

TEMA 12 – INFERENCIA ESTADÍSTICA. ESTIMACIÓN DE LA MEDIA

12.1 – DISTRIBUCIÓN NORMAL. REPASO DE TÉCNICAS BÁSICAS

UTILIZACIÓN DE LA TABLA DE LA NORMAL N(0,1)

En la distribución N(0,1), a la variable se le suele representar por la letra z. La tabla nos da las probabilidades $P[z \leq k]$ para valores de k de 0 a 4, de centésima en centésima. A estas probabilidades se las llama $\phi(k)$: $\phi(k) = P[z \leq k]$ z se distribuye N(0,1)

$\phi(k)$ es la función de distribución de esta variable aleatoria.

El valor de k se busca así:

- Unidades y décimas en la columna de la izquierda
- Centésimas en la fila de arriba
- El número que nos da la tabla es el valor de : $\phi(k) = P[z \leq k]$

Ejemplo 1: Sea $Z \sim N(0,1)$ calcula:

- a) $P[z \leq 0,83] = \phi(0,83) = 0,7967$
- b) $P[z \leq 2,3] = \phi(2,30) = 0,9893$
- c) $P[z \leq 1] = \phi(1,00) = 0,8413$

Ejemplo 2 : Sea $Z \sim N(0,1)$ calcula en valor de “k”:

- a) $P[z \leq k] = 0,7190 \Rightarrow k = 0,58$
- b) $P[z \leq k] = 0,8645 \Rightarrow k \approx 1,1$
- c) $P[z \leq k] = 0,5537 \Rightarrow k \approx \frac{0,13 + 0,14}{2} = 0,135$

Ejercicio 3 : Sea $Z \sim N(0,1)$ calcula:

- a) $P[z \leq 0,45] = \phi(0,45) = 0,6736$
- b) $P[z \leq 1,2] = \phi(1,2) = 0,8849$
- c) $P[z \leq 7] = 1$

Ejercicio 4 : Sea $Z \sim N(0,1)$ calcula en valor de “k”:

- a) $P[z \leq k] = 0,5040 \Rightarrow k = 0,01$
- b) $P[z \leq k] = 0,62 \Rightarrow k \approx 0,31$
- c) $P[z \leq k] = 0,9735 \Rightarrow k \approx \frac{1,93 + 1,94}{2} = 1,935$

CÁLCULO DE PROBABILIDADES EN UNA DISTRIBUCIÓN N(0,1)

- Si $k \geq 0$, las probabilidades $\phi(k) = P[z \leq k] = P[z < k]$ se encuentran directamente en la tabla.
- $P[z \geq k] = 1 - P[z < k] = 1 - \phi(k)$
- Para abscisas negativas: $P[z \leq -k] = P[z \geq k] = 1 - \phi(k)$
- $P[a \leq z \leq b] = P[z \leq b] - P[z \leq a]$

Ejemplo 5 : Sea $Z \sim N(0,1)$ calcula:

- a) $P[z \geq 1,86] = 1 - P[z < 1,86] = 1 - \phi(1,86) = 1 - 0,9686 = 0,0314$
- b) $P[0,18 < z \leq 1,29] = P[z \leq 1,29] - P[z \leq 0,18] = \phi(1,29) - \phi(0,18) = 0,9015 - 0,5714 = 0,3301$
- c) $P[-0,56 < z \leq 1,9] = P[z \leq 1,9] - P[z \leq -0,56] = \phi(1,9) - P[z \geq 0,56] = \phi(1,9) - [1 - P[z < 0,56]] = \phi(1,9) - 1 + \phi(0,56) = 0,9713 - 1 + 0,7123 = 0,6836$
- d) $P[-1,83 < z < -1] = P[z < -1] - P[z \leq -1,83] = P[z > 1] - P[z > 1,83] = [1 - P[z \leq 1]] - [1 - P[z \leq 1,83]] = 1 - \phi(1) + 1 + \phi(1,83) = -\phi(1) + \phi(1,83) = -0,8413 + 0,9664 = 0,1251$

Ejemplo 6 : Sea $Z \sim N(0,1)$ calcula “k”

- a) $P[z < k] = 0,8365 \Rightarrow k = 0,98$
 b) $P[z > k] = 0,8365 \Rightarrow P[z \leq k] = 0,1635$ (No está en la tabla, k es negativo) $k = -s$
 $P[z > -s] = 0,8365 \Rightarrow P[z < s] = 0,8365 \Rightarrow s = 0,98 \Rightarrow k = -0,98$
 c) $P[z < k] = 0,1894$ (No está en la tabla, k es negativo) $k = -s$
 $P[z < -s] = P[z > s] = 0,1894 \Rightarrow P[z \leq s] = 0,9106 \Rightarrow s \approx 1,34 \Rightarrow k \approx -1,34$

