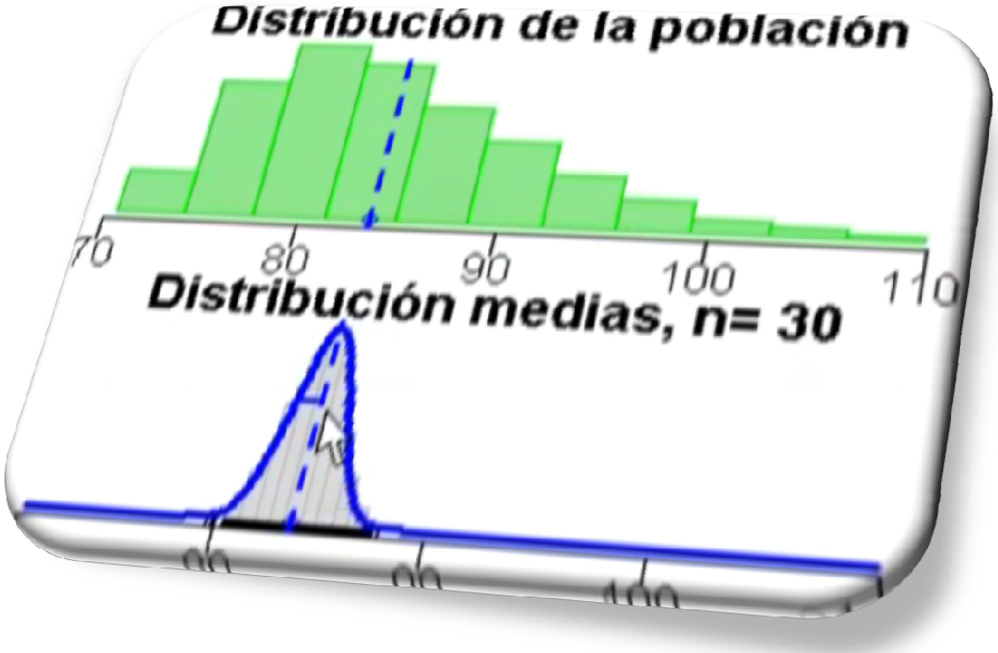
Asunto de Estado: Las muestras también se distribuyen



Distribución en el muestreo de una proporción

Supongamos que tenemos una variable aleatoria referida a una población en la que sólo puede tomar dos valores, que consideraremos éxito y fracaso.

Sea p la proporción de la población que corresponde al éxito. Llamaremos q=1-p a la proporción de individuos que corresponde al fracaso.

Tomamos muestras de tamaño n y para cada una de ellas hallamos la proporción de éxito, $\hat{p}$. Para cada muestra el valor de $\hat{p}$ será diferente. Estos valores dan lugar a una nueva variable aleatoria $\hat{P}$.

La variable aleatoria $\hat{P}$ se denomina **distribución en el muestreo de una proporción** y tiene las siguientes características:

1º Media: $\mu = p$

2º Desviación típica: $\sigma = \sqrt{\dfrac{p\cdot(1-p)}{n}}$

3º Para valores de $n \geq 30$ la distribución de $\hat{P}$ se aproxima a una $N\left(p, \sqrt{\dfrac{p\cdot(1-p)}{n}}\right)$ siempre $n\cdot p \geq 5$ y $n\cdot q \geq 5$.

Distribución en el muestreo de la media

Supongamos que tenemos una variable aleatoria cuantitativa referida a una población.

Tomamos muestras de tamaño n y para cada una de ellas hallamos la media, $\overline{X}$. Para cada muestra el valor de $\overline{X}$ será diferente. Estos valores dan lugar a una nueva variable aleatoria $\overline{X}$.

La variable aleatoria $\overline{X}$ se denomina **distribución en el muestreo de la media** y tiene las siguientes características:

1º Media: tiene la misma media que la población μ

2º Desviación típica: $\dfrac{\sigma}{\sqrt{n}}$

3º Para valores de n ≥ 30 la distribución de $\overline{X}$ se aproxima a una $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$

Educación Permanente

Distribución de las sumas muestrales

Supongamos que tenemos una variable aleatoria cuantitativa referida a una población.

Tomamos muestras de tamaño n y sumamos los resultados obtenidos al observar la variable en cada individuo de la muestra, t. Para cada muestra el valor de t será diferente. Estos valores dan lugar a una nueva variable aleatoria T .

La variable aleatoria **T** se denomina **distribución de las sumas muestrales** y tiene las siguientes características:

1º Media: $n \cdot \mu$

2º Desviación típica: $\sigma \cdot \sqrt{n}$

3º Para valores de $n \geq 30$ la distribución de T se aproxima a una $N\left(n \cdot \mu, \sigma \cdot \sqrt{n}\right)$

Distribución en el muestreo de la diferencia de las medias

Supongamos que tenemos una variable aleatoria cuantitativa referida a dos poblaciones. Estamos interesados en conocer si las medias de ambas poblaciones difieren.

Tomamos muestras de tamaño n_1 de la 1ª población y n_2 de la 2ª. Calculamos la media de cada muestra referida a cada población y las restamos, $\overline{X}_1 - \overline{X}_2$. Si repetimos este proceso sucesivamente, para cada muestra el valor de $\overline{X}_1 - \overline{X}_2$ será diferente. Estos valores dan lugar a una nueva variable aleatoria $\overline{X}_1 - \overline{X}_2$.

La variable aleatoria $\overline{X}_1 - \overline{X}_2$ se denomina **distribución en el muestreo de la diferencia de las medias** y tiene las siguientes características:

1º Media: $\mu_1 - \mu_2$

2º Desviación típica: $\sigma = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

3º Para valores de $n \geq 30$ la distribución de $\overline{X}_1 - \overline{X}_2$ se aproxima a una distribución $N\left(\mu_1 - \mu_2, \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}\right)$

Teorema central del Límite

Sea X una variable aleatoria de una población de media μ y desviación típica σ. Entonces se verifica que:

1º La distribución de las medias muestrales de tamaño n tiene una media μ y una desviación típica $\frac{\sigma}{\sqrt{n}}$

2º La distribución de las medias muestrales se aproxima a una normal si:

- La población de partida es normal, la distribución de las medias será normal, independientemente del tamaño de la muestra.
- Si la población de partida no es normal, la distribución de las medias se podrá aproximar por una normal cuando el tamaño de las muestras sea mayor o igual a 30 ($n \geq 30$).