Cátedra ESTADISTICA
TRABAJOS PRÁCTICOS
(SEGUNDA ETAPA)

2020

Facultad de Ingeniería
Universidad Nacional de La Patagonia S. J. B.
Comodoro Rivadavia

TRABAJO PRÁCTICO Nº 4
INFERENCIA ESTADÍSTICA
Distribuciones Muestrales-Intervalos de confianza

PRE - REQUISITOS:

Se requiere lectura previa y manejo conceptual de los siguientes conceptos trabajados en la primera etapa:

- ✓ Estadística Descriptiva – Variables, tipos - Medidas de posición y dispersión –
- ✓ Distribuciones de probabilidad, discretas y continuas – Propiedades, parámetros, usos, cálculos – Identificación de las variables aleatorias.
- ✓ Esperanza (media) y Varianza de las Distribuciones de probabilidad.
- ✓ Muestreo aleatorio -

Y análisis y comprensión de los contenidos que se están desarrollando en esta segunda etapa:

- ✓ Distribuciones muestrales.
- ✓ Teorema Central de Límite (T. C .L.) y sus consecuencias.
- ✓ Significado de ESTIMAR, diferencia entre estimación puntual y estimación por Intervalo de Confianza. ¿Cuál es realmente útil y por qué?

CONSIGNA PARTICULAR

Leer y comprender los conceptos que se presentan a continuación, recurrir a los textos propuestos, analizar los problemas resueltos.
Con esta lectura comprensiva y considerando también los otros pre-requisitos, usted está ya en condiciones de trabajar con sus compañeros.
Recuerde que, en esta segunda etapa de la cursada, usted debe ya ser capaz de trabajar en equipo y consultar sólo sus dudas.

Recuerde que es indispensable la **correcta interpretación en términos del problema, aún cuando no se lo señale específicamente en cada ejercicio.**

A continuación se transcriben párrafos del texto recomendado "**PROBABILIDAD Y ESTADÍSTICA PARA INGENIEROS"**, 4/E, de Miller, I. ; Freund, J. E.; Johnson, R. A., Edición 1992, Prentice-hall Hispanoamericana, S. A. , (capítulos 6 y 7), actualmente citado como: Johnson, Richard A.; Torre Marina Juan A. , **"Probabilidad y Estadística para Ingenieros de**

Miller y Freund”, Prentice-hall Hispanoamericana, S. A., 3ra. Edición, México, 1997 (Clas.: 519.5/J.26, en Biblioteca Central de la U. N. P. S. J. B., Comodoro Rivadavia)

. . . “se supondrá que estamos trabajando con una clase particular de muestra denominada muestra aleatoria” . . . “se debe a que permite generalizaciones válidas o lógicas de . . .”, . . . “veremos como ciertos estadísticos puede esperarse que varíen de muestra a muestra”. ..
“El concepto de distribución muestral es . . . “(pág. 187)

“Si una muestra aleatoria de tamaño “n” se elige de una población que tiene media **μ** y varianza **σ²**, entonces **x media** es una variable aleatoria cuya distribución tiene media **μ** y:

- para muestras tomadas de poblaciones infinitas la varianza de esta distribución es:

$$\boldsymbol{\sigma^2 / n}$$

- para muestras tomadas de poblaciones finitas de tamaño “N” la varianza de esta distribución es:

$$\left(\frac{\sigma^2}{n}\right) * \left(\frac{N-n}{N-1}\right) \qquad \text{(pág. 192)}$$

Si

$$x \approx N(E(x) = \mu, Var(x) = \sigma^2) \Rightarrow \qquad \frac{x-\mu}{\sigma} = z$$

$$\bar{x} \approx N(E(\bar{x}) = \mu, Var(\bar{x}) = \frac{\sigma^2}{n}) \Rightarrow \qquad \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = z$$

n grande

Si p=x/n (proporción)

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

Si σ² es desconocida

$$\frac{x-\mu}{s} = t_n$$

$$\frac{\bar{x}-\mu}{s/\sqrt{n}} = t_{n-1}$$

Distribución Siempre normal

Y de pág. 210...
“Cuando empleamos una media muestral para estimar la media de una población, sabemos que, aunque estamos utilizando un método de estimación con ciertas propiedades deseables, las posibilidades son escasas y prácticamente nulas de que la estimación en realidad sea igual a μ. En consecuencia, convendría acompañar la estimación puntual de μ con una afirmación de cuán cercana podemos razonablemente esperar que se encuentre la estimación.