Ejercicio 7 : Halla las siguientes probabilidades en una distribución $N(0,1)$

- a) $P[z > 2,8] = 1 - P[z \leq 2,8] = 1 - \phi(2,8) = 1 - 0,9974 = 0,0026$
 b) $P[z \leq -1,8] = P[z \geq 1,8] = 1 - P[z < 1,8] = 1 - \phi(1,8) = 1 - 0,9641 = 0,0359$
 c) $P[z > -1,8] = P[z < 1,8] = \phi(1,8) = 0,9641$
 d) $P[1,62 \leq z < 2,3] = P[z < 2,3] - P[z < 1,62] = \phi(2,3) - \phi(1,62) = 0,9893 - 0,9474 = 0,0419$
 e) $P[1 \leq z \leq 2] = P[z < 2] - P[z < 1] = \phi(2) - \phi(1) = 0,9772 - 0,8413 = 0,1359$
 f) $P[-0,61 \leq z \leq 1,4] = P[z < 1,4] - P[z < -0,61] = \phi(1,4) - P[z > 0,61] = \phi(1,4) - [1 - P[z \leq 0,61]] = \phi(1,4) - 1 + \phi(0,61) = 0,9192 - 1 + 0,7291 = 0,6483$
 g) $P[-1 \leq z \leq 2] = P[z \leq 2] - P[z < -1] = \phi(2) - P[z > 1] = \phi(2) - [1 - P[z \leq 1]] = \phi(2) - 1 + \phi(1) = 0,9772 - 1 + 0,8413 = 0,8185$
 h) $P[-2,3 < z < -1,7] = P[z < -1,7] - P[z < -2,3] = P[z > 1,7] - P[z > 2,3] = [1 - P[z \leq 1,7]] - [1 - P[z \leq 2,3]] = 1 - \phi(1,7) - 1 + \phi(2,3) = -0,9554 + 0,9893 = 0,0339$
 i) $P[-2 \leq z \leq -1] = P[z < -1] - P[z < -2] = P[z > 1] - P[z > 2] = [1 - P[z \leq 1]] - [1 - P[z \leq 2]] = 1 - \phi(1) - 1 + \phi(2) = -0,8413 + 0,9772 = 0,1359$

Ejercicio 8 : Calcula el valor de k (exacta o aproximadamente) en cada uno de los siguientes casos:

- a) $P[z \leq k] = 0,5 \Rightarrow k = 0$
 b) $P[z \leq k] = 0,8729 \Rightarrow k = 1,14$
 c) $P[z \leq k] = 0,9 \Rightarrow k \approx 1,28$
 d) $P[z \leq k] = 0,33 \Rightarrow (0,33 \text{ no está en la tabla, } k \text{ es negativo}) k = -s$
 $P[z \leq -s] = 0,33 \Rightarrow P[z \geq s] = 0,33 \Rightarrow P[z < s] = 0,67 \Rightarrow s = 0,44 \Rightarrow k = -0,44$
 e) $P[z \leq k] = 0,2 \Rightarrow (0,2 \text{ no está en la tabla, } k \text{ es negativo}) k = -s$
 $P[z \leq -s] = 0,2 \Rightarrow P[z \geq s] = 0,2 \Rightarrow P[z < s] = 0,8 \Rightarrow s = 0,84 \Rightarrow k = -0,84$
 f) $P[z > k] = 0,12 \Rightarrow P[z \leq k] = 0,88 \Rightarrow k = 1,175$
 g) $P[z \geq k] = 0,9971 \Rightarrow P[z < k] = 0,0029$ (no está en la tabla, k es negativo) $k = -s$
 $P[z \geq -s] = 0,9971 \Rightarrow P[z \leq s] = 0,9971 \Rightarrow s = 2,76 \Rightarrow k = -2,76$
 h) $P[z \geq k] = 0,6 \Rightarrow P[z < k] = 0,4$ (no está en la tabla, k es negativo) $k = -s$
 $P[z \geq -s] = 0,6 \Rightarrow P[z \leq s] = 0,6 \Rightarrow s = 0,25 \Rightarrow k = -0,25$

CALCULO DE PROBABILIDADES EN UNA DISTRIBUCIÓN $N(\mu, \sigma)$

Como ya sabemos, las probabilidades en dos distribuciones normales cualesquiera se reparten de forma análoga. Por tanto, para calcular probabilidades en una distribución $N(\mu, \sigma)$, la relacionaremos con la $N(0,1)$ para la cual disponemos del recurso de las tablas.

Si x es $N(\mu, \sigma)$, para calcular la probabilidad $P[b < x < k]$ se procede del siguiente modo: $P[b < x < k] = P\left[\frac{b-\mu}{\sigma} < z < \frac{k-\mu}{\sigma}\right]$. El cambio $k \Rightarrow \frac{k-\mu}{\sigma}$ se llama tipificación de la variable. La variable ya tipificada sigue una distribución $N(0,1)$

Ejemplo 9 : En una distribución $N(66,8)$, calcular:

$$a) P[x < 70] = P\left[\frac{x - \mu}{\sigma} < \frac{70 - 66}{8}\right] = P[z < 0,5] = \phi(0,5) = 0,6915$$

$$b) P[x > 80] = P\left[\frac{x - \mu}{\sigma} > \frac{80 - 66}{8}\right] = P[z > 1,75] = 1 - P[z < 1,75] = 1 - \phi(1,75) = 1 - 0,9599 = 0,0401$$

$$c) P[70 < x < 80] = P\left[\frac{70 - 66}{8} < \frac{x - \mu}{\sigma} < \frac{80 - 66}{8}\right] = P[0,5 < z < 1,75] = P[z < 1,75] - P[z \leq 0,5] = \phi(1,75) - \phi(0,5) = 0,9599 - 0,6915 = 0,2684$$

