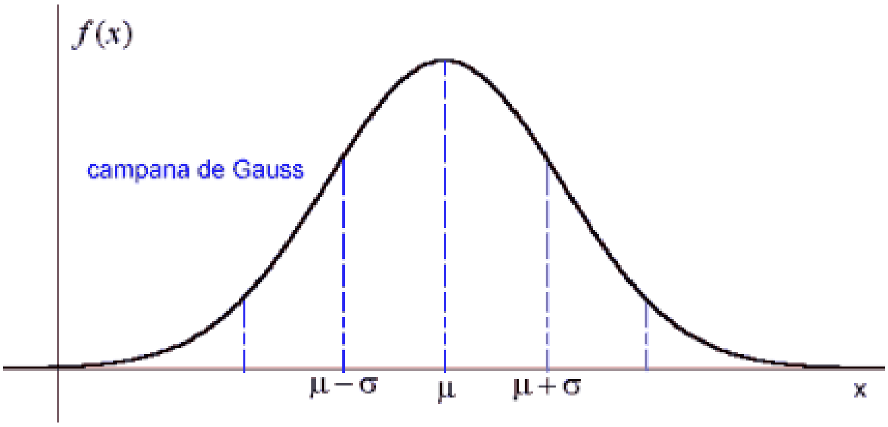
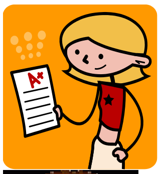
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
COLEGIO DE CIENCIAS Y HUMANIDADES
PLANTEL NAUCALPAN

GUÍA DE ESTUDIO DE ESTADÍSTICA Y PROBABILIDAD II





ELABORARON PROFESORAS:

- BEATRIZ ARELLANO SÁNCHEZ
- GUADALUPE RANGEL LÓPEZ

Febrero, 2010

INTRODUCCIÓN

Esta guía se ha realizado con la intención de apoyar a los alumnos que adeudan la materia de Estadística y Probabilidad II.

Se presentan en forma de contenidos temáticos o temario, las tres unidades del curso, así como los temas y subtemas de él.

Se pretende dar un panorama general de la materia, dando los elementos teóricos más importantes señalados en el programa de estudios vigente, seguidos de ejemplos resueltos para que el alumno tenga una idea más amplia de cómo aplicar los conceptos o las fórmulas, para resolver los ejercicios propuestos.

Los ejercicios propuestos pretenden contemplar, diferentes formas de planteamiento, así como diferentes formas de procedimiento para llegar a la misma solución.

Se dan las soluciones a algunos de los problemas planteados, sobre todo de los que tienen una solución numérica ya que los que contemplan un procedimiento por pasos, como es el caso de las pruebas de hipótesis, se darán sólo algunos valores que puedan orientar al alumno para llegar a la conclusión correcta.

Como todo trabajo, la presente guía es perfectible, por lo que de antemano agradecemos sus comentarios y aportaciones para mejorarla.

CONTENIDO

UNIDAD I: VARIABLE ALEATORIA.

1.1 Variable Aleatoria

1.1.1 Definición Variable Aleatoria

Si los valores que toma un símbolo tal como X están asociados con los eventos aleatorios simples de un experimento dado, y por lo tanto, dependen de ocurrencias aleatorias, al símbolo se le denomina variable aleatoria, y esta puede ser:

$$X: \quad Variable \quad aleatoria \quad \begin{cases} discreta \\ continua \end{cases}$$

1.1.2 Ejemplo

Sea el experimento aleatorio de lanzar 3 veces una moneda y observar el resultado.

a) Escribe el espacio muestral del experimento
b) ¿Cuántos posibles resultados tiene el experimento?
c) ¿Cómo se define la variable aleatoria?
d) Construye la tabla de distribución de probabilidad.

SOLUCIÓN

Ω = { aaa, aas, asa, saa, ssa, sas, ass, sss}

n (Ω) = 8

X: Número de águilas observadas en los tres lanzamientos

Tabla de distribución de probabilidad de X

X	P (X = x)
0	1/8
1	3/8
2	3/8
3	1/8
Σ	1

1.1.3 Ejercicios de Variable Aleatoria

1. Se lanzan al aire dos dados. Sea “X”: La suma de las caras de los dados.
   - Escribe el espacio muestral del experimento.
   - ¿Qué valores toma la variable “X”?
   - Construye la tabla de distribución de probabilidad de “X”.
2. Sea “X” el número de águilas obtenidas al lanzar cuatro monedas balanceadas, construye la tabla de distribución de probabilidad de “X”.
3. Cuatro esferas marcadas 2, 4, 6 y 8 se colocan en una caja y se mezclan antes de extraer una. Sea “X” la variable aleatoria que indica el número de la esfera que se extrae con reemplazo. Obtenga la distribución probabilística de “X”.

1.2 Esperanza Matemática o Valor Esperado de una variable aleatoria discreta.

1.2.1 Definiciones

Valor esperado o esperanza matemática de X, también se le conoce como el promedio a largo plazo y esta dado por:

$$E(X) = \sum_{i=1}^{m} x_i P(x_i)$$

La media de la distribución es igual al valor esperado.

$$\mu_x = E(X)$$

La varianza o variancia de X es:

$$\sigma_x^2 = E(X-\mu)^2 \qquad \text{ó} \qquad \sigma_x^2 = E(X^2) - \mu^2$$

1.2.2 Ejemplos

1. Considere un juego asociado al lanzamiento de una moneda, si cae águila usted gana $5.00, si la moneda cae sol usted pierde $5.00. ¿Cuál es el valor esperado del juego?

Solución: La tabla de distribución de probabilidad para X está dada por:

X	P(X)
$5	0.5
-$5	0.5
Σ	1.0

$$E(X) = \$5(0.5) - \$5(0.5) = 0$$

NOTA. Cuando $E(X) = 0$ se dice que el juego es justo.

2. Un corredor de bolsa, ofrece a un cliente un plan de inversión que podría dar como resultado uno de los siguientes beneficios, con las probabilidades indicadas.
3.

