

## **Razonamiento Aproximado. Ingeniería Informática. 2do Año.**

### **Tema III: Estimación y Dósimas de Hipótesis.**

L/T. Walpole (pág. 225-318); L/T. Miller (175-203; 204-221 (Media); 268-270 (Varianza); 278-282 (Proporción)

#### **Muestreo y distribuciones muestrales.**

Debemos indicar que hasta ahora hemos analizado los elementos probabilísticos en el tema I que nos servirán de base para los aspectos estadísticos desde el punto de vista inferencial que comenzaremos a ver precisamente a partir de este tema. Es importante recordar las características que comentamos acerca de la Estadística Descriptiva y la Inferencia Estadística.

A partir de este momento comenzaremos a trabajar precisamente esos elementos inferenciales partiendo como metodología de una muestra extraída de cierta población para poder realizar estas inferencias sobre ellas.

En la mayoría de los métodos de inferencia que se estudiarán en este libro, se supondrá que se trata con un tipo específico de muestra llamada *muestra aleatoria*.

En estadística, el uso del término **población** es una herencia de los días cuando la estadística se aplicaba principalmente a fenómenos sociológicos y económicos, ahora se aplica a *conjuntos o a colecciones de objetos, reales o conceptuales, pero sobre todo a conjuntos de números, mediciones u observaciones que se investigan*. Por ejemplo, si a uno le interesa determinar el número promedio de televisores por hogar en cierta región, la totalidad de dichos números de conjuntos, uno por cada hogar, constituye la población para este estudio. De igual modo, la población de la cual los inspectores extraen una muestra, para determinar alguna cualidad o característica de un producto manufacturado, pueden ser las mediciones correspondientes de todas las unidades en un lote dado; dependiendo de los objetivos de la inspección, también podría consistir de las mediciones correspondientes de todas las unidades que se fabriquen.

En algunos casos, como el ejemplo anterior concerniente al número de televisores por hogar, la población es **finita**. En otros casos, como en la determinación de alguna característica de todas las unidades pasadas, presentes y futuras que posiblemente se fabriquen mediante un proceso determinado, es conveniente pensar en la población como **infinita**.

Si una población es infinita, es imposible observar todos sus valores; incluso, si es finita quizá no sea práctico ni económico observarla en su totalidad. Entonces, por lo general es necesario usar una **muestra**, una parte de una población e inferir de ella los resultados para toda la población. Evidentemente, tales resultados serían útiles tan solo si la muestra en alguna forma es “representativa” de la población.

Para asegurar que una muestra sea representativa de la población de la que se obtiene, y proporcionar un marco de referencia para la aplicación de la teoría de probabilidad a problemas de muestreo, deberá limitarse la discusión a **muestras aleatorias**.

Comenzaremos con el problema del muestreo y para ello debemos hacer notar que el muestreo nos proporciona:

- Ahorro de costos.
- Ahorro de tiempo.

Partiendo de lo anterior podemos ver que el muestreo constituye un elemento importante y al hacerlo es usual plantear las siguientes interrogantes:

- ¿Tamaño de la muestra?
- ¿Cómo seleccionar los elementos?
- ¿Hasta qué punto fiarnos de los resultados?

En los siguientes elementos que veremos podremos encontrar respuestas a las mismas. Todos concuerdan que generalmente lo ideal sería realizar investigaciones totales, pero estas no siempre son posibles por:

- Razones teóricas (Ej. Lanzamiento de un dado).
  - Razones prácticas (Ej. Clasificación de peces).
  - Razones económicas (Ej. Proceso destructivos).
- El muestreo se realiza fundamentalmente con dos objetivos:
- Estimar parámetros de una población.
  - Verificar hipótesis relativas a una población determinada.

A partir de aquí comentaremos las preguntas hechas inicialmente y analizaremos que lo importante para nosotros ahora es garantizar la aleatoriedad del muestreo pues los demás aspectos lo determinan esencialmente los elementos vistos anteriormente y el investigador.