Ejemplo 10 : En una distribución $N(66,8)$, calcula k para que se den las siguientes igualdades:

$$a) P[x \leq k] = 0,9788 \Rightarrow P\left[\frac{x - \mu}{\sigma} \leq \frac{k - 66}{8}\right] = 0,9788 \Rightarrow \frac{k - 66}{8} = 2,3 \Rightarrow k = 8,2,3 + 66 = 84,4$$

$$b) P[x \leq k] = 0,15 \Rightarrow P\left[\frac{x - \mu}{\sigma} \leq \frac{k - 66}{8}\right] = 0,15 \text{ (0,15 no está en la tabla)} \frac{k - 66}{8} = -s$$

$$P[z \leq -s] = 0,15 \Rightarrow P[z \geq s] = 0,15 \Rightarrow P[z < s] = 0,85 \Rightarrow s = 1,04 \Rightarrow \frac{k - 66}{8} = -1,04 \Rightarrow k = -1,04 \cdot 8 + 66 = 57,68$$

$$c) P[x \geq k] = 0,6808 \Rightarrow P[x < k] = 0,3192 \Rightarrow P\left[\frac{x - \mu}{\sigma} < \frac{k - 66}{8}\right] = 0,3192 \text{ (no está en la tabla)} \frac{k - 66}{8} = -s$$

$$P[z < -s] = 0,3192 \Rightarrow P[z \geq s] = 0,3192 \Rightarrow P[z < s] = 0,6808 \Rightarrow s = 0,47 \Rightarrow \frac{k - 66}{8} = -0,47 \Rightarrow k = -0,47 \cdot 8 + 66 = 62,24$$

Ejercicio 11 : En una distribución $N(18,4)$, halla las siguientes probabilidades:

$$a) P[x \leq 20] = P\left[\frac{x - \mu}{\sigma} \leq \frac{20 - 18}{4}\right] = P[z \leq 0,5] = \phi(0,5) = 0,6915$$

$$b) P[x \geq 16,5] = P\left[\frac{x - \mu}{\sigma} \geq \frac{16,5 - 18}{4}\right] = P[z \geq -0,375] = P[z \leq 0,375] = \frac{0,6443 + 0,6480}{2} = 0,64615$$

$$c) P[x \leq 11] = P\left[\frac{x - \mu}{\sigma} \leq \frac{11 - 18}{4}\right] = P[z \leq -1,75] = P[z \geq 1,75] = 1 - P[z < 1,75] = 1 - \phi(1,75) = 1 - 0,9599 = 0,0401$$

$$d) P[19 \leq x \leq 23] = P\left[\frac{19 - 18}{4} \leq \frac{x - \mu}{\sigma} \leq \frac{23 - 18}{4}\right] = P[0,25 \leq z \leq 1,25] = P[z \leq 1,25] - P[z < 0,25] = \phi(1,25) - \phi(0,25) = 0,8944 - 0,5987 = 0,2957$$

$$e) P[11 \leq x < 25] = P\left[\frac{11 - 18}{4} \leq \frac{x - \mu}{\sigma} < \frac{25 - 18}{4}\right] = P[-1,75 \leq z \leq 1,75] = P[z \leq 1,75] - P[z < -1,75] = \phi(1,75) - P[z \geq 1,75] = \phi(1,75) - [1 + P[z < 1,75]] = \phi(1,75) - 1 + \phi(1,75) = 0,9599 \cdot 2 - 1 = 0,9198$$

Ejercicio 12 : En una distribución $N(6;0,9)$, calcula k para que se den las siguientes igualdades:

$$a) P[x \leq k] = 0,9772 \Rightarrow P\left[\frac{x - \mu}{\sigma} \leq \frac{k - 6}{0,9}\right] = 0,9772 \Rightarrow \frac{k - 6}{0,9} = 2 \Rightarrow k = 7,8$$

$$b) P[x \leq k] = 0,8 \Rightarrow P\left[\frac{x - \mu}{\sigma} \leq \frac{k - 6}{0,9}\right] = 0,8 \Rightarrow \frac{k - 6}{0,9} = 0,84 \Rightarrow k = 8,756$$

$$c) P[x \leq k] = 0,3 \Rightarrow P\left[\frac{x - \mu}{\sigma} \leq \frac{k - 6}{0,9}\right] = 0,3 \text{ (No está en la tabla)} \Rightarrow \frac{k - 6}{0,9} = -s$$

$$P[z \leq -s] = 0,3 \Rightarrow P[z \geq s] = 0,3 \Rightarrow P[z < s] = 0,7 \Rightarrow s = 0,52 \Rightarrow \frac{k - 6}{0,9} = -0,52 \Rightarrow k = 5,532$$

d) $P[x \geq k] = 0,6331 \Rightarrow P[x < k] = 0,3669 \Rightarrow P\left[\frac{x - \mu}{\sigma} \leq \frac{k - 6}{0,9}\right] = 0,3669$ (No está en la tabla) \Rightarrow

$$\frac{k - 6}{0,9} = -s \Rightarrow P[z \leq -s] = 0,3669 \Rightarrow P[z \geq s] = 0,3669 \Rightarrow P[z < s] = 0,6331 \Rightarrow s = 0,34 \Rightarrow \frac{k - 6}{0,9} = -0,34 \Rightarrow k = 5,694$$

12.2 – INTERVALOS CARACTERÍSTICOS

Si la variable x tiene una distribución de media μ , se llama **intervalo característico** correspondiente a una probabilidad p a un intervalo centrado en la media, $(\mu - k, \mu + k)$ tal que la probabilidad de que x pertenezca a dicho intervalo es p:

$$P[\mu - k < x < \mu + k] = p$$

INTERVALOS CARACTERÍSTICOS EN DISTRIBUCIONES N(0,1)

En una distribución normal N(0,1), si $(-k, k)$ es el intervalo característico correspondiente a una probabilidad p, es decir, si $P[-k < z < k] = p = 1 - \alpha$

$$\Rightarrow P[z > z_{\alpha/2}] = \alpha/2 \Rightarrow P[z \leq z_{\alpha/2}] = 1 - \alpha/2 \Rightarrow \text{Intervalo característico} = (-z_{\alpha/2}, z_{\alpha/2})$$

