6. DISTRIBUCIONES MUESTRALES

Dr. Edgar Acuna

http://academic.uprm.edu/eacuna

UNIVERSIDAD DE PUERTO RICO

RECINTO UNIVERSITARIO DE MAYAGUEZ

DISTRIBUCIONES MUESTRALES

Uno de los objetivos de la estadística es saber acerca del comportamiento de parámetros poblacionales tales como: la media (μ), la varianza (σ^2) o la proporción (p). Se extrae una muestra aleatoria de la población y se calcula el valor de un estadístico correspondiente, por ejemplo, la media muestral ($\bar{X}$), la varianza muestral (s^2) o la proporción muestral ($\hat{p}$). El valor del estadístico es aleatorio porque depende de los elementos elegidos en la muestra seleccionada. y, por lo tanto, el estadístico tiene una distribución de probabilidad la cual es llamada la Distribución Muestral del Estadístico.

6.1 Distribución de la Media Muestral cuando la población es normal

Se extraen muestras aleatorias de tamaño n de una población infinita Normal con media poblacional μ y varianza σ^2:

- La distribucion de las medias muestrales es también normal
- La media de las medias muestrales es igual a la media poblacional. Es decir, $\mu_{\bar{x}} = \mu$.
- La varianza de las medias muestrales es igual a la varianza poblacional dividida por n. En consecuencia la desviación estándar de las medias muestrales (llamada también el **error estándar** de la media muestral), esta dado por $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

Si la población fuera finita de tamaño N, se aplica el factor de correción: $\sqrt{\frac{N-n}{N-1}}$ al error estándar de la media muestral

6.2 El Teorema del Límite Central

De una población infinita con media μ y varianza σ^2 se extraen muestras aleatorias de tamaño n , entonces la media muestral se comporta aproximadamente como una variable aleatoria normal con media igual a la media poblacional y con varianza igual a la varianza poblacional dividida por el tamaño de la muestra, siempre que n sea grande. Esto es:

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, Estandarizando: $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

Ejemplo 6.1

Considerar una población que consiste de 3, 4, 6, 8, 10, 11, 12, 15, 20.

Solución:

1) Calculamos la media y desviación estándar de dicha población.

x=[3, 4, 6, 8, 10, 11, 12, 15, 20]

np. mean(x)

9.888

np.std(x,ddof=1)

5.418

2) Extraemos 50 muestras de tamaño 30 de dicha población,

a=np.random.choice(x, 1500, replace=True)

b=a.reshape(50,-1)

Ejemplo 6.1

3) Tercero, calculamos las medias de todas las 50 muestras usando:

medias1=b.mean(axis=1)

print medias1

9.996 9.96733333 9.42866667 9.95466667 9.821 9.17966667 10.11533333 ………….10.0022 10.5978 10.1598 10.8114 11.0782 10.1118 9.3902 9.7308

Las medidas estadísticas de la distribución de las medias muestrales son:

np. Mean(medias1)

9.723

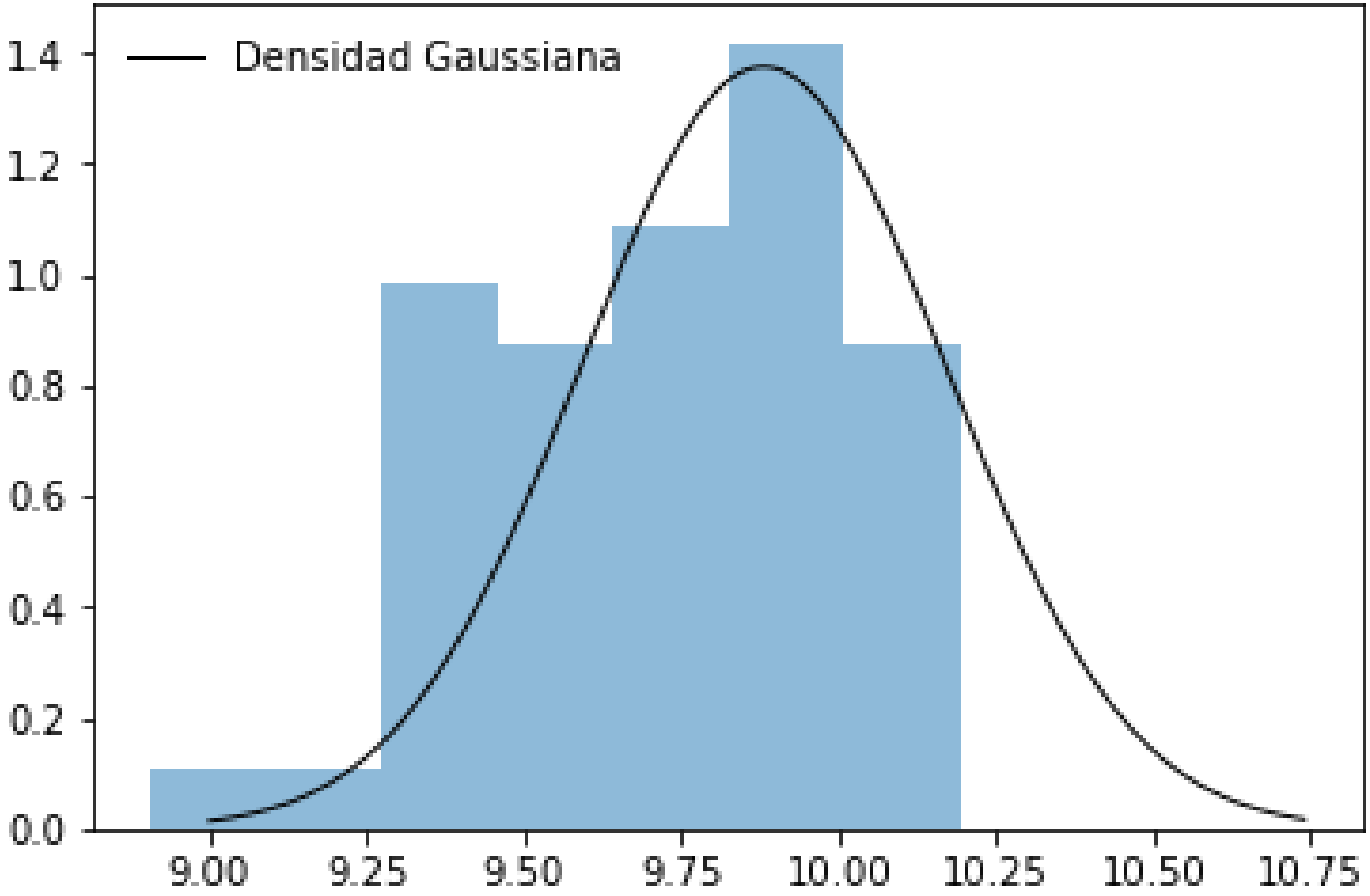
> sd(medias1,ddof=1)