Beneficio	Probabilidad
$1 millón	0.2
$2 millones	0.3
$3 millones	0.2
$4 millones	0.2
$5 millones	0.1
Σ	1.0

Solución:

X	P(X)	XP(X)	X^2	$X^2P(X)$
1	0.2	0.2	1	0.2
2	0.3	0.6	4	1.2
3	0.2	0.6	9	1.8
4	0.2	0.8	16	3.2
5	0.1	0.5	25	2.5
∑	1.0	2.7	–	8.9

$$E(X) = \mu = 2.7$$
$$\sigma^2 = E(X^2) - \mu^2$$
$$\sigma^2 = 8.9 - (2.7)^2$$
$$\sigma^2 = 8.9 - 7.29$$
$$\sigma^2 = 1.61$$

1.2.3 Ejercicios

1. Un billete de lotería tiene 0.00001 de probabilidad de ganar $100,000; 0.0002 de ganar $50,000 y 0.004 de ganar $25,000. ¿Cuál sería el precio justo que se debería pagar por el billete.
2. Si la distribución de probabilidad de los C.I. de todos los graduados de bachillerato es la siguiente. ¿Cuál es la media y la desviación estándar de sus C.I.?

Coeficiente de Inteligencia (C.I.)	Probabilidad
42.5- 57.5	0.01
57.5 – 72.5	0.02
72.5 – 87.5	0.05
87.5 – 102.5	0.40
102.5-117.5	0.30
117.5-132.5	0.10
132.5-147.5	0.08
147.5-162.5	0.04

4. En un estudio sobre la movilidad de los ejecutivos en el área de compras, se encontró que el número de compañías en las que un ejecutivo actualmente empleado ha prestado sus servicios como jefe de compras (Y) tiene aproximadamente la siguiente distribución probabilística.

Y	1	2	3	4	5
P(Y)	0.52	0.22	0.19	0.04	0.03

Encuentre la media y la desviación estándar de Y.

5. Un investigador en educación matemática desea conocer los índices de aprobación o reprobación en Matemáticas I. Donde al posible resultado “reprobado”, lo considera como “fracaso” y tiene una probabilidad de 54.72%. Elaborar la tabla de distribución de frecuencias para esta variable aleatoria; y encontrar la media y la desviación estándar de la variable.
6. Los datos de una escuela revelan que el 60% de los estudiantes aprueban su primer examen de álgebra. Seleccionando una muestra de 20 estudiantes al azar. ¿Cuál es la probabilidad de que hayan:
   - reprobado 6 de ellos?
   - Aprobado por lo menos 60% de los 20?
7. Un vendedor de autos nuevos observa que el 80% de autos vendidos son regresados al departamento de servicio para corregir diversos defectos de fabricación en los primeros 25 días después de su compra. De los 11 autos que se vendieron en los últimos días. ¿cuál es la probabilidad de que:

- 5 regresen en los próximos 25 días para recibir servicio?
- Todos regresen?

1.3 Distribuciones de Probabilidad de una Variable Aleatoria discreta

Función Probabilística: Una función probabilística o función de probabilidad es una regla que asigna una fracción probabilística a cada uno de los diferentes valores de la variable aleatoria.

Distribución Probabilística. Una distribución de probabilidades, cada una de las cuales está asociada con uno de los posibles valores diferentes de la variable aleatoria.

1.3.1 Modelo Bernoulli. El ensayo de Bernoulli, es el modelo probabilístico más simple y es la base de la distribución binomial.

W v.a. Bernoulli W = 0, 1

W v.a. discreta, tiene sólo dos resultados posibles:

Éxito (p) ⟶ 1 Fracaso (q) ⟶ 0, donde q = (1 – p)

Distribución de Probabilidad de W

W	P (X = x)
0	1 – p
1	P
$\sum$	1

El valor esperado y la varianza de W están dadas por: $E(W) = \mu = p \quad y \quad \sigma^2 = pq$

1.3.2 Distribución Binomial

Características de la variable aleatoria binomial (X)

X: Número de éxitos en n observaciones tomadas mediante un proceso Bernoulli

X: Variable Aleatoria Discreta

X= 0, 1, 2, 3, 4,…,n con n ≥ 2

La función probabilística binomial está dada por:

$$P(X = x) = {}_nC_x p^x q^{n-x} \qquad para\ X = 0,1,2,...,n$$

En donde:

${}_nC_x$ - El coeficiente para el evento de X éxitos en n ensayos.

Además ${}_nC_x = \frac{n!}{x!(n-x)!}$

p probabilidad de éxito

q probabilidad de fracaso.

1.3.2 .1 Valor Esperado, Varianza y Desviación Estándar

$\mu_x = E(X) = np$ *Valor esperado de X*

$\sigma_x{}^2 = npq$ *Varianza de X*

$\sigma_x = \sqrt{npq}$ *Desviación estándar de X*

1.3.2.2 Ejemplos

Ejemplo de probabilidades individuales

1. Se sabe que el 70% de todos los pacientes que toman cierta medicina, se recuperan. ¿Cuál es la probabilidad de que en una muestra aleatoria de 30 pacientes que han tomado la medicina, 20 se curen?

SOLUCION

n = 30

x = 20

p = 0.7

q = (1-0.7)=0.3

$$P(X = 20) = {}_{30}C_{20}(0.7)^{20}(0.3)^{10} = 0.14156$$

2. Para calcular la probabilidad binomial es necesario especificar n, que es el número de ensayos; x, que es el número de éxitos p, que es la probabilidad de éxito en cada ensayo. Supóngase que p = 0.80 (por lo tanto, q (fracaso) =0.20), y que se quiere determinar la probabilidad de obtener tres éxitos (y un fracaso) en cuatro observaciones. Cada una de las formas se muestra a continuación, junto con su respectiva probabilidad.

Ordenamiento	Probabilidad
EEEF	(0.8)(0.8)(0.8)(0.2) = 0.1024
EEFE	(0.8)(0.8)(0.2)(0.8) = 0.1024
EFEE	(0.8)(0.2)(0.8)(0.8) = 0.1024
FEEE	(0.2)(0.8)(0.8)(0.8) = 0.1024

	Suma 0.4096

La probabilidad de obtener tres éxitos y un fracaso, es la suma de todas las formas de obtener tres éxitos en cuatro observaciones. En éste caso, la suma es 0.4096. Se debe observar que cada situación tiene la misma probabilidad de ocurrir, ya que los mismos números se multiplican juntos: solamente su orden es diferente. Esto siempre se cumple.

Esta observación conduce a las siguientes pautas. Las probabilidades de obtener resultados binomiales se pueden determinar matemáticamente, tomando en consideración dos cosas: el número de formas en que puede ocurrir la situación y la probabilidad de una de dichas formas. Nótese también que la primera es realmente el número de combinaciones que se pueden distinguir,

$$\binom{n}{x}$$

Quizá el método más sencillo para determinar la probabilidad de una de las situaciones es considerar el caso en el que todos los éxitos ocurren primero y todos los fracasos después. De tener tres éxitos y un fracaso, (es decir, EEEF) y la probabilidad es entonces: (0.8)(0.8)(0.8)(0.2), o bien, $(0.8)^3$ (0.2). De este modo, P(x=3) = $4(0.8)^3 (0.2)$ = 0.496.