Ejemplo 13: Si $Z \sim N(0,1)$ calcular el intervalo característico con una probabilidad del 90%

$$P[Z \leq z_{\alpha/2}] = 1 - \frac{\alpha}{2} \Rightarrow \left[\begin{array}{l} 1 - \alpha = p \rightarrow 1 - \alpha = 0,9 \rightarrow \alpha = 0,1 \\ P[Z \leq z_{\alpha/2}] = 1 - \frac{0,1}{2} = 0,95 \xrightarrow{\text{TablaN(0,1)}} z_{\alpha/2} = 1,645 \end{array} \right.$$

Intervalo característico $(-1,645, 1,645)$

Significado: Z está en el intervalo $(-1,645, 1,645)$ con una probabilidad del 90%

Ejercicio 14: Si $Z \sim N(0,1)$ calcular el intervalo característico con una probabilidad del 99%

$$P[Z \leq z_{\alpha/2}] = 1 - \frac{\alpha}{2} \Rightarrow \left[\begin{array}{l} 1 - \alpha = p \rightarrow 1 - \alpha = 0,99 \rightarrow \alpha = 0,01 \\ P[Z \leq z_{\alpha/2}] = 1 - \frac{0,01}{2} = 0,995 \xrightarrow{\text{TablaN(0,1)}} z_{\alpha/2} = 2,575 \end{array} \right.$$

Intervalo característico $(-2,575, 2,575)$

Significado: Z está en el intervalo $(-2,575, 2,575)$ con una probabilidad del 99%

(Lógicamente el intervalo es mayor que el obtenido en el ejemplo 1)

INTERVALOS CARACTERÍSTICOS EN DISTRIBUCIONES N(μ, σ)

En una distribución normal N(μ, σ), si $(\mu - k, \mu + k)$ es el intervalo característico correspondiente a una probabilidad p, es decir, si $P[\mu - k < X < \mu + k] = p = 1 - \alpha$

$$\Rightarrow P[z > z_{\alpha/2}] = \alpha/2 \Rightarrow P[z \leq z_{\alpha/2}] = 1 - \alpha/2 \Rightarrow \text{Intervalo característico} = (\mu - z_{\alpha/2} \cdot \sigma, \mu + z_{\alpha/2} \cdot \sigma)$$

Ejemplo 15: Si $X \sim N(173,6)$ calcular el intervalo característico con una probabilidad del 90%

$$P[Z \leq z_{\alpha/2}] = 1 - \frac{\alpha}{2} \Rightarrow \left[\begin{array}{l} 1 - \alpha = p \rightarrow 1 - \alpha = 0,9 \rightarrow \alpha = 0,1 \\ P[Z \leq z_{\alpha/2}] = 1 - \frac{0,1}{2} = 0,95 \xrightarrow{\text{TablaN(0,1)}} z_{\alpha/2} = 1,645 \end{array} \right.$$

Intervalo característico $(173 - 1,645 \cdot 6, 173 + 1,645 \cdot 6) = (163,13; 182,87)$

Significado: X está en el intervalo $(163,13; 182,87)$ con una probabilidad del 90%

Ejercicio 16: Si $X \sim N(173,6)$ calcular el intervalo característico con una probabilidad del 95%

$$P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} \Rightarrow \begin{cases} 1 - \alpha = p \rightarrow 1 - \alpha = 0,95 \rightarrow \alpha = 0,05 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,05}{2} = 0,975 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 1,96 \end{cases}$$

Intervalo característico $(173 - 1,96 \cdot 6, 173 + 1,96 \cdot 6) = (161,24; 184,76)$

Significado: X está en el intervalo $(161,24; 184,76)$ con una probabilidad del 95%

Ejercicio 17: El peso de las cajas de leche siguen una distribución normal de media 1 Kg y desviación típica 50 gr. Hallar el intervalo característico con una probabilidad del 95%

$X \sim N(1; 0,05)$

$$P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} \Rightarrow \begin{cases} 1 - \alpha = p \rightarrow 1 - \alpha = 0,95 \rightarrow \alpha = 0,05 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,05}{2} = 0,975 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 1,96 \end{cases}$$

Intervalo característico $(1 - 1,96 \cdot 0,05, 1 + 1,96 \cdot 0,05) = (0,902; 1,098)$

Significado: Una caja de leche pesa entre 0,902 y 1,098 Kg con una probabilidad del 95%

12.3 – DISTRIBUCIÓN DE LAS MEDIAS MUESTRALES

DISTRIBUCIÓN DE LAS MEDIAS MUESTRALES

Dada una población de media μ y desviación típica σ , no necesariamente normal, la distribución de las medias de las muestras de tamaño n:

- Tiene la misma media, μ , que la población.
- Su desviación típica es σ/\sqrt{n} y, por consiguiente, disminuye al aumentar n.
- Cuando $n \geq 30$ es prácticamente normal.

$$\left. \begin{array}{l} \text{Si } X \approx N(\mu, \sigma) \\ \text{ó} \\ \text{Si } n > 30 \end{array} \right\} \Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$P[Z \leq z_{\alpha/2}] = 1 - \alpha/2 \Rightarrow \text{Intervalo característico} = \left(\mu - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \mu + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

Ejemplo 18: En una distribución $N(20,6)$, tomamos muestras de tamaño 64.

a) ¿Cuál es la distribución de las medias de la muestra?

$$\text{Como } X \sim N(20,6) \Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(20, \frac{6}{\sqrt{64}}\right) = N(20; 0,75)$$

Es decir, es una Normal de media 20 y desviación típica 0,75

b) ¿Cuál es la probabilidad de extraer una muestra cuya media esté comprendida entre 19 y 21?

