DISTRIBUCIONES EN EL MUESTREO

1.- INTRODUCCION

A partir de este tema vamos a desconocer los parámetros que definen a las poblaciones $(\mu, \sigma^2, \lambda, \pi, \sigma, ...)$.

Por lo tanto será necesario que calculemos unos estadísticos a partir de una m.a.s con los que realizaremos estimaciones sobre los parámetros desconocidos. También será necesario conocer, **la distribución de los estadísticos.**

Cuando las distribuciones de los colectivos sean desconocidas, siempre que las condiciones muestrales lo permitan, será de gran utilidad el **Teorema Central del Límite.**

- ✓ **Teorema Central del Límite:**

La suma n v.a. independientes, con distribución desconocida, se podrá aproximar a una distribución Normal, si el tamaño muestral "n" es lo suficientemente grande.

$$W = x_1 + x_2 + ... + x_n \xrightarrow{TCL} \mathbf{N}\,(\mu_W = \mu_1 + \mu_2 + ... + \mu_n\ ,\ \sigma_w^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2)$$

Si la muestra utilizada es una muestra aleatoria simple**, m.a.s. ,** las v.a. x_i, además de ser independientes, tendrán la misma media, varianza y distribución. Por tanto:

$$W = x_1 + x_2 + ... + x_n \xrightarrow{TCL} \mathbf{N}\,(\mu_W = n\mu\ ,\ \sigma_w^2 = n\sigma^2)$$

- ✓ **Media aritmética muestral,** $\bar{x} = \dfrac{x_1 + x_2 + ... + x_n}{n}$

Es el estadístico que vamos a utilizar para estimar la media poblacional μ. Por tanto, cuando utilicemos la $\bar{x}$ como estimador de la media, la podremos simbolizar de la siguiente manera: $\hat{\mu} = \bar{x}$

Salvo evidencia en contra, siempre que calculemos la media aritmética muestral **(** $\bar{x}$ **),** lo haremos utilizando un muestreo aleatorio simple (m.a.s.). Esto implica que las variables $(x_1, x_2, \dots\ x_n)$ son v.a. independientes, con la misma media y varianza (μ y σ^2), que el colectivo al que pertenecen. Por tanto :

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

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

Demostración de: $\mu_{\bar{x}}$ **y** $\sigma^2_{\bar{x}}$**:**

✓ $$\mu_{\bar{x}} = E_{\bar{x}} = \mathrm{E}\left(\frac{x_1 + x_2 + ... + x_n}{n}\right)$$

$$= \frac{E(x_1 + x_2 + ... + x_n)}{n}$$

$$= \frac{Ex_1 + Ex_2 + ... + Ex_n}{n}$$

$$= \frac{\mu + ... + \mu}{n} = \frac{n \cdot \mu}{n} = \mu$$

✓ $$\sigma^2_{\bar{x}} = Var_{\bar{x}} = \mathrm{Var}\left(\frac{x_1 + x_2 + ... + x_n}{n}\right)$$

$$= \frac{Var(x_1 + x_2 + ... + x_n)}{n^2}$$

$$= \frac{Varx_1 + Varx_2 + ... + Varx_n}{n^2}$$

$$= \frac{\sigma^2 + ... + \sigma^2}{n^2} = \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

2.- ESTADISTICOS Y SUS DISTRIBUCIONES

2.1. Estadístico $\bar{x}$ obtenido de un colectivo normal.

$$\bar{x} \in \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \in \mathrm{N}(0, 1)$$

2.2. Estadístico $\bar{x}$ obtenido de un colectivo no normal a través de una m.a.s. de tamaño ($n \geq 30$)

Como el tamaño muestral es lo suficientemente grande aplicaremos el T.C.L.

$$\bar{x} \xrightarrow{TCL} \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \xrightarrow{TCL} \mathrm{N}(0, 1)$$

2.3. Estadístico proporción muestral obtenido de un colectivo binario a través de una m.a.s. de tamaño ($n \geq 30$)

Sea un colectivo binario con parámetro **π**

P (x_i)	$1-\pi$	π
X	0	1

Sabemos: $\mu = \sum_{i=1}^{\infty} x_i \text{ P }(x_i) = \pi$

$$\sigma^2 = \sum_{i=1}^{\infty} (x_i - \mu)^2 \text{ P }(x_i) = \sum_{i=1}^{\infty} x_i^2 \text{ P }(x_i) - \mu_x^2 = \boldsymbol{\pi.(1-\pi)}$$

Si extraemos una m.a.s: (x_1, x_2, … x_n) de este colectivo, al estadístico $\bar{x}$ se le denomina: **Proporción muestral "P".** variable frecuencial, frecuencia muestral ó frecuencia relativa.

$$\mathbf{P} = \bar{x} = \frac{\sum x_i}{n} = \frac{x}{n} = \frac{n^o \, exitos}{n}$$

Por tanto, la distribución de la proporción muestral en muestras grandes, aplicando el **teorema central del límite** será:

$$\mathbf{P} \xrightarrow{TCL} \text{N}\left(\pi, \frac{\pi \cdot (1-\pi)}{n}\right)$$

$$\frac{\text{P} - \pi}{\sqrt{\frac{\pi.(1-\pi)}{n}}} \longrightarrow \text{N }(0 , 1)$$

Como podemos observar el estadístico "**P**", proporción muestral, es un caso particular de la media muestral $\bar{x}$, en un colectivo o población binaria