la regresión y en la correlación.

$$R_p = \frac{\sum Y_i b_i}{\sqrt{s^2(n-1) \sum b_i^2}}$$

Donde:

- Y_i = Observaciones ordenadas
- b_i = Puntuaciones normales de las observaciones ordenadas ($z = \frac{x-\mu}{\sigma}$)
- S^2 = Varianza de la muestra

Ejemplo resuelto:

Pruebe la hipótesis de si los siguientes datos se ajustan a una distribución normal

27	17	29	21	41	40
22	48	19	38	26	54
34	35	46	32	43	52

- Establecer las hipótesis:

H_0 : Los datos se ajustan a una distribución normal.

H_1 : Los datos no se ajustan a una distribución normal.

- Calcular la media y desviación muestral de los datos

$$\bar{x}=34.666 \quad s=11.39$$

- Ordenar los datos en la columna de la tabla para obtener los valores de la fórmula.

i	Yi	bi	yi*bi	bi^2	Media	Desviación
1	17				34.66	11.39
2	19					
3	21					
4	22					
5	26					
6	27					
7	29					
8	32					
9	34					
10	35					
11	38					
12	40					
13	41					
14	43					
15	46					
16	48					
17	52					
18	54					

- Calcular $B_i = \frac{Y_i - \bar{x}}{s}$

i	Yi	bi	yi*bi	bi^2	Media	Desviación
1	17	-1.55048288			34.66	11.39
2	19	-1.37489025				
3	21	-1.19929763				
4	22	-1.11150132				
5	26	-0.76031607				
6	27	-0.67251975				
7	29	-0.49692713				
8	32	-0.23353819				
9	34	-0.05794557				
10	35	0.029850746				
11	38	0.293239684				
12	40	0.468832309				
13	41	0.556628622				
14	43	0.732221247				
15	46	0.995610184				
16	48	1.171202809				
17	52	1.52238806				
18	54	1.697980685				
		0.010535558				

- Calcular (Yi)(Bi)

i	Yi	bi	yi*bi	bi^2	Media	Desviación
1	17	-1.55048288	-26.358209		34.66	11.39
2	19	-1.37489025	-26.1229148			
3	21	-1.19929763	-25.1852502			
4	22	-1.11150132	-24.453029			
5	26	-0.76031607	-19.7682177			
6	27	-0.67251975	-18.1580334			
7	29	-0.49692713	-14.4108867			
8	32	-0.23353819	-7.47322212			
9	34	-0.05794557	-1.97014925			
10	35	0.029850746	1.04477612			
11	38	0.293239684	11.143108			
12	40	0.468832309	18.7532924			
13	41	0.556628622	22.8217735			
14	43	0.732221247	31.4855136			
15	46	0.995610184	45.7980685			
16	48	1.171202809	56.2177349			
17	52	1.52238806	79.1641791			
18	54	1.697980685	91.690957			
		0.010535558	194.219491			

- Calcular Bi²

i	Yi	bi	yi*bi	bi^2	Media	Desviación
1	17	-1.55048288	-26.358209	2.40399716	34.66	11.39
2	19	-1.37489025	-26.1229148	1.890323212		
3	21	-1.19929763	-25.1852502	1.438314804		
4	22	-1.11150132	-24.453029	1.235435178		
5	26	-0.76031607	-19.7682177	0.578080521		
6	27	-0.67251975	-18.1580334	0.45228282		
7	29	-0.49692713	-14.4108867	0.246936572		
8	32	-0.23353819	-7.47322212	0.054540087		
9	34	-0.05794557	-1.97014925	0.003357689		
10	35	0.029850746	1.04477612	0.000891067		
11	38	0.293239684	11.143108	0.085989512		
12	40	0.468832309	18.7532924	0.219803734		
13	41	0.556628622	22.8217735	0.309835422		
14	43	0.732221247	31.4855136	0.536147954		
15	46	0.995610184	45.7980685	0.991239639		
16	48	1.171202809	56.2177349	1.371716021		
17	52	1.52238806	79.1641791	2.317665404		
18	54	1.697980685	91.690957	2.883138406		
		0.010535558	194.219491	17.0196952		

- Aplicar la fórmula:

$$R_p = \frac{\sum Y_i b_i}{\sqrt{s^2(n-1) \sum b_i^2}}$$

$$R_p = \frac{192.44}{\sqrt{(11.32)^2(18-1)(17.015)}} = 0.992$$

- Decisión:

Si el valor de R_p es mayor a α H_0 se acepta. se considera que los datos son normales; en este caso $R_p = 0.992 > \alpha = 0.05$ H_0 se acepta.

- Conclusión:

Los datos analizados provienen de una distribución normal

4.2.6 Prueba de Shappiro – Wilk.

La prueba de Shapiro–Wilk se usa para contrastar la normalidad de un conjunto de datos. Se plantea como hipótesis nula que una muestra proviene de una población normalmente distribuida. Fue publicado en 1965 por Samuel Shapiro y Martin Wilk.¹ Se considera uno de los test más potentes para el contraste de normalidad.

El estadístico de la prueba es:

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Donde

$X_{(i)}$ con el subíndice i es el número que ocupa la i -ésima posición en la muestra (con la ordenada de menor a mayor).

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad \text{media muestral.}$$

Las variables a se calculan:

$$(a_1, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}$$

Ejemplo resuelto:

Plantea la hipótesis nula que una muestra proviene de una distribución normal. Elegimos un nivel de significancia, por ejemplo 0,05, y tenemos una hipótesis alternativa que sostiene que la distribución no es normal.

Tenemos:

H0: La distribución es normal

H1: La distribución no es normal,

Ahora la prueba Shapiro-Wilks intenta rechazar la hipótesis nula a nuestro nivel de significancia. Para realizar el test usamos la función

```
Grupo.A = c(15, 12, 11, 18, 15, 15, 9, 19, 14, 13, 11, 12, 18, 15, 16, 14, 16, 17,  
15, 17, 13, 14, 13, 15, 17, 19, 17, 18, 16, 14)
```

Shapiro-Wilk normality test

```
data: Grupo.A
```

```
W = 0.97032, p-value = 0.548
```

```
Grupo.B = c (11, 16, 14, 18, 6, 8, 9, 14, 12, 12, 10, 15, 12, 9, 13, 16, 17, 12, 8,  
7, 15, 5, 14, 13, 13, 12, 11, 13, 11, 7)
```

Shapiro-Wilk normality test

```
data: Grupo.B
```

```
W = 0.97636, p-value = 0.7227
```

Vemos que en ambos casos el valor de probabilidad (p) es muy superior a nuestro nivel elegido (0,05), por lo que *no rechazamos la hipótesis nula*.

Ejercicios propuestos unidad 4:

1.- Las tallas, medida en metros, de nueve peces espada (*Xiphias gladius*) capturados por un palangrero fueron: 1.628, 1.352, 1.800, 1.420, 1.594, 2.132, 1.614, 1.924 y 1.692. Comprobar si los datos siguen una distribución Normal.

2.- Se quiere comprobar si el número de asignaturas aprobadas en una determinada convocatoria universitaria sigue una distribución de Poisson (nivel de significación 0,05; para ello disponemos de la siguiente información de 60 alumnos

número de aprobadas	n° alumnos
0	10
1	15
2	15
3	10
4	6
5	4

Turno/producción	mañana		tardes		noche		
	obs	teo	obs	teo	obs	teo	
100-200 piezas	12	10,73	12	11,84	13	14,43	37
		0,15		0,0021		0,14	
200-300 piezas	7	8,99	12	9,92	12	12,09	31
		0,44		0,43		0,00	
300-400 piezas	10	9,28	8	10,24	14	12,48	32
	0,05		0,49		0,18		
	29		32		39		100

3.- Desean saber si una muestra de datos pertenece a una población normalmente distribuida. Los datos (ordenados son):

$X = 1, 3, 4, 4, 5, 5, 6, 6, 6, 6$. Utilice un alfa de .05.

4.- Se efectuaron mediciones del nivel de glucemia de 36 hombres adultos en ayuno, no obesos y aparentemente sanos. Estas mediciones se muestran en la tabla que se presenta. Se pretende saber si es posible concluir que tales datos no pertenecen a una población que sigue una distribución normal, con una media de 80 y una desviación típica de 6. Emple un $\alpha = 0,05$.

Valores de glucemia en 36 varones sanos					
75	92	80	80	84	72
84	77	81	77	75	81
80	92	72	77	78	76
77	86	77	92	80	78
68	78	92	68	80	81
87	76	80	87	77	86

5.- Una surtidora automática fue utilizada para llenar con 16 ml de un medicamento mediante un muestreo aleatorio se seleccionaron 8 frascos y se les midió el volumen envasado, encontrándose los sig. resultados:

16	15.9	15.97	16.04	16.05	15.98	15.96	16.02
----	------	-------	-------	-------	-------	-------	-------

Se quiere saber si la variable volumen servido se distribuye normalmente. Procedimiento para la prueba de normal de wilk-Shapiro.

Unidad 5.

Regresión Lineal Simple



Instrumentación Didáctica

Competencia Específica	Competencia Genérica	Actividades de Aprendizaje
Utiliza el diagrama de dispersión de datos bivariados de un experimento para hacer una estimación en procesos de la industria y la logística aplicando los conceptos de regresión lineal simple.	• Capacidad de abstracción, análisis y síntesis. • Capacidad de aplicar los conocimientos en la práctica. • Capacidad para identificar, plantear y resolver problemas.	• Buscar información sobre conceptos relacionados con la regresión lineal simple. • Discutir en grupo sobre los conceptos investigados. • Utilizar un modelo de regresión para propósitos de estimación y predicción para analizar la relación entre dos variables. • Comprender la importancia del análisis de regresión lineal simple. • Aplicar las pruebas de hipótesis para evaluar su calidad de ajuste. • Utilizar TIC's para realizar la regresión lineal simple.

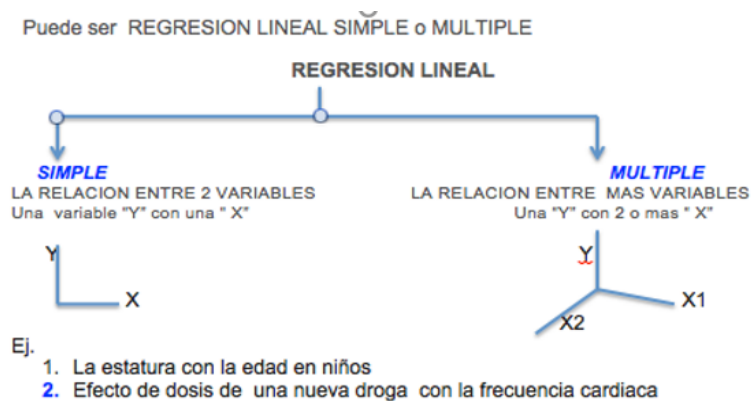
5.Regresión Lineal Simple.

El análisis de la regresión es un proceso estadístico para estimar las **relaciones entre variables**.

- Es un modelo matemático que modeliza la relación entre una variable dependiente Y, y las variables independientes X_i y un término aleatorio ξ

Se utiliza:

- ✓ Ampliamente para la predicción y previsión.
- ✓ Para Inferir relaciones causales entre una variable independiente y una dependiente
- ✓ Ayuda a entender como el valor de la variable dependiente varia al cambiar el valor de una de las variables independientes



Modelo matemático de una regresión lineal simple.

Ayuda a encontrar la recta de mejor ajuste para un conjunto de datos bivariado se realiza en forma matemática

$\hat{Y} = a + b x$; a y b son estimadores de los parámetros β_0 y β_1 entonces;

$$\hat{Y} = \beta_0 + \beta_1 x + \varepsilon$$

En donde:

β_0 = A la ordenada en el origen (el valor que toma Y cuando $X = 0$). EL ORIGEN DE LA RECTA

β_1 = Es la PENDIENTE de la recta (e indica como cambia Y al aumentar X en una unidad)

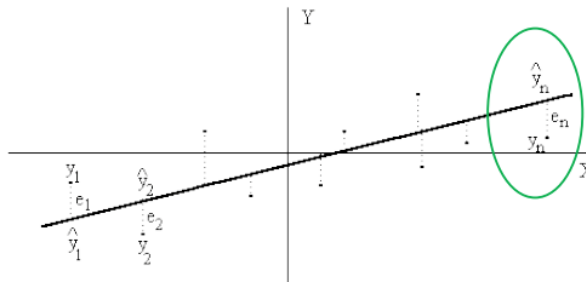
ε = Es una variable que incluye un conjunto grande de factores, cada uno influye en la respuesta sólo en pequeña magnitud al que le llamaremos ERROR.

Principio de mínimos cuadrados

Este método consiste en minimizar la suma de los cuadrados de los errores

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

es decir, la suma de los cuadrados de las diferencias entre los valores observados (y_i) y los valores estimados ($\hat{y}_i$).



Para hacer una estimación del modelo de regresión lineal simple, trataremos de hacer una forma de recta.

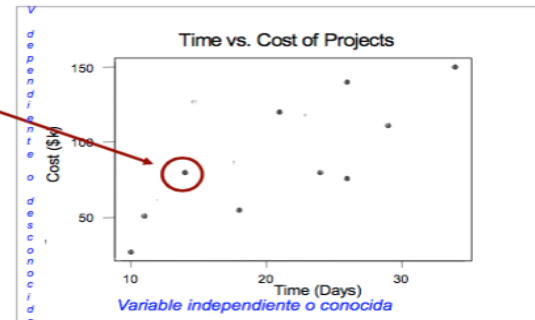
$$\hat{Y} = \beta_0 + \beta x = a + bx$$

Diagrama de dispersión:

Un diagrama de dispersión es una representación gráfica de datos pares (x,y).

“es un gráfico apropiado cuando queremos visualizar la correspondencia de valores de un conjunto de datos respecto a valores de otro conjunto de datos, de tal modo que podamos apreciar la posible relación entre ellos.

Project	Time (Days)	Cost (\$k)
1.	14	80
2.	29	111
3.	26	76
4.	10	27
5.	18	55
6.	11	51
7.	34	150
8.	26	140
9.	24	80
10.	21	120



Cuando aumenta la duración de un proyecto, los costos también aumentan.

“X” = **variable Independiente**, variable explicativa o variable CONOCIDA (es decir la primera que podemos conocer).

“Y” = **variable Dependiente**, variable de Respuesta (*porque nos va a responder con su valor*) o variable DESCONOCIDA (a partir de la conocida vamos a conocer su valor).

Tipos de relaciones entre dos variables:

- 1.- Directa
- 2.- Inversa

1.- RELACION DIRECTA: se da cuando “X” aumenta, y “Y” aumenta.



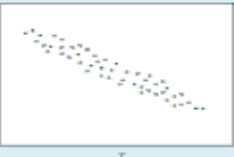
Ejemplo: Si aumentan las ventas de un producto específico, aumentarían también las ganancias económicas.

2.- RELACIÓN INVERSA: se da cuando “X” aumenta, y “Y” disminuye.

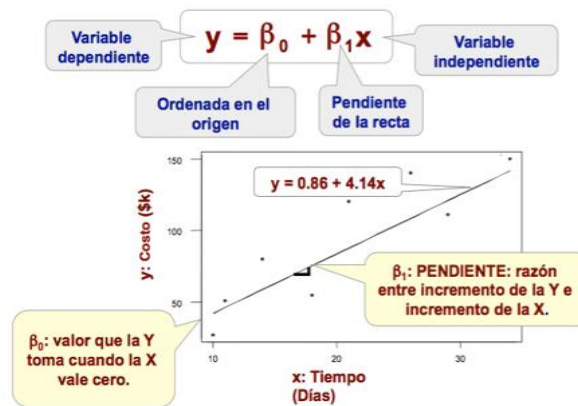


Por ejemplo: Si aumentan la implementación de anticontaminantes en una ciudad, entonces el % de contaminación ambiental disminuirá.

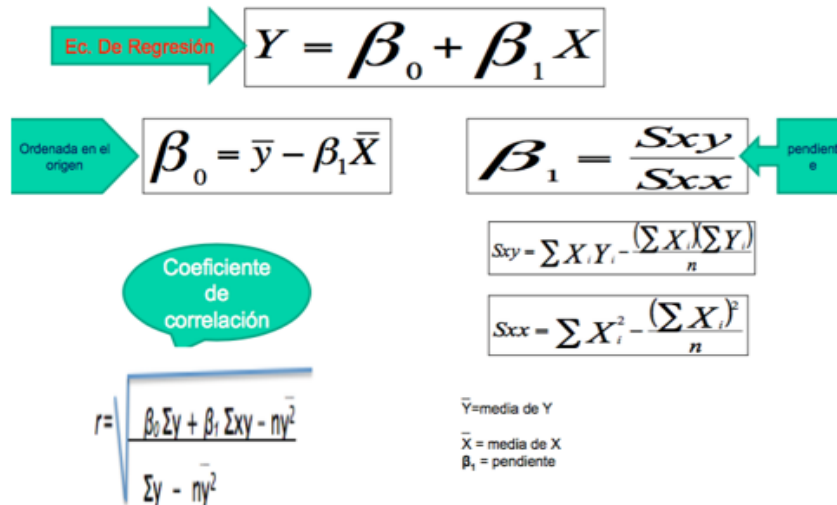
Representación Gráfica de los tipos de relaciones

	Baja correlación positiva. El valor de X aumenta ligeramente a medida que aumenta el valor de Y.
	Fuerte correlación negativa. El valor de X claramente disminuye a medida que aumenta el valor de Y.
	Débil correlación negativa. El valor de X disminuye ligeramente a medida que aumenta el valor de Y.
	Relación compleja. El valor de X parece estar relacionado con el valor de Y, pero esa relación no es fácil de establecer.

Determinación de la Ecuación de regresión:



Determinación manual de la ecuación de regresión mediante el método de MINIMOS CUADRADOS



Coeficiente de Determinación

Mide o Determina la calidad del modelo para replicar los resultados.

Se simboliza:

$$R^2 = r^2$$

Su valor:

$$-1 \leq r^2 \leq 1$$

$R^2 = 1$ o -1 , significa que la ecuación de regresión ofrece un perfecto ajuste para los datos de la muestra y acercándose al valor de -1 o 1 es una relación fuerte.

$$R^2 \text{ o } r^2 = \frac{(\beta_0)(\Sigma y) + \beta_1(\Sigma xy) - n(\bar{y}^2)}{\Sigma y^2 - n(\bar{y}^2)}$$

Ejemplo resuelto:

Ejemplo de REGRESION LINEAL SIMPLE. Por medio del método de MINIMOS CUADRADOS.

Determine:

- ecuación de regresión

- b) diagrama de dispersión
- c) trace la línea de mejor ajuste
- d) pronostique los gastos de reparación para un camión que tiene 10 años de antigüedad.
- e) cuál es la fuerza de relación entre las 2 variables (R)?
- f)Cuál es la calidad con que trabaja su modelo de pronóstico o ec. de regresión, (R²)?

Camión	edad	Gastos de reparación
101	5	7
102	3	7
103	3	6
104	1	4

CAMION	EDAD X	GASTOS DE REPARACIÓN Y	(X) (Y)	X ²	Y ²	
101	5	7	35	25	49	
102	3	7	21	9	49	
103	3	6	18	9	36	
104	1	4	4	1	16	
sumatoriasΣ	Σx=12	Σy=24	Σxy= 78	Σx ² = 44	ΣY ² = 150	
Media	$\bar{X} = 3$	$\bar{Y} = 6$				

$$S_{xy} = \sum X_i Y_i - \frac{(\sum X_i)(\sum Y_i)}{n}$$

$$S_{XY} = 78 - \frac{(12)(24)}{4} = \underline{6}$$

$$\beta_1 = \frac{S_{xy}}{S_{xx}}$$

$$\beta_1 = \frac{6}{8} = \underline{.75 \text{ pendiente}}$$

$$S_{XX} = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

$$\beta_0 = 6 - .75 (3) = \underline{3.75}$$

$$S_{XX} = 44 - \frac{12^2}{4} = \underline{8}$$

$$Y = \beta_0 + \beta_1 X$$

$$\beta_0 = \bar{y} - \beta_1 \bar{X}$$

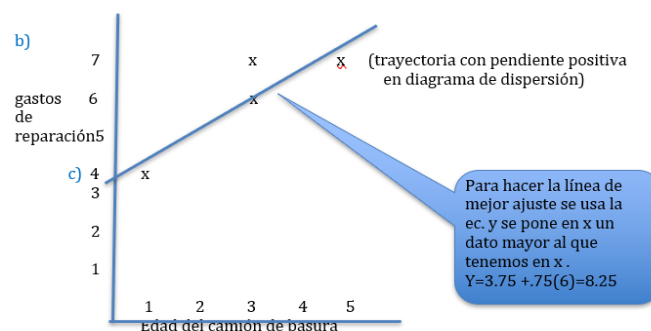
a) $Y = \underline{3.75 + .75 X}$

$$\beta_0 = 6 - .75 (3) = \underline{3.75}$$

Modelo de regresión
ó

Ordenada en el origen

Ec. de pronóstico ó Ec. de mínimos cuadrados.



e) $r = \sqrt{\frac{(3.75)(24) + (.75)(78) - 4(36)}{150 - (4)(36)}}$

$$r = .75$$

$r = .75$ coeficiente de correlación indica que tiene una fuerza de relación del 75% entre la edad del camión y los gastos de reparación en un camión de basura.

d) $Y = 3.75 + .75 (10) = 11.25$ miles de pesos de Pronóstico

f) $r^2 = (.75)^2 = .5625 = 56.25\%$

5.1 Prueba de hipótesis en la regresión lineal simple.

Cuando se miden 2 variables aleatorias en cada unidad experimental. En esta población bivariado, el coeficiente de correlación poblacional ρ (letra griega rho) se calcula e interpreta como tal en la muestra.