Uniendo estas dos ideas – número de formas y probabilidad de una de ellas – se obtiene lo siguiente:

$$P(X=x)=\binom{n}{x}P(\acute{e}xito)^{x}\cdot P(fracaso)^{n-x}$$

En donde

$$P(X=x)=\binom{n}{x}p^{x}q^{n-x}$$

$\binom{n}{x}$ = número de formas de obtener x éxitos y n-x fracasos en n ensayos.

Ejemplo de probabilidades acumuladas: El 35% de los internos de una institución correccional son reincidentes. Se selecciona, para una evaluación, una muestra aleatoria de 15 internos.

a) Hallar la probabilidad de que el número de reincidentes del grupo sea mayor que 10

Solución:

$p = 0.35$ $\quad$ $P(X > 10) = P(X \geq 11)$

$$\begin{aligned} P(X \geq 11) &= P(X=11)+P(X=12)+P(X=13)+P(X=14)+P(X=15) \\ &= {}_{15}C_{11}(0.35)^{11}(0.65)^{4} + {}_{15}C_{12}(0.35)^{12}(0.65)^{3} + \cdots \\ &= 0.00235+0.00042+0.00005+0 \\ &= 0.00282 \end{aligned}$$

En forma particular para n=10 p=0.4 , tenemos la tabla de probabilidad individuales

X	P(X=x)	P(X $\leq$ x)	Representación
0	0.00605	0.00605	P(X $\leq$ 0) = P(X = 0)
1	0.4031	0.4636	P(X $\leq$ 1) = P(X = 0) + P(X = 1)
2	0.12093	0.16729	P(X $\leq$ 2) = P(X = 0) + P(X = 1) + P(X = 2)
3	0.21499	0.38228	P(X $\leq$ 3) = P(X = 0) + P(X = 1) + P(X = 2) +...
4	0.25082	0.63310	P(X $\leq$ 4) = P(X = 0) + P(X = 1) + P(X = 2) +...
5	0.20066	0.83376	P(X $\leq$ 5) = P(X = 0) + P(X = 1) + P(X = 2) +...
6	0.11148	0.94524	P(X $\leq$ 6) = P(X = 0) + P(X = 1) + P(X = 2) +...
7	0.4247	0.98771	P(X $\leq$ 7) = P(X = 0) + P(X = 1) + P(X = 2) +...
8	0.1062	0.99833	P(X $\leq$ 8) = P(X = 0) + P(X = 1) + P(X = 2) +...
9	0.00157	0.9990	P(X $\leq$ 9) = P(X = 0) + P(X = 1) + P(X = 2) +...
10	0.0001	1.0000	P(X $\leq$ 10) = P(X = 0) + P(X = 1) + P(X = 2) +...

P(X > 6) = P(X $\geq$ 7)

= P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)

= 1 - P(X $\leq$ 6)

$= 1 – 0.94524$
$= 0.05476$

$P(X < 5) = P(X \leq 4)$
$= 0.63310$
$= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$P(6 \leq X \leq 9) = P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9)$
$= P(X \leq 9) - P(X \leq 5)$
$= 0.9999 – 0.83376$
$= 0.16614$

1.3.2.3 Ejercicios

1. La probabilidad de un nacimiento masculino es de 0.52. ¿Cuál es la probabilidad de que en una familia que tiene 3 hijos, haya:
   a) 3 varones?
   b) Ningún varón?
   c) Por lo menos un varón?

2. En el 15% de los hogares de un área metropolitana no hay nadie por la noche entre las 7:00 y la media noche. Una persona que está haciendo una encuesta telefónica selecciona al azar 10 hogares de esta zona y los encuesta entre las 7:00 de la noche y la media noche. ¿Cuál es la probabilidad de que la persona que está haciendo la encuesta no obtenga respuesta en:
   a) Todas las 10 llamadas
   b) Exactamente 5 llamadas?
   c) ¿Tres llamadas o menos?
   d) ¿Cuál es la probabilidad de que la persona obtenga una respuesta en cada una de las 10 llamadas?.

3. Un método modelo para enseñar una habilidad especial a ciertos individuos retardados resulta efectivo en el 50% de los casos. Se ensayó un nuevo método con 15 personas. Si el nuevo método no es mejor que el método modelo, ¿cuál es la probabilidad de que 11 o más aprendan la habilidad?

4. Supongamos que el 30% de los estudiantes de una Universidad se oponen a pagar una cuota para actividades estudiantiles. Hallar la probabilidad de que en una muestra aleatoria de 25 estudiantes el número de estudiantes que se opone a la cuota sea:
   a) Exactamente 5
   b) Mayor que 5
   c) Cinco o menos
   d) Un número comprendido entre 6 y 10 inclusive.

5. Suponga que la probabilidad de que al tirar un dado quede hacia arriba un número impar de puntos es de 0.4. ¿Cuál es la probabilidad de que en 5 tiradas del dado, el número de veces que aparezca un número impar de puntos sea:
   a) Menos de 2?
   b) Más de 2?
   c) Entre 2 y 4 inclusive?

6. Una compañía de explotación petrolera observa que en casi el 5% de los pozos de prueba que perfora, encuentra un depósito de gas natural. Si perfora 6 pozos obtenga la probabilidad de que al menos en uno se encuentre gas.

1.3.3 Distribución de Poisson

Esta distribución es útil para describir las probabilidades del número de acontecimientos con respecto a un intervalo continuo. Por ejemplo, los defectos por cm^2, accidentes por día, clientes por hora, llamadas telefónicas por minuto, cabezas de ganado por hectárea. Observe que la unidad de medida es continua, pero la variable aleatoria es discreta.

1.3.3.1 Características

El uso de la distribución de Poisson supone lo siguiente:

- El número de ocurrencias en cualquier intervalo es independiente.
- La probabilidad de una ocurrencia es la misma a través de todo el campo de observación.

- La probabilidad de más de una ocurrencia en cualquier punto único es aproximadamente cero.
- El límite inferior del número de ocurrencias es cero, y el límite superior teóricamente es infinito.

Cuando se dice que una variable aleatoria está distribuida mediante el método de Poisson, se quiere decir que la distribución de frecuencias de esa variable se puede aproximar razonablemente mediante una distribución de Poisson.

La distribución de Poisson se puede describir por medio de un sólo parámetro, la media (µ).

