

ESTADÍSTICA INFERENCIAL

1

ESTADÍSTICA INFERENCIAL

La estadística Inferencial, es el proceso por el cual se deducen (infieren) propiedades o características de una población a partir de una muestra significativa.

	Población	Muestra
Definición	Colección de elementos considerados	Parte o porción de la población seleccionada para su estudio
Características	"Parámetros"	"Estadísticos"
Símbolos	Tamaño de la población = N	Tamaño de la muestra = n
	Media de la población = μ	Media de la muestra = \bar{x}
	Desviación estándar de la población = σ	Desviación estándar de la muestra = s

2

MÉTODO DE MUESTREO

Métodos no probabilísticos.- Interviene la opinión del investigador para obtener cada elemento de la muestra.

Métodos probabilísticos.- Muestra que se selecciona de modo que cada integrante de la población en estudio tenga una probabilidad conocida (pero distinta de cero) de ser incluido en la muestra.

MUESTREO ALEATORIO SIMPLE

MUESTREO ALEATORIO SISTEMÁTICO

MUESTREO ALEATORIO ESTRATIFICADO

MUESTREO ALEATORIO POR CONGLOMERADO

3

MUESTREO ALEATORIO SIMPLE

Muestra seleccionada de manera que cada integrante de la población tenga la misma probabilidad de quedar incluido.

Ejemplo: un bingo, introduzco los números en una ánfora y selecciono una muestra al azar

MUESTREO ALEATORIO SISTEMÁTICO

Los integrantes o elementos de la población se ordenan en alguna forma (Ejemplo: alfabéticamente) se selecciona al azar un punto de partida y después se elige para la muestra cada k-ésimo elemento de la población.

Ejemplo: se desea establecer una muestra 100 empleados de los 3000 que tiene una empresa, para lo cual ordeno alfabéticamente a los empleados, divido $3000/100 = 30$ y selecciona a uno de cada treinta empleados

4

MUESTREO ALEATORIO ESTRATIFICADO

Una población se divide en subgrupos denominados estratos y se selecciona una muestra de cada uno

ESTRATO	EDADES	Nº DE EMPLEADOS	% DEL TOTAL	CANTIDAD MUESTREADA
1	MENOS DE 25 AÑOS	8	2	1
2	26-30 AÑOS	35	10	5
3	31-35 AÑOS	189	54	27
4	36-40 AÑOS	115	33	16
5	MÁS DE 41 AÑOS	5	1	1
TOTAL		352	100	50

MUESTREO ALEATORIO POR CONGLOMERADO

Se divide a la población en estratos (subunidades) se selecciona con que subunidades se va a trabajar y de las unidades seleccionadas, se toma una muestra aleatoriamente

EJEMPLO MPC

José Crespo y Castillo, Hermilio Valdizan, Rupa Rupa, Luyando, Mariano Damaso Beraun y Daniel Alomia Robles.

Con estas ciudades se cubre el 67% de la población urbana del país.

5

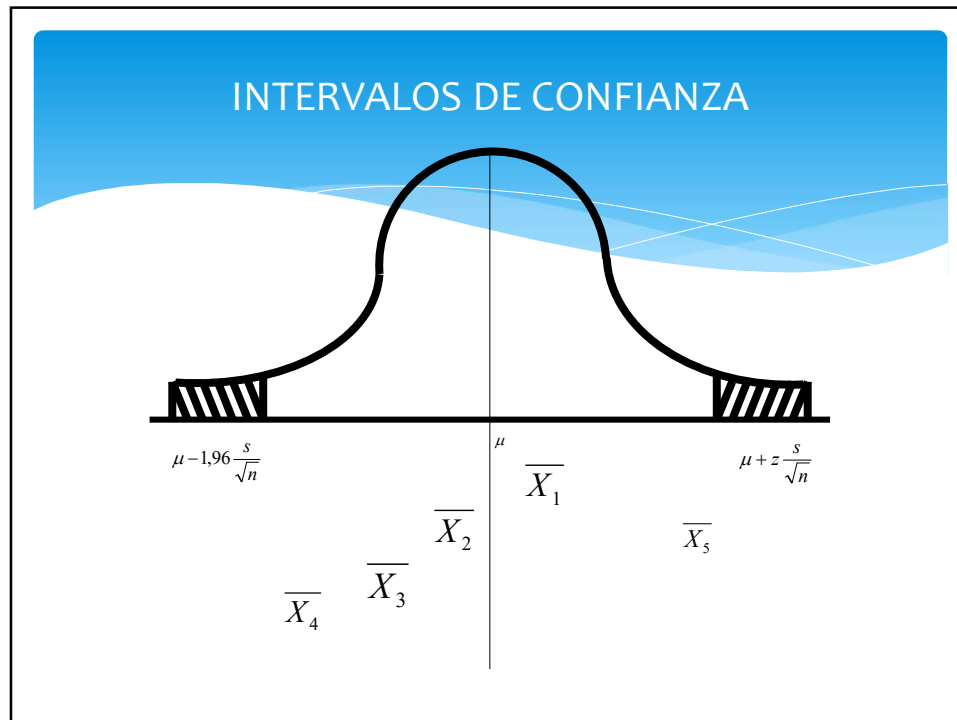
CONCEPTOS INICIALES

Estimación Puntual.- Estadístico calculado a partir de la información obtenida de la muestra y que se usa para estimar el parámetro poblacional