La prueba de hipótesis es para probar de que no hay ninguna correlación entre variables X y Y, y mediante un estadístico de prueba que equivale exactamente a la prueba de la β .

El establecimiento de las hipótesis es el siguiente:

$$\begin{array}{l} H_0: \beta = 0 \\ H_1: \beta \neq 0 \end{array} \Leftrightarrow \begin{array}{l} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{array}$$

El estadístico de prueba:

$$t = \frac{r \sqrt{n-2}}{\sqrt{1-r^2}}$$

región de rechazo:

1 cola

$$t > t_{\alpha}$$

2 colas

$$t > t_{\alpha/2} \text{ o } t < t_{\alpha/2}$$

$$t < -t_{\alpha} \text{ cuando } H_1 = \rho < 0, \text{ el valor de } t < \alpha$$

Ejemplo resuelto:

Las alturas y los pesos de 10 jugadores de futbol americano se eligen al azar del equipo estelar de una cd. Calcule el coeficiente de correlación para las alturas en pulgadas y pesos en lbs.

Jugador	Estatura x	Peso Y
1	73	185
2	71	175
3	75	200
4	72	210
5	72	190
6	75	195
7	67	150
8	69	170
9	71	180
10	69	175

$$\sum x = 714$$

$$\beta_0 = -204.73$$

$$\sum y = 1830$$

$$\beta_1 = 5.43$$

$$\sum xy = 130990$$

$$r = .8261$$

$$\sum x^2 = 51040$$

$$r^2 = .6824$$

$$\sum y^2 = 337500$$

La correlación r difiere de 0 en forma significativa

$$H_0 = \rho = 0$$

$$H_1 = \rho \neq 0$$

$$t = 0.8261 \sqrt{\frac{10-2}{1-.8261^2}} = 4.15$$

con 8 gl y $\alpha = .01$; $\alpha/2 = .005$; $t_{\text{tabla}} = 3.35$

Ho se rechaza porque $t_{\text{calculada}}$ es mayor que t_{tabla}

Conclusión: la diferencia es significativa. Entre las 2 variables; por lo que sí hay correlación entre estatura y peso en el jugador

5.2 Calidad del ajuste en regresión lineal simple.

La calidad del ajuste de una regresión lineal simple permite verificar la calidad con la que el modelo planteado permite hacer estimaciones. Se necesita conocer qué tanta variabilidad en Y fue explicada por el modelo, si se cumplen los supuestos de normalidad en los residuos y si la variación no tiene ningún patrón fuera de lo usual.

La presentación de varios criterios para evaluar la calidad del modelo tiene el propósito de destacar que los buenos modelos se construyen a medida que cumplen más criterios de calidad de ajustes. El no cumplimiento de alguno de los criterios no hará necesariamente inviable el modelo desde el punto de vista práctico. Entre los criterios que utilizamos para la evaluación de la calidad del ajuste, se encuentran:

- Verificación de los supuestos: evaluación de la normalidad de los residuales por shapiro- Wilks y una combinación de pruebas gráficas.
- Verificación del supuesto de Independencia: mediante gráfico de los residuos contra el orden en que se obtuvieron los datos.

5.3 Estimación y predicción por intervalo en regresión lineal simple.

Intervalo de confianza para β_1

Fórmula para un intervalo de confianza de β_1 es el siguiente:

$$b_1 \pm t_{\alpha/2} S_{b_1}$$

El estimador puntual es b_1 y el margen de error $t_{\alpha/2} S_{b_1}$. El coeficiente de confianza para este intervalo es $1 - \alpha$ y $t_{\alpha/2}$ es el valor de t que proporciona un área $\alpha/2$ en la cola superior de la distribución t con $n-2$ grados de libertad.

Ejemplo resuelto:

se recolectan datos de una muestra de 10 restaurantes Armand's Pizza Parlors ubicados todos cerca de campus universitarios. Se presentan los valores de X_i y Y_i en esta muestra de 10 restaurantes. Como se ve, el restaurante 1, para el que $X_1 = 2$ y $Y_1 = 58$, está cerca de un campus de 2000 estudiantes y sus ventas trimestrales son de \$58 000. El restaurante 2, para el que $X_2 = 6$ y $Y_2 = 105$, está

cerca de un campus de 6000 estudiantes y sus ventas trimestrales son de \$105 000. El valor mayor es el que corresponde a ventas del restaurante 10, el cual está cerca de un campus de 26 000 estudiantes y sus ventas trimestrales son de \$202 000.

Restaurante i	Población de estudiantes (miles) x_i	Ventas trimestrales (miles de \$) y_i
1	2	58
2	6	105
3	8	88
4	8	118
5	12	117
6	16	137
7	20	157
8	20	169
9	22	149
10	26	202

Armands Pizza Parlots se desea obtener una estimación de β_1 mediante un intervalo de 99% de confianza.

Con $\alpha/2 = 0.01/2 = .005$

$gl = 10 - 8 = 2$

$t_{\text{tablas}} = 3.355$.

$$b_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{2840}{568} = 5$$

$s = 13.829$.

$$\sum(x_i - \bar{x})^2 = 568$$

$$Sb_1 = \frac{13.829}{\sqrt{568}} = 0.5803$$

por lo tanto, la estimación mediante un intervalo del 99% de confianza es:

$$b_1 \pm t_{\alpha/2} Sb_1$$

$$5 \pm 3.355(.5803) =$$

$$5 \pm 1.95 =$$

$$\text{de } \underline{3.05} \text{ a } \underline{6.95}$$

Intervalo de Predicción.

Supóngase que en lugar de que lo que interese sea estimar el valor medio de las ventas de todos los restaurantes Armand's que se encuentran cerca de campus de 10 000 estudiantes, se deseen estimar las ventas de un solo restaurante que

se encuentra cerca de Talbot College, una escuela de 10 000 estudiantes. Como ya se indicó, la estimación puntual de y_p , el valor de y que corresponde a un valor dado x_p , se obtiene mediante la ecuación de regresión $\bar{y}_p = b_0 + b_1 x_p$. En el caso del restaurante cerca de Talbot College, como $x_p = 10$, las ventas trimestrales pronosticadas serán $\bar{y}_p = 60 + 5(10) = 110$ o \$110 000. Obsérvese que este valor es el mismo que el obtenido como estimación puntual de la media de las ventas en los restaurantes que se encuentran cerca de campus de 10 000 estudiantes. Para obtener un intervalo de predicción, es necesario determinar primero la varianza correspondiente al uso de como estimación de un valor individual de y cuando a $x = x_p$. Esta varianza está formada por la suma de los dos componentes siguientes.

1. La varianza de los valores individuales de y respecto a la media $E(y_p)$, para la cual una estimación está dada por s^2
2. La varianza correspondiente al uso de $\bar{y}_p$ para estimar $E(y_p)$, para la cual una estimación está dada por La fórmula para estimar la varianza de un valor individual de y_p que se denota s_{ind}^2 , es:

$$s_{ind} = s \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$$

sustituyendo

$$\begin{aligned} s_{ind} &= 13.829 \sqrt{1 + \frac{1}{10} + \frac{(10 - 14)^2}{568}} \\ &= 13.829 \sqrt{1.1282} \\ &= 14.69 \end{aligned}$$

y la fórmula general de predicción es:

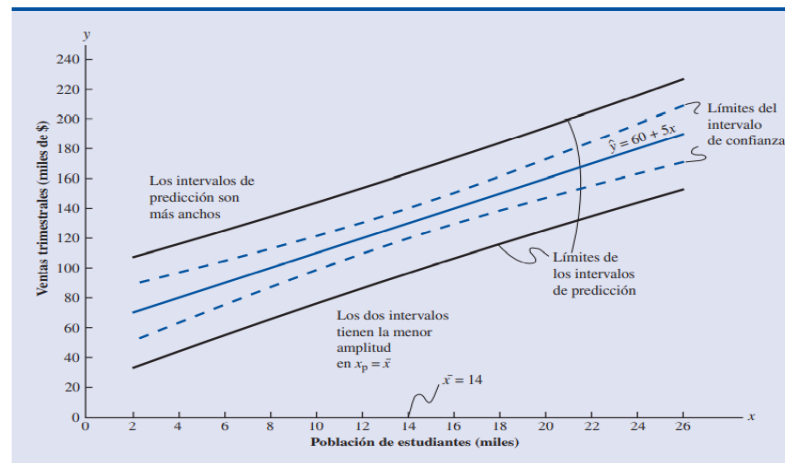
$$\bar{Y}_p \pm t_{\alpha/2} S_{ind}$$

Donde el coeficiente de confianza es $1 - \alpha$ y $t_{\alpha/2}$ es un valor de la distribución t para $n-2$ grados de libertad.

Ejemplo resuelto:

El intervalo de predicción de las ventas trimestrales del restaurante situado cerca de Talbot Collage se encuentra empleando $t_{0.025} = 2.306$ y $S_{ind} = 14.69$. Por lo tanto, como $\bar{Y}_p = 110$ y el margen de error es $t_{\alpha/2} S_{ind} = 2.306(14.69) = 33.875$, el intervalo de predicción de 95% de confianza es:

$$110 \pm 33.875$$



Intervalos de confianza y predicción para las ventas y que corresponden a valores dados x del tamaño de la población de estudiantes.

En dólares, el intervalo de predicción es

$$\$110,000 \pm \$33,875.$$

es decir:

$$\$76,125 \text{ a } \$143,875.$$

El intervalo de predicción para un solo restaurante que se encuentra cerca de un campo de 10,000 estudiantes es más amplio que el intervalo de confianza para la media de las ventas de todos los restaurantes que se encuentran cerca de campus de 10,000 estudiantes. Esta diferencia refleja el hecho de que se puede estimar con mas precisión la media de y que un solo valor individual de y .

Tanto las estimaciones mediante un intervalo de confianza como las estimaciones mediante un intervalo de predicción son más precisas cuando el valor de la variable independiente es $X_p = \bar{x}$. La figura muestra la forma general de los intervalos de confianza y de los intervalos de predicción que son más anchos.

Ejercicios Propuestos Unidad 5:

1.- Por medio de un procedimiento químico llamado “polarografía diferencial de pulso” un químico midió la corriente máxima (en microamperes), que se genera al agregar a una solución amortiguadora otra que contiene una cierta cantidad de níquel (en partes por c/1000 billones)

Ni (ppb)	CORRIENTE MAXIMA (m.a.)			
34.1	.445			
53.1	.524			
72.3	.606			
91.2	.698			
110	.779			
129	.85			
146	.93			
165	1.001			
185	1.072			

- Identificar cual es la variable Independiente (X).
 - Identificar cual es la variable Dependiente (Y).
 - ¿Qué tipo de relación se observa en el diagrama de dispersión?
 - Determine la ec. De regresión
 - Pronostique la corriente máxima para 150 de níquel
 - Determine el coeficiente de correlación
 - Determine el coeficiente de determinación.
- 2.- El inventor de un nuevo material aislante quiere determinar la magnitud de la compresión que se producirá en una pieza de 2 pulgadas de espesor cuando se somete a diferentes cantidades de presión Para ello prueba 5 piezas de material

bajo diferentes presiones. Los pares de valores observados (x, y) se muestran en la siguiente tabla:

Pieza	Presión	Compresión
1	1	1
2	2	1
3	3	2
4	4	2
5	5	4

- Identificar cual es la variable Independiente (X).
- Identificar cual es la variable Dependiente (Y).
- ¿Qué tipo de relación se observa en el diagrama de dispersión?
- Determine la ec. De regresión
- Pronostique la compresión para una presión de 8
- Determine el coeficiente de correlación.
- Determine el coeficiente de determinación.

3.- En el departamento de prototipos de una empresa manufacturera de arneses para auto desean saber que tan directa es la relación entre las variables de experiencia y productividad, por lo que muestrearon a varios empleados que tienen diferencia en experiencia y registraron sus producciones en 6 meses.

empleado	experiencia	Producción
1	5	20
2	11	40
3	4	30
4	5	34
5	3	35
6	2	20

- Identificar cual es la variable Independiente (X).
- Identificar cual es la variable Dependiente (Y).
- ¿Qué tipo de relación se observa en el diagrama de dispersión?
- Determine la ec. De regresión
- Pronostique la producción para una experiencia de 9 años
- Determine el coeficiente de correlación
- Determine el coeficiente de determinación.

4.-Los datos siguientes son :

X_i	1	2	3	4	5
Y_i	3	7	5	11	14

- Estime la desviación estándar $\bar{Y}_p$ cuando $x=4$
- Obtenga un intervalo de confianza del 95% para el valor esperado de y cuando $x=4$
- estime la desviación estándar de un valor de y cuando $x = 4$.
- obtener un intervalo de predicción de 95% para y cuando $x = 4$

5.-

X_i	2	6	9	13	20
Y_i	7	18	9	26	23

Obtenga los intervalos de confianza y de predicción del 95% para $x = 12$. Explique por qué son diferentes estos dos intervalos

6.- En los datos siguientes, y corresponde a los salarios mensuales y x es el promedio obtenido por los estudiantes que terminaron la licenciatura de administración con especialidad en sistemas de información. La ecuación de regresión estimada obtenida con estos datos es $\bar{y} = 1790.5 + 581.1X$

Promedio	Salario mensual (\$)
2.6	3300
3.4	3600
3.6	4000
3.2	3500
3.5	3900
2.9	3600

7.- La siguiente tabla de contingencia 3 X 3 contiene las frecuencias observadas en una muestra de 240. Pruebe la independencia de la variable de los renglones y la variable en las columnas usando la prueba X^2 con $\alpha = 0.05$.

	Variable de las columnas		
Variable de los renglones	A	B	C
P	20	30	20

Q	30	60	25
R	10	15	30

a partir de los datos de temperatura ($^{\circ}\text{F}$) = x y precio (\$) = y de 12 sacos de dormir, fabricados por Bergans of Noway, se obtuvo la ecuación de regresión $\bar{Y} = 359.2668 - 5.2772x$. Para estos datos $s = 37.9372$.

Dé una estimación puntual del precio de un saco de dormir cuya temperatura sea 30

- Dé un intervalo de 95% de confianza para el precio medio de todos los sacos de dormir cuya temperatura sea 30.
- Suponga que Bergans elabora un nuevo modelo cuya temperatura es 30.
- Dé un intervalo de predicción de 95% para el precio de este nuevo modelo.
- Explique la diferencia entre sus respuestas a los incisos b) y c).

8.- Fidelity Magellan es un fondo mutualista de capital grande y Fidelity Small Cap Stock es un fondo mutualista de capital pequeño (Morningstar Funds 500, 2006). La desviación estándar de ambos fondos se calculó empleando muestras aleatorias de tamaño 26. La desviación estándar muestral de Fidelity Magellan fue 8.89% y la desviación estándar muestral de Fidelity Small Cap Stock fue 13.03%. Los analistas financieros suelen usar la desviación estándar como una medida del riesgo. Realice una prueba de hipótesis para determinar si los fondos de capital pequeño son más riesgosos que los fondos de capital grande. Use $\alpha = 0.05$ como nivel de significancia.

La ecuación de regresión estimada $\bar{Y} = 1246.67 + 7.6x$ se obtuvo empleando los datos de volumen de producción x y costos totales y de una determinada operación de fabricación.

- En el plan de producción de la empresa se ve que el mes próximo deberán producirse 500 unidades. Dé la estimación puntual de los costos totales
- Dé un intervalo de predicción de 99% para el costo total de producción de las 500 unidades, el mes próximo.
- Si al final del mes próximo, el informe de costos de un contador indica que en ese mes los costos reales de producción fueron \$6000, ¿debería preocupar a los gerentes el haber incurrido ese mes en costos totales tan altos? Analice.

REFERENCIAS BIBLIOGRAFICAS

Anderson, D. R. (2008). Estadística para administración y economía. (10ª. ed.) México : Cengage Learning.

Applets relacionados con la representación de una distribución normal.
<http://psych.colorado.edu/~mcclella/java/normal/accurateNormal.html>

Berenson, M. (2006). Estadística para administración. (4ª. ed.) México : Pearson Educación.

Carot, V. (2006). Control estadístico de la calidad. España :Alfaomega.

Descripción y applet del Teorema Central del Límite
http://www.ideamas.cl/cursoProb/javaEstat/central_limit_theorem/clt.html

Devore, J. L. (2012) Probabilidad y estadística para ingenierías y ciencia. (8ª. ed.) México : Cengage Learning.

Gutiérrez, P. H. (2012). Análisis y diseño de experimentos. (3ª. ed.) México : McGraw-Hill

Gutiérrez, P. H. (2009). Control estadístico de calidad y seis sigma. (2ª. ed) México : McGrawHill.

Johnson, R. A. (2012) Probabilidad y estadística para ingenieros. (8ª. ed.) México : Pearson Educación.

Levine, D. M. (2010) Estadística para administración y economía. (7ª. ed.) México : Pearson Educación.

Mendenhall, W. (2010). Introducción a la Probabilidad y Estadística. (13ª. ed.) México: Cengage Learning.

Montgomery, D. C. (2011). Probabilidad y estadística aplicadas a la ingeniería. (2ª. ed.) México : Limusa : Wiley.

Ross, S. M. (2002). Probabilidad y estadística para ingenieros. México : McGraw-Hill.

Selección de recursos en Internet para la enseñanza-aprendizaje de la Estadística

<http://www.udc.es/dep/mate/recursos.html>

Soporte Minitab 18

<https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/how-to/normality-test/before-you-start/example/>

Vidal Bajaña Silvia María (2019). “Aplicación de Herramientas de Calidad para Evaluar Características Nutricionales de Alimento Balanceado para Aves Guayaquil - Ecuador.

Walpole, R. E. (2012). Probabilidad y estadística para ingeniería y ciencias. (9ª. ed) México : Pearson Educación.

ANEXOS

ANEXO 1

RESPUESTA A LOS PROBLEMAS PROPUESTOS:

UNIDAD 1: DISTRIBUCIÓN DE MUESTREO

1.
 - a) 0.9960
 - b) 0.2
2.
 - a) 0.3221
 - b) 0.3122
3.
 - a) 0.2150
 - b) 0.0064
 - c) 0.4504
4.
 - a) 0.2227
 - b) 0.2848
5.
 - a) 6
 - b) 9
 - c) 2
 - d) 12
 - e) $\mu_x = 22.4$, $\sigma_x = 0.008$
 - f) $\mu_x = 22.4$, $\sigma_x =$ ligeramente menor que 0.008

6.

d) 0.6898

7.

e) 0.0013

8.

$$t_{\text{calculada}}=2.25 \quad t_{\text{tabla}}=.02$$

Es probable que el fabricante concluya que el proceso produce un mejor producto del que piensa. $P(-2.977 < t < -1.761) = 0.045$

9.

$$P(-t_{0.025} < t < t_{0.05}) = 0.925$$

10.

$X^2_{\text{cal.}}=32$ El valor de 32 se busca adentro de la tabla en el renglón

de 16 grados de libertad y se encuentra que a este valor le corresponde un área a la derecha de 0.01. En consecuencia, el valor de la probabilidad es $P(s^2 > 2)$.

11.

valor de ji-cuadrada correspondiente a $s^2=2$

a) $X^2=36.4$ Con el valor de 32 y 16 gl.en la tabla corresponde a un área derecha de 0.01, Por lo que la $P(s^2 > 2)$.

b) $X^2_1=13.847$ y $X^2_2=42.98$

Con 24 gl. Buscar en la tabla 13.846^a la derecha= 0.95. el valor 42.98 a la derecha da un área de 0.01 entonces entre 2 valores se resta el área de 0.95 menos 0.01 quedando 0.94

Por lo tanto, la $P(3.462 \leq X^2_{10.745}) = 0.94$

UNIDAD 2: ESTIMACIÓN

1.

a) $32 \pm 1.645\left(\frac{6}{\sqrt{50}}\right)$

32 ± 1.4 ; 30.6 a 33.4

b) $32 \pm 1.96\left(\frac{6}{\sqrt{50}}\right)$
 32 ± 1.66 ; 30.34 a 33.66

c) $32 \pm 2.576\left(\frac{6}{\sqrt{50}}\right)$
 32 ± 2.19 ; 29.81 a 34.19

2.

a) $1.96\left(\frac{5}{\sqrt{49}}\right)=1.40$
b) 24.80 ± 1.40 ; 23.40 a 26.20

3.

8.1 a 8.9

4.

- a) Que la población es por lo menos aproximadamente normal
b) 3.1
c) 4.1

5.

$1300 \pm 2.58 (184.4)$
 $824.2 \leq \mu_1 - \mu_2 \leq 1775.8$ millas / hr.
Si hay diferencia entre los 2 tipos de marcas de neumáticos

6.

