6. DISTRIBUCIONES MUESTRALES

Dr. Edgar Acuna

http://math.uprm.edu/~edgar

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 con media poblacional μ y varianza σ^2:

- 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), es igual a la deviación estándar poblacional dividida por la raíz cuadrada de n .Es decir $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$.

Si la población fuera finita de tamaño N, se aplica el factor de correción: $\sqrt{\dfrac{N-n}{N-1}}$ 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.

```
Descriptive Statistics

Variable          N      Mean    Median   Tr Mean    StDev  SE Mean
C1                9      9.89     10.00      9.89     5.42     1.81

Variable        Min       Max        Q1        Q3
C1             3.00     20.00      5.00     13.50
```

Notar que $\mu = 9.89$ y $\sigma = 5.42$.

2) Extraemos 30 muestras de tamaño 4 de dicha población, ejecutando 4 veces la siguiente secuencia **Calc ▸ Random Data ▸ Sample from columns.** Guardar cada una de las 4 observaciones de las muestras en 4 columnas distintas: *Obs1*, *Obs2*, *Obs3*, *y Obs4*.

Ejemplo 6.1

3) Tercero, calculamos las medias de todas esas muestras usando la opción **Row Statistics** del menú **Calc** y tratamos de ver gráficamente al menos si hay acercamiento a Normalidad.

Se eligen las 30 muestras.

Las medidas estadísticas de la media muestral son:

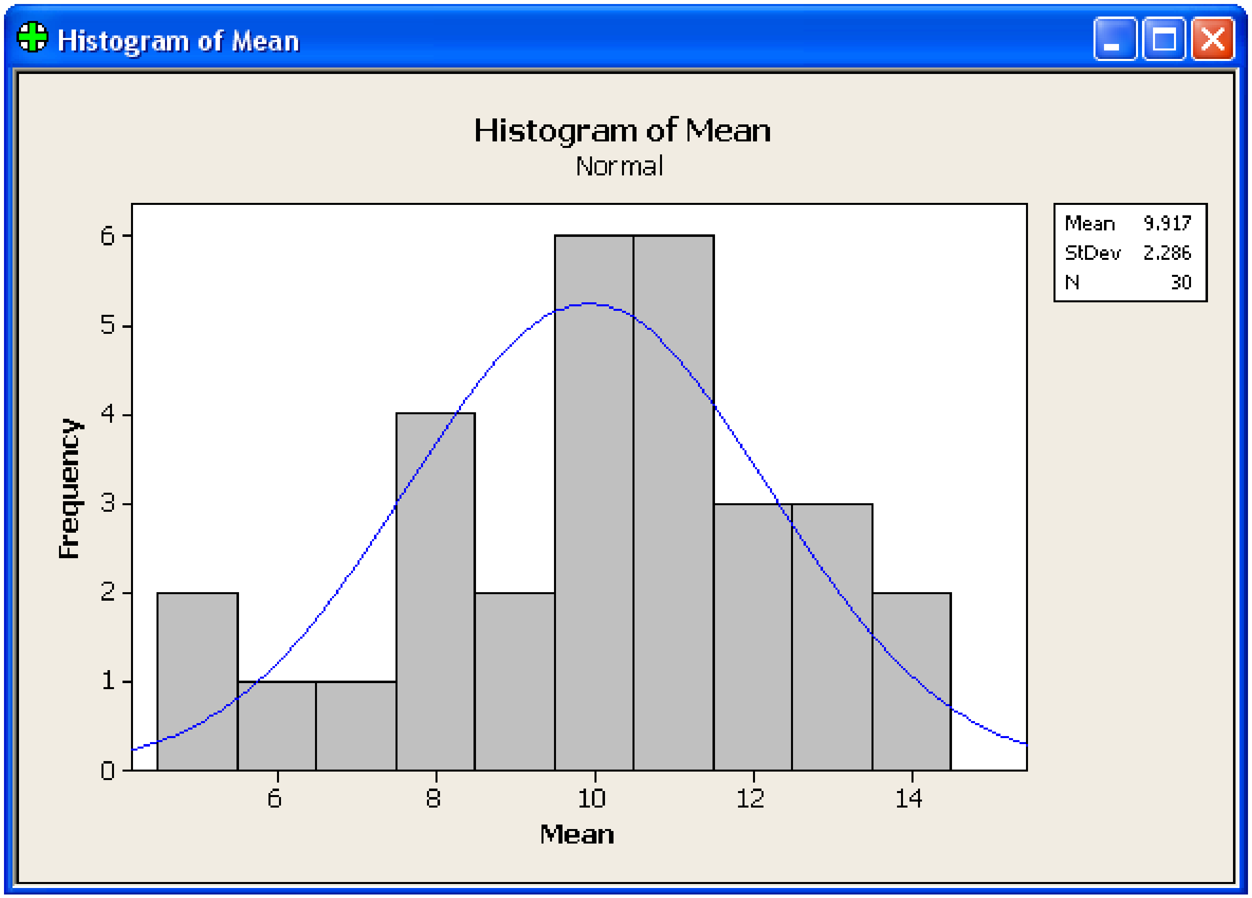
```
Variable          N      Mean    Median  Tr Mean     StDev   SE Mean
media            30    10.108    10.125   10.019     2.806     0.512

Variable        Min       Max        Q1       Q3
media         5.250    16.750     7.938   11.875
```

- ***Interpretación**: Notar que la media de las medias muestrales es* $\mu_{\bar{x}} = 10.108$
- *que está bien cerca de la media poblacional* $\mu = 9.89$ *. Además la desviación estándar de la media muestral es 2.806 mientras que* $\sigma/\sqrt{n}$ *es igual a 5.42/2=2.71 ambos valores también están relativamente cerca. El histograma si está un poco alejado de la normalidad.*

Si se incrementa el tamaño de las muestras se puede notar una mejor aproximación a la Normal.

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

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

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

1-.9730 = .0630.