Los tipos de Muestreo Aleatorios se clasifican fundamentalmente:

- **Muestreo Aleatorio Simple.**
- **Muestreo Aleatorio Estratificado.**
- **Muestreo Aleatorio Sistemático.**
- **Muestreo Aleatorio por Conglomerados.**
- **Muestreo Complejo.**

Como indicamos anteriormente nos dedicaremos al M.A.S. (Poblaciones finitas).

- Pasos para realizar el M.A.S.
  - a) Se numera la población de 1 a N.
  - b) Se toman n números aleatorios de tantas cifras como tenga N. (Tablas de números aleatorios L/T)
  - c) El valor del número aleatorio indicará el elemento a seleccionar.

A partir de aquí estamos en condiciones de comenzar el estudio de las distribuciones muestrales que constituyen la base probabilística de los elementos de inferencia que veremos más adelante.

**Distribuciones muestrales:** Si se consideran todas las posibles muestras de tamaño n de una población y a todas ellas se les calcula una determinada característica (Ej. La media), pudiéndose obtener un resultado numérico o función para cada muestra, la distribución de frecuencia de estos valores constituye la distribución muestral de esa característica llamada también distribución del estadígrafo l..

El conjunto de todas las muestras constituye el espacio muestral.

Veremos a continuación las distribuciones muestrales de algunos estadísticos o estadígrafos.

### 1) Distribución muestral de la media.

a) Si  $X \rightarrow N(\mu, \sigma^2)$  entonces  $\bar{X} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$

b) Para cualquier distribución X a medida que n “crece” se tienen  $\bar{X} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$ . (Teorema central del límite)

En general sabemos que:  $\frac{\bar{X} - \mu}{\sigma} \rightarrow N(0,1)$

De aquí entonces:  $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$

## 2) Distribución de la diferencia de medias.

(Es frecuente en estadística la necesidad de comparar dos conjuntos de datos similares y realizar esta comparación en cuanto a su media.

Sean  $X_1$  y  $X_2$  dos variables, entonces a medida que “crece”  $n_1$  y  $n_2$  tenemos:

$$\overline{X}_1 - \overline{X}_2 \rightarrow N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

por tanto:

$$\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightarrow N(0,1)$$

**Nota:** En presencia de normalidad las exigencias a  $n_1$  y  $n_2$  desaparecen.

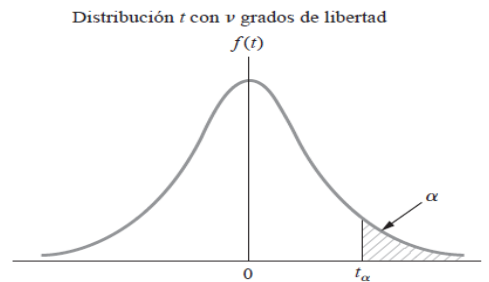
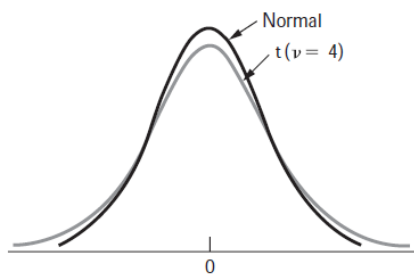
## 3) Distribución t-student. (1er Decenio del siglo, William S. Gosset.)

La aplicación de la teoría anterior requiere conocimiento de la desviación estándar poblacional  $\sigma$ . Si  $n$  es grande, esto no plantea problema alguno aun cuando  $\sigma$  sea desconocida, pues en dicho caso es razonable sustituirla por la desviación estándar muestral  $s$ . Sin embargo se sabe muy poco acerca de su distribución muestral exacta para valores pequeños de  $n$ , a menos que se haga la suposición de que la muestra proviene de una población normal. Con esta suposición, uno probaría lo siguiente:

Se notan deficiencias en  $\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$  cuando sustituimos  $s$  por  $\sigma$ , especialmente en muestras pequeñas entonces:

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} \rightarrow T_{(n-1)}$$

La notación minúscula  $t$  ayuda a diferenciar este estadístico importante de otros. Como se observa en la figura, la forma general de una **distribución  $t$**  es similar a la de una distribución normal: ambas tienen forma de campana y son simétricas en torno la media. Como la distribución normal estándar, la distribución  $t$  tiene la media 0, pero su varianza depende del parámetro  $\nu$  ( $nu$ ), llamado número de **grados de libertad**. La varianza de la distribución  $t$  supera 1, pero tiende a 1 conforme  $n \rightarrow \infty$ . De hecho, puede demostrarse que la distribución  $t$  con  $\nu$  grados de libertad tiende a la distribución normal estándar conforme  $\nu \rightarrow \infty$ .



Las tablas al final de los libros contiene valores seleccionados de  $t_\alpha$  para diversos valores de  $\nu$ , donde  $t_\alpha$  es tal que el área bajo la distribución  $t$  a su derecha es igual a  $\alpha$ . En esta tabla, la columna izquierda contiene valores de  $\nu$ , los encabezados de columna son áreas  $\alpha$  en la cola derecha de la distribución  $t$ , y las entradas son valores de  $t_\alpha$ . (Véase también la figura)

#### 4) Distribución $\chi^2$ (Chi-Cuadrado).

Para estudiar la distribución muestral de la varianza es necesario estudiar la distribución teórica de probabilidad llamada:

- Distribución ji-cuadrado (Chi-cuadrado)

(Fue enunciada por Karl Pearson- 1900 y juega un papel importante en la Estadística)

Esta distribución en su función de densidad depende de un parámetro llamado “grados de libertad” de ahí que se hable de la distribución chi-cuadrado con determinado grado de libertad y se representa  $\chi^2_{g.l.}$ .

El cociente entre dos formas cuadráticas se distribuye según una distribución chi-cuadrado, con tantos grados de libertad como sumandos menos los parámetros que haya que estimar, por lo tanto:

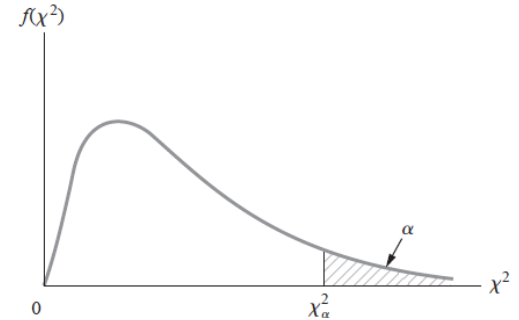
$\sum_{i=1}^n (x_i - \mu)^2 / \sigma^2$  se distribuye según  $\chi^2$  con n grados de libertad.

$\left[ (\bar{x} - \mu) / (\sigma / \sqrt{n}) \right]^2$  se distribuye según  $\chi^2$  con 1 grado de libertad.

En general en una muestra de n elementos se tiene:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \rightarrow \chi^2_{(n-1)}$$

Distribución  $\chi^2$  con  $\nu$  grados de libertad



#### 5) Distribución muestral de la proporción.

En el caso de la proporción de forma similar se tiene:

$$E(\hat{p}) = P \quad \text{Var}(\hat{p}) = \frac{P(1-P)}{n}$$

Nota: Por tanto, aunque la distribución de P se ajusta al modelo binomial con parámetros n y P, la distribución binomial se aproxima a la normal a medida que el tamaño de la muestra va aumentando (teorema central del límite), lo que, a medida que el tamaño de la muestra aumenta la distribución muestral del estadístico proporción tiende a la normal con parámetros P y  $\sqrt{P(1-P)/n}$ .