$$P[19 \leq \bar{X} \leq 21] \underset{\bar{X} \sim N(20,0,75) \text{ Tipificamos}}{=} P\left[\frac{19-20}{0,75} \leq Z \leq \frac{21-20}{0,75}\right] = P[-1,33 \leq Z \leq 1,33] =$$

$$= P[Z \leq 1,33] - P[Z < -1,33] = P[Z \leq 1,33] - P[Z > 1,33] = P[Z \leq 1,33] - (1 - P[Z \leq 1,33]) =$$

$$= 0,9082 - 1 + 0,9082 = 0,8164$$

Miramos en la tabla de la $N(0,1)$

Ejercicio 19 : Los parámetros de una variable son $\mu = 16,4$ $\sigma = 4,8$. Nos disponemos a extraer una muestra de $n = 400$ individuos.

a) Halla el intervalo característico para la media muestral correspondiente a una probabilidad $p = 0,99$.

$$P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} \Rightarrow \begin{cases} 1 - \alpha = p \rightarrow 1 - \alpha = 0,99 \rightarrow \alpha = 0,01 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,01}{2} = 0,995 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 2,575 \end{cases}$$

$$(\mu - Z_{\alpha/2}\sigma, \mu + Z_{\alpha/2}\sigma)$$

Intervalo característico $(16 - 2,575 \cdot 4,8, 16 + 2,575 \cdot 4,8) = (3,64; 28,36)$

b) Calcula $P [16 < \bar{x} < 17]$

Como $n = 400 > 30 \Rightarrow \bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}}) = N(16, \frac{4,8}{\sqrt{400}}) = N(16; 0,24)$

$$P[16 \leq \bar{X} \leq 17] \underset{x \sim N(16; 0,24) \text{ Tipificamos}}{=} P\left[\frac{16-16}{0,24} \leq Z \leq \frac{17-16}{0,24}\right] = P[0 \leq Z \leq 4,17] = P[Z \leq 4,17] - P[Z < 0] = \phi(4,17) - \phi(0) \underset{\text{Miramos en la tabla de la } N(0,1)}{=} 1 - 0,5 = 0,5$$

DISTRIBUCIÓN DE LA SUMA DE TODOS LOS INDIVIDUOS DE LA MUESTRA

Puesto que $\sum_{i=1}^n x_i = n\bar{x}$, sabemos que $\sum_{i=1}^n x_i$ se distribuye normal de media $n\mu$ y desviación típica

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

$$\left. \begin{array}{l} \text{Si } X \approx N(\mu, \sigma) \\ \text{ó} \\ \text{Si } n > 30 \end{array} \right\} \Rightarrow \sum_{i=1}^n x_i \sim N(n\mu, \sigma\sqrt{n})$$

$$P[z \leq z_{\alpha/2}] = 1 - \alpha/2 \Rightarrow \text{Intervalo característico} = (n\mu - z_{\alpha/2} \cdot \sigma\sqrt{n}, n\mu + z_{\alpha/2} \cdot \sigma\sqrt{n})$$

Ejemplo 20 : Los sueldos, en euros, de los empleados de una fábrica se distribuyen $N(1200, 400)$. Se elige al azar una muestra de 25 de ellos. ¿Cuál es la probabilidad de que la suma de sus sueldos sea superior a 35.000 euros?

$$\text{Si } X \approx N(1200, 400) \Rightarrow \sum_{i=1}^n x_i \sim N(25 \cdot 1200, 400\sqrt{25}) = N(30.000, 2000)$$

$$P\left[\sum_{i=1}^n x_i > 35000\right] = P\left[Z > \frac{35000 - 30000}{2000}\right] = P[Z > 2,5] = 1 - P[Z \leq 2,5] = 1 - \phi(2,5) = 1 - 0,9938 = 0,0062$$

Halla el intervalo característico para las sumas de 25 individuos, correspondientes a una probabilidad del 0,9.

$$P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} \Rightarrow \begin{cases} 1 - \alpha = p \rightarrow 1 - \alpha = 0,9 \rightarrow \alpha = 0,1 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,1}{2} = 0,95 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 1,645 \end{cases}$$

$$(n\mu - Z_{\alpha/2}\sigma\sqrt{n}, n\mu + Z_{\alpha/2}\sigma\sqrt{n})$$

$$\text{Intervalo característico } (25 \cdot 1200 - 1,645 \cdot 400\sqrt{25}, 25 \cdot 1200 + 1,645 \cdot 400\sqrt{25}) = (26710; 33290)$$

Significado: La suma de los sueldos de 25 empleados está entre 26710 y 33290 euros con una probabilidad del 90%

Ejemplo 21: Las bolsas de azúcar envasadas por una cierta máquina tienen de media 500 gr y desviación típica 35 gr. Las bolsas se empaquetan en cajas de 100 unidades.

a) Calcular la probabilidad de que la media de los pesos de las bolsas de un paquete sea menor que 495 gr.

$$X \sim N(500, 35)$$

$$P[\bar{X} < 495] \underset{\text{Como } X \sim N(500,35) \Rightarrow \bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}}) = N(500, \frac{35}{\sqrt{100)}) = N(500;3,5) \Rightarrow \text{Tipificamos}}{=} P[Z < \frac{495 - 500}{3,5}] = P[Z < -1,43]$$

$$= P[Z > 1,43] = 1 - P[Z \leq 1,43] \underset{\text{Miramos en la tabla de la } N(0,1)}{=} 1 - 0,9236 = 0,0764$$

b) Hallar el intervalo característico de la media muestral para una probabilidad del 95%

$$P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2}$$

$$\left. \begin{matrix} (\mu - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \mu + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) \end{matrix} \right\} \Rightarrow \left[\begin{matrix} 1 - \alpha = p \rightarrow 1 - \alpha = 0,95 \rightarrow \alpha = 0,05 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,05}{2} = 0,975 \xrightarrow{\text{Tabla } N(0,1)} Z_{\alpha/2} = 1,96 \end{matrix} \right.$$

$$\text{Intervalo característico } (500 - 1,96 \cdot \frac{35}{\sqrt{100}}; 500 + 1,96 \cdot \frac{35}{\sqrt{100}}) = (493,1; 506,9)$$

