

### **3. Pruebas de Hipótesis**

# Pruebas de Hipótesis

## Introducción

La experiencia sobre el comportamiento de algún índice de un proceso, o la exigencia del cumplimiento de alguna norma nos lleva a realizar proposiciones sobre el valor de algún parámetro estadístico.

Estas proposiciones se deben contrastar con la realidad (mediante el muestreo de datos) para tomar una decisión entre aceptar o rechazar la proposición

Estas proposiciones se denominan Hipótesis y el procedimiento para decidir si se aceptan o se rechazan se denomina Prueba de Hipótesis

Una prueba de hipótesis es una herramienta de análisis de datos que puede en general formar parte de un experimento comparativo más completo

# Pruebas de Hipótesis

## Introducción

Una hipótesis Estadística es una proposición sobre los parámetros de una población o sobre la distribución de probabilidad de una variable aleatoria

Ejemplo: Se tiene interés en la rapidez de combustión de un agente propulsor de la salida de emergencia en aeronaves (esta rapidez es una variable aleatoria con alguna distribución de probabilidad). Especialmente interesa la rapidez de combustión promedio (que es un parámetro ( $\mu$ ) de dicha distribución). De manera más específica, interesa decidir si esta rapidez promedio es o no 50 cm/seg.

El planteamiento formal de la situación se realiza en términos de una Hipótesis Nula (que es la proposición que se quiere poner a prueba) y una Hipótesis Alternativa, la cual se aceptará si se rechaza la hipótesis nula:

Hipótesis Nula:  $H_0: \mu = 50 \text{ cm/seg}$

Hipótesis Alternativa:  $H_1: \mu \neq 50 \text{ cm/seg}$

- En el ejemplo se tiene una Hipótesis Alternativa Bilateral, ya que se verifica para valores de  $\mu$  a ambos lados de 50 cm/seg.

# Pruebas de Hipótesis

## Introducción

En ocasiones interesa una Hipótesis Alternativa Unilateral, Por ejemplo:

$$\begin{array}{ccc} H_0: \mu = 50 \text{ cm/seg} & & H_0: \mu = 50 \text{ cm/seg} \\ H_1: \mu < 50 \text{ cm/seg} & \text{ó} & H_1: \mu > 50 \text{ cm/seg} \end{array}$$

¿De donde puede surgir una Hipótesis Nula sobre un parámetro?

¿Cuál sería el interés dependiendo del origen de la hipótesis?

- 1) Origen: Experiencia, pruebas pasadas o conocimiento del proceso. Interés: averiguar si ha cambiado el parámetro
- 2) Origen: Alguna teoría o modelo sobre el funcionamiento del proceso. Interés: Verificar la validés de dicha teoría
- 3) Origen: Especificaciones de diseño, obligaciones contractuales, normas a cumplir o solicitudes del cliente. Interés: probar el cumplimiento o incumplimiento de las especificaciones.

➤ La verdad o falsedad de la hipótesis NO puede conocerse con total seguridad a menos que pueda examinarse toda la población

# Pruebas de Hipótesis

## Introducción

Procedimiento General para la prueba de una hipótesis

Tomar un muestra aleatoria

Calcular un estadístico basado en la muestra

Usar el estadístico y sus propiedades para tomar una decisión sobre la  
Hipótesis Nula

# Pruebas de Hipótesis

## Introducción

Ejemplo: Consideremos el ejemplo anterior de la rapidez de combustión.

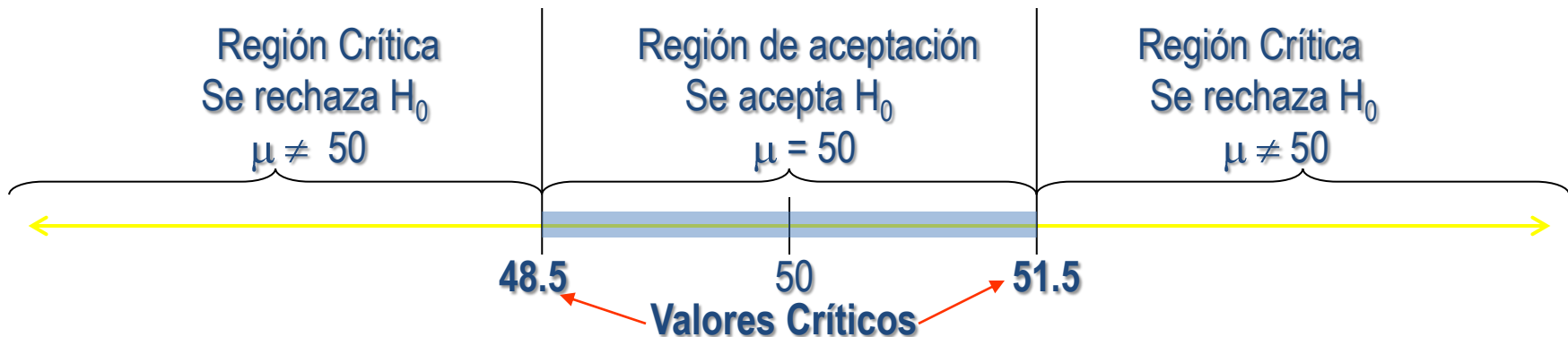
Aquí se tenía:

$$H_0: \mu = 50 \text{ cm/seg}$$

$$H_1: \mu \neq 50 \text{ cm/seg}$$

Aceptación de  $H_0$ .- Un valor de la media muestral  $\bar{x}$  “muy cercano” a 50 cm/seg es una evidencia que apoya a la hipótesis nula, sin embargo es necesario introducir un criterio para decidir que tanto es muy cercano, para el ejemplo este criterio pudiera ser:  $48.5 \leq \bar{x} \leq 51.5$ , si esto ocurre se acepta  $H_0$

De lo contrario, es decir, si  $\bar{x} < 48.5$  o  $\bar{x} > 51.5$ , se acepta  $H_1$



# Pruebas de Hipótesis

## Errores Tipo I y Tipo II

El procedimiento anterior puede llevarnos a una de dos conclusiones erróneas:

Error Tipo I.- Se rechaza  $H_0$  cuando ésta es verdadera

Error Tipo II.- Se acepta  $H_0$  cuando ésta es falsa

En el ejemplo se cometerá un error de tipo I cuando  $\mu=50$ , pero  $\bar{x}$  para la muestra considerada cae en la región crítica

Y se cometerá un error de tipo II cuando  $\mu \neq 50$  pero  $\bar{x}$  para la muestra considerada cae en la región de aceptación

Decisión \ Condición real	$H_0$ verdadera	$H_0$ falsa
	Rechazar $H_0$	Aceptar $H_0$
Rechazar $H_0$	Error Tipo I	OK
Aceptar $H_0$	OK	Error Tipo II

# Pruebas de Hipótesis

## Error Tipo I

A la probabilidad de cometer un error de Tipo I se denota por  $\alpha$ , y se le llama el nivel o tamaño de significancia de la prueba es decir

$$\alpha = P(\text{error Tipo I}) = P(\text{rechazar } H_0 \mid H_0 \text{ es verdadera})$$

Ejemplo: Calcular  $\alpha$  para el ejemplo de la rapidez de combustión para una muestra de  $N=10$  datos, suponiendo que la desviación estándar de la rapidez de combustión es  $\sigma=2.5$  cm/seg.

Solución: en este caso  $\alpha = P(\bar{x} \text{ caiga en la región crítica} \mid \mu=50)$ , es decir:

$$\alpha = P(\bar{x} < 48.5) + P(\bar{x} > 51.5)$$

Recordando que La distribución de  $\bar{x}$  es Normal con media  $\mu=50$  y desviación estándar  $\sigma/\sqrt{N}=0.79$ , por lo tanto:

$$\begin{aligned}\alpha &= \Phi((48.5-50)/0.79) + [1 - \Phi((51.5-50)/0.79)] \\ &= 0.288 + 0.288 = 0.0576\end{aligned}$$

donde  $\Phi(z)$  es el área bajo la curva de densidad normal (0,1) hasta  $z$ .

❖ Esto significa que el 5.76% de las muestras de tamaño 10 conducirán al rechazo de la Hipótesis  $H_0: \mu=50$  cm/seg, cuando ésta es verdadera.

# Pruebas de Hipótesis

## Error Tipo I

Es claro que  $\alpha$  se puede reducir de dos maneras:

- Aumentando la región de aceptación
- Aumentando el tamaño de la muestra

Ejemplo: recalcular  $\alpha$  del ejemplo anterior para a) los nuevos límites de la región de aceptación 48 y 52. b) Para  $N=16$  con los límites originales c) con ambas modificaciones

Solución:

$$a) \alpha = \Phi((48-50)/0.79) + [1 - \Phi((52-50)/0.79)] = 0.0114$$

$$b) \alpha = \Phi((48.5-50)/0.625) + [1 - \Phi((51.5-50)/0.625)] = 0.0164$$

$$c) \alpha = \Phi((48-50)/0.625) + [1 - \Phi((52-50)/0.625)] = 0.0014$$

# Pruebas de Hipótesis

## Error tipo II

Para evaluar un experimento de prueba de hipótesis también se requiere calcular la probabilidad del error de Tipo II, denotada por  $\beta$ , es decir