$.01 \pm 1.64 (.00712)$
 $-.00168 \leq P_1 - P_2 \leq .0216$ no hay diferencia entre los 2
Procesos.

7.

$8.399 \leq \sigma^2 \leq 39.828$

8.

$.000218 \leq \sigma^2 \leq .0035$

9.

$0.0162 \leq \sigma_1^2/\sigma_2^2 \leq .4537$ si hay diferencia

10.

- a) Valor planeado para $\sigma = \frac{\text{rango}}{4} = \frac{36}{4} = 9$
 b) $n = \left(\frac{(1.96)(9)}{3} \right)^2 = 34.57$, use $n=35$
 c) $n = \left(\frac{(1.96)(9)}{2} \right)^2 = 77.79$; use $n = 78$

11.

- a) 18
 b) 35
 c) 97

12.

- a) 0.1790
 b) 0.0738, 0.5682 a 0.7158
 c) 354

13.

- a) $n = \frac{(1.96)^2 (.156)(1-.156)}{(.03)^2} = 562$
 b) $n = \frac{(2.576)^2 (.156)(1-.156)}{(.03)^2} = 970.77$; use 971

UNIDAD 3. PRUEBA DE HIPÓTESIS

1. $X^2 = 27.075$ es mayor a el valor critico $X^2_{\text{tablas}} = 24.99$ H_0 se rechaza.
2. $H_0: \sigma^2 \leq .0004$
 $H_a: \sigma^2 > .0004$
 $X^2 = (29)(.0005)/.0004 = 36.25$. para 29 grados de libertad, el valor-p es mayor que 0.010.
 Valor-p > 0.05; no rechazar H_0 .
 Las especificaciones del producto no parecen estarse violando.
3. $H_0: \sigma^2 \leq 331.24$
 $H_a: \sigma^2 > 331.24$

$$X^2=52.07, gl=35$$

El valor-p está entre 0.025 y 0.05 Rechazar H_0 .

4. a) 0.8106

b) $X^2=9.49$ El valor-p es mayor a 0.20. No rechazar H_0

5. $H_0: P_1=P_2$

$H_1: P_1 \neq P_2$

$z=2.40$; sí, $P=0.01$ si hay diferencia significativa entre los que están a favor en los residentes urbanos y los suburbanos.

6. $P<0.0001$, concluir que menos de 1/5 de las casas se calientan con petróleo.

7.

a) $z=2.50$; se rechaza H_0 .

b) $P=0.0062$.

c) 0

8.

a) Se Rechaza H_0 , $z=-6.32$

b) 0

c) 0.248

9.

valor $P=0.6$. No se rechaza H_0 , $z=0.93$, valor de $P=0.1762$

Rechazar H_0 . Valor $P=0.0207$

a) 0.0358

b) 0.0802

c) 0.5824

d) 0.1586

e) 0

10.

a) $z=2.89$, no se debe usar el acero de alta pureza o se no se rechaza H_0 .

b) 0.2981

11.

a) Valor $P=0.4247$, no se rechaza H_0 .

UNIDAD 4. PRUEBAS DE BONDAD DE AJUSTE Y PRUEBAS NO PARAMÉTRICAS.

1.

Ho: La distribución de los defectos de tarjeta de circuitos tiene una distribución poisson con un promedio de 0.75

Ha: La distribución de los defectos de tarjeta de circuitos tiene una distribución poisson con un promedio de 0.75

$$X^2_{\text{calculada}} = 3.40 \quad X^2_{\text{tabla}} = 7.81 \quad \text{por lo tanto;}$$

$$X^2_{\text{calculada}} < X^2_{\text{tabla}} \quad \text{Ho no se rechaza}$$

Conclusión: La distribución de los defectos de tarjeta de circuitos tiene una distribución poisson con un promedio de 0.75.

2.

Ho: el proceso de entrevistas con prob. de 40% de tener una calif. positiva tiene una distribución binomial

Ha: el proceso de entrevistas con prob. de 40% de tener una calif. positiva no tiene una distribución binomial

$$X^2 = 5.050 \quad X^2_{\text{tabla}} = 9.23, \quad \text{entonces; } X^2 > X^2_{\text{tabla}}, \quad \text{por lo tanto}$$

Conclusión: el proceso de entrevistas con prob. de 40% de tener una calif. positiva no tiene una distribución binomial

3.

Ho: La preferencia del método de evaluación es independiente de la región de aplicación. (*verdadera*)

H₁: La preferencia del método de evaluación es dependiente de la región de aplicación. (*falsa*)

$$X^2_c = 2.71$$

$$gl = (\#R - 1)(\#C - 1) = 3 \text{ gl} \quad \text{y } \alpha = .05 \quad X^2_{\text{tabla}} = 7.81$$

$$X^2_c < X^2_{\text{tabla}} \quad \text{por lo tanto Ho no se rechaza}$$

CONCLUSION: La preferencia del método de evaluación es independiente de la región de aplicación

4.

Ho: Los defectos son independientes del turno que trabaja el mueble

Hi: Los defectos son dependientes del turno que trabaja el mueble.

$$X^2_c = 19.18$$

$gl=(\#R-1)(\#C-1)=6$ gl y $\alpha = .05$ $X^2_{\text{tabla}} = 7.81$

$X^2_c > X^2_{\text{tabla}}$ por lo tanto H_0 se rechaza

CONCLUSION: Los defectos son dependientes del turno que trabaja el mueble.

5.

H_0 : la práctica de un deporte es independiente de la depresión en los jóvenes.

H_1 : la práctica de un deporte es dependiente de la depresión en los jóvenes.

$X^2_c = 5.82$

$gl=(\#R-1)(\#C-1)=1$ gl y $\alpha = .05$ $X^2_{\text{tabla}} = 3.84$

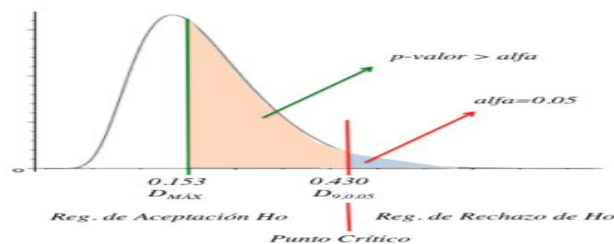
$X^2_c > X^2_{\text{tabla}}$ por lo tanto H_0 se rechaza

CONCLUSION: la práctica de un deporte es dependiente de la depresión en los jóvenes.

Pruebas no paramétricas:

1.

$\bar{x} = 1,684$						
$s = 0,242$						
x_i	F_i	z_i	$\Phi(z_i)$	$ F_i - \Phi(z_i) $	$ F_{i-1} - \Phi(z_i) $	
1.352	1/9	-1.373	0.085	0.026	0.085	
1.420	2/9	-1.092	0.137	0.085	0.026	
1.594	3/9	-0.372	0.355	0.022	0.0133	
1.614	4/9	-0.289	0.386	0.058	0.053	
1.628	5/9	-0.232	0.408	0.147	0.036	
1.692	6/9	0.033	0.513	0.153	0.042	
1.800	7/9	0.480	0.684	0.094	0.018	
1.924	8/9	0.992	0.840	0.049	0.062	
2.132	1	1.852	0.968	0.032	0.079	



$0,153 < D_{0,05} = 0,430 \Rightarrow$ No Rechazo H_0 para $\alpha = 0,05$.

$p\text{-valor} > \alpha$, los datos nos indican que NO debemos rechazar o no hay evidencia estadística suficiente para rechazar H_0 , es decir la variable X se ajusta a una distribución normal.

2.

apr	$n_{0,i}$	$P(X_i) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$	$F_o(x_i)$	$F_T(x_i)$	$D = F_T(x_i) - F_o(x_i) $
0	10	0,1376	10/60=0,16	0,13	0,03 máximo
1	15	0,272	25/60=0,41	0,4	0,01
2	15	0,270	40/60=0,66	0,67	0,01
3	10	0,1789	50/60=0,83	0,85	0,02
4	6	0,088	56/60=0,93	0,94	0,01
5	4	0,0351	1	0,98	0,02

El máximo resultado de la diferencia es .03, y con un $\alpha=.05$ y $n=60$ el valor es mayor al .03, no rechazamos H_0 . Los aprobados siguen una distribución Poisson.

3.

z Normal	$F(X)_N(*)$	$F(X)_{obs}$	diferencias
1.28->	1	1.0	0
.84-1.28	.9	1.0	.1
.52-.84	.8	.6	.2
.25-.52	.7	.6	.1
0-.25	.6	.6	0
-.25-0	.5	.4	.1
-.52-(-.25)	.4	.4	0
-.84-(-.52)	.3	.2	.1
-1.28-(-.84)	.2	.2	0
<-1.28	.1	.1	0

$D=0.2$

Significación del estadístico de contraste =0.81.

H_0 se acepta, ya que la significación del estadístico el contraste es mayor que el nivel previamente establecido) $\alpha=(.05)$

4.

valor de Kolmogórov-Smirnov =0,981.

el valor de $p = 0,291. > 0,05$, no se rechaza la hipótesis nula

Conclusión: que hay evidencias suficientes para pensar que la muestra proviene de la distribución normal, con $\alpha = 5\%$.

5.

H_0 : la variable aleatoria volumen servido sigue una distribución normal.

H_1 : la variable aleatoria volumen servido no sigue una distribución normal

1	2	3	4	5	6	7
i	X _i	X _(n-1+1)	X _(n-1+1) - X _i	a _i	a _i [X _(n-1+1) - X _i]	(X _i - $\bar{x}$) ²
1	15.90	16.05	0.15	0.6052	0.090780	0.0081
2	15.96	16.04	0.08	0.3164	0.025312	0.0009
3	15.97	16.02	0.05	0.1743	0.008715	0.0004
4	15.98	16	0.02	0.0561	0.001122	0.0001
5	16	15.98	-0.02	0	0	0.0001
6	16.02	15.97	-0.05	0	0	0.0009
7	16.04	15.96	-0.08	0	0	0.0025
8	16.05	15.9	-0.15	0	0	0.0036
total					0.0159	0.0166
$\bar{x}$ =15.99						

$$W = \frac{0.0159^2}{0.0166} = 0.0152$$

$W(\alpha, n) = 0.0152 < 0.818$ ho se acepta, se concluye que la variable volumen servidor se distribuye normalmente.

UNIDAD 5: REGRESIÓN LINEAL SIMPLE

1.

- a) Níquel
- b) corriente máxima
- c) relación directa
- d) $Y = .300 + .0042x$
- e) $Y = .300 + .0042(150) = .93$
- f) 95.46%
- g) 97.70%

2.

- a) Presión
- b) Compresión
- c) relación directa
- d) $y = -.10 + .70x$
- e) $Y = -.10 + .70(8) = 5.5$
- f) 81.66%
- g) 90.36%

3.

- a) Experiencia
- b) Producción
- c) relación directa
- d) $Y = 21.83 + 1.6x$
- e) $Y = 21.83 + 1.6(9) = 36.23$
- f) 61.31%
- g) 37.59%

4.

- a) $S = 2.033$
 $\bar{x} = 10$
 $S_{yp} = 1.11$
- c) 10.6 ± 3.53 o 7.07 a 14.13

- c) $S_{ind} = 2.32$
- d) 10.6 ± 7.38 ó 3.22 a 17.98

5.

I.C.: 8.65 a 21.15

I.P.: -4.50 a 41.41.30

6.

$s_{yp}=68.54$

a) \$3343.53 a 3,724.07

b) \$3086.34 a \$3981.26

7.

a) \$201

b) 167.25 a 234.65

c) 108.75 a 293.15

8.

a) \$5046.67

b) \$3815.19 a \$293.15

c) No son fuera de lo común.

ANEXO 2

RECURSOS DE EVALUACIÓN POR UNIDAD

<https://docs.google.com/presentation/d/1dmlQVAwaNFDF8gfoetNv6Yz2d4KumUL5/edit?usp=sharing&ouid=111493099739666903937&rtpof=true&sd=true>

ANEXO 3

USO DE SOFTWARE ESTADÍSTICO MINITAB®

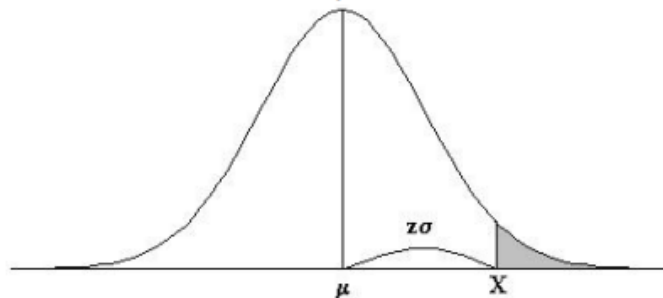
<https://docs.google.com/presentation/d/1nksay7BuPsYjtOpRV6zIIgf6CTePMNyk/edit?usp=sharing&ouid=111493099739666903937&rtpof=true&sd=true>

ANEXO 4

TABLAS DE DISTIBUCIONES ESTADÍSTICAS

TABLA 1: DISTRIBUCIÓN NORMAL

Áreas bajo la curva normal



Ejemplo:

$$Z = \frac{X - \mu}{\sigma}$$

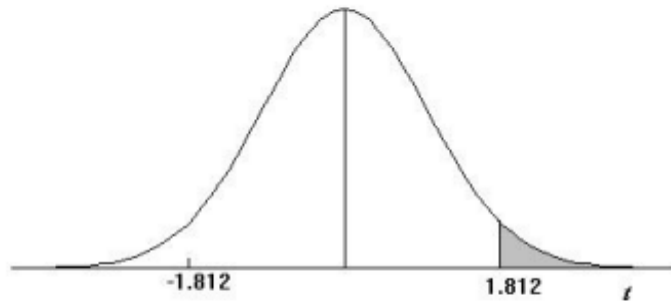
$$P[Z > 1] = 0.1587$$

$$P[Z > 1.96] = 0.0250$$

Desv. normal x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010

TABLA 2: DISTRIBUCIÓN t DE STUDENT

Puntos de porcentaje de la distribución t



Ejemplo

Para $\phi = 10$ grados de libertad:

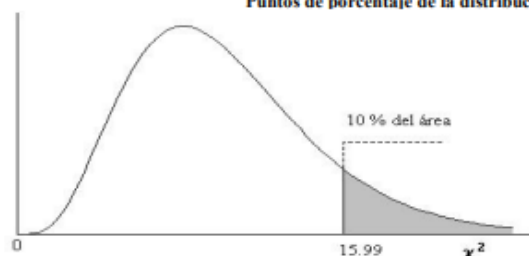
$$P[t > 1.812] = 0.05$$

$$P[t < -1.812] = 0.05$$

α r	0,25	0,2	0,15	0,1	0,05	0,025	0,01	0,005	0,0005
1	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,656	636,578
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,600
3	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,768
24	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,689
28	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,660
30	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,679	0,848	1,045	1,296	1,671	2,000	2,390	2,660	3,460
120	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,290

TABLA 3: DISTRIBUCIÓN χ^2

Puntos de porcentaje de la distribución χ^2



Ejemplo:

Para $\phi = 10$ grados de libertad

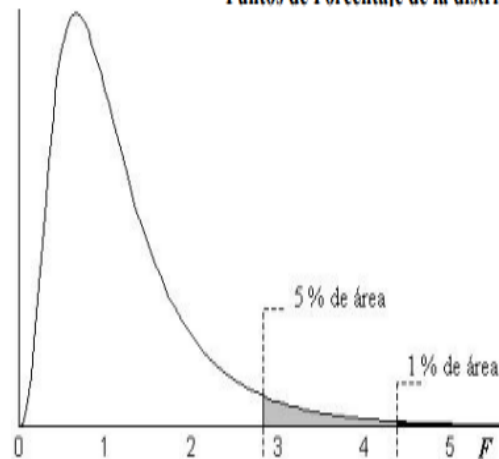
$$P[\chi^2 > 15.99] = 0.10$$

π ϕ	0.995	0.99	0.975	0.95	0.9	0.75	0.5	0.25	0.1	0.05	0.025	0.01	0.005	π ϕ
1	3.93E-05	1.57E-04	9.82E-04	3.93E-03	1.58E-02	0.102	0.455	1.323	2.71	3.84	5.02	6.63	7.88	1
2	1.00E-02	2.01E-02	5.06E-02	0.103	0.211	0.575	1.386	2.77	4.61	5.99	7.38	9.21	10.60	2
3	7.17E-02	0.115	0.216	0.352	0.584	1.213	2.37	4.11	6.25	7.81	9.35	11.34	12.84	3
4	0.207	0.297	0.484	0.711	1.064	1.923	3.36	5.39	7.78	9.49	11.14	13.28	14.86	4
5	0.412	0.554	0.831	1.145	1.610	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75	5
6	0.676	0.872	1.237	1.635	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55	6
7	0.989	1.239	1.690	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.3	7
8	1.344	1.647	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.1	22.0	8
9	1.735	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.7	23.6	9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.5	23.2	25.2	10
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.9	24.7	26.8	11
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.0	23.3	26.2	28.3	12
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.4	24.7	27.7	29.8	13
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.1	23.7	26.1	29.1	31.3	14
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.3	25.0	27.5	30.6	32.8	15
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.5	26.3	28.8	32.0	34.3	16
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.5	24.8	27.6	30.2	33.4	35.7	17
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.6	26.0	28.9	31.5	34.8	37.2	18
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.7	27.2	30.1	32.9	36.2	38.6	19
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.8	28.4	31.4	34.2	37.6	40.0	20
21	8.03	8.90	10.28	11.59	13.24	16.34	20.3	24.9	29.6	32.7	35.5	38.9	41.4	21
22	8.64	9.54	10.98	12.34	14.04	17.24	21.3	26.0	30.8	33.9	36.8	40.3	42.8	22
23	9.26	10.20	11.69	13.09	14.85	18.14	22.3	27.1	32.0	35.2	38.1	41.6	44.2	23
24	9.89	10.86	12.40	13.85	15.66	19.04	23.3	28.2	33.2	36.4	39.4	43.0	45.6	24
25	10.52	11.52	13.12	14.61	16.47	19.94	24.3	29.3	34.4	37.7	40.6	44.3	46.9	25
26	11.16	12.20	13.84	15.38	17.29	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	26
27	11.81	12.88	14.57	16.15	18.11	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	27
28	12.46	13.56	15.31	16.93	18.94	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	28
29	13.12	14.26	16.05	17.71	19.77	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	29
30	13.79	14.95	16.79	18.49	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	30
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8	40
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5	50
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0	60
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100.4	104.2	70
80	51.2	53.5	57.2	60.4	64.3	71.1	79.3	88.1	96.6	101.9	106.6	112.3	116.3	80
90	59.2	61.8	65.6	69.1	73.3	80.6	89.3	98.6	107.6	113.1	118.1	124.1	128.3	90
100	67.3	70.1	74.2	77.9	82.4	90.1	99.3	109.1	118.5	124.3	129.6	135.8	140.2	100
Z_α	-2.58	-2.33	-1.96	-1.64	-1.28	-0.674	0.000	0.674	1.282	1.645	1.96	2.33	2.58	Z_α

Para $\phi > 100$ tómese $\chi^2 = \frac{1}{2}(Z_\alpha + \sqrt{2\phi - 1})^2$. Z_α es la desviación normal estandarizada correspondiente al nivel de significancia y se muestra en la parte superior de la tabla.