Significado: La media de los pesos están en el intervalo (493,1;506,9) con una probabilidad del 95%

c) Calcular la probabilidad de que una caja de 100 bolsas pese más de 51 Kg.

$$P[\sum X_i > 51.000] \underset{\text{Como } X \sim N(500,35) \Rightarrow \sum X_i \sim N(n\mu, \sigma \cdot \sqrt{n}) = N(100 \cdot 500, 35 \cdot \sqrt{100}) = N(50.000; 350) \Rightarrow \text{Tipificamos}}{=} P[Z > \frac{51000 - 50000}{35}] = P[Z > 2,86] = 1 - P[Z \leq 2,86] \underset{\text{Miramos en la tabla de la } N(0,1)}{=} 1 - 0,9979 = 0,0021$$

d) ¿Cuál es la probabilidad de que una caja pese más de 550 gr?

$$P[X > 550] \underset{X \sim N(500,35) \rightarrow \text{Tipificamos}}{=} P[Z > \frac{550 - 500}{35}] = P[Z > 1,43] = 1 - P[Z \leq 1,43] \underset{\text{Miramos en la tabla de la } N(0,1)}{=} 1 - 0,9236 = 0,0764$$

Ejercicio 22 : Los pesos en kilogramos de los soldados de una promoción siguen una distribución normal N(69,8). Las guardias en un regimiento están formadas por 12 soldados.

a) Hallar la probabilidad de que la media de los pesos de los soldados de una guardia sea superior a 71 Kg.

$$\text{Como } X \sim N(69,8) \Rightarrow \bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}}) = N(69, \frac{8}{\sqrt{12}}) = N(69; 2,31)$$

$$P[\bar{X} > 71] = P[Z > \frac{71 - 69}{2,31}] = P[Z > 0,87] = 1 - P[Z \leq 0,87] = 1 - \phi(0,87) = 1 - 0,8078 = 0,1922$$

b) Obtener el intervalo característico para \bar{x} correspondiente a una probabilidad de 0,9

$$P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2}$$

$$\left. \begin{matrix} (\mu - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \mu + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) \end{matrix} \right\} \Rightarrow \left[\begin{matrix} 1 - \alpha = p \rightarrow 1 - \alpha = 0,9 \rightarrow \alpha = 0,1 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,1}{2} = 0,95 \xrightarrow{\text{Tabla } N(0,1)} Z_{\alpha/2} = 1,645 \end{matrix} \right.$$

$$\text{Intervalo característico } (69 - 1,645 \cdot \frac{8}{\sqrt{12}}, 69 + 1,645 \cdot \frac{8}{\sqrt{12}}) = (65,2; 72,8)$$

Significado: La media de los pesos están en el intervalo (65,2; 72,8) con una probabilidad del 90%

c) ¿Cuál es la probabilidad de que la suma de los pesos de los soldados de una guardia sea menor que 800 Kgs?

$$\text{Si } X \approx N(69,8) \Rightarrow \sum_{i=1}^n x_i \sim N(12 \cdot 69, 8\sqrt{12}) = N(828, 27,71)$$

$$P\left[\sum_{i=1}^n x_i < 800\right] = P\left[Z < \frac{800 - 828}{27,71}\right] = P[Z < -1,01] = P[Z > 1,01] = 1 - P[Z \leq 1,01] = 1 - \phi(1,01) = 1 - 0,8438 = 0,1562$$

d) ¿Cuál es la probabilidad de que un miembro de la guardia, elegido al azar, pese más de 93 Kg?

$$P[X > 93] = P\left[Z > \frac{93 - 69}{8}\right] = P[Z > 3] = 1 - P[Z \leq 3] = 1 - \phi(3) = 1 - 0,9987 = 0,0013$$

12.4 – EN QUÉ CONSISTE LA ESTADÍSTICA INFERENCIAL

ESTIMACIÓN PUNTUAL Y ESTIMACIÓN POR INTERVALOS

Los parámetros de la población se pueden estimar a partir de los de la muestra. Así:

La media muestral, \bar{x} , sirve para estimar la media poblacional, μ .

La desviación típica muestral, s , es una estimación de la desviación típica poblacional, σ

La estimación puntual (el valor de μ es aproximadamente \bar{x}), sirve de poco mientras desconozcamos cuál es el grado de aproximación de \bar{x} a μ .

La estimación por intervalos: A partir de una muestra de tamaño n podemos estimar el valor de un parámetro de la población del siguiente modo:

- Dando un intervalo dentro del cual confiamos que esté el parámetro. Se llama intervalo de confianza.
- Hallando la probabilidad de que tal cosa ocurra. A dicha probabilidad se le llama nivel de confianza.

Cuanto mayor sea el tamaño de la muestra, mayor eficacia tendremos en nuestra estimación.

La eficacia de esta estimación se manifiesta de dos formas:

- En el tamaño del intervalo (cuanto más pequeño, más precisos estamos siendo)
- En el nivel de confianza (más nivel de confianza significa más seguridad en la estimación).

Tamaño de la muestra, longitud del intervalo y nivel de confianza son tres variables estrechamente relacionadas. Conocidas dos de ellas obtendremos la tercera.