El error, $\bar{x} - \mu$, es la diferencia entre la estimación y la cantidad que se supone que estima.

A fin de examinar este error, utilizaremos el hecho de que

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

es un valor de una variable aleatoria que tiene aproximadamente la distribución normal estándar. En consecuencia, podemos “asegurar” con una probabilidad de 1-α que la desigualdad

$$-z_{\alpha/2} \leq \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}$$

será satisfecha o que

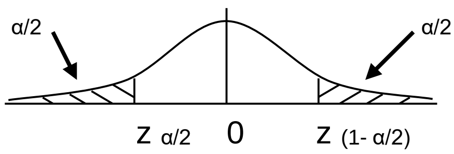
$$\frac{|\bar{x}-\mu|}{\sigma/\sqrt{n}} \leq z_{\alpha/2}$$

donde $z_{\alpha/2}$ es tal que el área bajo la curva normal a su derecha es α/2. Si ahora indicamos con E el valor máximo de$|\bar{x}-\mu|$, o sea el error máximo de estimación, tenemos

$$E = z_{\alpha/2} * \sigma/\sqrt{n}$$

con una probabilidad (1- α). En otras palabras, si intentamos estimar μ con la media de una muestra aleatoria, podemos afirmar con una probabilidad de (1- α) que el error$|\bar{x}-\mu|$, será a lo sumo $z_{\alpha/2} * \sigma/\sqrt{n}$. Los valores de mayor uso para (1- α) son 0.95 y 0.99 y los valores correspondientes de $z_{\alpha/2}$ son $z_{0.025} =$ 1.96 y $z_{0.005} =$ 2.575.

De acuerdo con el teorema central del límite, es de esperar que la distribución muestral de $\bar{X}$ tenga una distribución aproximadamente normal con media $\mu_x = \mu$ y desviación estándar$\sigma_{\bar{x}} = \sigma/\sqrt{n}$.Al escribir $z_{\alpha/2}$ para el valor z, sobre el cual se encuentra un área de α/2, a partir de la figura siguiente se advierte que: $P(-z_{\alpha/2}. < z < z_{\alpha/2}) = 1-\alpha$



donde,

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

De aquí que

$$P\left(-z_{\alpha/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha.$$

Al multiplicar cada término de la desigualdad por $\sigma/\sqrt{n}$, y después restar de cada término y multiplicar por –1 (para invertir el sentido de las desigualdades), se obtiene:

$$P(\bar{x} - z_{\alpha/2} * \sigma/\sqrt{n} < \mu < \bar{x} + z_{\alpha/2} * \sigma/\sqrt{n}) = 1-\alpha$$

Se selecciona una muestra aleatoria de tamaño n de una población cuya variancia σ^2se conoce y se calcula la media muestral $\bar{x}$ para obtener el intervalo de confianza (1- α)100%.

Intervalo de confianza de μ, conociendo σ²

Si $\bar{x}$ es la media de una muestra aleatoria de tamaño n de una población con una variancia conocida σ^2, el intervalo de confianza de (1- α)100% para μ es,

$$\bar{x} - z_{\alpha/2} * \sigma/\sqrt{n} < \mu < \bar{x} + z_{\alpha/2} * \sigma/\sqrt{n}$$

donde $z_{\alpha/2}$ es el valor de z a la derecha del cual se tiene un área de α/2.

Ejemplo:

Una muestra aleatoria de tamaño n = 100 se extrae de una población con $\sigma^2 = (5.1)^2$. Dado que la media muestral es $\bar{x} = 21.6$, construir un intervalo de confianza del 95% para la media de la población μ .

Solución: Sustituyendo los valores dados de n, $\bar{x}$, σ^2 y $z_{0.025} = 1.96$ en la fórmula para el intervalo de confianza, obtenemos

$$21.6 - 1.96 * \frac{5.1}{\sqrt{100}} < \mu < 21.6 + 1.96 * \frac{5.1}{\sqrt{100}}$$

ó $20.6 < \mu < 22.6$. Por supuesto, el intervalo de 20.6 a 22.6 contiene la media de la población μ o no, pero tenemos una confianza del 95% de que así sea. Esto significa que el método con que se obtuvo el intervalo "funciona" el 95% de las veces. En otras palabras, en repetidas aplicaciones de la fórmula del intervalo de confianza cabe esperar que el 95% de los intervalos contenga las medias de las poblaciones respectivas.