Si una variable aleatoria se describe mediante una distribución de Poisson, la probabilidad de la variable se puede obtener mediante:

P(x) = $P(x)=e^{x}(\mu)^{x}/x!$ $(e^{-\lambda t}(\lambda t)^{x})/x!$

Donde:

x : Número de ocurrencias

e = 2.7182818......

λ : Razón media por unidad

t: Número de unidades λ

Además $\mu=\lambda t$ $\mu=\lambda t$ por lo que:

$P(x)=e^{x}(\mu)^{x}/x!$

1.3.3.2 Ejemplos

1. Mediante un proceso mecánico se producen alfombras de lana que presentan un promedio de dos defectos por metro. Encuentre la probabilidad de que un metro cuadrado tenga exactamente un defecto, suponiendo que el proceso puede ser aproximado mediante una distribución de Poisson.

Se sabe que $\mu=2$, y a partir de la tabla del apéndice, se observa que $e^{-2}=0.135$

$e^{-2} = 0.135$

De tal manera que:

P(x = 1) = $(e^{-1}(1)^3)/3!\,(e^{-2}(2)^t)/1!$

= (0.135)(2)/1

= 0.270

2. Suponga que los barcos arriban a un puerto a razón de $\lambda = 2$ barcos por hora, y que esta proporción esta bien aproximada mediante el proceso de Poisson. Si se observa este proceso durante un periodo de media hora (t=1/2), encuentre la probabilidad de que:

   a) no arribe ningún barco

   b) arriben tres barcos

Solución:

Primeramente, se encuentra $\mu = \lambda t$

μ = 2(1/2)

μ = 1

De la tabla se obtiene e^{-1} = 0.368

a) P(x = 0) = $(e^{-1}(1)^0)/0!$

= 0.368

b) P(x = 3) = $(e^{-1}(1)^3)/3!$

= 0.368 / (3*2*1)

= 0.061

1.3.3.3 Ejercicios

1. El número promedio de radios que vende una compañía por día se aproxima mediante el método de Poisson, con una media de 1.5. Calcule la probabilidad de que la compañía venda por lo menos cuatro radios durante un:

   a) Periodo de dos días

   b) Periodo de tres días

   c) Periodo de cuatro días.

2.Los defectos en un rollo fotográfico de color promedian 0.1 defectos/rollo, y la distribución que se sigue para determinar el número de defectos es la de Poisson. Obtenga la probabilidad de que cualquier rollo fotográfico de color presente uno o más defectos.

3. Los clientes llegan a una exhibición a razón de 6.5 clientes/hora (Poisson). Calcule la probabilidad de que en cualquier hora dada:

a) No llegue ningún cliente
b) Por lo menos lleguen cinco
c) Llegue más de uno
d) Lleguen exactamente 6.5

4. El número de quejas que una oficina de boletos de tren recibe por día es una variable aleatoria que tiene la distribución de Poisson con λ =3,6 ¿Cuál es la probabilidad de que reciba sólo dos quejas un día determinado?.

5. El número de llamadas de emergencia que un servicio de ambulancias recibe por día es una v.a. que se dio mediante el modelo de Poisson con c ¿Cuál es la probabilidad de que cualquier día determinado reciba sólo 4 llamadas de emergencia?

6. El número de solicitudes de préstamo que un banco recibe por día es una v.a. que tiene una distribución de Poisson con λ = 7.5 Encuentre la probabilidad de que cualquier día el banco reciba:

a) Exactamente 6 solicitudes de préstamo
b) Máximo 4 solicitudes de préstamo
c) Mínimo 8 solicitudes de préstamo
d) Entre 5 y 10 solicitudes de préstamo.

1.3.4 Distribución Normal

La distribución normal también es conocida como “Distribución de Gauss”, es continua y es sumamente importante debido a que puede utilizarse como una aproximación para otras distribuciones.

1.3.4.1 Definición: Una variable aleatoria X tiene una **Distribución Normal** si su densidad de probabilidad está dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad para -\infty < x < \infty$$

$$donde \quad \sigma > 0.$$

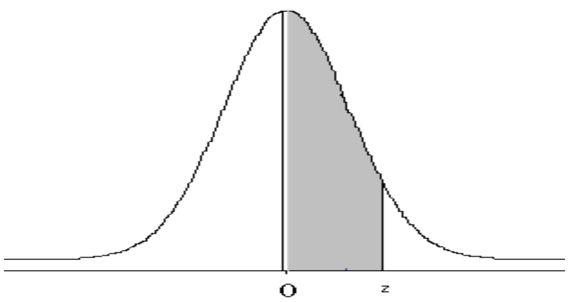
Además; la distribución normal con $\mu = 0 \quad y \quad \sigma = 1$ se conoce como la **Distribución Normal Estándar.**

1.3.4.2 Características.

- La curva normal tiene forma de campana.
- Es simétrica con respecto a la media de la distribución.
- Se extiende de $-\infty \quad a \quad +\infty$.
- Cada distribución normal es completamente especificada por su media (μ) y su desviación estándar (σ); existe una distribución normal diferente para cada combinación de madia y desviación estándar.
- El área total bajo la curva normal es igual a 1 (100%).
- El área bajo la curva entre dos puntos es la probabilidad de que una variable distribuida normalmente asuma un valor entre ellos.
- Dado que existe un número ilimitado de valores en el intervalo que va de $-\infty \quad a \quad +\infty$, la probabilidad de que una v. a. distribuida normalmente sea exactamente igual a cualquier valor dado es aproximadamente cero. Por lo tanto, las probabilidades siempre serán para un intervalo de valores.
- El área bajo la curva entre la media y cualquier otro punto es una función del número de desviaciones estándar que el punto dista de la media.

1.3.4.3 Cálculo de Áreas Bajo la Curva Normal (Uso de Tablas)

El área sombreada en la siguiente figura



Son los valores $\int_0^z f(x) = \int_0^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$, que se encuentran en las tablas que se anexan.

La distribución normal estándar se define como: $\mathbf{Z} = \frac{X-\mu}{\sigma}$

Donde: X es la variable aleatoria normal
µ es la media
σ es la desviación estándar
Z es la variable aleatoria normal estándar}

2.3.4.4 Ejemplos

1. Sea X una variable aleatoria normal, utilice la regla empírica para determinar la probabilidad de X quede:

a) a la izquierda de μ + λ

b) a la izquierda de μ + 2 λ

c) a la izquierda de λ μ + 3 λ

d) a la derecha de μ + λ

e) a la derecha de λ μ - 2 λ

f) a la derecha de λ μ - 3 λ

g) entre μ - λ

h) entre μ - 2 λ y μ

i) entre μ - 3 λ y μ

j) a la izquierda de μ - λ o a la derecha de μ + λ

- Buscar en la tabla los valores de Z $\pm$ 1, 2, 3 y otros.