Intervalo de confianza.- es un conjunto de valores obtenido a partir de los datos muestrales en el que hay una determinada probabilidad de que se encuentre el parámetro, a esta probabilidad se le conoce como el nivel de significancia

Error de muestreo.- Diferencia entre un valor estadístico de muestra y su parámetro de población correspondiente.

6



7

INTERVALOS DE CONFIANZA

INTERVALO DE CONFIANZA PARA MUESTRAS MAYORES A 30 ELEMENTOS

$$\bar{X} \pm z \frac{s}{\sqrt{n}}$$

INTERVALO DE CONFIANZA PARA MUESTRAS MENORES A 30 ELEMENTOS

$$\bar{X} \pm t \frac{s}{\sqrt{n}}$$

DESVIACIÓN ESTÁNDAR

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}}$$

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

8

PROPORCIONES

PROPORCIÓN.- Fracción, razón o porcentaje que indica la parte de la muestra o población que tiene una característica determinada

PROPORCIÓN MUESTRAL: $p = \frac{x}{n}$

INTERVALOS DE CONFIANZA PARA UNA PROPORCIÓN POBLACIONAL

$$p \pm z\sigma_p$$

ERROR ESTÁNDAR DE LA PROPORCIÓN MUESTRAL

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

9

EJERCICIO

Suponga que se toma una muestra de 30 empleados de los cuales reciben en promedio 349\$ y una desviación estándar de 110\$. ¿Cuál es el intervalo de confianza?

$$\bar{X} \pm z \frac{s}{\sqrt{n}}$$

$$349 \pm 39.36$$

$$349 \pm 1.96 \frac{110}{\sqrt{30}}$$

$$[349 - 39.36 < \mu < 349 + 39.36]$$

$$[309.64 < \mu < 388.36]$$

Suponga que se toma una muestra de 20 empleados de los cuales reciben en promedio 346\$ y una desviación estándar de 126\$. ¿Cuál es el intervalo de confianza?

$$\bar{X} \pm t \frac{s}{\sqrt{n}}$$

$$346 \pm 58.97$$

$$346 \pm 2.093 \frac{126}{\sqrt{20}}$$

$$[346 - 58.97 < \mu < 346 + 58.97]$$

$$[287.03 < \mu < 404.97]$$

10

EJERCICIO - PROPORCIONES

En una muestra aleatoria de 2000 miembros de sindicato, se tiene que 1600 están a favor de fusionarse con otra empresa ¿Cuál es el valor estimado de la proporción poblacional? ¿Cuál es el intervalo de confianza al 95% de confianza?

$$p = \frac{x}{n} \quad p = \frac{1600}{2000} = 0,80$$

$$p \pm z \sqrt{\frac{p(1-p)}{n}}$$

$$0,80 \pm 1,96 \sqrt{\frac{0,80(1-0,80)}{2000}} = 0,80 \pm 0,018$$

$$[0.8 - 0.018 < P < 0.8 + 0.018]$$

$$[0.782 < P < 0.818]$$

11

PRUEBAS DE HIPÓTESIS PARA UNA MUESTRA

HIPÓTESIS.- Es una afirmación sobre una población, que puede someterse a pruebas al extraer una muestra aleatoria.

PRUEBA DE HIPÓTESIS.- Formular una teoría y luego contrastarla

PASOS PARA PROBAR UNA HIPÓTESIS

1. PRUEBA DE HIPÓTESIS
2. SELECCIONAR EL NIVEL DE SIGNIFICANCIA
3. CALCULAR EL VALOR ESTADÍSTICO DE PRUEBA
4. FORMULAR LA REGLA DE DECISIÓN
5. DECIDIR

12

PASO 1. PLANTEAR H_0 Y H_1

$$H_0: \mu = \mu_0$$

$$H_0: \mu > \mu_0$$

Hipótesis nula: Afirmación acerca del valor de un parámetro poblacional

$$H_1: \mu \neq \mu_0$$

$$H_1: \mu < \mu_0$$

Hipótesis Alternativa: Afirmación que se aceptará si los datos muestrales aseguran que es falsa H_0

Paso 2. Seleccionar el nivel de significancia

Generalmente son del 5% o 1% (Error de tipo I y Error de tipo II)

ERROR DE TIPO I.- Rechazar la hipótesis nula, H_0 cuando es verdadera

ERROR DE TIPO II.- Aceptar la hipótesis nula, H_0 cuando es Falsa

13

Paso 3. Calcular el valor estadístico de prueba.