Ejemplo:

Un investigador quiere determinar el tiempo promedio que un mecánico tarda en rotar los neumáticos de un automóvil, y además desea poder asegurar con una confianza del 95% que el error de estimación de la media de su muestra sea a lo sumo de 0,50 minutos. Se puede asumir por experiencia que $\sigma^2 = (1,6)^2$ minutos^2. ¿Qué tamaño debe tener la muestra?

Solución: Sustituyendo E = 0,50, $\sigma^2 = (1,6)^2$, y $z_{0,025} = 1,96$ en la fórmula para n obtenemos

$$n = \left[\frac{1.96 * 1.6}{0.5}\right]^2 = 39.3$$

ó 40, redondeado al entero más cercano. Así, el investigador tendrá que observar 40 operaciones de la tarea de rotar los neumáticos de un automóvil.

Los métodos expuestos hasta ahora exigen conocer σ ó que pueda ser estimada mediante la desviación estándar muestral S, requiriendo así que n sea relativamente grande. Bajo el supuesto de estar muestreando en una población normal, podemos fundamentar nuestro argumento en el teorema central del límite, es decir, en el hecho de que

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

es un valor de una variable aleatoria que tiene la distribución *t* con n-1 grados de libertad. Con un razonamiento similar al empleado para la variable *z,* llegamos al resultado de que podemos asegurar con una probabilidad de (1-α) que el error en que incurrimos al emplear para estimar μ será a lo sumo de

$$E = t_{\alpha/2} * S/\sqrt{n}$$

llamado *Error máximo de estimación*

y obtenemos la fórmula del intervalo con un nivel de confianza del (1-α) por ciento cuando no se conoce σ^2.

$$P\left(\bar{x} - t_{\alpha/2} * S/\sqrt{n} < \mu < \bar{x} + t_{\alpha/2} * S/\sqrt{n}\right) = 1 - \alpha$$

INFERENCIA ESTADÍSTICA: TEOREMA CENTRAL DEL LÍMITE y DISTRIBUCIONES MUESTRALES

1. - Cierta población tiene media y varianza iguales a 500 y 900 respectivamente. Si se seleccionan muchas muestras de tamaño 36 y se calculan las medias muestrales:

a) ¿Qué valor es de esperar que tenga la media de todas esas medias?

b) ¿Qué valor se espera para la varianza y la desviación estándar de todas las medias muestrales?

c) ¿Qué forma es de esperar que tenga la distribución de todas las medias muestrales? Si es necesario, plantee los supuestos adecuados.

2.- Se toma una muestra aleatoria de ~~tamaño~~ 49 de tubos de metal y se les toma las medidas de diámetro. Se considera que provienen de una población ~~considerada~~ infinita, con media μ=102 mm y varianza σ^2=121 mm^2. Usar el teorema central del límite para encontrar la probabilidad de obtener $\bar{x}$ >104,5 mm.

3.- En un servicio de atención al cliente, el tiempo de espera hasta recibir atención es una variable que se distribuye en forma normal con media 10 minutos y desviación típica 2 minutos.

a.-Si se toma una muestra de 16 llamadas, ¿cuál es la probabilidad de que el tiempo promedio de espera no supere los 9 minutos?

b.- Si se toman muestras aleatorias del tiempo de espera de los clientes que llegan un día concreto. ¿Cuál es la probabilidad de que el tiempo medio de espera de una muestra de 25 clientes no supere los 9 minutos?

c.- ¿Cuál es la distribución de la media muestral, si se toman muestras aleatorias de 36 clientes? Especificar sus parámetros.

d.- Para una muestra de 9 clientes. Calcule y luego comente, analizando los resultados obtenidos.

c1) P (x > 12) c2) P (x> 10,5)

c3) P ($\bar{x}$ > 12) c4) P ($\bar{x}$ > 10,5).

4.- Supongamos que en una empresa que fabrica rulemanes la proporción de defectuosos en la producción diaria es de 0,1. ¿Cuál es la probabilidad de que, si se extraen 100 rulemanes, la proporción muestral de rulemanes fallados sea menor al 6%?

5.- Supongamos que el líquido (en litros) que vierte una máquina embotelladora tiene una distribución normal con σ^2 = 1. Si se elige una muestra aleatoria de 10 botellas y se mide la cantidad de líquido que contiene cada una

a- . ¿Cuál es la probabilidad de que la varianza muestral sea mayor que 1,2?

b- ¿Entre que valores simétricos de la varianza se encuentran el 90 % de las observaciones?