TABLA 4: DISTRIBUCIÓN F DE FISHER

Puntos de Porcentaje de la distribución F



Ejemplo:

Para $n_1 = 9, n_2 = 12$ grados de libertad:

$P[F > 2.80] = 0.05$

$P[F > 4.39] = 0.01$

n ₂	5 % (normal) y 1 % (negritas) puntos para la distribución de F																											n ₂
	n1 grados de libertad (para el mayor cuadrado medio)																											
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞				
1	161	199	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254	254	254	1		
	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6083	6107	6143	6170	6209	6234	6260	6286	6302	6324	6334	6350	6360	6366				
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41	19.42	19.43	19.45	19.45	19.46	19.47	19.48	19.48	19.49	19.49	19.49	19.50	2			
	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.41	99.42	99.43	99.44	99.45	99.46	99.47	99.48	99.48	99.49	99.49	99.49	99.50	99.50				
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.59	8.58	8.56	8.55	8.54	8.53	8.53	3			
	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.13	27.05	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.28	26.24	26.18	26.15	26.13				
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.87	5.84	5.80	5.77	5.75	5.72	5.70	5.68	5.66	5.65	5.64	5.63	4			
	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.45	14.37	14.25	14.15	14.02	13.93	13.84	13.75	13.69	13.61	13.58	13.52	13.49	13.46				
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.41	4.39	4.37	4.37	5			
	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.96	9.89	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.08	9.04	9.02				
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.73	3.71	3.69	3.68	3.67	6			
	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.60	7.52	7.40	7.31	7.23	7.14	7.09	7.02	6.99	6.93	6.90	6.88				
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.53	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.27	3.25	3.24	3.23	7			
	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47	6.36	6.28	6.16	6.07	5.99	5.91	5.86	5.79	5.75	5.70	5.67	5.65				
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.24	3.20	3.15	3.12	3.08	3.04	3.02	2.99	2.97	2.95	2.94	2.93	8			
	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67	5.56	5.48	5.36	5.28	5.20	5.12	5.07	5.00	4.96	4.91	4.88	4.86				
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.03	2.99	2.94	2.90	2.86	2.83	2.80	2.77	2.76	2.73	2.72	2.71	9			
	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11	5.01	4.92	4.81	4.73	4.65	4.57	4.52	4.45	4.41	4.36	4.33	4.31				
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.86	2.83	2.77	2.74	2.70	2.66	2.64	2.60	2.59	2.56	2.55	2.54	10			
	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91				

n ₂	5 % (normal) y 1 % (negritas) puntos para la distribución de F																										n ₁
	n1 grados de libertad (para el mayor cuadrado medio)																										
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞			
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.51	2.47	2.46	2.43	2.42	2.40	2.40	11	
12	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40	4.29	4.21	4.10	4.02	3.94	3.86	3.81	3.74	3.71	3.66	3.62	3.60	12		
	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.64	2.60	2.54	2.51	2.47	2.43	2.40	2.37	2.35	2.32	2.31	2.30	12		
13	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16	4.05	3.97	3.86	3.78	3.70	3.62	3.57	3.50	3.47	3.41	3.38	3.36	13		
	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.22	2.21	13		
14	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.86	3.78	3.66	3.59	3.51	3.43	3.38	3.31	3.27	3.22	3.19	3.17	14		
	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13	14		
15	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.70	3.62	3.51	3.43	3.35	3.27	3.22	3.15	3.11	3.06	3.03	3.00	15		
	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.42	2.38	2.33	2.29	2.25	2.20	2.18	2.14	2.12	2.10	2.08	2.07	15		
16	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.56	3.49	3.37	3.29	3.21	3.13	3.08	3.01	2.98	2.92	2.89	2.87	16		
	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.37	2.33	2.28	2.24	2.19	2.15	2.12	2.09	2.07	2.04	2.02	2.01	16		
17	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.62	3.55	3.45	3.37	3.26	3.18	3.10	3.02	2.97	2.90	2.86	2.81	2.78	2.75	17		
	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.10	2.08	2.04	2.02	1.99	1.97	1.96	17		
18	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.46	3.35	3.27	3.16	3.08	3.00	2.92	2.87	2.80	2.76	2.71	2.68	2.65	18		
	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.06	2.04	2.00	1.98	1.95	1.93	1.92	18		
19	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.43	3.37	3.27	3.19	3.08	3.00	2.92	2.84	2.78	2.71	2.68	2.62	2.59	2.57	19		
	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.26	2.21	2.16	2.11	2.07	2.03	2.00	1.96	1.94	1.91	1.89	1.88	19		
20	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30	3.19	3.12	3.00	2.92	2.84	2.76	2.71	2.64	2.60	2.55	2.51	2.49	20		
	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.22	2.18	2.12	2.08	2.04	1.99	1.97	1.93	1.91	1.88	1.86	1.84	20		
21	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23	3.13	3.05	2.94	2.86	2.78	2.69	2.64	2.57	2.54	2.48	2.44	2.42	21		
	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.16	2.10	2.05	2.01	1.96	1.94	1.90	1.88	1.84	1.83	1.81	21		
22	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.24	3.17	3.07	2.99	2.88	2.80	2.72	2.64	2.58	2.51	2.48	2.42	2.38	2.36	22		
	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.17	2.13	2.07	2.03	1.98	1.94	1.91	1.87	1.85	1.82	1.80	1.78	22		
23	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12	3.02	2.94	2.83	2.75	2.67	2.58	2.53	2.46	2.42	2.36	2.33	2.31	23		
	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20	2.15	2.11	2.05	2.01	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76	23		
24	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07	2.97	2.89	2.78	2.70	2.62	2.54	2.48	2.41	2.37	2.32	2.28	2.26	24		
	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.13	2.09	2.03	1.98	1.94	1.89	1.86	1.82	1.80	1.77	1.75	1.73	24		
25	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03	2.93	2.85	2.74	2.66	2.58	2.49	2.44	2.37	2.33	2.27	2.24	2.21	25		
	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.11	2.07	2.01	1.96	1.92	1.87	1.84	1.80	1.78	1.75	1.73	1.71	25		
26	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	3.06	2.99	2.89	2.81	2.70	2.62	2.54	2.45	2.40	2.33	2.29	2.23	2.19	2.17	26		
	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.09	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.73	1.71	1.69	26		
27	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.02	2.96	2.86	2.78	2.66	2.58	2.50	2.42	2.36	2.29	2.25	2.19	2.16	2.13	27		
	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.08	2.04	1.97	1.93	1.88	1.84	1.81	1.76	1.74	1.71	1.69	1.67	27		
28	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.99	2.93	2.82	2.75	2.63	2.55	2.47	2.38	2.33	2.26	2.22	2.16	2.12	2.10	28		
	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.82	1.79	1.75	1.73	1.69	1.67	1.65	28		
29	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.96	2.90	2.79	2.72	2.60	2.52	2.44	2.35	2.30	2.23	2.19	2.13	2.09	2.06	29		
	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.01	1.94	1.90	1.85	1.81	1.77	1.73	1.71	1.67	1.65	1.64	29		
30	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.93	2.87	2.77	2.69	2.57	2.49	2.41	2.33	2.27	2.20	2.16	2.10	2.06	2.03	30		
	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.70	1.66	1.64	1.62	30		
32	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84	2.74	2.66	2.55	2.47	2.39	2.30	2.25	2.17	2.13	2.07	2.03	2.01	32		
	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.10	2.07	2.01	1.97	1.91	1.86	1.82	1.77	1.74	1.69	1.67	1.63	1.61	1.59	32		
	7.50	5.34	4.46	3.97	3.65	3.43	3.26	3.13	3.02	2.93	2.86	2.80	2.70	2.62	2.50	2.42	2.34	2.25	2.20	2.12	2.08	2.02	1.98	1.96			

n ₂	5 % (normal) y 1 % (negritas) puntos para la distribución de F																										n ₁
	n1 grados de libertad (para el mayor cuadrado medio)																										
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞			
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.08	2.05	1.99	1.95	1.89	1.84	1.80	1.75	1.71	1.67	1.65	1.61	1.59	1.57	34		
36	7.44	5.29	4.42	3.93	3.61	3.39	3.22	3.09	2.98	2.89	2.82	2.76	2.66	2.58	2.46	2.38	2.30	2.21	2.16	2.08	2.04	1.98	1.94	1.91	36		
	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.07	2.03	1.98	1.93	1.87	1.82	1.78	1.73	1.69	1.65	1.62	1.59	1.56	1.55			
38	7.40	5.25	4.38	3.89	3.57	3.35	3.18	3.05	2.95	2.86	2.79	2.72	2.62	2.54	2.43	2.35	2.26	2.18	2.12	2.04	2.00	1.94	1.90	1.87	38		
	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.96	1.92	1.85	1.81	1.76	1.71	1.68	1.63	1.61	1.57	1.54	1.53			
40	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.92	2.83	2.75	2.69	2.59	2.51	2.40	2.32	2.23	2.14	2.09	2.01	1.97	1.90	1.86	1.84	40		
	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51			
42	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66	2.56	2.48	2.37	2.29	2.20	2.11	2.06	1.98	1.94	1.87	1.83	1.81	42		
	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.03	1.99	1.94	1.89	1.83	1.78	1.73	1.68	1.65	1.60	1.57	1.53	1.51	1.49			
44	7.28	5.15	4.29	3.80	3.49	3.27	3.10	2.97	2.86	2.78	2.70	2.64	2.54	2.46	2.34	2.26	2.18	2.09	2.03	1.95	1.91	1.85	1.80	1.78	44		
	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.92	1.88	1.81	1.77	1.72	1.67	1.63	1.59	1.56	1.52	1.49	1.48			
46	7.25	5.12	4.26	3.78	3.47	3.24	3.08	2.95	2.84	2.75	2.68	2.62	2.52	2.44	2.32	2.24	2.15	2.07	2.01	1.93	1.89	1.82	1.78	1.75	46		
	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.15	2.09	2.04	2.00	1.97	1.91	1.87	1.80	1.76	1.71	1.65	1.62	1.57	1.55	1.51	1.48	1.46			
48	7.22	5.10	4.24	3.76	3.44	3.22	3.06	2.93	2.82	2.73	2.66	2.60	2.50	2.42	2.30	2.22	2.13	2.04	1.99	1.91	1.86	1.80	1.76	1.73	48		
	4.04	3.19	2.80	2.57	2.41	2.29	2.21	2.14	2.08	2.03	1.99	1.96	1.90	1.86	1.79	1.75	1.70	1.64	1.61	1.56	1.54	1.49	1.47	1.45			
50	7.19	5.08	4.22	3.74	3.43	3.20	3.04	2.91	2.80	2.71	2.64	2.58	2.48	2.40	2.28	2.20	2.12	2.02	1.97	1.89	1.84	1.78	1.73	1.70	50		
	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95	1.89	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44			
55	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.63	2.56	2.46	2.38	2.27	2.18	2.10	2.01	1.95	1.87	1.82	1.76	1.71	1.68	55		
	4.02	3.16	2.77	2.54	2.38	2.27	2.18	2.11	2.06	2.01	1.97	1.93	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.53	1.50	1.46	1.43	1.41			
60	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75	2.66	2.59	2.53	2.42	2.34	2.23	2.15	2.06	1.97	1.91	1.83	1.78	1.71	1.67	1.64	60		
	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.82	1.75	1.70	1.65	1.59	1.56	1.51	1.48	1.44	1.41	1.39			
65	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	2.39	2.31	2.20	2.12	2.03	1.94	1.88	1.79	1.75	1.68	1.63	1.60	65		
	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.03	1.98	1.94	1.90	1.85	1.80	1.73	1.69	1.63	1.58	1.54	1.49	1.46	1.42	1.39	1.37			
70	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.80	2.69	2.61	2.53	2.47	2.37	2.29	2.17	2.09	2.00	1.91	1.85	1.77	1.72	1.65	1.60	1.57	70		
	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.57	1.53	1.48	1.45	1.40	1.37	1.35			
80	7.01	4.92	4.07	3.60	3.29	3.07	2.91	2.78	2.67	2.59	2.51	2.45	2.35	2.27	2.15	2.07	1.98	1.89	1.83	1.74	1.70	1.62	1.57	1.54	80		
	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.43	1.38	1.35	1.33			
100	6.96	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.64	2.55	2.48	2.42	2.31	2.23	2.12	2.03	1.94	1.85	1.79	1.70	1.65	1.58	1.53	1.50	100		
	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85	1.79	1.75	1.68	1.63	1.57	1.52	1.48	1.42	1.39	1.34	1.31	1.28			
125	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50	2.43	2.37	2.27	2.19	2.07	1.98	1.89	1.80	1.74	1.65	1.60	1.52	1.47	1.43	125		
	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.96	1.91	1.87	1.83	1.77	1.73	1.66	1.60	1.55	1.49	1.45	1.40	1.36	1.31	1.27	1.25			
150	6.84	4.78	3.94	3.47	3.17	2.95	2.79	2.66	2.55	2.47	2.39	2.33	2.23	2.15	2.03	1.94	1.85	1.76	1.69	1.60	1.55	1.47	1.41	1.37	150		
	3.90	3.06	2.66	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.48	1.44	1.38	1.34	1.29	1.25	1.22			
200	6.81	4.75	3.91	3.45	3.14	2.92	2.76	2.63	2.53	2.44	2.37	2.31	2.20	2.12	2.00	1.92	1.83	1.73	1.66	1.57	1.52	1.43	1.38	1.33	200		
	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80	1.74	1.69	1.62	1.57	1.52	1.46	1.41	1.35	1.32	1.26	1.22	1.19			
400	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27	2.17	2.09	1.97	1.89	1.79	1.69	1.63	1.53	1.48	1.39	1.33	1.28	400		
	3.86	3.02	2.63	2.39	2.24	2.12	2.03	1.96	1.90	1.85	1.81	1.78	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.17	1.13			
1000	6.70	4.66	3.83	3.37	3.06	2.85	2.68	2.56	2.45	2.37	2.29	2.23	2.13	2.05	1.92	1.84	1.75	1.64	1.58	1.48	1.42	1.32	1.25	1.19	1000		
	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08			
∞	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.27	2.20	2.10	2.02	1.90	1.81	1.72	1.61	1.54	1.44	1.38	1.28	1.19	1.12	∞		
	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75	1.69	1.64	1.57	1.52	1.46	1.39	1.35	1.28	1.24	1.17	1.11	1.00			
	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18	2.08	2.00	1.88	1.79	1.70	1.59	1.52	1.42	1.36	1.25	1.15	1.00			

TABLA 5: PROBABILIDADES BINOMIALES

n	k	p									
		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
1	0	0.9500	0.9000	0.8500	0.8000	0.7500	0.7000	0.6500	0.6000	0.5500	0.5000
1	1	0.0500	0.1000	0.1500	0.2000	0.2500	0.3000	0.3500	0.4000	0.4500	0.5000
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
2	1	0.0950	0.1800	0.2550	0.3200	0.3750	0.4200	0.4550	0.4800	0.4950	0.5000
2	2	0.0025	0.0100	0.0225	0.0400	0.0625	0.0900	0.1225	0.1600	0.2025	0.2500
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
3	1	0.1354	0.2430	0.3251	0.3840	0.4219	0.4410	0.4436	0.4320	0.4084	0.3750
3	2	0.0071	0.0270	0.0574	0.0960	0.1406	0.1890	0.2389	0.2880	0.3341	0.3750
3	3	0.0001	0.0010	0.0034	0.0080	0.0156	0.0270	0.0429	0.0640	0.0911	0.1250
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
4	1	0.1715	0.2916	0.3685	0.4096	0.4219	0.4116	0.3845	0.3456	0.2995	0.2500
4	2	0.0135	0.0486	0.0975	0.1536	0.2109	0.2646	0.3105	0.3456	0.3675	0.3750
4	3	0.0005	0.0036	0.0115	0.0256	0.0469	0.0756	0.1115	0.1536	0.2005	0.2500
4	4	0.0000	0.0001	0.0005	0.0016	0.0039	0.0081	0.0150	0.0256	0.0410	0.0625
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0313
5	1	0.2036	0.3281	0.3915	0.4096	0.3955	0.3602	0.3124	0.2592	0.2059	0.1563
5	2	0.0214	0.0729	0.1382	0.2048	0.2637	0.3087	0.3364	0.3456	0.3369	0.3125
5	3	0.0011	0.0081	0.0244	0.0512	0.0879	0.1323	0.1811	0.2304	0.2757	0.3125
5	4	0.0000	0.0005	0.0022	0.0064	0.0146	0.0284	0.0488	0.0768	0.1128	0.1563
5	5	0.0000	0.0000	0.0001	0.0003	0.0010	0.0024	0.0053	0.0102	0.0185	0.0313
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
6	1	0.2321	0.3543	0.3993	0.3932	0.3560	0.3025	0.2437	0.1866	0.1359	0.0938
6	2	0.0305	0.0984	0.1762	0.2458	0.2966	0.3241	0.3280	0.3110	0.2780	0.2344
6	3	0.0021	0.0146	0.0415	0.0819	0.1318	0.1852	0.2355	0.2765	0.3032	0.3125
6	4	0.0001	0.0012	0.0055	0.0154	0.0330	0.0595	0.0951	0.1382	0.1861	0.2344
6	5	0.0000	0.0001	0.0004	0.0015	0.0044	0.0102	0.0205	0.0369	0.0609	0.0938
6	6	0.0000	0.0000	0.0000	0.0001	0.0002	0.0007	0.0018	0.0041	0.0083	0.0156
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
7	1	0.2573	0.3720	0.3960	0.3670	0.3115	0.2471	0.1848	0.1306	0.0872	0.0547
7	2	0.0406	0.1240	0.2097	0.2753	0.3115	0.3177	0.2985	0.2613	0.2140	0.1641
7	3	0.0036	0.0230	0.0617	0.1147	0.1730	0.2269	0.2679	0.2903	0.2918	0.2734
7	4	0.0002	0.0026	0.0109	0.0287	0.0577	0.0972	0.1442	0.1935	0.2388	0.2734
7	5	0.0000	0.0002	0.0012	0.0043	0.0115	0.0250	0.0466	0.0774	0.1172	0.1641
7	6	0.0000	0.0000	0.0001	0.0004	0.0013	0.0036	0.0084	0.0172	0.0320	0.0547
7	7	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0006	0.0016	0.0037	0.0078
8	0	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084	0.0039
8	1	0.2793	0.3826	0.3847	0.3355	0.2670	0.1977	0.1373	0.0896	0.0548	0.0313
8	2	0.0515	0.1488	0.2376	0.2936	0.3115	0.2965	0.2587	0.2090	0.1569	0.1094
8	3	0.0054	0.0331	0.0839	0.1468	0.2076	0.2541	0.2786	0.2787	0.2568	0.2188
8	4	0.0004	0.0046	0.0185	0.0459	0.0865	0.1361	0.1875	0.2322	0.2627	0.2734
8	5	0.0000	0.0004	0.0026	0.0092	0.0231	0.0467	0.0808	0.1239	0.1719	0.2188
8	6	0.0000	0.0000	0.0002	0.0011	0.0038	0.0100	0.0217	0.0413	0.0703	0.1094
8	7	0.0000	0.0000	0.0000	0.0001	0.0004	0.0012	0.0033	0.0079	0.0164	0.0313
8	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0007	0.0017	0.0039

TABLA 5 (CONTINUACIÓN)

n	k	p									
		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020
9	1	0.2985	0.3874	0.3679	0.3020	0.2253	0.1556	0.1004	0.0605	0.0339	0.0176
9	2	0.0629	0.1722	0.2597	0.3020	0.3003	0.2668	0.2162	0.1612	0.1110	0.0703
9	3	0.0077	0.0446	0.1069	0.1762	0.2336	0.2668	0.2716	0.2508	0.2119	0.1641
9	4	0.0006	0.0074	0.0283	0.0661	0.1168	0.1715	0.2194	0.2508	0.2600	0.2461
9	5	0.0000	0.0008	0.0050	0.0165	0.0389	0.0735	0.1181	0.1672	0.2128	0.2461
9	6	0.0000	0.0001	0.0006	0.0028	0.0087	0.0210	0.0424	0.0743	0.1160	0.1641
9	7	0.0000	0.0000	0.0000	0.0003	0.0012	0.0039	0.0098	0.0212	0.0407	0.0703
9	8	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0013	0.0035	0.0083	0.0176
9	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0008	0.0020
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
10	1	0.3151	0.3874	0.3474	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098
10	2	0.0746	0.1937	0.2759	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439
10	3	0.0105	0.0574	0.1298	0.2013	0.2503	0.2668	0.2522	0.2150	0.1665	0.1172
10	4	0.0010	0.0112	0.0401	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051
10	5	0.0001	0.0015	0.0085	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461
10	6	0.0000	0.0001	0.0012	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051
10	7	0.0000	0.0000	0.0001	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172
10	8	0.0000	0.0000	0.0000	0.0001	0.0004	0.0014	0.0043	0.0106	0.0229	0.0439
10	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098
10	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010
11	0	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0088	0.0036	0.0014	0.0005
11	1	0.3293	0.3835	0.3248	0.2362	0.1549	0.0932	0.0518	0.0266	0.0125	0.0054
11	2	0.0867	0.2131	0.2866	0.2953	0.2581	0.1998	0.1395	0.0887	0.0513	0.0269
11	3	0.0137	0.0710	0.1517	0.2215	0.2581	0.2568	0.2254	0.1774	0.1259	0.0806
11	4	0.0014	0.0158	0.0536	0.1107	0.1721	0.2201	0.2428	0.2365	0.2060	0.1611
11	5	0.0001	0.0025	0.0132	0.0388	0.0803	0.1321	0.1830	0.2207	0.2360	0.2256
11	6	0.0000	0.0003	0.0023	0.0097	0.0268	0.0566	0.0985	0.1471	0.1931	0.2256
11	7	0.0000	0.0000	0.0003	0.0017	0.0064	0.0173	0.0379	0.0701	0.1128	0.1611
11	8	0.0000	0.0000	0.0000	0.0002	0.0011	0.0037	0.0102	0.0234	0.0462	0.0806
11	9	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0018	0.0052	0.0126	0.0269
11	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0007	0.0021	0.0054
11	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0005
12	0	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0057	0.0022	0.0008	0.0002
12	1	0.3413	0.3766	0.3012	0.2062	0.1267	0.0712	0.0368	0.0174	0.0075	0.0029
12	2	0.0988	0.2301	0.2924	0.2835	0.2323	0.1678	0.1088	0.0639	0.0339	0.0161
12	3	0.0173	0.0852	0.1720	0.2362	0.2581	0.2397	0.1954	0.1419	0.0923	0.0537
12	4	0.0021	0.0213	0.0683	0.1329	0.1936	0.2311	0.2367	0.2128	0.1700	0.1208
12	5	0.0002	0.0038	0.0193	0.0532	0.1032	0.1585	0.2039	0.2270	0.2225	0.1934
12	6	0.0000	0.0005	0.0040	0.0155	0.0401	0.0792	0.1281	0.1766	0.2124	0.2256
12	7	0.0000	0.0000	0.0006	0.0033	0.0115	0.0291	0.0591	0.1009	0.1489	0.1934
12	8	0.0000	0.0000	0.0001	0.0005	0.0024	0.0078	0.0199	0.0420	0.0762	0.1208
12	9	0.0000	0.0000	0.0000	0.0001	0.0004	0.0015	0.0048	0.0125	0.0277	0.0537
12	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0008	0.0025	0.0068	0.0161
12	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0.0029
12	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002