12.5 – INTERVALO DE CONFIANZA PARA LA MEDIA

Se desea estimar la media, μ , de una población cuya desviación típica, σ , es conocida.

Para ello se recuerde a una muestra de tamaño n en la cual se obtiene una media muestral, \bar{x} .

Si la población de partida es normal, o si el tamaño de la muestra es $n \geq 30$, entonces el **intervalo de confianza** de μ con un nivel de confianza de $(1-\alpha) \cdot 100\%$ es: $\left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$

Ejemplo 23 : Deseamos valorar el grado de conocimientos en historia de una población de varios miles de alumnos. Sabemos que la desviación típica es 2,3. Nos proponemos estimar la media poblacional pasando una prueba a 100 alumnos.

a) Calcular el intervalo característico para la media muestral correspondiente a una probabilidad de 0,95.

Como $n = 100 > 30$

$$\left. \begin{aligned} P[Z \leq Z_{\alpha/2}] &= 1 - \frac{\alpha}{2} \\ (\mu - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \mu + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) \end{aligned} \right\} \Rightarrow \left[\begin{aligned} 1 - \alpha = p \rightarrow 1 - \alpha = 0,95 \rightarrow \alpha = 0,05 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,05}{2} = 0,975 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 1,96 \end{aligned} \right.$$

Intervalo característico $(\mu - 1,96 \cdot \frac{2,3}{\sqrt{100}}; \mu + 1,96 \cdot \frac{2,3}{\sqrt{100}}) = (\mu - 0,45; \mu + 0,45)$

Significado: La media muestral dista menos de 0,45 de la poblacional con una probabilidad del 95%

b) Una vez realizada la prueba a 100 alumnos concretos, se ha obtenido una media de 6,32. Hallar el intervalo de confianza con un nivel de significación del 95%

$$\left. \begin{aligned} P[Z \leq Z_{\alpha/2}] &= 1 - \frac{\alpha}{2} \\ (\bar{X} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) \end{aligned} \right\} \Rightarrow \left[\begin{aligned} 1 - \alpha = p \rightarrow 1 - \alpha = 0,95 \rightarrow \alpha = 0,05 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,05}{2} = 0,975 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 1,96 \end{aligned} \right.$$

Intervalo de confianza $(6,32 - 1,96 \cdot \frac{2,3}{\sqrt{100}}; 6,32 + 1,96 \cdot \frac{2,3}{\sqrt{100}}) = (5,87; 6,77)$

Significado: La media poblacional está en el intervalo (5,87;6,77) con un nivel de confianza del 95%

Ejemplo 24 : Para estimar la media de los resultados que obtendrán al resolver un cierto test los alumnos de 4º de ESO de toda una comunidad autónoma, se les pasa dicho test a 400 de ellos escogidos al azar. Los resultados obtenidos vienen recogidos en la siguiente tabla

x_i	1	2	3	4	5
f_i	24	80	132	101	63

A partir de ellos, estima con un nivel de confianza del 95% el valor de la media poblacional.

Como $n = 400 > 30$ El intervalo de confianza $\mu \in (\bar{X} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$

Hallamos al media y la desviación típica de la muestra:

x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	24	24	24
2	80	160	320
3	132	396	1180
4	101	404	1616
5	63	315	1575
	$N = 400$	1299	4723

$$\bar{X} = \frac{\sum x_i \cdot f_i}{N} = \frac{1299}{400} = 3,2475$$

$$\sigma = \sqrt{\frac{\sum x_i^2 \cdot f_i}{N} - \bar{X}^2} = \sqrt{\frac{4723}{400} - 3,2475^2} = \sqrt{1,26} = 1,12$$

$$\left. \begin{aligned} P[Z \leq Z_{\alpha/2}] &= 1 - \frac{\alpha}{2} \\ (\bar{X} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) \end{aligned} \right\} \Rightarrow \left[\begin{aligned} 1 - \alpha = 0,95 \rightarrow \alpha = 0,05 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,05}{2} = 0,975 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 1,96 \end{aligned} \right.$$

$\mu \in (3,2475 - 1,96 \cdot \frac{1,12}{\sqrt{400}}; 3,2475 + 1,96 \cdot \frac{1,12}{\sqrt{400}}) = (3,14; 3,36)$

Significado: Tenemos una confianza del 95% de que la nota media de la población total esté comprendida entre (3,14; 3,36).

Ejercicio 25 : De una variable estadística, conocemos la desviación típica 8, pero desconocemos la media. Para estimarla, extraemos una muestra de tamaño $n = 60$ cuya media es 37. Estima la media poblacional mediante un intervalo de confianza del 99%.

Como $n = 60 > 30$

$$\left. \begin{aligned} P[Z \leq Z_{\alpha/2}] &= 1 - \frac{\alpha}{2} \\ (\bar{X} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) \end{aligned} \right\} \Rightarrow \begin{cases} 1 - \alpha = p \rightarrow 1 - \alpha = 0,99 \rightarrow \alpha = 0,01 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{0,01}{2} = 0,995 \xrightarrow{\text{TablaN}(0,1)} Z_{\alpha/2} = 2,575 \end{cases}$$

Intervalo de confianza $(37 - 2,575 \cdot \frac{8}{\sqrt{60}} ; 37 + 2,575 \cdot \frac{8}{\sqrt{60}}) = (34,34; 39,66)$

Significado: La media poblacional está en el intervalo (34,34;39,66) con un nivel de confianza del 99%

12.6 – RELACIÓN ENTRE NIVEL DE CONFIANZA, ERROR ADMISIBLE Y TAMAÑO DE LA MUESTRA

El valor $E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ se llama **error máximo admisible**.