INFERENCIA ESTADÍSTICA: INTERVALOS DE CONFIANZA

6.- En una localidad se está haciendo un estudio sobre el número de días que dura un contrato temporal. Se sabe que la desviación estándar de los contratos es igual a 57 días. Con una confianza del 95% calcular el número mínimo de contratos en los que se ha de mirar su duración para que el intervalo tenga una amplitud no mayor de 10 días.

7. - Si un intervalo de confianza para μ está dado por 13,2 ± 1,0; con un coeficiente de confianza del 75% ¿Cuál es el coeficiente de confianza para que el intervalo sea 13,2 ± 2,0?

8. – Al estudiar la vida útil de unas pilas nuevas que se van a lanzar al mercado se obtuvo un intervalo de confianza para μ dado por 963 ± 1,92 horas; para lo cual se tomó una muestra de 40 pilas ¿Cuáles son los límites del intervalo con una muestra de 80 pilas con la misma confianza, la misma media y la misma varianza? ¿Por qué?

9.- En un programa de control de precios se toma una muestra de los precios de un mismo producto en 16 comercios elegidos al azar, y se obtienen los siguientes resultados:

95	108	97	112	99	106	105	100	99	98	104	110	107	111	103	110

Cálculos: $\Sigma x=1664$ $\Sigma x^2=173504$

(a)-Suponiendo que los precios se distribuyen normalmente con varianza 25; calcule el intervalo de confianza, al 95%, para la media poblacional:

- Seleccione la variable pivotal; fundamente su elección.
- Escriba la expresión teórica del intervalo.
- Construya el intervalo. Muestre sus cálculos.
- Interprete en términos del problema.

(b)- Repita (a) suponiendo que la varianza poblacional σ^2 es desconocida. Interprete en términos del problema. Compare los resultados con el inciso anterior.

10.-Se desea conocer entre que valores se encuentran los parámetros de la viscosidad del petróleo de una zona acotada. Para tal fin se extraen un total de 30 muestras de diferentes pozos y se calcula su viscosidad. La media obtenida fue de 1,53 Poise y la desviación estándar fue de 0,31 Poise.

a) Construya un intervalo de confianza del 95% para la verdadera media.
b) Construya un intervalo de confianza del 90% para la varianza.
c) En ambos casos interprete en términos del problema, compare y saque conclusiones.

11.- En la fabricación de alimentos congelados, interesa conocer la uniformidad del llenado de envases. Se toma una muestra de 20 envases, y se obtiene una media de 1013 gramos y una varianza de 2104 (gramos)^2. Encuentre el intervalo de confianza del 95% para estimar la variabilidad de llenado e interprete en términos del problema.

12.- Un fabricante de impresoras utiliza un conjunto de pruebas amplias para evaluar la función eléctrica de su producto. Todas las impresoras deben pasar todas las pruebas antes de venderse. En una muestra aleatoria de 500 impresoras, 15 fallaron en una o más pruebas. Encuentre un intervalo de confianza de 90% para la proporción de impresoras de la población que no pasarían todas las pruebas.

13.- Al tomar una muestra al azar de 120 estudiantes de una Universidad, se encontró que 54 de ellos hablaban inglés. Estime, con un nivel de confianza del 90%, la proporción de estudiantes que hablan el idioma inglés en esa Universidad.

14. A.- Al estimar la media poblacional mediante un intervalo de confianza del 99% se obtiene: Límite inferior y superior 270 y 330 respectivamente. Si dichos valores se calcularon en función de 41 observaciones, ¿cuál es el valor de la media muestral y de la varianza muestral utilizados para determinar el intervalo?

15. B.- Si un valor de 52 tiene un margen de error de $\pm$ 5, ¿cómo están relacionados estos números con:

-La estimación puntual?
-El ancho del intervalo?
-Los límites del intervalo?

16.- La longitud de los cables de acero fabricados por una máquina se distribuyen normalmente. Se desea estimar la media de la longitud de dichos cables, para lo cuál un operario toma una muestra de 12 cables obteniendo las siguientes mediciones de la longitud (en metros):

9,2 9,7 9,8 10,2 10,4 10 9,4 9,5 9,5 10,3 9,9 9,7

a) ¿Cuál es la variable en estudio?
b) Escriba la expresión teórica del intervalo.
c) Construya el intervalo solicitado con una confianza del 95% mostrando sus cálculos.
d) Interprete en términos del problema.