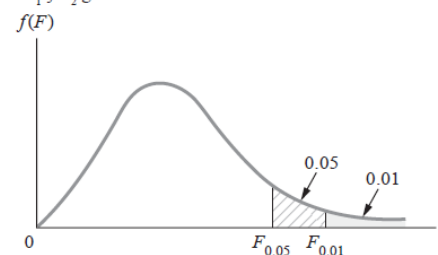
#### 6) Distribución F de Fisher.

Otra distribución que utilizaremos más adelante es la F de Fisher. (3era Década del presente siglo, R.A. Fisher).

En sentido general sean  $X_1^2$  y  $X_2^2$  dos variable aleatorias con distribución de probabilidad  $\chi^2_{(n_1)}$  y  $\chi^2_{(n_2)}$  respectivamente entonces si  $X_1^2$  y  $X_2^2$  son independientes:

$$\frac{X_1^2/n_1}{X_2^2/n_2} \rightarrow F(n_1; n_2)$$

Distribución F con  $\nu_1$  y  $\nu_2$  grados de libertad



Todas estas distribuciones se encuentran tabuladas al final de ambos libros de texto.

**Estimación puntual y por intervalos para la media. Error máximo permisible. Tamaño de muestra.**

Aquí veremos uno de los grandes problemas de la Inferencia Estadística, que es el problema de Estimación. Para ello partiremos de definir en qué consiste la estimación.

Básicamente, la **estimación puntual** trata de la elección de un estadístico; esto es, un solo número calculado a partir de datos muestrales (y acaso otra información) para el cual se tiene cierta esperanza, o seguridad, de que está razonablemente cercana al parámetro que se supone estima. Explicar lo que aquí se entiende por razonablemente cercano no es tarea sencilla: primero, el valor del parámetro es desconocido; y segundo, el valor del estadístico es desconocido hasta después de obtener la muestra. Por ende, solo puede preguntarse si, en el muestreo repetido, la distribución del estadístico tiene ciertas propiedades deseables comparables con la cercanía. Por tanto podemos definir la **estimación puntual**:

***Estimación puntual:***

Estimación del parámetro  $\theta$  del espacio muestral en el espacio paramétrico tal que:

$$\begin{aligned}\hat{\theta}: \Psi &\rightarrow \mathfrak{R} \\ (x_1, x_2, \dots, x_n) &\rightarrow \hat{\theta}(x_1, x_2, \dots, x_n)\end{aligned}$$

Una función que a cada muestra le hace corresponder un solo valor del parámetro.

**Ejemplo:** Sea la Población siguiente:  $X \Rightarrow (3, 4, 5, 3, 2, 4, 5, 3)$  y  $\theta = \mu = 3,625$ .

Sacaremos muestras de tamaño 3:  $m_1(3, 4, 5)$ ;  $m_2(3, 2, 4)$ ;  $m_3(4, 5, 3)$

Y definimos como estimadores:  $\hat{\theta}_1 = \frac{x_1+x_2+x_3}{3}$ ,  $\hat{\theta}_2 = \frac{x_1+x_2}{2}$ .

Si analizamos las estimaciones para ambos casos, ¿Cuál escoger?

Existen diversos métodos para seleccionar el mejor estimador, pero en sentido general se buscan aquellos que mejor cumplan las siguientes propiedades:

- Insesgado:  $E(\hat{\theta}) = \theta \quad \forall \theta \in \mathfrak{R}$
- Eficiente:  $V(\hat{\theta}_1) \leq V(\hat{\theta}_2) \Rightarrow \hat{\theta}_1$  es mas eficiente que  $\hat{\theta}_2$
- Consistencia:  $\lim_{n \rightarrow +\infty} P(|\hat{\theta} - \theta| < \varepsilon) = 1$

Partiendo de lo anterior, en sentido general tenemos los siguientes estimadores:

$\mu$  (media poblacional)-----  $\bar{X} = \sum \frac{x_i}{n}$ . (media muestral).

$\sigma^2$  (varianza poblacional)-----  $S^2 = \sum \frac{(x_i - \bar{x})^2}{n-1}$  (varianza muestral).