Z $\approx$ 0.00 a dos decimales

3. Se considera que las calificaciones de cierta prueba de aprendizaje aplicada a los graduados de educación superior se distribuyen normalmente en μ = 70 y λ = 10 puntos. ¿Cuál es la probabilidad de que una calificación seleccionada aleatoriamente sea mayor a 90 puntos?

$P(X > 90) = P(Z > 2)$

$0.5 – 0.4772 = 0.0228$

- Buscar en la tabla los valores de Z $\pm$ 1, 2, 3 y otros.

$Z \approx$ 0.00 a dos decimales

1.3.4.5 Ejercicios

1. Obtener la probabilidad de que:

a) $Z < 1.64$

b) $Z < 1.64$

c) $Z > -1.65$

d) $0 < Z < 1.96$

e) $-1.96 < Z < 1.96$

f) $Z < -1.96$

g) $Z < 1.96$

2. Obtenga el área bajo la curva normal estándar de que quede:

a) A la izquierda de Z = 1.56

b) A la derecha de Z = -1.25

c) A la izquierda de Z = -1.69

d) Entre Z = -1.35 y Z = 2

e) Entre Z = -1.35 y Z= 2

3. Una variable aleatoria X se aproxima normalmente con μ =100 y σ = 15. Obtenga la probabilidad de que:

a) $X < 80.5$

b) $X < 112$

c) $85 \leq X \leq 97$

d) $X > 116.5$

4. Una población normalmente distribuida tiene una media de 40 y una desviación estándar de 3. Calcule los valores reales para los siguientes valores de z:

a) 0.10

b) 2

c) - 3

d) - 3.20

UNIDAD II. DISTRIBUCIONES MUESTRALES

2.1 Población y Muestra

En una población los parámetros son:

P $\rightarrow$ proporción

$\sigma \rightarrow$ desviación estándar $\sigma = \sqrt{\frac{\Sigma(x-P)^2}{N}}$

N $\rightarrow$ Tamaño de la población

x $\rightarrow$ Variable Aleatoria

$\bar{p} \rightarrow$ Media de las proporciones

$\bar{p}$ = E(p)

$\bar{p}$ = P donde p es la proporción de la muestra

$\sigma_x = \frac{\sigma}{\sqrt{n}}$ Error estándar de la distribución de probabilidad

Poblaciones infinitas

$$\sigma_p = \sqrt{\frac{P \cdot Q}{n}}$$

Poblaciones finitas

$$\sigma_p = \sqrt{\frac{P \cdot Q}{n}}\sqrt{\frac{N-n}{N-1}}$$

Cuando n $\rightarrow \infty$

Se puede aplicar el Teorema del Límite Central a la distribución muestral de proporciones, ya que ésta es simétrica si p = 0.5 y tiene una tendencia normal si n aumenta, que dando la estandarización:

$$Z = \frac{p - P}{\sigma p}$$

P Es la proporción de la población

p Es la proporción de la muestra

σp Desviación estándar de la distribución muestral de proporción.

2.2 Parámetros y Estadísticos

Un estimador puntual de algún parámetro poblacional ϕ es un simple valor numérico $\theta\hat{}$ de un estadístico. El estadístico θ es llamado estimador puntual.

Por ejemplo, suponiendo que la variable aleatoria X esta normalmente distribuída con una media desconocido μ. El ejemplo de la media es un estimador puntual de una media poblacional desconocida μ. Esto es, el estimador de la media poblacional es igual $\bar{X}$. Posteriormente el valor numérico del promedio es un estimador puntual de la media. Si

$$x_1 = 25, x_2 = 30, x_3 = 29, x_4 = 31 \, el \; estimador \; puntual \; de \; \mu \, es$$
$$\bar{X} = \frac{25 + 30 + 29 + 31}{4} = 28.75$$

A continuación se presenta un cuadro de estimadores puntuales

Parámetro	Estimador	Tamaño de las muestras	Desviación Estándar del Estimador
μ	$\bar{X}$	n	$\frac{\sigma^2}{\sqrt{n}}$
P	p	n	$\sqrt{\frac{pq}{n}}$

2.3 Teorema del Límite Central

Proporciona información relativa a las la media, desviación típica o dispersión de la distribución y el patrón de la distribución de la media muestral. Si se extraen todas las posibles muestras de igual tamaño a partir de una población dada, entonces:

1- La media de $\overline{X}$ es igual a la media de x; esto es, $\mu_{\overline{x}} = \mu$

2- El error típico de $\overline{X}$ es igual a la desviación típica de la población dividida entre la raíz cuadrada de n; esto es, $\sigma_{\overline{x}} = \sigma_x / n$.

3- La distribución muestral de la media muestral es casi normal, independientemente del patrón de la distribución de la población.

Si X es una variable aleatoria normal con media μ y una desviación típica σ entonces la media $\overline{X}$ de n observaciones tomadas a partir de la población de X también será una variable aleatoria normal con media μ y desviación típica $\frac{\sigma}{\sqrt{n}}$. Esto es verdadero sin importar el tamaño de la muestra.

Sin embargo, algunas variables aleatorias de distribuciones no tienen un patrón parecido a la normal. Por lo que el teorema del límite central enuncia lo siguiente:

> *Si X es cualquier variable aleatoria con media μ y desviación típica σ, la distribución muestral de la media muestral $\overline{X}$ será apoximadamente normal con media μ y desviación típica* $\frac{\sigma}{\sqrt{n}}$ sin importar la forma de la distribución de X, siempre y cuando el tamaño de la muestra sea suficientemente grande.

Conforme a lo anterior, al determinar las probabilidades relativas a X, el valor normal estándar Z se expresa así:

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

La puntuación Z se expresa como la diferencia entre las medias muestrales de la población, en términos del error típico.

2.3.1 Ejemplo

Las puntuaciones en cierto examen estándar se distribuye normalmente con media 60 y varianza 256. Se selecciona una muestra aleatoria de 16 puntuaciaciones y se calcula la media muestral, ¿Cuál es la probabilidad de que $\overline{X}$ sea mayor de 70?

P($\bar{X}$ > 70) = P (Z > $\frac{70-60}{16/\sqrt{16}}$)

= P(Z >2.5)

= 1- P(Z<2.5)

=1-0.9938

= 0.0062

2.3.2 Ejercicios

1. El registro académico de cierta universidad tiene establecido que el 40% de los estudiantes llevan un promedio de 7 o más. Si se toma una muestra aleatoria simple de tamaño 100, ¿Cuál es la probabilidad de que en esa muestra, la proporción de estimación con 7 o más sea:

   a) De 0.35 o menos

   b) De 0.50 o menos

   c) De más de 0.52

   d) De más de 0.32

   e) Entre 0.38 y 0.42 inclusive

X_i = No de estimaciones

n = 100

P = 0.4

Q = 1 - 0.