TABLA 5 (CONTINUACIÓN)

n	k	p									
		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
13	0	0.5133	0.2542	0.1209	0.0550	0.0238	0.0097	0.0037	0.0013	0.0004	0.0001
13	1	0.3512	0.3672	0.2774	0.1787	0.1029	0.0540	0.0259	0.0113	0.0045	0.0016
13	2	0.1109	0.2448	0.2937	0.2680	0.2059	0.1388	0.0836	0.0453	0.0220	0.0095
13	3	0.0214	0.0997	0.1900	0.2457	0.2517	0.2181	0.1651	0.1107	0.0660	0.0349
13	4	0.0028	0.0277	0.0838	0.1535	0.2097	0.2337	0.2222	0.1845	0.1350	0.0873
13	5	0.0003	0.0055	0.0266	0.0691	0.1258	0.1803	0.2154	0.2214	0.1989	0.1571
13	6	0.0000	0.0008	0.0063	0.0230	0.0559	0.1030	0.1546	0.1968	0.2169	0.2095
13	7	0.0000	0.0001	0.0011	0.0058	0.0186	0.0442	0.0833	0.1312	0.1775	0.2095
13	8	0.0000	0.0000	0.0001	0.0011	0.0047	0.0142	0.0336	0.0656	0.1089	0.1571
13	9	0.0000	0.0000	0.0000	0.0001	0.0009	0.0034	0.0101	0.0243	0.0495	0.0873
13	10	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0022	0.0065	0.0162	0.0349
13	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0012	0.0036	0.0095
13	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0016
13	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
14	0	0.4877	0.2288	0.1028	0.0440	0.0178	0.0068	0.0024	0.0008	0.0002	0.0001
14	1	0.3593	0.3559	0.2539	0.1539	0.0832	0.0407	0.0181	0.0073	0.0027	0.0009
14	2	0.1229	0.2570	0.2912	0.2501	0.1802	0.1134	0.0634	0.0317	0.0141	0.0056
14	3	0.0259	0.1142	0.2056	0.2501	0.2402	0.1943	0.1366	0.0845	0.0462	0.0222
14	4	0.0037	0.0349	0.0998	0.1720	0.2202	0.2290	0.2022	0.1549	0.1040	0.0611
14	5	0.0004	0.0078	0.0352	0.0860	0.1468	0.1963	0.2178	0.2066	0.1701	0.1222
14	6	0.0000	0.0013	0.0093	0.0322	0.0734	0.1262	0.1759	0.2066	0.2088	0.1833
14	7	0.0000	0.0002	0.0019	0.0092	0.0280	0.0618	0.1082	0.1574	0.1952	0.2095
14	8	0.0000	0.0000	0.0003	0.0020	0.0082	0.0232	0.0510	0.0918	0.1398	0.1833
14	9	0.0000	0.0000	0.0000	0.0003	0.0018	0.0066	0.0183	0.0408	0.0762	0.1222
14	10	0.0000	0.0000	0.0000	0.0000	0.0003	0.0014	0.0049	0.0136	0.0312	0.0611
14	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0010	0.0033	0.0093	0.0222
14	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0019	0.0056
14	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0009
14	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
15	0	0.4633	0.2059	0.0874	0.0352	0.0134	0.0047	0.0016	0.0005	0.0001	0.0000
15	1	0.3658	0.3432	0.2312	0.1319	0.0668	0.0305	0.0126	0.0047	0.0016	0.0005
15	2	0.1348	0.2669	0.2856	0.2309	0.1559	0.0916	0.0476	0.0219	0.0090	0.0032
15	3	0.0307	0.1285	0.2184	0.2501	0.2252	0.1700	0.1110	0.0634	0.0318	0.0139
15	4	0.0049	0.0428	0.1156	0.1876	0.2252	0.2186	0.1792	0.1268	0.0780	0.0417
15	5	0.0006	0.0105	0.0449	0.1032	0.1651	0.2061	0.2123	0.1859	0.1404	0.0916
15	6	0.0000	0.0019	0.0132	0.0430	0.0917	0.1472	0.1906	0.2066	0.1914	0.1527
15	7	0.0000	0.0003	0.0030	0.0138	0.0393	0.0811	0.1319	0.1771	0.2013	0.1964
15	8	0.0000	0.0000	0.0005	0.0035	0.0131	0.0348	0.0710	0.1181	0.1647	0.1964
15	9	0.0000	0.0000	0.0001	0.0007	0.0034	0.0116	0.0298	0.0612	0.1048	0.1527
15	10	0.0000	0.0000	0.0000	0.0001	0.0007	0.0030	0.0096	0.0245	0.0515	0.0916
15	11	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0024	0.0074	0.0191	0.0417
15	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0016	0.0052	0.0139
15	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0.0032
15	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005
15	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

n	k	0.05	0.10	0.15	0.20	p	0.25	0.30	0.35	0.40	0.45	0.50
16	0	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0010	0.0003	0.0001	0.0000	0.0000
16	1	0.3706	0.3294	0.2097	0.1126	0.0535	0.0228	0.0087	0.0030	0.0009	0.0002	0.0002
16	2	0.1463	0.2745	0.2775	0.2111	0.1336	0.0732	0.0353	0.0150	0.0056	0.0018	0.0018
16	3	0.0359	0.1423	0.2285	0.2463	0.2079	0.1465	0.0888	0.0468	0.0215	0.0085	0.0085
16	4	0.0061	0.0514	0.1311	0.2001	0.2252	0.2040	0.1553	0.1014	0.0572	0.0278	0.0278
16	5	0.0008	0.0137	0.0555	0.1201	0.1802	0.2099	0.2008	0.1623	0.1123	0.0667	0.0667
16	6	0.0001	0.0028	0.0180	0.0550	0.1101	0.1649	0.1982	0.1983	0.1684	0.1222	0.1222
16	7	0.0000	0.0004	0.0045	0.0197	0.0524	0.1010	0.1524	0.1889	0.1969	0.1746	0.1746
16	8	0.0000	0.0001	0.0009	0.0055	0.0197	0.0487	0.0923	0.1417	0.1812	0.1964	0.1964
16	9	0.0000	0.0000	0.0001	0.0012	0.0058	0.0185	0.0442	0.0840	0.1318	0.1746	0.1746
16	10	0.0000	0.0000	0.0000	0.0002	0.0014	0.0056	0.0167	0.0392	0.0755	0.1222	0.1222
16	11	0.0000	0.0000	0.0000	0.0000	0.0002	0.0013	0.0049	0.0142	0.0337	0.0667	0.0667
16	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0040	0.0115	0.0278	0.0278
16	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0008	0.0029	0.0085	0.0085
16	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0018	0.0018
16	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0002
16	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
17	0	0.4181	0.1668	0.0631	0.0225	0.0075	0.0023	0.0007	0.0002	0.0000	0.0000	0.0000
17	1	0.3741	0.3150	0.1893	0.0957	0.0426	0.0169	0.0060	0.0019	0.0005	0.0001	0.0001
17	2	0.1575	0.2800	0.2673	0.1914	0.1136	0.0581	0.0260	0.0102	0.0035	0.0010	0.0010
17	3	0.0415	0.1556	0.2359	0.2393	0.1893	0.1245	0.0701	0.0341	0.0144	0.0052	0.0052
17	4	0.0076	0.0605	0.1457	0.2093	0.2209	0.1868	0.1320	0.0796	0.0411	0.0182	0.0182
17	5	0.0010	0.0175	0.0668	0.1361	0.1914	0.2081	0.1849	0.1379	0.0875	0.0472	0.0472
17	6	0.0001	0.0039	0.0236	0.0680	0.1276	0.1784	0.1991	0.1839	0.1432	0.0944	0.0944
17	7	0.0000	0.0007	0.0065	0.0267	0.0668	0.1201	0.1685	0.1927	0.1841	0.1484	0.1484
17	8	0.0000	0.0001	0.0014	0.0084	0.0279	0.0644	0.1134	0.1606	0.1883	0.1855	0.1855
17	9	0.0000	0.0000	0.0003	0.0021	0.0093	0.0276	0.0611	0.1070	0.1540	0.1855	0.1855
17	10	0.0000	0.0000	0.0000	0.0004	0.0025	0.0095	0.0263	0.0571	0.1008	0.1484	0.1484
17	11	0.0000	0.0000	0.0000	0.0001	0.0005	0.0026	0.0090	0.0242	0.0525	0.0944	0.0944
17	12	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0024	0.0081	0.0215	0.0472	0.0472
17	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0021	0.0068	0.0182	0.0182
17	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0016	0.0052	0.0052
17	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0.0010
17	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
17	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
18	0	0.3972	0.1501	0.0536	0.0180	0.0056	0.0016	0.0004	0.0001	0.0000	0.0000	0.0000
18	1	0.3763	0.3002	0.1704	0.0811	0.0338	0.0126	0.0042	0.0012	0.0003	0.0001	0.0001
18	2	0.1683	0.2835	0.2556	0.1723	0.0958	0.0458	0.0190	0.0069	0.0022	0.0006	0.0006
18	3	0.0473	0.1680	0.2406	0.2297	0.1704	0.1046	0.0547	0.0246	0.0095	0.0031	0.0031
18	4	0.0093	0.0700	0.1592	0.2153	0.2130	0.1681	0.1104	0.0614	0.0291	0.0117	0.0117
18	5	0.0014	0.0218	0.0787	0.1507	0.1988	0.2017	0.1664	0.1146	0.0666	0.0327	0.0327
18	6	0.0002	0.0052	0.0301	0.0816	0.1436	0.1873	0.1941	0.1655	0.1181	0.0708	0.0708
18	7	0.0000	0.0010	0.0091	0.0350	0.0820	0.1376	0.1792	0.1892	0.1657	0.1214	0.1214
18	8	0.0000	0.0002	0.0022	0.0120	0.0376	0.0811	0.1327	0.1734	0.1864	0.1669	0.1669
18	9	0.0000	0.0000	0.0004	0.0033	0.0139	0.0386	0.0794	0.1284	0.1694	0.1855	0.1855
18	10	0.0000	0.0000	0.0001	0.0008	0.0042	0.0149	0.0385	0.0771	0.1248	0.1669	0.1669
18	11	0.0000	0.0000	0.0000	0.0001	0.0010	0.0046	0.0151	0.0374	0.0742	0.1214	0.1214
18	12	0.0000	0.0000	0.0000	0.0000	0.0002	0.0012	0.0047	0.0145	0.0354	0.0708	0.0708
18	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0012	0.0045	0.0134	0.0327	0.0327
18	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0039	0.0117	0.0117
18	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0009	0.0031	0.0031
18	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0006
18	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
18	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLA 5 (CONTINUACIÓN)

n	k	0.05	0.10	0.15	0.20	p 0.25	0.30	0.35	0.40	0.45	0.50
19	0	0.3774	0.1351	0.0456	0.0144	0.0042	0.0011	0.0003	0.0001	0.0000	0.0000
19	1	0.3774	0.2852	0.1529	0.0685	0.0268	0.0093	0.0029	0.0008	0.0002	0.0000
19	2	0.1787	0.2852	0.2428	0.1540	0.0803	0.0358	0.0138	0.0046	0.0013	0.0003
19	3	0.0533	0.1796	0.2428	0.2182	0.1517	0.0869	0.0422	0.0175	0.0062	0.0018
19	4	0.0112	0.0798	0.1714	0.2182	0.2023	0.1491	0.0909	0.0467	0.0203	0.0074
19	5	0.0018	0.0266	0.0907	0.1636	0.2023	0.1916	0.1468	0.0933	0.0497	0.0222
19	6	0.0002	0.0069	0.0374	0.0955	0.1574	0.1916	0.1844	0.1451	0.0949	0.0518
19	7	0.0000	0.0014	0.0122	0.0443	0.0974	0.1525	0.1844	0.1797	0.1443	0.0961
19	8	0.0000	0.0002	0.0032	0.0166	0.0487	0.0981	0.1489	0.1797	0.1771	0.1442
19	9	0.0000	0.0000	0.0007	0.0051	0.0198	0.0514	0.0980	0.1464	0.1771	0.1762
19	10	0.0000	0.0000	0.0001	0.0013	0.0066	0.0220	0.0528	0.0976	0.1449	0.1762
19	11	0.0000	0.0000	0.0000	0.0003	0.0018	0.0077	0.0233	0.0532	0.0970	0.1442
19	12	0.0000	0.0000	0.0000	0.0000	0.0004	0.0022	0.0083	0.0237	0.0529	0.0961
19	13	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0024	0.0085	0.0233	0.0518
19	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0024	0.0082	0.0222
19	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0022	0.0074
19	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0018
19	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003
19	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
19	19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0002	0.0000	0.0000	0.0000
20	1	0.3774	0.2702	0.1368	0.0576	0.0211	0.0068	0.0020	0.0005	0.0001	0.0000
20	2	0.1887	0.2852	0.2293	0.1369	0.0669	0.0278	0.0100	0.0031	0.0008	0.0002
20	3	0.0596	0.1901	0.2428	0.2054	0.1339	0.0716	0.0323	0.0123	0.0040	0.0011
20	4	0.0133	0.0898	0.1821	0.2182	0.1897	0.1304	0.0738	0.0350	0.0139	0.0046
20	5	0.0022	0.0319	0.1028	0.1746	0.2023	0.1789	0.1272	0.0746	0.0365	0.0148
20	6	0.0003	0.0089	0.0454	0.1091	0.1686	0.1916	0.1712	0.1244	0.0746	0.0370
20	7	0.0000	0.0020	0.0160	0.0545	0.1124	0.1643	0.1844	0.1659	0.1221	0.0739
20	8	0.0000	0.0004	0.0046	0.0222	0.0609	0.1144	0.1614	0.1797	0.1623	0.1201
20	9	0.0000	0.0001	0.0011	0.0074	0.0271	0.0654	0.1158	0.1597	0.1771	0.1602
20	10	0.0000	0.0000	0.0002	0.0020	0.0099	0.0308	0.0686	0.1171	0.1593	0.1762
20	11	0.0000	0.0000	0.0000	0.0005	0.0030	0.0120	0.0336	0.0710	0.1185	0.1602
20	12	0.0000	0.0000	0.0000	0.0001	0.0008	0.0039	0.0136	0.0355	0.0727	0.1201
20	13	0.0000	0.0000	0.0000	0.0000	0.0002	0.0010	0.0045	0.0146	0.0366	0.0739
20	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0012	0.0049	0.0150	0.0370
20	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0013	0.0049	0.0148
20	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0013	0.0046
20	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011
20	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002
20	19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
20	20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLA 6: PROBABILIDADES DE POISSON

k	0.005	0.01	0.02	0.03	0.04	λ	0.05	0.06	0.07	0.08	0.09
0	0.9950	0.9900	0.9802	0.9704	0.9608		0.9512	0.9418	0.9324	0.9231	0.9139
1	0.0050	0.0099	0.0196	0.0291	0.0384		0.0476	0.0565	0.0653	0.0738	0.0823
2	0.0000	0.0000	0.0002	0.0004	0.0008		0.0012	0.0017	0.0023	0.0030	0.0037
3	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0001	0.0001	0.0001

k	0.1	0.2	0.3	0.4	0.5	λ	0.6	0.7	0.8	0.9	1.0
0	0.9048	0.8187	0.7408	0.6703	0.6065		0.5488	0.4966	0.4493	0.4066	0.3679
1	0.0905	0.1637	0.2222	0.2681	0.3033		0.3293	0.3476	0.3595	0.3659	0.3679
2	0.0045	0.0164	0.0333	0.0536	0.0758		0.0988	0.1217	0.1438	0.1647	0.1839
3	0.0002	0.0011	0.0033	0.0072	0.0126		0.0198	0.0284	0.0383	0.0494	0.0613
4	0.0000	0.0001	0.0003	0.0007	0.0016		0.0030	0.0050	0.0077	0.0111	0.0153
5	0.0000	0.0000	0.0000	0.0001	0.0002		0.0004	0.0007	0.0012	0.0020	0.0031
6	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0001	0.0002	0.0003	0.0005
7	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0001

k	1.1	1.2	1.3	1.4	1.5	λ	1.6	1.7	1.8	1.9	2.0
0	0.3329	0.3012	0.2725	0.2466	0.2231		0.2019	0.1827	0.1653	0.1496	0.1353
1	0.3662	0.3614	0.3543	0.3452	0.3347		0.3230	0.3106	0.2975	0.2842	0.2707
2	0.2014	0.2169	0.2303	0.2417	0.2510		0.2584	0.2640	0.2678	0.2700	0.2707
3	0.0738	0.0867	0.0998	0.1128	0.1255		0.1378	0.1496	0.1607	0.1710	0.1804
4	0.0203	0.0260	0.0324	0.0395	0.0471		0.0551	0.0636	0.0723	0.0812	0.0902
5	0.0045	0.0062	0.0084	0.0111	0.0141		0.0176	0.0216	0.0260	0.0309	0.0361
6	0.0008	0.0012	0.0018	0.0026	0.0035		0.0047	0.0061	0.0078	0.0098	0.0120
7	0.0001	0.0002	0.0003	0.0005	0.0008		0.0011	0.0015	0.0020	0.0027	0.0034
8	0.0000	0.0000	0.0001	0.0001	0.0001		0.0002	0.0003	0.0005	0.0006	0.0009
9	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0001	0.0001	0.0001	0.0002

k	2.1	2.2	2.3	2.4	2.5	λ	2.6	2.7	2.8	2.9	3.0
0	0.1225	0.1108	0.1003	0.0907	0.0821		0.0743	0.0672	0.0608	0.0550	0.0498
1	0.2572	0.2438	0.2306	0.2177	0.2052		0.1931	0.1815	0.1703	0.1596	0.1494
2	0.2700	0.2681	0.2652	0.2613	0.2565		0.2510	0.2450	0.2384	0.2314	0.2240
3	0.1890	0.1966	0.2033	0.2090	0.2138		0.2176	0.2205	0.2225	0.2237	0.2240
4	0.0992	0.1082	0.1169	0.1254	0.1336		0.1414	0.1488	0.1557	0.1622	0.1680
5	0.0417	0.0476	0.0538	0.0602	0.0668		0.0735	0.0804	0.0872	0.0940	0.1008
6	0.0146	0.0174	0.0206	0.0241	0.0278		0.0319	0.0362	0.0407	0.0455	0.0504
7	0.0044	0.0055	0.0068	0.0083	0.0099		0.0118	0.0139	0.0163	0.0188	0.0216
8	0.0011	0.0015	0.0019	0.0025	0.0031		0.0038	0.0047	0.0057	0.0068	0.0081
9	0.0003	0.0004	0.0005	0.0007	0.0009		0.0011	0.0014	0.0018	0.0022	0.0027
10	0.0001	0.0001	0.0001	0.0002	0.0002		0.0003	0.0004	0.0005	0.0006	0.0008
11	0.0000	0.0000	0.0000	0.0000	0.0000		0.0001	0.0001	0.0001	0.0002	0.0002
12	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0001

k	3.1	3.2	3.3	3.4	3.5	λ	3.6	3.7	3.8	3.9	4.0
0	0.0450	0.0408	0.0369	0.0334	0.0302		0.0273	0.0247	0.0224	0.0202	0.0183
1	0.1397	0.1304	0.1217	0.1135	0.1057		0.0984	0.0915	0.0850	0.0789	0.0733
2	0.2165	0.2087	0.2008	0.1929	0.1850		0.1771	0.1692	0.1615	0.1539	0.1465
3	0.2237	0.2226	0.2209	0.2186	0.2158		0.2125	0.2087	0.2046	0.2001	0.1954
4	0.1733	0.1781	0.1823	0.1858	0.1888		0.1912	0.1931	0.1944	0.1951	0.1954
5	0.1075	0.1140	0.1203	0.1264	0.1322		0.1377	0.1429	0.1477	0.1522	0.1563
6	0.0555	0.0608	0.0662	0.0716	0.0771		0.0826	0.0881	0.0936	0.0989	0.1042
7	0.0246	0.0278	0.0312	0.0348	0.0385		0.0425	0.0466	0.0508	0.0551	0.0595
8	0.0095	0.0111	0.0129	0.0148	0.0169		0.0191	0.0215	0.0241	0.0269	0.0298
9	0.0033	0.0040	0.0047	0.0056	0.0066		0.0076	0.0089	0.0102	0.0116	0.0132
10	0.0010	0.0013	0.0016	0.0019	0.0023		0.0028	0.0033	0.0039	0.0045	0.0053
11	0.0003	0.0004	0.0005	0.0006	0.0007		0.0009	0.0011	0.0013	0.0016	0.0019
12	0.0001	0.0001	0.0001	0.0002	0.0002		0.0003	0.0003	0.0004	0.0005	0.0006
13	0.0000	0.0000	0.0000	0.0000	0.0001		0.0001	0.0001	0.0001	0.0002	0.0002
14	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0001