[1] 0.291

Notar que la media de todas las medias muestras esta cerca de la media poblacional y de que su desviacion estandar esta cerca de

$$\frac{\sigma}{\sqrt{n}} = 5.42 / sqrt(30) = .290$$

Figura 6.1 Histograma de la distribución de las medias maestrales del Ejemplo 6.1



6.3 Distribución de la Proporción Muestral

Si de una población distribuida Binomialmente con probabilidad de éxito *p*, se extrae una muestra aleatoria de tamaño *n*, entonces se puede mostrar que la media de *X*: número de éxitos en la muestra, es $\mu = np$ y que su varianza es $\sigma^2 = npq$. En consecuencia la proporción muestral $\hat{p} = \frac{X}{n}$ tiene media *p*, y varianza $\frac{pq}{n}$. Entonces: por el Teorema del Limite Central, cuando *n* es grande se tiene:

$$z = \frac{X - np}{\sqrt{npq}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Fórmulas de aproximación Normal a la Binomial.

Si X es una Binomial con parámetros n y p, entonces

i) $$P(X = k) \cong P(k - .5 < X < k + .5) = P\left(\frac{k - .5 - np}{\sqrt{npq}} < Z < \frac{k + .5 - np}{\sqrt{npq}}\right)$$

ii) $$P(a < X < b) = P(a + .5 < X < b - .5) = P\left(\frac{a + .5 - np}{\sqrt{npq}} < Z < \frac{b - .5 - np}{\sqrt{npq}}\right)$$

iii) $$P(a \le X \le b) = P(a - .5 < X < b + .5) = P\left(\frac{a - .5 - np}{\sqrt{npq}} < Z < \frac{b + .5 - np}{\sqrt{npq}}\right)$$

Ejemplo 6.4.

Según reportes del centro nacional para estadísticas de salud, alrededor del 20 % de la población masculina adulta de los Estados Unidos es obesa. Se elige al azar una muestra de 150 hombres adultos en los Estados Unidos. ¿Cuál es la probabilidad de que:

a) Haya a lo más 25 personas obesas?

b) Haya más de 22 pero menos de 35 obesos?

c) Haya por lo menos un 25% de obesos en la muestra?

Ejemplo 6.4.

Solución:

Sea X el número de personas obesas en la muestra.

Usando aproximación normal a la Binomial se tiene que:

a) $$P(X \leq 25) \cong P(X < 25.5) = P\left(Z < \frac{25.5 - 30}{\sqrt{24}}\right) = P(Z < -0.91) = 0.1814$$

b) $$P(22 < X < 35) \cong P(22.5 < x < 34.5) = P\left(\frac{22.5 - 30}{\sqrt{24}} < Z < \frac{34.5 - 30}{\sqrt{24}}\right) =$$

$$P(-1.53 < Z < 0.91) = 0.8186 - 0.0063 = 0.8123$$

c) $$P(\hat{p} \geq .25) = P(X \geq 37.5) = P(Z > \frac{37.5 - 30}{\sqrt{24}}) =$$ *P(Z>1.53) = 1-P(Z<1.53) =*

1-.9730 = .0630.