σp = 0.0490

Corrección por continuidad $\frac{1}{2n}$ entonces

$$Z = \frac{p \pm \frac{1}{2n} - P}{\sigma p}$$

Por ser $X \sim b(n, p)$ se debe considerar si está o no incluida

a) $+\frac{1}{2n} \Rightarrow P(p \leq 0.35) = 0.1788$

b) 0.9838

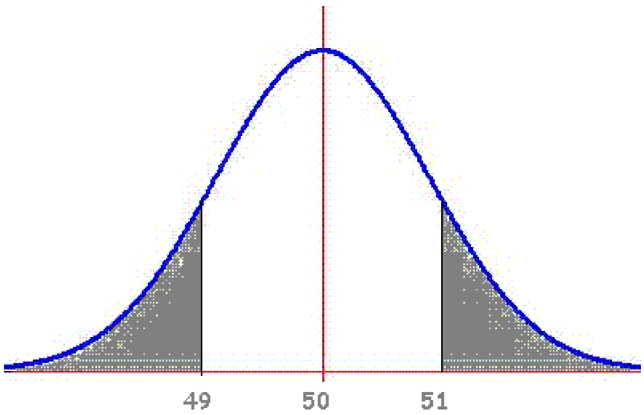
c) 0.9370

d) 0.9906

e) 0.39

2. Un fabricante de acumuladores asegura que su producto tiene una (promedio) esperada de 50 meses. Mediante estudios realizados por esta compañía se sabe que la desviación estándar de la vida del acumulador es de cuatro meses, ¿qué porcentaje de muestras de 36 observaciones tendrán una vida media promedio que vaya de 1 a 50 meses, suponiendo que 50 es el promedio de vida verdadera de los acumuladores? ¿Cuál es la respuesta si se toma una muestra de 64 observaciones?

Se sabe que como n>30 observaciones, la distribución de los valores medios será aproximadamente normal, con una media igual a la media de la población y una desviación estándar igual a la de la población dividida entre la raíz cuadrada del tamaño de la muestra. Gráficamente queda:



La solución consiste en determinar el número de desviaciones estándar a las que se encuentra 49 y 51 meses de la media y posteriormente, en referirnos a la tabla de áreas bajo la curva normal para obtener las probabilidades necesarias.
Primero se determina la desviación estándar de la distribución de muestreo:

$$\sigma_x = \frac{\sigma_x}{\sqrt{n}}$$

Para n = 36 $\sigma_x = \frac{4}{\sqrt{36}}$

= 0.67

Para n = 64 $\sigma_x = \frac{4}{\sqrt{64}}$

= 0.50

A continuación, se calcula la diferencia relativa del valor esperado

$$\frac{Valor\ estadistico - parámetro}{\sigma_x} \quad \frac{Valor\ estadistico - parámetro}{\sigma_x}$$

Para n = 36 :

$$= -1.5\sigma_x \qquad = 1.5\sigma_x$$

Para n = 64: $\frac{49-50}{0.50} = -\frac{1}{0.50}$ $\quad \frac{51-50}{0.50} = \frac{1}{0.50}$

$$= -2\sigma_x \qquad = 2\sigma_x$$

Ahora, se determinan las áreas utilizando la tabla

$n = 36: \; Z = 1.5$ área = 0.4332

$$P(49 < \bar{X} < 51) = 0.4332 + 0.4332$$
$$= 0.8664$$

$n = 64 : Z = 2.0$ área=0.4773

$$P(49 < \bar{X} < 51) = 0.4773 + 0.4773$$
$$= 0.9546$$

La probabilidad de obtener la media de la muestra en el intervalo dado es mayor para muestras de 64 observaciones que de 36, debido al hecho de que la desviación estándar de la distribución de muestreo disminuye a medida que n aumenta.

3. Con la información del ejemplo anterior ¿cuál sería la probabilidad de obtener un valor medio de muestra menor de 49.8 meses,

UNIDAD III INFERENCIA ESTADÍSTICA

3.1 Estimación.

Esta unidad iniciará con los conceptos básicos de acuerdo a Johnson (2004 p.305) Estimación Puntual de un Parámetro “es el número que se designa para estimar un parámetro cuantitativo de una población. Es el valor muestral de la estadística de la muestra correspondiente”.

3.1.1 Tipos de estimación.

$$estimación \begin{cases} Puntual\ (valor\ único) \\ Por\ \text{int}\,ervalo\ (Conjunto\ de\ valores) \end{cases}$$

3.1.2 Ejemplos.

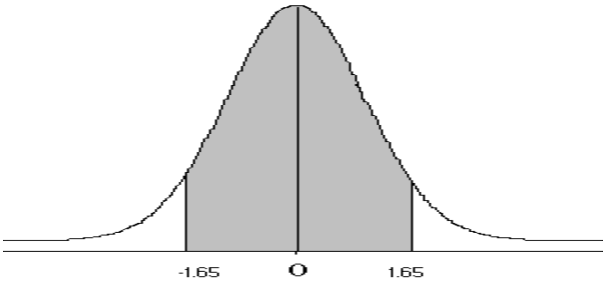
– En un estudio del flujo de estudiantes universitarios que reciben clases, se encontró que en promedio, una muestra de 36 estudiantes llegaban 17 minutos tarde. Una investigación previa había demostrado que la desviación típica del tiempo en llegar tarde era $\sigma = 15$ minutos.

a) Determine un intervalo de confianza del 90% para estimar la media poblacional.

b) ¿Con qué coeficiente de confianza puede afirmarse que el tiempo medio de llegadas tarde esté entre 12 y 22 minutos?

SOLUCIÓN:

a) El intervalo de confianza del 90% para estimar μ es: $\overline{X} \pm Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$



$(1-\alpha) = 0.90 \Rightarrow \alpha = 0.10 \ y \ \alpha/2 = 0.05$, por lo que de tablas tenemos que

$Z_{\alpha/2} = 1.65$
$\sigma = 15$
$n = 36$
$\overline{X} = 17$

sustituyendo en el intervalo tenemos

$$17 \pm 1.65\left(\frac{15}{\sqrt{36}}\right); \quad ó \quad (12.875, 21.125)$$

b) Si se estima que $\mu \in (12, 22) \Rightarrow Z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}} = 5$

Sustituyendo tenemos

$$Z_{\alpha/2}\cdot\frac{15}{\sqrt{36}} = 5 \Rightarrow 2.5Z_{\alpha/2} = 5 \Rightarrow Z_{\alpha/2} = \frac{5}{2.5}$$
$$Z_{\alpha/2} = 2$$

Un valor de $Z_{\alpha/2} = 2$ corresponde a 0.4772, de tablas de la distribución normal, por lo que el coeficiente de confianza es $(1-\alpha) = 0.9544$, para estimar la media poblacional.