Por todo lo anterior si queremos estimar puntualmente, a partir de una muestra, el valor de un parámetro solamente tenemos que evaluar el estadígrafo en la muestra.

Ahora estamos en condiciones de analizar la estimación por Intervalos:

Antes de definir este problema, debemos hacer notar que:

- Cuando se habla de estimación por intervalos, el intervalo que se brinda se considera que probablemente contenga el parámetro estimado.
- El término probablemente nos indica que una estimación por intervalos viene acompañada de una probabilidad.
- Una estimación por intervalo expresa, de alguna manera el nivel de confianza con que se espera que esté el valor del parámetro en el intervalo, por lo que suele llamarse intervalo de confianza.

**Estimación por Intervalos:** Si queremos estimar un parámetro  $\theta$  es necesario buscar un intervalo  $[\theta_1(x); \theta_2(x)]$  tal que  $\theta_1(x)$  y  $\theta_2(x)$  sean estadígrafos y  $P(\theta_1(x) \leq \theta \leq \theta_2(x)) = C$  donde  $C = 1 - \alpha$  es el grado de confianza requerido.

El intervalo  $[\theta_1(x); \theta_2(x)]$  es llamado intervalo de confianza  $(1 - \alpha)100\%$  para  $\theta$ . ( $1 - \alpha$  es llamado nivel de confianza). Los valores  $\theta_1(x)$  y  $\theta_2(x)$  son llamados límites de confianza inferior y superior respectivamente.

### 1) Estimación por Intervalos para $\mu$ con $\sigma^2$ conocida.

Haremos el análisis solo para este caso, pues el resto es similar:

Partimos de:  $P(Z_1 < Z < Z_2) = 1 - \alpha$

Es equivalente a:  $P(-Z_{1-\alpha/2} < Z < Z_{1-\alpha/2}) = 1 - \alpha$

De aquí:  $P\left(-Z_{1-\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < Z_{1-\alpha/2}\right) = 1 - \alpha$

Despejando obtenemos al final:

$$P\left(\bar{X} - Z_{1-\alpha/2} \sigma/\sqrt{n} < \mu < \bar{X} + Z_{1-\alpha/2} \sigma/\sqrt{n}\right) = 1 - \alpha$$

El intervalo quedaría:

$$\left(\bar{X} - Z_{1-\alpha/2} \sigma/\sqrt{n}; \bar{X} + Z_{1-\alpha/2} \sigma/\sqrt{n}\right)$$

### 2) Intervalo para $\mu$ con $\sigma^2$ desconocida y muestra pequeña.

$$\left(\bar{X} - T_{\alpha/2, n-1} S/\sqrt{n}; \bar{X} + T_{\alpha/2, n-1} S/\sqrt{n}\right)$$

### 3) Intervalo para $\mu$ con $\sigma^2$ desconocida y muestra grande.

$$\left(\bar{X} - Z_{1-\alpha/2} S/\sqrt{n}; \bar{X} + Z_{1-\alpha/2} S/\sqrt{n}\right)$$

Como podemos apreciar en la estimación de la media sumamos y restamos, en ambos casos, una misma cantidad a ambos lados, a esta cantidad le llamaremos  $d$  y constituye:

$d$ -----error máximo permisible.

En ocasiones el problema consiste en calcular el tamaño de muestra necesario para realizar una cierta estimación, en estos casos tenemos:

$$\text{Caso 1: } n = \left( \frac{Z_{1-\alpha/2} \cdot \sigma}{d} \right)^2$$

$$\text{Caso 2: } n = \left( \frac{T_{\alpha/2; M-1} \cdot S_M}{d} \right)^2$$

De esta manera estamos en condiciones de calcular intervalos para la media de una distribución normal y además poder analizar el tamaño de muestra necesario para obtener cierta precisión. En el caso práctico que no se conozca la varianza poblacional tenemos que estimar la  $S$  y esta se utiliza partiendo de una muestra piloto con la cual se parte para realizar el análisis.