k	4.1	4.2	4.3	4.4	4.5	λ	4.6	4.7	4.8	4.9	5.0
0	0.0166	0.0150	0.0136	0.0123	0.0111		0.0101	0.0091	0.0082	0.0074	0.0067
1	0.0679	0.0630	0.0583	0.0540	0.0500		0.0462	0.0427	0.0395	0.0365	0.0337
2	0.1393	0.1323	0.1254	0.1188	0.1125		0.1063	0.1005	0.0948	0.0894	0.0842
3	0.1904	0.1852	0.1798	0.1743	0.1687		0.1631	0.1574	0.1517	0.1460	0.1404
4	0.1951	0.1944	0.1933	0.1917	0.1898		0.1875	0.1849	0.1820	0.1789	0.1755
5	0.1600	0.1633	0.1662	0.1687	0.1708		0.1725	0.1738	0.1747	0.1753	0.1755
6	0.1093	0.1143	0.1191	0.1237	0.1281		0.1323	0.1362	0.1398	0.1432	0.1462
7	0.0640	0.0686	0.0732	0.0778	0.0824		0.0869	0.0914	0.0959	0.1002	0.1044
8	0.0328	0.0360	0.0393	0.0428	0.0463		0.0500	0.0537	0.0575	0.0614	0.0653
9	0.0150	0.0168	0.0188	0.0209	0.0232		0.0255	0.0281	0.0307	0.0334	0.0363
10	0.0061	0.0071	0.0081	0.0092	0.0104		0.0118	0.0132	0.0147	0.0164	0.0181
11	0.0023	0.0027	0.0032	0.0037	0.0043		0.0049	0.0056	0.0064	0.0073	0.0082
12	0.0008	0.0009	0.0011	0.0013	0.0016		0.0019	0.0022	0.0026	0.0030	0.0034
13	0.0002	0.0003	0.0004	0.0005	0.0006		0.0007	0.0008	0.0009	0.0011	0.0013
14	0.0001	0.0001	0.0001	0.0001	0.0002		0.0002	0.0003	0.0003	0.0004	0.0005
15	0.0000	0.0000	0.0000	0.0000	0.0001		0.0001	0.0001	0.0001	0.0001	0.0002

k	5.1	5.2	5.3	5.4	5.5	λ	5.6	5.7	5.8	5.9	6.0
0	0.0061	0.0055	0.0050	0.0045	0.0041		0.0037	0.0033	0.0030	0.0027	0.0025
1	0.0311	0.0287	0.0265	0.0244	0.0225		0.0207	0.0191	0.0176	0.0162	0.0149
2	0.0793	0.0746	0.0701	0.0659	0.0618		0.0580	0.0544	0.0509	0.0477	0.0446
3	0.1348	0.1293	0.1239	0.1185	0.1133		0.1082	0.1033	0.0985	0.0938	0.0892
4	0.1719	0.1681	0.1641	0.1600	0.1558		0.1515	0.1472	0.1428	0.1383	0.1339
5	0.1753	0.1748	0.1740	0.1728	0.1714		0.1697	0.1678	0.1656	0.1632	0.1606
6	0.1490	0.1515	0.1537	0.1555	0.1571		0.1584	0.1594	0.1601	0.1605	0.1606
7	0.1086	0.1125	0.1163	0.1200	0.1234		0.1267	0.1298	0.1326	0.1353	0.1377
8	0.0692	0.0731	0.0771	0.0810	0.0849		0.0887	0.0925	0.0962	0.0998	0.1033
9	0.0392	0.0423	0.0454	0.0486	0.0519		0.0552	0.0586	0.0620	0.0654	0.0688
10	0.0200	0.0220	0.0241	0.0262	0.0285		0.0309	0.0334	0.0359	0.0386	0.0413
11	0.0093	0.0104	0.0116	0.0129	0.0143		0.0157	0.0173	0.0190	0.0207	0.0225
12	0.0039	0.0045	0.0051	0.0058	0.0065		0.0073	0.0082	0.0092	0.0102	0.0113
13	0.0015	0.0018	0.0021	0.0024	0.0028		0.0032	0.0036	0.0041	0.0046	0.0052
14	0.0006	0.0007	0.0008	0.0009	0.0011		0.0013	0.0015	0.0017	0.0019	0.0022
15	0.0002	0.0002	0.0003	0.0003	0.0004		0.0005	0.0006	0.0007	0.0008	0.0009
16	0.0001	0.0001	0.0001	0.0001	0.0001		0.0002	0.0002	0.0002	0.0003	0.0003
17	0.0000	0.0000	0.0000	0.0000	0.0000		0.0001	0.0001	0.0001	0.0001	0.0001

TABLA 6 (CONTINUACION)

k	6.1	6.2	6.3	6.4	6.5	λ	6.6	6.7	6.8	6.9	7.0
0	0.0022	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0011	0.0010	0.0009	0.0009
1	0.0137	0.0126	0.0116	0.0106	0.0098	0.0090	0.0082	0.0076	0.0070	0.0064	0.0064
2	0.0417	0.0390	0.0364	0.0340	0.0318	0.0296	0.0276	0.0258	0.0240	0.0223	0.0223
3	0.0848	0.0806	0.0765	0.0726	0.0688	0.0652	0.0617	0.0584	0.0552	0.0521	0.0521
4	0.1294	0.1249	0.1205	0.1162	0.1118	0.1076	0.1034	0.0992	0.0952	0.0912	0.0912
5	0.1579	0.1549	0.1519	0.1487	0.1454	0.1420	0.1385	0.1349	0.1314	0.1277	0.1277
6	0.1605	0.1601	0.1595	0.1586	0.1575	0.1562	0.1546	0.1529	0.1511	0.1490	0.1490
7	0.1399	0.1418	0.1435	0.1450	0.1462	0.1472	0.1480	0.1486	0.1489	0.1490	0.1490
8	0.1066	0.1099	0.1130	0.1160	0.1188	0.1215	0.1240	0.1263	0.1284	0.1304	0.1304
9	0.0723	0.0757	0.0791	0.0825	0.0858	0.0891	0.0923	0.0954	0.0985	0.1014	0.1014
10	0.0441	0.0469	0.0498	0.0528	0.0558	0.0588	0.0618	0.0649	0.0679	0.0710	0.0710
11	0.0244	0.0265	0.0285	0.0307	0.0330	0.0353	0.0377	0.0401	0.0426	0.0452	0.0452
12	0.0124	0.0137	0.0150	0.0164	0.0179	0.0194	0.0210	0.0227	0.0245	0.0263	0.0263
13	0.0058	0.0065	0.0073	0.0081	0.0089	0.0099	0.0108	0.0119	0.0130	0.0142	0.0142
14	0.0025	0.0029	0.0033	0.0037	0.0041	0.0046	0.0052	0.0058	0.0064	0.0071	0.0071
15	0.0010	0.0012	0.0014	0.0016	0.0018	0.0020	0.0023	0.0026	0.0029	0.0033	0.0033
16	0.0004	0.0005	0.0005	0.0006	0.0007	0.0008	0.0010	0.0011	0.0013	0.0014	0.0014
17	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006
18	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001

k	7.1	7.2	7.3	7.4	7.5	λ	7.6	7.7	7.8	7.9	8.0
0	0.0008	0.0007	0.0007	0.0006	0.0006	0.0005	0.0005	0.0004	0.0004	0.0003	0.0003
1	0.0059	0.0054	0.0049	0.0045	0.0041	0.0038	0.0035	0.0032	0.0029	0.0027	0.0027
2	0.0208	0.0194	0.0180	0.0167	0.0156	0.0145	0.0134	0.0125	0.0116	0.0107	0.0107
3	0.0492	0.0464	0.0438	0.0413	0.0389	0.0366	0.0345	0.0324	0.0305	0.0286	0.0286
4	0.0874	0.0836	0.0799	0.0764	0.0729	0.0696	0.0663	0.0632	0.0602	0.0573	0.0573
5	0.1241	0.1204	0.1167	0.1130	0.1094	0.1057	0.1021	0.0986	0.0951	0.0916	0.0916
6	0.1468	0.1445	0.1420	0.1394	0.1367	0.1339	0.1311	0.1282	0.1252	0.1221	0.1221
7	0.1489	0.1486	0.1481	0.1474	0.1465	0.1454	0.1442	0.1428	0.1413	0.1396	0.1396
8	0.1321	0.1337	0.1351	0.1363	0.1373	0.1381	0.1388	0.1392	0.1395	0.1396	0.1396
9	0.1042	0.1070	0.1096	0.1121	0.1144	0.1167	0.1187	0.1207	0.1224	0.1241	0.1241
10	0.0740	0.0770	0.0800	0.0829	0.0858	0.0887	0.0914	0.0941	0.0967	0.0993	0.0993
11	0.0478	0.0504	0.0531	0.0558	0.0585	0.0613	0.0640	0.0667	0.0695	0.0722	0.0722
12	0.0283	0.0303	0.0323	0.0344	0.0366	0.0388	0.0411	0.0434	0.0457	0.0481	0.0481
13	0.0154	0.0168	0.0181	0.0196	0.0211	0.0227	0.0243	0.0260	0.0278	0.0296	0.0296
14	0.0078	0.0086	0.0095	0.0104	0.0113	0.0123	0.0134	0.0145	0.0157	0.0169	0.0169
15	0.0037	0.0041	0.0046	0.0051	0.0057	0.0062	0.0069	0.0075	0.0083	0.0090	0.0090
16	0.0016	0.0019	0.0021	0.0024	0.0026	0.0030	0.0033	0.0037	0.0041	0.0045	0.0045
17	0.0007	0.0008	0.0009	0.0010	0.0012	0.0013	0.0015	0.0017	0.0019	0.0021	0.0021
18	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009	0.0009
19	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003	0.0004	0.0004
20	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001

TABLA 7: TABLA DE NÚMEROS AL AZAR

Primera serie																				
	1 - 4		5 - 8		9 - 12		13 - 16		17 - 20		21 - 24		25 - 28		29 - 32		33 - 36		37 - 40	
1	20	77	81	43	63	92	68	61	70	79	88	81	05	47	63	07	13	10	46	19
2	09	42	70	58	14	76	27	02	29	73	87	66	47	73	31	59	02	96	41	07
3	93	04	55	07	83	92	26	76	50	57	05	97	12	85	01	30	82	45	52	08
4	95	99	93	67	54	96	35	98	84	64	80	88	29	39	07	00	97	95	59	24
5	40	82	40	08	05	80	60	50	33	93	68	58	83	62	06	09	20	56	91	36
6	78	11	44	01	19	42	06	02	32	19	99	23	94	02	29	27	29	38	17	82
7	56	41	30	34	77	26	83	55	26	08	69	53	66	16	19	43	77	69	70	77
8	76	42	48	52	69	11	70	01	23	21	99	22	30	75	42	61	99	20	91	90
9	41	14	93	39	41	11	56	76	60	04	24	75	18	06	14	42	91	25	31	92
10	17	42	51	18	60	28	10	87	61	25	88	92	04	30	90	80	32	26	91	22
11	96	66	80	87	48	97	22	47	84	24	58	51	41	10	54	26	93	19	90	20
12	24	81	91	42	70	40	96	75	48	30	48	66	21	54	20	98	12	00	86	61
13	78	65	68	07	07	95	15	50	67	10	01	62	36	75	93	76	40	54	97	68
14	29	27	78	63	25	00	14	51	15	18	18	14	03	96	63	08	85	49	16	14
15	34	16	38	45	71	04	00	72	44	03	63	46	49	56	50	76	57	32	84	43
16	82	76	24	97	43	39	05	39	93	69	61	80	25	47	90	15	70	06	74	13
17	18	93	50	05	65	07	39	37	51	99	78	42	52	78	82	86	81	17	69	09
18	46	84	90	64	55	19	12	20	32	11	56	30	00	54	75	95	54	22	80	38
19	59	52	94	41	54	33	08	80	51	39	35	64	22	90	59	82	79	76	23	22
20	38	12	76	09	53	32	80	07	19	34	18	55	60	86	33	22	36	15	79	85
21	14	72	18	71	55	19	09	25	27	36	10	35	60	87	96	55	74	86	08	54
22	44	29	94	19	34	91	62	94	56	81	35	00	79	15	62	92	66	16	67	29
23	50	10	67	79	43	27	66	85	52	00	97	65	07	58	31	74	90	09	24	75
24	11	19	88	34	80	11	94	03	56	28	53	52	86	83	51	38	97	02	50	20
25	12	16	81	62	90	38	45	23	13	08	18	57	67	45	15	75	86	07	77	57

Segunda serie																				
	1 - 4		5 - 8		9 - 12		13 - 16		17 - 20		21 - 24		25 - 28		29 - 32		33 - 36		37 - 40	
1	83	79	21	68	54	51	23	50	78	17	73	51	94	71	91	31	29	97	13	16
2	27	98	04	97	52	48	19	33	82	16	00	34	30	67	58	00	80	80	92	26
3	75	57	42	55	59	91	72	75	66	75	39	70	55	24	09	19	70	22	42	10
4	60	73	07	27	71	94	68	70	48	81	40	91	16	24	45	24	54	03	18	03
5	27	42	15	67	94	44	48	62	37	53	83	15	90	90	60	19	78	62	44	70
6	57	27	77	75	38	18	24	32	91	53	78	91	33	38	31	95	85	11	33	87
7	22	79	42	39	32	65	83	60	74	63	56	77	47	21	88	36	43	10	19	41
8	61	94	32	42	83	81	92	40	99	00	05	66	33	61	32	10	43	15	49	25
9	24	28	53	76	13	61	73	22	50	51	75	08	10	90	71	58	22	42	46	83
10	06	75	46	13	77	68	97	05	56	73	34	86	42	22	37	75	94	87	57	72
11	83	81	98	41	67	38	97	30	12	31	87	76	81	07	32	88	42	29	94	58
12	33	66	09	91	21	26	52	57	47	14	27	75	07	84	50	96	95	12	25	01
13	63	14	10	59	10	68	27	91	00	17	36	79	01	79	65	43	13	98	52	21
14	84	43	66	38	65	72	14	55	93	78	24	57	38	10	54	53	92	41	82	56
15	36	75	92	36	76	77	89	27	06	57	37	70	36	09	99	90	66	44	91	89
16	17	23	20	39	81	03	49	79	68	20	94	45	95	92	63	06	55	29	20	82
17	57	09	57	60	40	64	00	77	31	05	83	44	96	62	56	42	42	68	46	42
18	33	49	47	96	50	21	75	68	28	12	46	25	72	64	50	72	75	16	67	00
19	80	52	82	00	02	12	67	61	43	23	14	53	10	66	16	29	06	60	23	20
20	45	29	44	64	44	70	22	10	70	26	43	49	28	51	69	52	85	95	98	58
21	16	04	54	67	60	53	51	60	82	45	56	26	08	89	42	25	93	76	69	73
22	81	09	97	91	19	60	91	96	53	66	45	33	75	09	67	42	59	02	18	97
23	20	61	83	96	58	12	64	99	78	24	94	36	99	60	07	70	27	40	97	15
24	43	17	00	05	86	39	95	58	12	35	84	31	97	75	50	52	06	88	27	79
25	27	84	00	90	41	44	05	43	36	93	01	37	91	93	24	78	00	12	04	12

TABLA7 (CONTINUACION)

Tercera serie																				
	01 - 04		05 - 08		09 - 12		13 - 16		17 - 20		21 - 24		25 - 28		29 - 32		33 - 36		37 - 40	
1	35	87	68	11	29	27	78	34	74	92	86	13	33	22	34	75	59	44	61	63
2	16	58	21	36	34	35	92	56	72	85	03	91	31	33	26	62	08	30	95	09
3	44	20	51	20	95	87	87	37	37	05	06	43	61	75	51	99	52	03	23	29
4	36	04	26	10	46	13	47	92	34	02	53	54	97	57	24	18	37	62	98	35
5	92	15	17	81	81	67	17	00	76	64	77	34	20	18	88	01	88	31	18	40
6	29	12	08	74	26	94	48	03	85	60	23	69	65	52	53	18	26	72	47	62
7	19	37	78	71	07	12	90	53	13	34	19	49	81	36	85	96	19	81	68	65
8	14	92	75	64	31	51	00	20	86	47	07	04	40	26	57	62	34	95	57	17
9	02	19	07	98	20	98	21	10	69	15	39	55	80	23	80	26	63	33	81	26
10	31	18	25	46	66	31	15	27	09	80	13	44	72	58	42	93	71	72	69	17
11	51	96	70	26	99	94	81	55	72	29	57	37	29	36	99	55	35	99	22	79
12	11	78	29	90	66	21	63	60	39	87	77	34	55	26	56	94	74	43	71	04
13	93	10	34	79	41	18	07	79	30	84	77	27	07	40	54	24	90	96	43	63
14	19	32	42	52	58	31	29	38	39	31	00	68	71	25	28	05	27	55	56	18
15	26	84	50	36	62	73	35	95	56	00	76	61	16	40	08	80	71	87	20	48
16	05	50	36	40	59	22	90	33	76	82	58	66	30	65	68	23	38	98	39	92
17	45	14	18	03	19	34	32	93	93	37	97	11	37	04	64	62	60	53	86	33
18	71	44	13	65	14	74	67	80	11	11	01	90	47	58	36	86	13	68	85	64
19	47	46	52	39	25	41	96	96	98	22	11	51	90	91	82	16	96	43	11	05
20	33	95	58	94	59	65	09	99	47	45	70	06	60	66	75	22	43	93	79	62
21	28	74	81	01	28	31	47	96	55	75	26	61	48	51	10	20	50	43	83	28
22	33	68	25	37	96	38	70	64	36	19	60	48	71	41	05	53	77	56	81	78
23	85	38	32	52	85	14	64	87	82	11	85	11	19	29	80	35	09	99	54	23
24	04	29	07	21	28	33	85	69	93	51	57	30	02	14	58	12	84	34	30	71
25	47	20	21	03	47	88	47	25	51	49	70	21	07	71	07	03	08	69	69	35

Cuarta serie																				
	01 - 04		05 - 08		09 - 12		13 - 16		17 - 20		21 - 24		25 - 28		29 - 32		33 - 36		37 - 40	
1	72	43	49	72	61	99	36	48	80	20	10	31	66	31	84	83	87	39	93	25
2	34	71	89	58	13	15	56	47	41	24	57	70	96	69	59	17	28	25	55	05
3	94	75	59	95	98	99	88	03	83	67	75	75	92	73	56	60	91	80	59	53
4	28	97	88	17	22	80	49	80	81	22	07	83	41	63	15	27	25	72	00	01
5	75	89	87	84	83	13	56	59	46	28	26	11	62	28	81	61	42	86	35	50
6	77	44	16	92	59	93	41	27	08	86	09	45	47	43	35	19	66	71	81	08
7	46	06	66	76	61	89	59	75	61	68	68	50	47	98	47	21	84	05	52	95
8	53	46	28	21	43	74	40	07	15	60	52	15	39	94	67	29	90	83	08	61
9	55	23	44	40	08	34	48	33	12	12	28	12	69	81	14	58	25	01	00	13
10	38	44	74	56	06	24	85	13	22	04	07	70	00	18	43	99	03	53	77	98
11	08	12	50	56	81	86	61	59	77	00	11	71	00	47	29	62	68	87	25	30
12	34	27	77	14	64	22	20	77	22	41	50	92	51	67	70	54	14	26	54	47
13	19	98	90	19	27	12	80	34	87	97	73	09	98	80	66	74	77	59	11	54
14	39	05	72	32	69	87	95	50	82	76	50	79	82	18	72	77	88	60	92	11
15	38	96	83	88	05	76	23	20	09	33	08	02	10	74	93	22	09	75	83	23
16	72	94	21	37	57	90	48	48	43	96	66	75	33	08	09	65	62	00	94	38
17	86	61	52	23	80	46	97	81	15	83	88	88	71	98	73	73	43	63	93	74
18	14	78	87	24	89	77	62	94	19	26	16	08	78	08	97	24	19	03	47	46
19	24	17	92	19	81	85	71	91	02	41	45	08	44	99	72	36	84	23	59	51
20	92	52	08	30	47	44	31	36	80	12	73	65	98	06	15	69	59	18	38	58
21	94	47	16	55	89	88	64	29	02	59	48	90	06	90	57	14	65	55	75	97
22	47	21	92	90	36	20	85	34	96	73	11	69	71	36	65	16	49	89	41	78
23	76	89	16	29	20	24	00	47	33	89	74	11	59	62	84	53	03	14	74	77
24	90	93	86	13	61	32	15	62	99	30	63	47	88	87	86	12	59	30	69	44
25	44	93	41	83	33	43	31	74	85	47	44	45	91	36	40	48	56	84	10	09

ANEXO 5

INSTRUMENTACIÓN DIDÁCTICA COMPLETA

INSTRUMENTACIÓN DIDÁCTICA PARA LA FORMACIÓN Y DESARROLLO DE COMPETENCIAS PROFESIONALES

Subdirección Académica

Periodo: AGOSTO-DICIEMBRE 2022



Nombre de la asignatura: ESTADÍSTICA INFERENCIAL I
Plan de estudios: INGENIERÍA INDUSTRIAL
INGENIERÍA EN LOGÍSTICA
Clave de asignatura: AEF – 1024
Horas teoría–horas prácticas–créditos: 3-2-5

1. Caracterización de la asignatura

Esta asignatura aporta al perfil del Ingeniero en Logística e Ingeniero Industrial los elementos básicos para hacer análisis a partir del estadístico de la muestra y conceptos de la estimación estadística. Le permite establecer inferencias sobre una población y conclusiones a partir de la información que arrojan las pruebas estadísticas. Estadística Inferencial I es una asignatura integradora para la competencia específica de Probabilidad y Estadística y a su vez provee las competencias previas para Estadística Inferencial II, por lo que se plantea como una asignatura básica de la carrera de Ingeniería en Logística e Industrial. Con la asignatura de Investigación de Operaciones II, así como en la de Simulación tienen relación los temas de intervalos de confianza y pruebas de hipótesis. El tema de regresión lineal simple es competencia previa para la asignatura de Administración de Operaciones I. En Control Estadístico de la Calidad es previo el tema de pruebas estadísticas. Tienen relación los métodos estadísticos en la asignatura de Gestión de los Sistemas de Calidad, así como en Administración del Mantenimiento. Por lo que se pueden realizar proyectos integradores con cualquiera de esas asignaturas.