– El gerente de una planta eléctrica quiere estimar la cantidad de carbón que necesitará para el presente año y tomó una muestra aleatoria al medir el consumo de carbón durante 10 semanas, obteniendo $\overline{X} = 11,400ton \quad y \quad S = 700ton$

Si desea su estimación con un coeficiente del 95%, construir el intervalo de confianza.

SOLUCIÓN:

Como $n<25$ y σ desconocida, se usa la distribución t de Student. Estimamos σ por medio de la muestra $\Rightarrow \sigma = S = 700$ y

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{700}{\sqrt{10}} = 221.36$$

t para (10-1) y $\alpha = 0.05$ es igual a 2.262, por lo tanto el intervalo está dado por: $\left(11400 - 2.262(221.36), 11400 + 2.262(221.36)\right) \Rightarrow (10899.28, 11900.72)$

3.1.3 Resumen de los límites de confianza

ESTIMACION DE μ	**POBLACION FINITA**	**POBLACION INFINITA**
Cuando σ es conocida	$\overline{X} \pm z\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$	$\overline{X} \pm z\frac{\sigma}{\sqrt{n}}$
Cuando no se conoce σ ($\sigma = S$) y $n>25$	$\overline{X} \pm z\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$	$\overline{X} \pm z\frac{S}{\sqrt{n}}$
Cuando no se conoce σ; $n \leq 25$ y la población se distribuye en forma normal o $\approx$ normal	$\overline{X} \pm t\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$	$\overline{X} \pm t\frac{S}{\sqrt{n}}$

ESTIMACION DE p	**POBLACIÓN FINITA**	**POBLACIÓN INFINITA**
Cuando $n>25$; $\sigma_p = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$	$\overline{p} \pm z\sigma_p\sqrt{\frac{N-n}{N-1}}$	$\overline{p} \pm z\sigma_p$

3.1.4 Ejercicios.

1. Defina brevemente que es la estimación.
2. Por qué son diferentes la estimación de punto y la estimación de intervalo?
3. ¿Qué entiende por error muestral?
4. Investigue cuáles son las características deseables de un estimador. Escriba una descripción breve de cada una de ellas.
5. Se tomo una muestra aleatoria de 100 estudiantes del total de aspirantes que presentaron examen de admisión a la UNAM. El promedio obtenido fue de 62 puntos, con una desviación estándar de la muestra igual a 15 puntos:
   a. Determinar un intervalo de confianza de 85% para estimar la media poblacional.
   b. ¿Con qué grado de confianza puede afirmarse que la media poblacional este entre 58.29 y 65.71?

6. Durante cierta semana el número de personas que acudieron a cierta biblioteca fue de 1444, de estos el 65% solicitaron préstamo de material a domicilio, si consideramos el número de usuarios como una muestra, encuentre el intervalo de confianza del 90% para estimar la verdadera proporción de usuarios de esa biblioteca que solicitaron préstamo a domicilio.
7. Una compañía desea estimar el promedio de gastos de representación semanales de sus agentes de ventas. Para esto, toma una muestra aleatoria de 30 cuentas, los auditores encontraron un gasto promedio de $220.00 con una desviación típica de $20.00.
   a. Obtenga una estimación puntual para el promedio de gastos de todos sus agentes de ventas.
   b. Establezca un intervalo del 99% de confianza, para estimar el promedio de gastos de todos sus agentes.
8. En general los estudios de factibilidad de proyectos requieren de una medida de la demanda para determinar la rentabilidad potencial de un bien o servicio. En un estudio para determinar la factibilidad de aumentar programas de telenovelas, una agencia procesadora de estadísticas, encontró que 66 de los 120 hogares seleccionados al azar que tienen televisión, veían telenovelas.
   a. determine un intervalo de confianza del 90% para estimar la verdadera proporción de hogares que ven telenovelas.
   b. A qué nivel de confianza podría decirse que la proporción poblacional de hogares que ven telenovelas este entre $0.4382< p <0.6672$?
9. El director de una escuela metropolitana desea conocer el coeficiente de inteligencia promedio de sus alumnos. Con este motivo desea hacer una investigación estadística. Determinar cuál debe ser el tamaño de la muestra si se tienen las siguientes restricciones: a). el error muestral máximo permisible no debe ser mayor que 0.5, b) la desviación estándar poblacional es de 1.9 según archivos y c) se desea un coeficiente de confianza del 99%.
10. Qué tamaño de muestra sería necesario para obtener un intervalo de confianza del 95% para la proporción de la población si el error muestral es de 0.08?

3.2 Pruebas de Hipótesis

3.2.1 El contraste de hipótesis estadísticas es un procedimiento de cuatro pasos:

Paso 1. Formular la hipótesis nula (H_0) y alternativa (H_a). Se debe considerar los siguientes resultados en un contraste de hipótesis.

Paso 2. Determinar el criterio de prueba o contraste. Consiste en la especificación de:

1) Nivel de significación. Probabilidad de cometer error tipo I (α).
2) Estadístico de prueba o contraste (z). Variable aleatoria utilizada para tomar la decisión “no se rechaza H_o” o bien “se rechaza H_o”.
3) La región critica. Conjunto de valores de la estadística de contraste que causan el rechazo de la hipótesis nula.
4) Los valores críticos. Valores “frontera” en la región crítica.

Paso 3. Obtener los datos muestrales y calcular el valor de la estadística de prueba.

Paso 4. Tomar la decisión e interpretarla.

1) Regla de decisión: Si el valor calculado de la estadística de prueba queda localizado dentro de la región crítica, se rechazará H_o. En caso contrario, no se podrá rechazar H_o.
2) Conclusión. Si la decisión es “rechazar H_o”, entonces la conclusión podría ser redactada como “Existe evidencia suficiente al nivel de significación α para indicar que… (el significado de la hipótesis alternativa)”. Si la decisión es “no se puede rechazar H_o”, entonces la conclusión podría ser redactada como “No existe suficiente evidencia al nivel de significación α que indique que… (el significado de la hipótesis alternativa)”.