Depende de α y de n del siguiente modo:

- Cuanto mayor sea el tamaño de la muestra, menor es E (más estrecho es el intervalo, es decir, más afinaremos en la estimación).
- Cuanto mayor sea $1 - \alpha$ (es decir, cuanto más seguros queramos estar de nuestra estimación), mayor es E.

Cuanto mayor es $1 - \alpha$, mayor es $z_{\alpha/2}$ y, por tanto, mayor es E.

E, n y α son tres variables estrechamente relacionadas. Conocidas dos de ellas obtendremos la

tercera despejando de la fórmula: $E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

Ejemplo 26 : La desviación típica de los resultados de las distintas mediciones que se realizan para calcular la duración de un proceso es 0,5 sg. ¿Cuál es el número de medidas que hay que realizar para que con un 99% de confianza, el error de la estimación no exceda de 0,1 sg.?

$$\left[\begin{array}{l} \sigma = 0,5 \\ n = ? \\ 1 - \alpha = 0,99 \\ E < 0,1 \end{array} \right] \Rightarrow \left[\begin{array}{l} E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} = 1 - \frac{0,01}{2} = 0,995 \rightarrow Z_{\alpha/2} = 2,575 \end{array} \right.$$

$$0,1 = 2,575 \cdot \frac{0,5}{\sqrt{n}} \Rightarrow \sqrt{n} = 12,875 \Rightarrow n = 165,77 \Rightarrow n = 166 \Rightarrow \text{Debemos tomar 166 medidas}$$

Ejemplo 27 : Al medir el tiempo de reacción, un psicólogo sabe que la desviación típica de las distintas mediciones del mismo es 0,5 sg. Desea estimar el tiempo medio de reacción con un error máximo de 0,1 sg, para lo cual realiza 100 experiencias. Dí con que nivel de confianza podrá dar el intervalo $(\bar{X} - 0,1; \bar{X} + 0,1)$

$$\left[\begin{array}{l} \sigma = 0,5 \\ n = 100 \\ 1 - \alpha = ? \\ E = 0,1 \end{array} \right] \Rightarrow \left[\begin{array}{l} E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow 0,1 = Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{100}} \Rightarrow Z_{\alpha/2} = 2 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} \\ P[Z \leq 2] = 0,9772 \end{array} \right. \text{Miramos en la tabla de la } N(0,1)$$

$$1 - \frac{\alpha}{2} = 0,9772 \Rightarrow \alpha = 0,0456 \Rightarrow 1 - \alpha = 0,9544 \Rightarrow 95,44\% \Rightarrow \text{El nivel de confianza es del } 95,44\%$$

Ejemplo 28 : Un ganadero de reses bravas quiere estimar el peso medio de los toros de su ganadería con un nivel de confianza del 95%. Para ello toma una muestra de 30 toros y los pesa. Obtiene una media de 507 Kg y una desviación típica de 32 Kg. ¿Cuál es el error cometido?

$$\left[\begin{array}{l} \sigma = 32 \\ n = 30 \\ 1 - \alpha = 0,95 \\ E = ? \end{array} \right] \Rightarrow \left[\begin{array}{l} E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} = 1 - \frac{0,05}{2} = 0,975 \rightarrow Z_{\alpha/2} = 1,96 \end{array} \right]$$

$$E = 1,96 \cdot \frac{32}{\sqrt{30}} = 11,45 \text{ Kg} \Rightarrow \text{El error es de 11,45 Kg.}$$

Ejercicio 29: Un coronel desea estimar la estatura media de todos los soldados de su regimiento con un error menor que 0,5 cm utilizando una muestra de 30 soldados. Sabiendo que $\sigma = 5,3$ cm. ¿Cuál será el nivel de confianza con el que se realiza la estimación?

$$\left[\begin{array}{l} \sigma = 5,3 \\ n = 30 \\ 1 - \alpha = ? \\ E = 0,5 \end{array} \right] \Rightarrow \left[\begin{array}{l} E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow 0,5 = Z_{\alpha/2} \cdot \frac{5,3}{\sqrt{30}} \Rightarrow Z_{\alpha/2} = 0,52 \\ P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} \\ P[Z \leq 0,52] \underset{\text{Miramos la tabla de } N(0,1)}{=} 0,6985 \end{array} \right]$$

$$1 - \frac{\alpha}{2} = 0,6985 \Rightarrow \alpha = 0,603 \Rightarrow 1 - \alpha = 0,397 \Rightarrow 39,7\%$$

El nivel de confianza es del 39,7%

Ejercicio 30 : El cociente intelectual de un cierto colectivo tiene una media μ desconocida y una desviación típica $\sigma = 8$. ¿De qué tamaño debe ser la muestra con la cual se estime la media con un nivel de confianza del 99% y un error admisible $E = 3$?

$$\left[\begin{array}{l} \sigma = 8 \\ n = ? \\ 1 - \alpha = 0,99 \Rightarrow \alpha = 0,01 \\ E = 3 \end{array} \right] \Rightarrow \left[\begin{array}{l} P[Z \leq Z_{\alpha/2}] = 1 - \frac{\alpha}{2} = 1 - \frac{0,01}{2} = 0,995 \Rightarrow Z_{\alpha/2} = 2,575 \\ E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow 3 = 2,575 \cdot \frac{8}{\sqrt{n}} \Rightarrow \sqrt{n} = 6,87 \Rightarrow n = 47,15 \end{array} \right]$$

El tamaño de la muestra debe ser de 48.

EJERCICIOS REPASO (Libro páginas 294-297)