1

2. Intención didáctica

Se organiza la materia de Estadística Inferencial I en cinco temas: El tema uno, distribuciones fundamentales para el muestreo, introduce al estudiante en los conceptos, teoremas y contexto de la teoría del muestreo probabilístico y no probabilístico, así como las distribuciones fundamentales para el muestreo. En el tema dos sobre estimación le brinda al estudiante los conceptos de estimación puntual y por intervalo de la media, proporción, varianzas, diferencia de medias y proporciones y relación de varianzas. El tema tres contiene el procedimiento para realizar pruebas de hipótesis, tomando en cuenta la confiabilidad y la eficacia de los errores tipo I y tipo II así como la determinación de potencia de la prueba a que se somete una muestra del experimento de interés respecto a una población de referencia. El tema cuatro introduce al estudiante a realizar pruebas de bondad y ajuste y pruebas no paramétricas de una muestra en un experimento de interés respecto a una población de referencia, verificando la adecuación del modelo probabilístico. El tema cinco de regresión lineal simple introduce al estudiante al análisis de la relación lineal entre dos variables, la aplicación de la teoría de mínimos cuadrados y el modelo matemático resultante del caso de estudio y sus límites de validez. El enfoque de la asignatura se presenta para que el estudiante desarrolle las competencias aplicando las bases estadísticas obtenidas en las asignaturas anteriores, de tal forma que establezca el problema a resolver con el diseño y análisis de experimentos más conveniente a una situación real. Identifica, variables a controlar y registrar los elementos que le permitan diseñar los problemas de manera más autónoma. La lista de actividades de aprendizaje no es exhaustiva, se sugieren sobre todo las necesarias para hacer más significativo y efectivo el aprendizaje. Algunas de las actividades sugeridas pueden hacerse como actividad extra clase y comenzar el diseño en clase a partir de la discusión de los resultados de las observaciones. Se busca que el estudiante realice una investigación de campo donde identifique alguna característica de su entorno y recopile la información correspondiente, haga análisis estadístico, corra un experimento y que interprete los resultados. En el transcurso de las actividades es importante que el estudiante aprenda a valorar las actividades que lleva a cabo y entienda que está construyendo su hacer futuro y en consecuencia actúe de una manera profesional; de igual manera, aprecie la importancia del conocimiento y los hábitos de trabajo; desarrolle la precisión y la curiosidad, la puntualidad, el entusiasmo y el interés, la tenacidad, la flexibilidad y la autonomía. El docente de Estadística Inferencial I debe mostrar y objetivar su conocimiento y experiencia en el área para construir escenarios de aprendizaje significativo en los estudiantes que inician su formación profesional. El docente enfatiza el desarrollo de las actividades de aprendizaje de esta asignatura a fin de que ellas refuercen los aspectos formativos: incentivar la curiosidad, el entusiasmo, la puntualidad, la constancia, el interés por mejorar, el respeto y la tolerancia hacia sus compañeros y docentes, a sus ideas y enfoques y considerar también la responsabilidad social y el respeto al medio ambiente. El docente de Estadística Inferencial I debe de estar realizando actividades de investigación.

3. Competencia de la asignatura

Emplea los métodos de muestreo adecuados para la obtención de la muestra experimental con la finalidad de realizar inferencias sobre la población y el desarrollo de pruebas estadísticas.

4. Análisis por competencias específicas

Competencia No.: 1

Descripción:

Distribuciones Fundamentales para el Muestreo: El alumno deberá recopilar, organizar, medir, analizar e interpretar estadísticamente conjuntos de datos tomados de una situación real. Y los representará gráficamente

Temas y subtemas para desarrollar	Actividades de aprendizaje	Actividades de Enseñanza	Desarrollo de competencias genéricas	Horas teórico-práctica
1.1 Introducción a la estadística (breve historia de la estadística, concepto de estadística, paramétrica y no paramétrica, estadística descriptiva, estadística inferencia, población muestra).	El alumno analizará el concepto de estadística inferencial y sus aplicaciones y elaborará un esquema.	El docente abordará el concepto de estadística inferencial y sus aplicaciones y solicitará al alumno la elaboración de un esquema del tema.	Capacidad de abstracción, análisis y síntesis.	12 horas teoría – 8 horas práctica – 20 horas totales
1.2 Componentes de una investigación estadística.	El alumno investigará, analizará y elaborará un esquema sobre las	El docente solicitará al estudiante investigue las componentes de una investigación estadística y mediante	Capacidad de abstracción, análisis y síntesis.	

Rev. 1

	componentes de una investigación documental.	una mesa de diálogo se discutirán resultados. Pedirá la elaboración de un esquema acerca del tema.		
1.3 Muestreo: Introducción al muestreo y tipos de muestreo.	El alumno analizará el concepto de muestreo y sus tipos y elaborará un esquema.	El docente abordará el muestreo y sus diferentes tipos y solicitará la elaboración de un esquema del tema.	Capacidad de abstracción, análisis y síntesis.	
1.4 Teorema del límite central.	El alumno investigará y analizará el Teorema del Límite Central.	El docente solicitará al estudiante investigue el Teorema del Límite Central y mediante una mesa de diálogo se discutirán los puntos más sobresalientes para llegar a conclusiones.	Capacidad de abstracción, análisis y síntesis.	
1.5 Distribuciones fundamentales para el muestreo.	El alumno investigará y analizará las distribuciones fundamentales para el muestreo.	El docente solicitará al estudiante investigue las distribuciones fundamentales para el muestreo y sus aplicaciones y mediante una mesa de diálogo se discutirán resultados.	Capacidad de abstracción, análisis y síntesis.	
1.5.1 Distribución muestral de la media.	El alumno analizará la distribución muestral de la media y resolverá ejercicios del tema.	El docente explicará la distribución muestral de la media. Daré solución y propondrá ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
1.5.2 Distribución muestral de la diferencia de medias.	El alumno analizará la distribución muestral de la diferencia de medias y resolverá ejercicios del tema.	El docente explicará la distribución muestral de la diferencia de medias. Daré solución y propondrá ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver	

Rev. 1

			problemas.	
1.5.3 Distribución muestral de la proporción.	El alumno analizará la distribución muestral de la proporción y resolverá ejercicios del tema.	El docente explicará la distribución muestral de la proporción. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
1.5.4 Distribución muestral de la diferencia de proporciones.	El alumno analizará la distribución muestral de la diferencia de proporciones y resolverá ejercicios del tema.	El docente explicará la distribución muestral de la diferencia de proporciones. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
1.5.5 Distribución t-student.	El alumno analizará la distribución muestral t-student y resolverá ejercicios del tema.	El docente explicará la distribución muestral t-student. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	

Rev. 1

1.5.6 Distribución muestral de la varianza.	El alumno analizará la distribución muestral de la varianza y resolverá ejercicios del tema.	El docente explicará la distribución muestral de la varianza. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
1.5.7 Distribución muestral de la relación de varianzas.	El alumno analizará la distribución muestral de la razón de varianzas y resolverá ejercicios del tema.	El docente explicará la distribución muestral de la razón de varianzas. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	

Estrategias de evaluación (Matriz)

Evidencia de aprendizaje	%	Indicador de alcance						Método de Evaluación			
		A	B	C	D	E	F	Instrumento	P	C	A
Tareas selladas	30	X	X					Guía de observación	x		x
Examen escrito	70	X			x		x	cuestionario		X	

Niveles de desempeño:

Rev. 1

Desempeño	Nivel de desempeño	Indicadores de alcance	Valoración numérica
Competencia alcanzada	Excelente	Cumple al menos cinco de los siguientes indicadores a) Se adapta a situaciones y contextos complejos. Puede trabajar en equipo, reflejar sus conocimientos en la interpretación de la realidad. Inferir comportamientos o consecuencias de los fenómenos o problemas en estudio. Incluir más variables en dichos casos de estudio. b) Hace aportaciones a las actividades académicas desarrolladas. Pregunta integrando conocimientos de otras asignaturas o de casos anteriores de la misma asignatura. Presenta otros puntos de vista que complementan al presentado en la clase. Presenta fuentes de información adicionales (Internet, documentales), usa más bibliografía, consulta fuentes en un segundo idioma, etc. c) Propone y/o explica soluciones o procedimientos no vistos en clase (creatividad). Ante problemas o casos de estudio propone perspectivas diferentes, para abordarlos y sustentarlos correctamente. Aplica procedimientos aprendidos en otra asignatura o contexto para el problema que se está resolviendo. d) Introduce recursos y experiencias que promueven un pensamiento crítico; (por ejemplo el uso de las tecnologías de la información estableciendo previamente un criterio). Ante temas de una asignatura, introduce cuestionamientos de tipo ético, ecológico, histórico, político, económico, etc.; que deben tomarse en cuenta para comprender mejor, o a futuro dicho tema. Se apoya en foros, autores, bibliografía, documentales, etc. para sustentar su punto de vista. e) Incorpora conocimientos y actividades interdisciplinarias en su aprendizaje. En el desarrollo de los temas de la asignatura, incorpora conocimientos y actividades desarrollados en otras asignaturas para lograr la competencia. f) Realiza su trabajo de manera autónoma y autorregulada. Es capaz de organizar su tiempo y trabajar sin necesidad de una supervisión estrecha y/o coercitiva. Aprovecha la planeación de la asignatura presentada por el (la) profesor(a) (instrumentación didáctica) para presentar propuestas de mejora de la temática vista durante el curso. Realiza actividades de investigación para participar activamente durante el curso.	95-100
	Notable	Cumple cuatro de los indicadores diferidos en desempeño excelente	85-94
	Bueno	Cumple tres de los indicadores diferidos en desempeño excelente	75-84
	Suficiente	Cumple dos de los indicadores diferidos en desempeño excelente	70-74
Competencia no alcanzada	Insuficiente	No se cumple con el 100% de evidencias conceptuales, procedimentales y actitudinales de los indicadores diferidos en el desempeño excelente.	NA (no alcanzada)

Rev. 1

CompetenciaNo.: 2

Descripción:

Estimación: Capacidad de abstracción, análisis y síntesis.
Capacidad de aplicar los conocimientos en la práctica.
Capacidad para identificar, plantear y resolver problemas.

Temas y subtemas para desarrollar	Actividades de aprendizaje	Actividades de Enseñanza	Desarrollo de competencias genéricas	Horas teórico-práctica
2.1 Introducción.	El alumno investigará, analizará y elaborará un esquema acerca de los conceptos básicos de estimación.	El docente solicitará al estudiante la investigación de los conceptos relacionados con estimación y mediante una mesa de diálogo se discutirán los resultados. Pedirá la elaboración de un esquema del tema.	Capacidad de abstracción, análisis y síntesis.	12 horas teoría – 8 horas práctica – 20 horas totales
2.2 Características de un estimador.	El alumno investigará, analizará y elaborará un esquema acerca de las características de un estimador.	El docente solicitará al estudiante la investigación de las características de un estimador y mediante una mesa de diálogo se discutirán los resultados. Pedirá la elaboración de un esquema del tema.	Capacidad de abstracción, análisis y síntesis.	
2.3 Estimación puntual.	El alumno analizará la estimación puntual, ventajas y desventajas y elaborará u esquema.	El docente abordará la estimación puntual, sus ventajas y desventajas y solicitará la elaboración de un esquema del tema.	Capacidad de abstracción, análisis y síntesis.	
2.4 Estimación por intervalos.	El alumno analizará la estimación por intervalos, ventajas y desventajas y elaborará u esquema.	El docente abordará la estimación por intervalos, sus ventajas y desventajas y solicitará la elaboración de un	Capacidad de abstracción, análisis y síntesis.	

Rev. 1

		esquema del tema.		
2.4.1 Intervalo de confianza para la media.	El alumno analizará el intervalo de confianza para la media y resolverá ejercicios del tema.	El docente explicará el intervalo de confianza para la media. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
2.4.2 Intervalo de confianza para la diferencia de medias.	El alumno analizará el intervalo de confianza para la diferencia de medias y resolverá ejercicios del tema.	El docente explicará el intervalo de confianza para la diferencia de medias. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
2.4.3 Intervalos de confianza para la proporción.	El alumno analizará los intervalos de confianza para la proporción y resolverá ejercicios del tema.	El docente explicará los intervalos de confianza para la proporción. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
2.4.4 Intervalos de confianza para la diferencia de proporciones.	El alumno analizará los intervalos de confianza para la diferencia de proporciones y resolverá ejercicios del tema.	El docente explicará los intervalos de confianza para la diferencia de proporciones. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica.	

Rev. 1

	tema.	Daré solución y propondré ejercicios del tema.	conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
2.4.5 Intervalos de confianza para la varianza.	El alumno analizará los intervalos de confianza para la varianza y resolverá ejercicios del tema.	El docente explicará los intervalos de confianza para la varianza. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
2.4.6 Intervalos de confianza para la razón de varianzas.	El alumno analizará los intervalos de confianza para la razón de varianzas y resolverá ejercicios del tema.	El docente explicará los intervalos de confianza para la razón de varianzas. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
2.5 Determinación del tamaño de muestra.	El alumno analizará la determinación del tamaño de una muestra y su importancia.	El docente abordará la determinación del tamaño de una muestra y su importancia.	Capacidad de abstracción, análisis y síntesis.	
2.5.1 Basado en la media de la Población.	El alumno analizará la determinación del tamaño de una muestra basada en la media de la población y resolverá ejercicios del tema.	El docente explicará la determinación del tamaño de una muestra basada en la media de la población. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica.	

		del tema.	práctica. Capacidad para identificar, plantear y resolver problemas.	
2.5.2 Basado en la proporción de la Población.	El alumno analizará la determinación del tamaño de una muestra basada en la proporción de la población y resolverá ejercicios del tema.	El docente explicará la determinación del tamaño de una muestra basada en la proporción de la población. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
2.5.3 Basado en la diferencia entre las medias de la población	El alumno analizará la determinación del tamaño de una muestra basada en la diferencia entre las medias de la población y resolverá ejercicios del tema.	El docente explicará la determinación del tamaño de una muestra basada en la diferencia entre las medias de la población. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	

Estrategias de evaluación (Matriz)

Evidencia de aprendizaje	%	Indicador de alcance						Método de Evaluación			
		A	B	C	D	E	F	Instrumento	P	C	A
Tareas selladas	30	X	X					Guía de observación	x		x

Rev. 1

Examen escrito	70	X			x		x	cuestionario		X	
----------------	----	---	--	--	---	--	---	--------------	--	---	--

Niveles de desempeño:

Desempeño	Nivel de desempeño	Indicadores de alcance	Valoración numérica
Competencia alcanzada	Excelente	Cumple al menos cinco de los siguientes indicadores a) Se adapta a situaciones y contextos complejos. Puede trabajar en equipo, reflejar sus conocimientos en la interpretación de la realidad. Inferir comportamientos o consecuencias de los fenómenos o problemas en estudio. Incluir más variables en dichos casos de estudio. b) Hace aportaciones a las actividades académicas desarrolladas. Preguntar integrando conocimientos de otras asignaturas o de casos anteriores de la misma asignatura. Presenta otros puntos de vista que complementan al presentado en la clase. Presenta fuentes de información adicionales (Internet, documentales), usa más bibliografía, consulta fuentes en un segundo idioma, etc. c) Propone y/o explica soluciones o procedimientos no vistos en clase (creatividad). Ante problemas o casos de estudio propone perspectivas diferentes, para abordarlos y sustentarlos correctamente. Aplica procedimientos aprendidos en otra asignatura o contexto para el problema que se está resolviendo. d) Introduce recursos y experiencias que promueven un pensamiento crítico; (por ejemplo el uso de las tecnologías de la información estableciendo previamente un criterio). Ante temas de una asignatura, introduce cuestionamientos de tipo ético, ecológico, histórico, político, económico, etc.; que deben tomarse en cuenta para comprender mejor, o a futuro dicho tema. Se apoya en foros, autores, bibliografía, documentales, etc. para sustentar su punto de vista. e) Incorpora conocimientos y actividades interdisciplinarias en su aprendizaje. En el desarrollo de los temas de la asignatura, incorpora conocimientos y actividades desarrollados en otras asignaturas para lograr la competencia. f) Realiza su trabajo de manera autónoma y autorregulada. Es capaz de organizar su tiempo y trabajar sin necesidad de una supervisión estrecha y/o coercitiva. Aprovecha la planeación de la asignatura presentada por el (la) profesor(a) (instrumentación didáctica) para presentar propuestas de mejora de la temática vista durante el curso. Realiza actividades de investigación para participar activamente durante el curso.	95-100
	Notable	Cumple cuatro de los indicadores diferidos en desempeño excelente	85-94
	Bueno	Cumple tres de los indicadores diferidos en desempeño excelente	75-84

	Suficiente	Cumple dos de los indicadores diferidos en desempeño excelente	70-74
Competencia no alcanzada	Insuficiente	No se cumple con el 100% de evidencias conceptuales, procedimentales y actitudinales de los indicadores diferidos en el desempeño excelente.	NA (no alcanzada)

CompetenciaNo.: 3

Descripción:

Pruebas de hipótesis. Realiza pruebas de hipótesis para comparar si los valores de los estadísticos obtenidos de una muestra tienen una diferencia significativa con un valor supuesto asumiendo cierto nivel de confianza y tomando en cuenta los criterios de aceptación o rechazo en problemas de la industria la logística que involucren errores tipo I o tipo II

Temas y subtemas para desarrollar	Actividades de aprendizaje	Actividades de Enseñanza	Desarrollo de competencias genéricas	Horas teórico-práctica
3.1 Introducción.	El alumno investigará, analizará y elaborará un esquema acerca de los conceptos básicos de prueba de hipótesis.	El docente solicitará al estudiante la investigación de los conceptos relacionados con pruebas de hipótesis y mediante una mesa de diálogo se discutirán los resultados. Pedirá la elaboración de un esquema del tema.	Capacidad de abstracción, análisis y síntesis.	12 horas teoría – 8 horas práctica – 20 horas totales
3.2 Confiabilidad y significancia.	El alumno investigará, analizará y elaborará un esquema acerca de confiabilidad y significancia.	El docente solicitará al estudiante la investigación de los conceptos de confiabilidad y significancia y mediante una mesa de diálogo se discutirán los resultados. Pedirá la elaboración de un esquema del tema.	Capacidad de abstracción, análisis y síntesis.	