3.2.2 Ejemplo.

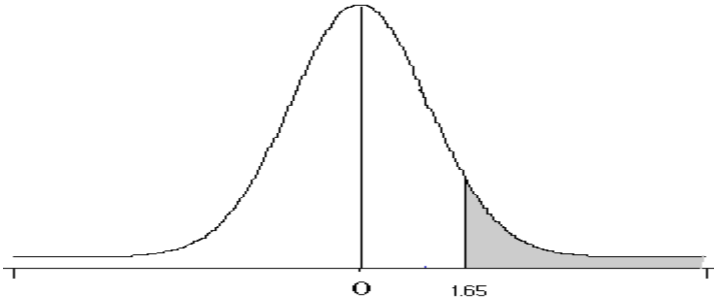
– Suponga que cierta compañía de aviación exige al fabricante de sus aparatos que utilice remaches cuya resistencia promedio a la ruptura exceda de 120 libras. Todo manufacturero de remaches que desee venderle al fabricante de los aviones, debe demostrar que sus remaches cumplen la especificación requerida, es decir, la μ de todas las resistencias a la ruptura de las piezas debe ser mayor que 120 libras. Si se sabe que $\sigma = 12$.

SOLUCIÓN.

Paso 1: $H_o: \quad \mu = 120$, se trabaja con el supuesto de que ésta hipótesis es verdadera.

$$H_a: \quad \mu > 120$$

Paso 2: Especificación del nivel de significación $\alpha = 0.05$; el estadístico de prueba será la distribución normal estándar (z); la región crítica,



Paso 3: Se examina una muestra de 36 remaches y las medidas resultantes producen una media igual a 124-4. Calculando z

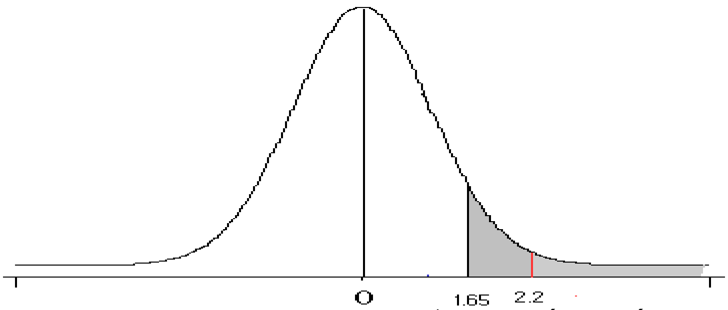
$$z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{124.4 - 120}{12 / \sqrt{26}} = \frac{4.4}{2.0} = 2.2$$

$$z = 2.2$$

Se tiene ahora el “valor calculado” del estadístico de prueba.

De hecho se tienen dos valores de z, el primero $z_{(0.05)} = 1.65$ y el segundo es $z = 2.2$

Paso 4: Tomar la decisión.



REGIÓN CRÍTICA

Puesto que el valor de $z = 2.2$, está dentro de la región crítica, se llega a la conclusión de rechazar H_o.

3.2.3 Ejercicios.

1. ¿Qué error puede cometerse en la decisión, si la hipótesis nula es verdadera?
2. Si se toma la decisión “no se puede rechazar Ho”, ¿qué error puede cometerse?
3. ¿Qué se está diciendo acerca del error tipo I, si a α se le asigna el valor 0.05?
4. ¿Qué se está diciendo acerca del error tipo II, si a β se le asigna el valor 0.001?
5. ¿Qué decisión debe tomarse si el valor de la estadística de prueba
    a) cae en la región crítica?
    b) no cae en la región crítica?

6. Un supervisor docente, afirma que el coeficiente medio de inteligencia de los niños mexicanos es 100. Si se sabe que $\sigma = 10$. Sin embargo, un psicólogo cree que dicho promedio es mayor que el que dice el supervisor. Supongamos que el psicólogo al tomar una muestra aleatoria de 400 niños y medir su inteligencia obtuvo 102 como media de la muestra. Se debe probar si es cierta la afirmación del supervisor, utilizando un nivel de significación del 5%.
7. Se diseña cierta dimensión crítica de 5 pulgadas para que una pieza manufacturada encaje en otras. Se sabe que la desviación típica del proceso manufacturero es de 0.08 pulgadas. Un grupo de estudiantes del Instituto Tecnológico Industrial toma una muestra aleatoria de 36 piezas, obteniendo una $\overline{X} = 5.05$ pulgadas. Probar la hipótesis de que se ha diseñado una buena dimensión crítica, al nivel de significación del 1%.
8. Un estudio muestral de 30 familias de una comunidad rural dio un ingreso promedio familiar de $210.00 semanales, con una $\sigma = 38$. Pero un sociólogo afirma que el salario promedio para esa comunidad es de $\mu = 230$ pesos. Pruebe la afirmación del sociólogo con un nivel de significación del 5%.
9. Cierto tipo de tubos de cobre tienen un diámetro externo de 5cm con una desviación típica de 0.18cm. se desea probar $H_0 : \mu = 5$ cm, frente a $H_a : \mu \neq 5$ cm. Si una muestra aleatoria de 36 tubos, da una media de $\overline{X} = 4.95$ cm, determine el intervalo de confianza del 99%, para estimar la media poblacional. ¿Indica este intervalo, el rechazo o la aceptación de la hipótesis nula $H_0 : \mu = 5$ cm? Use $\alpha = 0.01$.
10. Una muestra aleatoria de 64 estudiantes de cierta universidad dio un promedio de estatura de $\overline{X} = 1.65$ m, con una desviación típica de 0.08m. ¿Son estas observaciones consistentes con el supuesto de que la media de estatura para todos los estudiantes de esa universidad es de $\mu = 1.68$ m. Use un error α de 5%.

BIBLIOGRAFIA.

Chao, L. *Introducción a la Estadística.* Compañía Editorial Continental. México, 2006.

Johnson, R. y Kuby P. *Estadística Elemental.* Tomson Learning. México, 2004.

Stevenson, W. *Estadística para Administración y Economía*, Harla. México, 1981.

Mendenhall, W y Reinmuth, J. *Estadística para Administración y Economía,* Wadsworth Internacional/Iberoamérica. U.S.A. 1978.

Daniel, W. *Bioestadística.* Limusa, México 1983.

Freund, J. y Simon, G. *Estadística Elemental,* Prentice Hall, México, 1992.

Daniel, W. *Estadística con Aplicaciones a las Ciencias Sociales y a la Educación.* Mc Graw Hill. México, 1993.

Freund, J et al. *Estadística Matemática con aplicaciones.* Prentice Hall, México, 2000.