Estadísticos de pruebas como: Z, t de Student, F y Ji cuadrado

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

Para muestras grandes

$$t = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Para muestras pequeñas

$$Z = \frac{P - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Para proporciones

Paso 4: Formular la regla de decisión

Son las condiciones según las que se acepta o rechaza la hipótesis nula

Paso 5: Tomar una decisión

El valor observado de la estadística muestral se compara con el valor de estadística de prueba

14

EJEMPLO: PRUEBA DE HIPÓTESIS

La producción diaria en una planta industrial registrada durante $n = 30$ días tiene una media Muestral de 990 toneladas y una desviación estándar de 20 toneladas, pruebe la hipótesis de que el promedio de la producción diaria es diferente a 1000 toneladas por día.

PASO 1: ESTABLECER HIPÓTESIS

$$H_0: \mu = 1000 - \text{toneladas}$$

$$H_1: \mu \neq 1000 - \text{toneladas}$$

PASO 2: Nivel de significancia (0.05)

PASO 3: Valor estadístico de prueba

$$\bar{x} = 990 \text{ toneladas}$$

$$\mu_0 = 1000 \text{ toneladas} \quad n = 30 \text{ días}$$

$$\sigma = 20 \text{ toneladas}$$

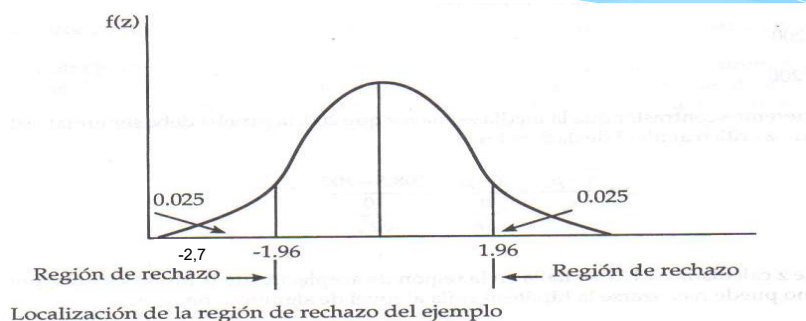
$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad z = \frac{990 - 1000}{\frac{20}{\sqrt{30}}} = -2,7$$

15

PASO 4: FORMULAR LA REGLA DE DECISIÓN

Para un nivel de significancia de 0.05, la región de rechazo es $z > 1.96$ o $z < -1.96$

PASO 5: TOMAR UNA DECISIÓN



Se rechaza H_0 μ no es igual a 1000 toneladas

16

EJEMPLO: PRUEBA DE HIPÓTESIS

El gerente de ventas de una empresa editora de libros, afirma que cada uno de sus representantes realiza 40 visitas por semana a profesores. Varios vendedores dicen que esa estimación es muy baja. Para investigar lo anterior, una muestra aleatoria de 28 representantes de ventas reveló que el número medio de visitas realizadas la semana pasada fue de 42. Se calculó que la desviación estándar de la muestra fue de 2.1 visitas. Al nivel de significancia de 0.05, ¿se puede concluir que el número medio de visitas realizadas por vendedor y por semana es mayor que 40?

PASO 1: ESTABLECER HIPÓTESIS

$$H_0 : \mu \leq 40$$

$$H_1 : \mu > 40$$

17

PASO 2: NIVEL DE SIGNIFICANCIA (0.05)**PASO 3: ESTADÍSTICO DE PRUEBA**

En este caso es T de student

$$\mu = 40 \text{ visitas}$$

$$\bar{x} = 42 \text{ visitas}$$

$$s = 2.1 \text{ visitas}$$

$$n = 28$$

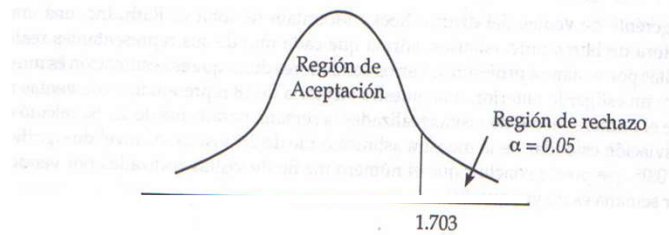
$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{42 - 40}{2.1 / \sqrt{28}} = 5.04$$

18

PASO 4: REGLA DE DECISIÓN

Rechazo H_0 SI t calculado es mayor a 1.703

PASO 5: TOMAR DECISIÓN

T calculado = 5.04 cae en la región de rechazo. Por lo tanto, rechazamos H_0 . El número medio de visitas realizadas por vendedor y por semana es mayor que 40