### **Estudio Independiente:**

Ejercicios L/T. Miller pág. 212. (Estimación de la media)

L/T. Walpole pág. 280 (Límites de tolerancia).

Ejercicios L/T. Walpole pág. 282. (Estimación de la media)

Ejercicios Guía Tema III.

### **Estimación puntual y por intervalos para la varianza y la proporción.**

En la mayoría de las aplicaciones prácticas, la estimación por intervalos para  $\sigma$  o  $\sigma^2$  se basan en la desviación estándar muestral o en la varianza muestral. En muestras aleatorias de poblaciones normales, utilizaremos la expresión:

$$\frac{(n-1)S^2}{\sigma^2}$$

correspondiente a la expresión de tipificación de la distribución muestral de la varianza. Según el cual es un valor de una variable aleatoria que tiene la distribución ji-cuadrada con  $n-1$  grados de libertad, de tal manera podemos asegurar con una probabilidad de  $1-\alpha$  que se satisface la desigualdad:

$$\chi^2_{1-\alpha/2} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}$$

Una vez que los datos han sido obtenidos, hacemos la misma afirmación con  $(1-\alpha)$  100% de nivel de confianza. Resolviendo esta desigualdad para  $\sigma^2$ , obtenemos el siguiente IC:

$$\left( \frac{S^2(n-1)}{\chi^2_{1-\alpha/2; n-1}} ; \frac{S^2(n-1)}{\chi^2_{\alpha/2; n-1}} \right) \Rightarrow \text{IC para } \sigma^2$$

Si extraemos la raíz cuadrada de cada miembro de la desigualdad, obtenemos un intervalo correspondiente con un nivel de confianza de  $(1-\alpha)$  100% para  $\sigma$ .

$$\left( \sqrt{\frac{S^2(n-1)}{\chi^2_{1-\alpha/2; n-1}}} ; \sqrt{\frac{S^2(n-1)}{\chi^2_{\alpha/2; n-1}}} \right) \Rightarrow \text{IC para } \sigma$$

Nótese que los IC para  $\sigma$  o  $\sigma^2$  considerando “colas iguales”, como en la expresión anterior en realidad no dan los IC más reducidos, debido a que la distribución ji-cuadrada no es simétrica. Sin embargo nos servimos de ellos en la mayoría de las aplicaciones a fin de evitar cálculos complicados.

**Ejemplo:**

Se tiene la variable X: resistencia de un tejido. Estime un IC para la variabilidad de la resistencia del tejido.

Datos:

$$x \sim N(? ; ?) \quad n = 4 \quad \bar{x} = 176 \quad S = 4$$

Calculando  $1 - \alpha/2$  y  $\alpha/2$

$$1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow \alpha/2 = 0.025 \rightarrow 1 - \alpha/2 = 0.975$$

Calculando (LI; LS)

$$(LI; LS) = \left( \frac{3 \cdot 16}{\chi_{0.975;3}^2} ; \frac{3 \cdot 16}{\chi_{0.025;3}^2} \right) = \left( \frac{48}{93.348} ; \frac{48}{0.216} \right) = (5.13 ; 222.2)$$

R/ La variabilidad de la resistencia de tejido se encuentra en el intervalo (5.13 ; 222.2) con un nivel de confianza del 95%.

***IC para el parámetro binomial p en muestras grandes***

Cuando  $n$  es grande (*mayor que 30*), podemos construir IC aproximados para el parámetro binomial  $p$  valiéndonos de la aproximación normal a la distribución binomial. Así pues, podemos asegurar con una probabilidad de  $1 - \alpha$  que la desigualdad:

$$-Z_{1-\alpha/2} < \frac{x - np}{\sqrt{np(1-p)}} < Z_{1-\alpha/2}$$

se cumpla. Resolviendo esta desigualdad cuadrática para  $p$ , podemos obtener un conjunto correspondientes de límites de confianza aproximados para  $p$ , obteniendo:

$$p - Z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}} < p < p + Z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