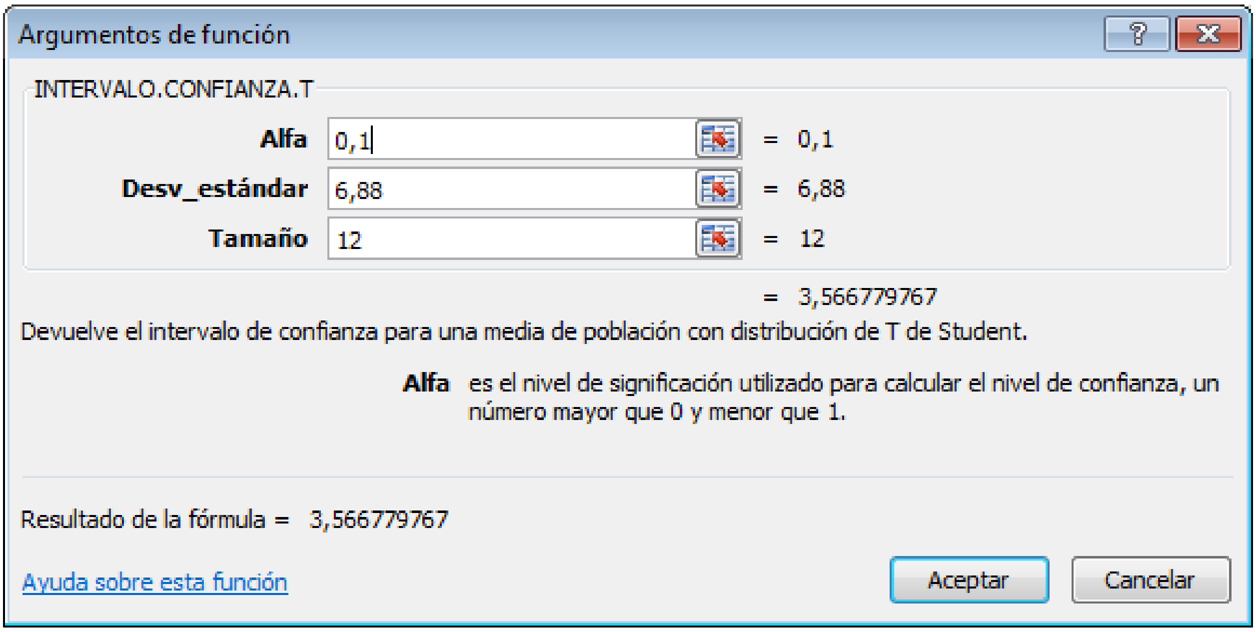
SALIDA DE PC

17. - Se sabe que la producción diaria de una refinería de petróleo está normalmente distribuida. De año en año la producción diaria usualmente se incrementa, pero la desviación estándar permanece constante.

La siguiente información representa la producción diaria de días seleccionados aleatoriamente durante 2014.

74.45	85.20	88.20	86.10	82.60	90.10	93.90	75.30	79.80	73.20	77.90	89.80

Se encontró un intervalo de confianza del 90% para la producción media diaria del año 2014. Calcule por su cuenta, compare con la salida siguiente e interprete.



PRUEBA DE CONCEPTOS

1) Coloque una cruz en la variable que se emplea para construir un intervalo de confianza para la varianza poblacional σ^2.

☐ $= \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$ ☐ $= \dfrac{X - \mu}{\sigma}$ ☐ $= \dfrac{(n-1)s^2}{\sigma^2}$ ☐ $= \dfrac{(n-1)\sigma^2}{\sigma^2}$ ☐Ninguna de las anteriores

2) Señale la expresión correcta para referirse a la interpretación de un intervalo de confianza.
   a) El parámetro está dentro del intervalo..........
   b) El intervalo encierra al valor del parámetro.
   c) El parámetro varía dentro del intervalo.................

3) Indique la expresión correcta.
   a) $P(10 < \mu < 25) = 0,95$
   b) $P (\mu = 25) = 0,95$
   c) $10 < \mu < 25$, con coeficiente de confianza 0,95.
   d) $10 < \mu < 25$, con nivel de significación 0,95.
   e) Ninguna de las anteriores.

4) Seleccione la respuesta correcta:

¿Cuál intervalo de confianza preferiría

4,32 ± 0,30? 4,32 ± 0,13?

Nota: ambos intervalos están calculados con el mismo coeficiente de confianza.