Rev. 1

3.3 Errores tipo I y tipo II.	El alumno analizará la metodología de prueba de hipótesis para errores de tipo I y tipo II. Formulará y resolverá ejercicios del tema.	El docente abordará la metodología de prueba de hipótesis para la variable que se esté midiendo y en donde se obtenga el tamaño de muestra para errores tipo I y tipo II. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
3.4 Potencia de la prueba.	El alumno analizará la metodología de potencia de la prueba. Formulará y resolverá ejercicios del tema.	El docente abordará la metodología de potencia de la prueba. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
3.5 Formulación de Hipótesis estadísticas.	El alumno analizará la metodología de para la formulación de hipótesis estadísticas. Formulará y resolverá ejercicios del tema.	El docente abordará la metodología para la formulación de hipótesis estadísticas. Daré solución y propondré ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
3.6 Prueba de hipótesis para la media.	El alumno simulará casos en los cuales se generen hipótesis cuyo interés sea la media de acuerdo a lo siguiente ✓ Generación de datos. ✓ Prueba de hipótesis.	El docente desarrollará ejercicios de casos en los cuales se generen hipótesis cuyo interés sea la media y solicitará a los estudiantes la simulación de casos.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la	

Rev. 1

3.10 Prueba de hipótesis para la varianza.	El alumno simulará casos en los cuales se generen hipótesis cuyo interés sea la varianza de acuerdo a lo siguiente ✓ Generación de datos. ✓ Prueba de hipótesis. ✓ Obtención de conclusiones. ✓ Cambio en el tamaño de la muestra y demostración de su impacto.	El docente desarrollará ejercicios de casos en los cuales se generen hipótesis cuyo interés sea la varianza y solicitará a los estudiantes la simulación de casos.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
3.11 Prueba de hipótesis para la razón de varianzas.	El alumno simulará casos en los cuales se generen hipótesis cuyo interés sea la razón de varianzas de acuerdo a lo siguiente ✓ Generación de datos. ✓ Prueba de hipótesis. ✓ Obtención de conclusiones. ✓ Cambio en el tamaño de la muestra y demostración de su impacto.	El docente desarrollará ejercicios de casos en los cuales se generen hipótesis cuyo interés sea la razón de varianzas y solicitará a los estudiantes la simulación de casos.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
3.12 Uso de software estadístico minitab .	El alumno realizará una práctica de laboratorio en Minitab en donde se involucren las pruebas de hipótesis	El docente propondrá una práctica de laboratorio en Minitab en donde se involucren las pruebas de hipótesis.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	

Estrategias de evaluación (Matriz)

Evidencia de aprendizaje	%	Indicador de alcance	Método de Evaluación
--------------------------	---	----------------------	----------------------

Rev. 1

3.10 Prueba de hipótesis para la varianza.	El alumno simulará casos en los cuales se generen hipótesis cuyo interés sea la varianza de acuerdo a lo siguiente ✓ Generación de datos. ✓ Prueba de hipótesis. ✓ Obtención de conclusiones. ✓ Cambio en el tamaño de la muestra y demostración de su impacto.	El docente desarrollará ejercicios de casos en los cuales se generen hipótesis cuyo interés sea la varianza y solicitará a los estudiantes la simulación de casos.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
3.11 Prueba de hipótesis para la razón de varianzas.	El alumno simulará casos en los cuales se generen hipótesis cuyo interés sea la razón de varianzas de acuerdo a lo siguiente ✓ Generación de datos. ✓ Prueba de hipótesis. ✓ Obtención de conclusiones. ✓ Cambio en el tamaño de la muestra y demostración de su impacto.	El docente desarrollará ejercicios de casos en los cuales se generen hipótesis cuyo interés sea la razón de varianzas y solicitará a los estudiantes la simulación de casos.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
3.12 Uso de software estadístico minitab .	El alumno realizará una práctica de laboratorio en Minitab en donde se involucren las pruebas de hipótesis	El docente propondrá una práctica de laboratorio en Minitab en donde se involucren las pruebas de hipótesis.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	

Estrategias de evaluación (Matriz)

Evidencia de aprendizaje	%	Indicador de alcance	Método de Evaluación
--------------------------	---	----------------------	----------------------

Rev. 1

		A	B	C	D	E	F	Instrumento	P	C	A
Tareas selladas	30	X	X					Guía de observación	x		x
Examen escrito	70	X			x		x	cuestionario		X	

Niveles de desempeño:

Desempeño	Nivel de desempeño	Indicadores de alcance	Valoración numérica
Competencia alcanzada	Excelente	Cumple al menos cinco de los siguientes indicadores a) Se adapta a situaciones y contextos complejos. Puede trabajar en equipo, reflejar sus conocimientos en la interpretación de la realidad. Inferir comportamientos o consecuencias de los fenómenos o problemas en estudio. Incluir más variables en dichos casos de estudio. b) Hace aportaciones a las actividades académicas desarrolladas. Pregunta integrando conocimientos de otras asignaturas o de casos anteriores de la misma asignatura. Presenta otros puntos de vista que complementan al presentado en la clase. Presenta fuentes de información adicionales (Internet, documentales), usa más bibliografía, consulta fuentes en un segundo idioma, etc. c) Propone y/o explica soluciones o procedimientos no vistos en clase (creatividad). Ante problemas o casos de estudio propone perspectivas diferentes, para abordarlos y sustentarlos correctamente. Aplica procedimientos aprendidos en otra asignatura o contexto para el problema que se está resolviendo. d) Introduce recursos y experiencias que promueven un pensamiento crítico; (por ejemplo el uso de las tecnologías de la información estableciendo previamente un criterio). Ante temas de una asignatura, introduce cuestionamientos de tipo ético, ecológico, histórico, político, económico, etc.; que deben tomarse en cuenta para comprender mejor, o a futuro dicho tema. Se apoya en foros, autores, bibliografía, documentales, etc. para sustentar su punto de vista. e) Incorpora conocimientos y actividades interdisciplinarias en su aprendizaje. En el desarrollo de los temas de la asignatura, incorpora conocimientos y actividades desarrollados en otras asignaturas para lograr la competencia. f) Realiza su trabajo de manera autónoma y autorregulada. Es capaz de organizar su tiempo y trabajar sin necesidad de una supervisión estrecha y/o coercitiva. Aprovecha la planeación de la asignatura presentada por el (la) profesor(a) (instrumentación didáctica) para presentar propuestas de mejora de la temática vista durante el curso. Realiza actividades de	95-100

Rev. 1

		investigación para participar activamente durante el curso.	
	Notable	Cumple cuatro de los indicadores diferidos en desempeño excelente	85-94
	Bueno	Cumple tres de los indicadores diferidos en desempeño excelente	75-84
	Suficiente	Cumple dos de los indicadores diferidos en desempeño excelente	70-74
Competencia no alcanzada	Insuficiente	No se cumple con el 100% de evidencias conceptuales, procedimentales y actitudinales de los indicadores diferidos en el desempeño excelente	NA (no alcanzada)

CompetenciaNo.: 4

Descripción:

Pruebas de bondad de ajuste y pruebas no paramétricas: Realiza pruebas de bondad de ajuste y no paramétricas para determinar si el comportamiento de un experimento se adecua a una distribución determinada en procesos de la industria y la logística.

Temas y subtemas para desarrollar	Actividades de aprendizaje	Actividades de Enseñanza	Desarrollo de competencias genéricas	Horas teórico-práctica
4.1 Bondad de ajuste.	El alumno investigará y elaborará un esquema sobre la prueba de bondad de ajuste.	El docente solicitará al alumno la investigación y elaboración de un esquema la prueba de bondad de ajuste.	Capacidad de abstracción, análisis y síntesis.	6 horas teoría - 4 horas práctica - 10 horas totales
4.1.1 Pruebas de independencia y tablas de contingencia.	El alumno analizará y resolverá ejercicios de pruebas de independencia y tablas de contingencia.	El docente abordará la metodología para las pruebas de independencia y tablas de contingencia. Propondrá ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica.	

Rev. 1

			Capacidad para identificar, plantear y resolver problemas.	
4.1.2 Prueba de bondad de ajuste con ji cuadrada.	El alumno analizará y resolverá ejercicios de prueba de bondad de ajuste con Ji Cuadrada.	El docente abordará la metodología para prueba de bondad de ajuste con Ji Cuadrada. Propondrá ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
4.1.3 Uso de software estadístico.	El alumno analizará y resolverá ejercicios de prueba de bondad de ajuste mediante un software estadístico.	El docente abordará la metodología para la prueba de bondad de ajuste mediante un software estadístico. Propondrá ejercicios del tema.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
4.2 Pruebas no paramétricas.	El alumno investigará y elaborará un esquema sobre la prueba no paramétrica.	El docente solicitará al alumno la investigación y elaboración de un esquema la prueba no paramétrica.	Capacidad de abstracción, análisis y síntesis.	
4.2.1 Escala de medición.	El alumno investigará y elaborará un esquema sobre las escalas de medición para pruebas no paramétricas.	El docente solicitará al alumno la investigación y elaboración de un esquema sobre las escalas de medición para pruebas no paramétricas.	Capacidad de abstracción, análisis y síntesis.	
4.2.2 Métodos estadísticos contra no paramétricos.	El alumno investigará y elaborará un esquema sobre los métodos estadísticos contra no paramétricos.	El docente solicitará al alumno la investigación y elaboración de un esquema sobre los métodos estadísticos contra no paramétricos.	Capacidad de abstracción, análisis y síntesis.	
4.2.3 Prueba de	El alumno resolverá ejercicios aplicando la Prueba de Kolmogorov-Smirnov.	El docente desarrollará y propondrá	Capacidad de abstracción,	

Rev. 1

Kolmogorov – Smirnov.		ejercicios aplicando la Prueba de Kolmogorov-Smirnov.	análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
4.2.4 Prueba de Anderson – Darling.	El alumno resolverá ejercicios aplicando la Prueba de Anderson-Darling.	El docente desarrollará y propondrá ejercicios aplicando la Prueba de Anderson-Darling.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
4.2.5 Prueba de Ryan – Joiner.	El alumno resolverá ejercicios aplicando la Prueba de Ryan-Joiner.	El docente desarrollará y propondrá ejercicios aplicando la Prueba de Ryan-Joiner.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
4.2.6 Prueba de Shappiro – Wilk.	El alumno resolverá ejercicios aplicando la Prueba de Shapiro-Wilk.	El docente desarrollará y propondrá ejercicios aplicando la Prueba de Shapiro-Wilk.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	

Rev. 1

4.2.7 Uso de software estadístico	El alumno resolverá ejercicios de pruebas no paramétricas aplicando un software estadístico.	El docente desarrollará y propondrá ejercicios de pruebas no paramétricas mediante el uso de un software estadístico.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
-----------------------------------	--	---	--	--

Estrategias de evaluación (Matriz)

Evidencia de aprendizaje	%	Indicador de alcance						Método de Evaluación				
		A	B	C	D	E	F	Instrumento		P	C	A
Tareas selladas	30	X	X					Guía de observación		x		x
Examen escrito	70	X			x		x	cuestionario			X	

Niveles de desempeño:

Desempeño	Nivel de desempeño	Indicadores de alcance	Valoración numérica
Competencia alcanzada	Excelente	Cumple al menos cinco de los siguientes indicadores a) Se adapta a situaciones y contextos complejos. Puede trabajar en equipo, reflejar sus conocimientos en la interpretación de la realidad. Inferir comportamientos o consecuencias de los fenómenos o problemas en estudio. Incluir más variables en dichos casos de estudio. b) Hace aportaciones a las actividades académicas desarrolladas. Preguntas integrando conocimientos de otras asignaturas o de casos anteriores de la misma asignatura. Presenta otros puntos de vista que complementan al presentado en la clase. Presenta fuentes de información adicionales (Internet, documentales), usa más bibliografía, consulta fuentes en un segundo idioma, etc. c) Propone y/o explica soluciones o procedimientos no vistos en clase (creatividad). Ante problemas o casos de estudio propone perspectivas diferentes, para abordarlos y	

Rev. 1

		sustentarlos correctamente. Aplica procedimientos aprendidos en otra asignatura o contexto para el problema que se está resolviendo. d) Introduce recursos y experiencias que promueven un pensamiento crítico; (por ejemplo el uso de las tecnologías de la información estableciendo previamente un criterio). Ante temas de una asignatura, introduce cuestionamientos de tipo ético, ecológico, histórico, político, económico, etc.; que deben tomarse en cuenta para comprender mejor, o a futuro dicho tema. Se apoya en foros, autores, bibliografía, documentales, etc. para sustentar su punto de vista. e) Incorpora conocimientos y actividades interdisciplinarias en su aprendizaje. En el desarrollo de los temas de la asignatura, incorpora conocimientos y actividades desarrollados en otras asignaturas para lograr la competencia. f) Realiza su trabajo de manera autónoma y autorregulada. Es capaz de organizar su tiempo y trabajar sin necesidad de una supervisión estrecha y/o coercitiva. Aprovecha la planeación de la asignatura presentada por el (la) profesor(a) (instrumentación didáctica) para presentar propuestas de mejora de la temática vista durante el curso. Realiza actividades de investigación para participar activamente durante el curso.	95-100
	Notable	Cumple cuatro de los indicadores diferidos en desempeño excelente	85-94
	Bueno	Cumple tres de los indicadores diferidos en desempeño excelente	75-84
	Suficiente	Cumple dos de los indicadores diferidos en desempeño excelente	70-74
Competencia no alcanzada	Insuficiente	No se cumple con el 100% de evidencias conceptuales, procedimentales y actitudinales de los indicadores diferidos en el desempeño excelente.	NA (no alcanzada)

CompetenciaNo.: 5

Descripción:

Regresión Lineal Simple: Utiliza el diagrama de dispersión de datos bivariados de un experimento para hacer una estimación en procesos de la industria y la logística aplicando los conceptos de regresión lineal simple.

Temas y subtemas para desarrollar	Actividades de aprendizaje	Actividades de Enseñanza	Desarrollo de competencias genéricas	Horas teórico-práctica
5.1 Introducción. Conceptos básicos, análisis de regresión y correlación.	El alumno analizará y elaborará un esquema acerca de los conceptos relacionados con la regresión lineal simple.	El docente explicará los conceptos relacionados con la regresión lineal simple y solicitará al alumno la elaboración de un esquema.	Capacidad de abstracción, análisis y síntesis.	8 horas teoría – 4 horas práctica – 10 horas totales

Rev. 1

5.2 Supuestos en la Regresión Lineal Simple	El alumno analizará y elaborará un esquema de los supuestos en la regresión lineal simple.	El docente abordará los supuestos en la regresión lineal simple y solicitará al estudiante la elaboración de un esquema.	Capacidad de abstracción, análisis y síntesis.	
5.3 Diagrama de Dispersión	El alumno analizará y elaborará un esquema de la importancia y la interpretación de los diagramas de dispersión.	El docente explicará la importancia y la interpretación de los diagramas de dispersión y solicitará al estudiante la elaboración de un esquema.	Capacidad de abstracción, análisis y síntesis.	
5.4 Modelo de Regresión lineal simple	El alumno analizará el modelo de regresión lineal simple (método de mínimos cuadrados).	El docente solicitará al estudiante una investigación sobre el modelo de regresión lineal simple (método de mínimos cuadrados) y mediante una mesa de diálogo se llegará a conclusiones.	Capacidad de abstracción, análisis y síntesis.	
5.4.1 Método de mínimos cuadrados	El alumno resolverá ejercicios de método de mínimos cuadrados para la regresión lineal simple.	El docente desarrollará y propondrá ejercicios en donde intervenga el método de mínimos cuadrados para la regresión lineal simple.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
5.5 Estimación de la recta de regresión	El alumno resolverá ejercicios de estimación de la recta de regresión.	El docente desarrollará y propondrá ejercicios de estimación de la recta	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los	

Rev. 1

		de regresión.	conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
5.6 Coeficientes de determinación y correlación	El alumno resolverá ejercicios de coeficientes de determinación y correlación.	El docente desarrollará y propondrá ejercicios de coeficientes de determinación y correlación.	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	
5.7 Pruebas de hipótesis en regresión lineal simple (significación de la regresión)	El alumno resolverá ejercicios de pruebas de hipótesis en regresión lineal simple (significación de la regresión).	El docente desarrollará y propondrá ejercicios de pruebas de hipótesis en regresión lineal simple (significación de la regresión).	Capacidad de abstracción, análisis y síntesis. Capacidad de aplicar los conocimientos en la práctica. Capacidad para identificar, plantear y resolver problemas.	

Estrategias de evaluación (Matriz)

Evidencia de aprendizaje	%	Indicador de alcance						Método de Evaluación			
		A	B	C	D	E	F	Instrumento	P	C	A
Tareas selladas	30	X	X					Guía de observación	x		x
Examen escrito	70	X			x		x	cuestionario		X	

Rev. 1

Niveles de desempeño:

Desempeño	Nivel de desempeño	Indicadores de alcance	Valoración numérica
Competencia alcanzada	Excelente	Cumple al menos cinco de los siguientes indicadores a) Se adapta a situaciones y contextos complejos. Puede trabajar en equipo, reflejar sus conocimientos en la interpretación de la realidad. Inferir comportamientos o consecuencias de los fenómenos o problemas en estudio. Incluir más variables en dichos casos de estudio. b) Hace aportaciones a las actividades académicas desarrolladas. Pregunta integrando conocimientos de otras asignaturas o de casos anteriores de la misma asignatura. Presenta otros puntos de vista que complementan al presentado en la clase. Presenta fuentes de información adicionales (Internet, documentales), usa más bibliografía, consulta fuentes en un segundo idioma, etc. c) Propone y/o explica soluciones o procedimientos no vistos en clase (creatividad). Ante problemas o casos de estudio propone perspectivas diferentes, para abordarlos y sustentarlos correctamente. Aplica procedimientos aprendidos en otra asignatura o contexto para el problema que se está resolviendo. d) Introduce recursos y experiencias que promueven un pensamiento crítico; (por ejemplo el uso de las tecnologías de la información estableciendo previamente un criterio). Ante temas de una asignatura, introduce cuestionamientos de tipo ético, ecológico, histórico, político, económico, etc.; que deben tomarse en cuenta para comprender mejor, o a futuro dicho tema. Se apoya en foros, autores, bibliografía, documentales, etc. para sustentar su punto de vista. e) Incorpora conocimientos y actividades interdisciplinarias en su aprendizaje. En el desarrollo de los temas de la asignatura, incorpora conocimientos y actividades desarrollados en otras asignaturas para lograr la competencia. f) Realiza su trabajo de manera autónoma y autorregulada. Es capaz de organizar su tiempo y trabajar sin necesidad de una supervisión estrecha y/o coercitiva. Aprovecha la planeación de la asignatura presentada por el (la) profesor(a) (instrumentación didáctica) para presentar propuestas de mejora de la temática vista durante el curso. Realiza actividades de investigación para participar activamente durante el curso.	95-100
	Notable	Cumple cuatro de los indicadores diferidos en desempeño excelente	85-94
	Bueno	Cumple tres de los indicadores diferidos en desempeño excelente	75-84
	Suficiente	Cumple dos de los indicadores diferidos en desempeño excelente	70-74
Competencia	Insuficiente	No se cumple con el 100% de evidencias conceptuales, procedimentales y	NA (no

Rev. 1

no alcanzada	actitudinales de los indicadores diferidos en el desempeño excelente.	alcanzada)
--------------	---	------------

5. Fuentes de información y apoyos didácticos

Fuentes de Información	Apoyos didácticos
- Anderson, D. R. (2008). Estadística para administración y economía. (10ª. ed.) México : Cengage Learning. - Box, G. E. P. (2008). Estadística para investigadores : Diseño, innovación y descubrimiento. (2ª. Ed.). España : Reverté - Berenson, M. (2006). Estadística para administración. (4ª. ed.) México : Pearson Educación. - Carot, V. (2006). Control estadístico de la calidad. España : Alfaomega. - Devore, J. L. (2012). Probabilidad y estadística para ingenierías y ciencia. (8ª. ed.) México : Cengage Learning. - Fernández, A. M. (2006). Ejercicios de econometría. (2006). España : McGraw-Hill. - Gamiz, B. E. (2012). Probabilidad y estadística con prácticas en Excel. (3ª. ed.) México : JIT Press. - Gujarati, D. (2010). Econometría. (5ª. Ed.). México : McGraw-Hill. - Gutiérrez, P. H. (2012). Análisis y diseño de experimentos. (3ª. ed.) México : McGraw-Hill - Gutiérrez, P. H. (2009). Control estadístico de calidad y seis sigma. (2ª. ed) México : McGrawHill.	Pizarrón, marcador, computadora, canon, cinta métrica. Diversos objetos a medir, salón de clase

11. Hines, W. (2009) Probabilidad y estadística para ingeniería (4ª. ed.) México : CECSA : Grupo Editorial Patria.
12. Johnson, R. A. (2012) Probabilidad y estadística para ingenieros. (8ª. ed.) México : Pearson Educación.
13. Kazmier, L. (2006). Estadística aplicada a administración y economía. (4ª. ed.) México : McGraw-Hill.
14. Larson, H. J. (1992). Introducción a la teoría de probabilidades e inferencia estadística. México : Limusa.
15. Levine, D. M. (2010) Estadística para administración y economía. (7ª. ed.) México : Pearson Educación.
16. Mendenhall, W. (2010). Introducción a la Probabilidad y Estadística. (13ª. ed.) México: Cengage Learning.
17. Montgomery, D. C. (2011). Probabilidad y estadística aplicadas a la ingeniería. (2ª. ed.) México : Limusa : Wiley.
18. Quezada, L. (2010). Estadística para ingenieros. México : Empresa Editora Macro.
19. Rodríguez, F. J. (2008). Estadística para administración. México : Grupo Editorial Patria.
20. Ross, S. M. (2002). Probabilidad y estadística para ingenieros. México : McGraw-Hill.
21. Salvatore, D. (2004). Estadística y econometría. (2ª. Ed.). España : McGraw-Hill.
22. Spiegel, M. (2010). Fórmulas y tablas de matemática aplicada. (3ª. ed.) México : McGraw-Hill

Rev. 1

23. Spiegel, M. (2010). Teoría y problemas de Probabilidad y estadística. (3ª. Ed.) México : McGraw-Hill.
24. Wackerly, D. D. (2010). Estadística matemática con aplicaciones. (7ª. ed.) México : Cengage Learning.
25. Walpole, R. E. (2012). Probabilidad y estadística para ingeniería y ciencias. (9ª. ed) México : Pearson Educación.

6. Calendarización de evaluación en semanas

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
TP				EF1				EF2				EF3		EF4		EF5
TR				ES1				ES2				ES3		ES4		ES5
SD				RI				SD1					SD2			RF

TP=tiempo planeado
ED=evaluación diagnóstica

TR=tiemporeal
EFn=evaluaciónformativa(competenciaespecífica)

SD=seguimientodepartamental
ES=evaluaciónsumativa

Fecha de elaboración: AGOSTO 2022

MA. CONCEPCION FUENTES MORALES
Nombre del profesor(a)

LIC. SILVANA CERA
Nombre del Jefe(a) de Departamento Académico

Rev. 1