Aquí se presenta un problema, y es que el valor de  $\sqrt{\frac{p(1-p)}{n}}$  nunca podrá ser conocido, pues precisamente estamos estimando  $p$ , por lo que realizaremos una aproximación más al sustituir  $\frac{x}{n}$  por  $p$ . Lo cual quedaría:

$$\frac{x}{n} - Z_{1-\alpha/2} \sqrt{\frac{\frac{x}{n} \left(1 - \frac{x}{n}\right)}{n}} < p < \frac{x}{n} + Z_{1-\alpha/2} \sqrt{\frac{\frac{x}{n} \left(1 - \frac{x}{n}\right)}{n}}$$

Conocemos que  $\frac{x}{n} = \hat{p}$ , correspondiente a la estimación puntual del parámetro binomial  $p$ ; por lo que de manera general, la expresión del **IC para el parámetro binomial  $p$  sería:**

$$\left( \hat{p} \pm Z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

Es necesario analizar que el parámetro binomial  $p$  está siendo estimado mediante una aproximación a la distribución normal que a su vez es una distribución continua; por lo que habría que comprobar los supuestos correspondientes para que la aproximación sea la más adecuada.

Para ello se pueden tener en cuenta varios aspectos, nosotros utilizaremos el llamado “Criterio de Cochran”, el cual establece el tamaño de muestra mínimo necesario para una proporción muestral dada que garantice una adecuada aproximación.

### Criterio de Cochran

Si $\hat{p}$ es:	$n$ mínima necesaria:
0.5	30
0.4 ó 0.6	50
0.3 ó 0.7	80
0.2 ó 0.8	200
0.1 ó 0.9	600
0.05 ó 0.95	1400

Una vez satisfecho tal supuesto del tamaño de muestra, habría que analizar qué valor de  $\hat{p}$  tomar. Esto puede resolverse sustituyendo  $\hat{p}$  por un valor cualquiera siempre que esté entre 0.3 y 0.7. La aproximación al valor real de  $p$  es buena para esos casos.

Puede suceder que el valor de  $\hat{p}$  no esté en tal intervalo, pero siempre que se cumpla el criterio de Cochran será necesario tomar  $\hat{p} = 0.5$  donde se obtendría el intervalo de máxima longitud.

### Ejemplo:

Se quiere lanzar un producto al mercado y para ello se midió el grado de aceptación del mismo a través de 600 personas que asistieron al mercado. De los 600 encuestados 400 afirmaron que les gustaba el producto. De un estimado por intervalos del porcentaje de personas que aceptaron el producto con un nivel de confianza del 95%.

X: # de personas que aceptaron el producto.

$X \sim B(600 ; ?)$

$X = 400$

$\hat{p} = \frac{400}{600} = 0.7$

Comprobando el criterio de Cochran:

Si  $\hat{p} = 0.7$   $n$  mínima debe ser 80  $\Rightarrow$  Se cumple el criterio de Cochran

Dado que  $\hat{p}$  se encuentra entre 0.3 y 0.7, entonces el valor a tomar sería  $\hat{p} = 0.7$

Calculando el IC

$$\left( \hat{p} \pm Z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right) = \left( 0.7 \pm 1.96 \sqrt{\frac{0.7 \cdot 0.3}{600}} \right) = (0.7 \pm 0.036)$$

$(LI; LS) = (0.664 ; 0.736)$

R/ La proporción de personas que aceptaron el producto se encuentran entre (0.664 ; 0.736) con un nivel de confianza del 95%.

### Estudio Independiente:

Ejercicios L/T. Miller pág. 270. (Estimación de la varianza)

Ejercicios L/T. Walpole pág. 307. (Estimación de la varianza)

Ejercicios L/T. Miller pág. 282. (Estimación de la proporción)

Ejercicios L/T. Walpole pág. 302. (Estimación de la proporción)

Ejercicios Guía Tema III.