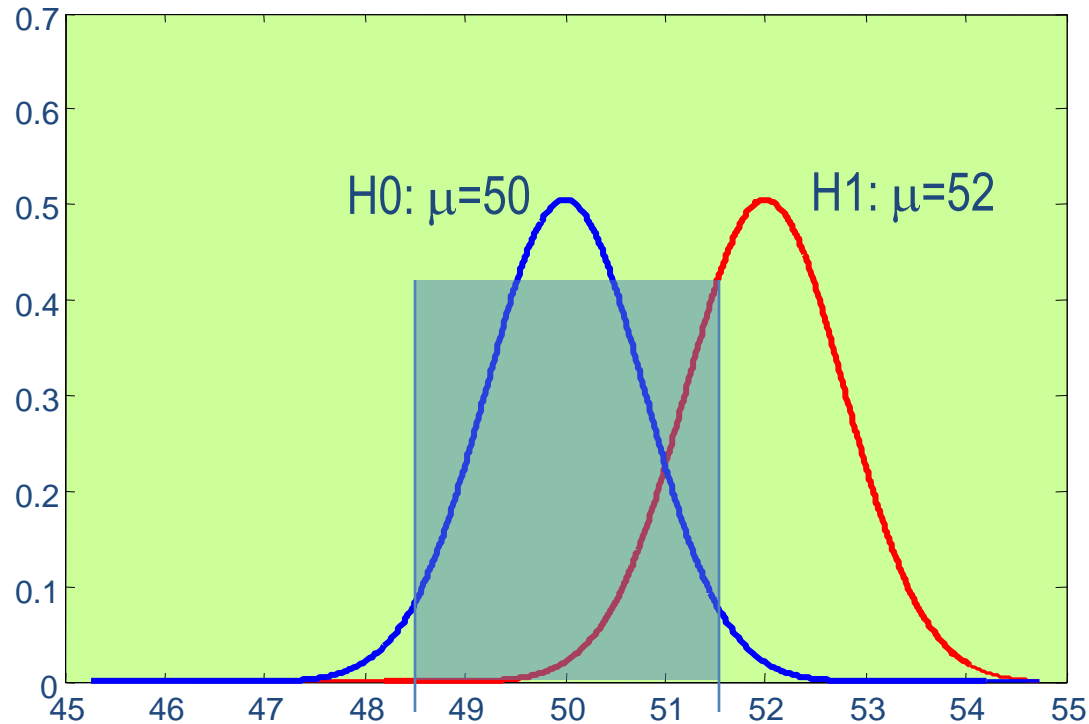
$$\beta = P(\text{error Tipo II}) = P(\text{aceptar } H_0 \mid H_0 \text{ es falsa})$$

Sin embargo, no es posible calcular  $\beta$  si no se tiene una hipótesis alternativa específica, es decir, un valor particular del parámetro bajo prueba en lugar de un rango de valores

Por ejemplo, supongamos que es importante rechazar  $H_0$  si la rapidez promedio de combustión  $\mu$  es mayor que 52 cm/seg o menor que 48 cm/seg. Dada la simetría sólo se requiere evaluar la probabilidad de aceptar  $H_0$ :  $\mu=50$  cuando el valor verdadero es  $\mu=52$ .

# Pruebas de Hipótesis

## Error tipo II



De acuerdo a la figura:  $\bar{\beta} = P(48.5 \leq x \leq 51.5 \mid \mu=52)$

Es decir,

$$\beta = \Phi((51.5-52)/0.79) - \Phi((48.5-52)/0.79) = 0.2643$$

# Pruebas de Hipótesis

## Error tipo II

La probabilidad de obtener un error de tipo II aumenta muy rápido a medida que el valor verdadero  $\mu$  tiende al valor hipotético, por ejemplo, si suponemos que  $\mu=50.5$ , y recalculamos  $\beta$ , obtenemos

Usando Matlab:

$$\beta = \Phi((51.5-50.5)/0.79) - \Phi((48.5-50.5)/0.79) = 0.8923$$

$\beta$  también depende del tamaño de la muestra, por ejemplo, si  $N=16$  obtenemos en el ejemplo cuando  $\mu=52$ :  $\sigma=0.625$ , por lo tanto

$$\beta = \Phi((51.5-52)/0.625) - \Phi((48.5-52)/0.625) = 0.2119$$

Es decir,  $\beta$  disminuye cuando  $N$  aumenta, excepto si el valor real de  $\mu$  está muy cerca del hipotético bajo  $H_0$

# Pruebas de Hipótesis

## Conclusiones Fuerte y Débil

Como uno puede elegir los valores críticos del intervalo de aceptación uno controla el valor de  $\alpha$ . Uno puede entonces controlar la probabilidad de rechazar de manera errónea  $H_0$ .

Es por eso que el rechazo de  $H_0$  siempre se considera como una Conclusión Fuerte. (los datos aportan fuerte evidencia de que  $H_0$  es falsa)

La decisión de aceptar  $H_0$  se considera una Conclusión Débil, a menos que se sepa que  $\beta$  es considerablemente pequeño.

Por esto en lugar de decir “se acepta  $H_0$ ” se prefiere decir “incapaz de rechazar  $H_0$ ”, es decir, no se ha encontrado evidencia suficiente para rechazar  $H_0$ . O sea, no quiere decir que exista gran evidencia de que  $H_0$  sea cierta sino que no hay gran evidencia de que sea falsa.

# Pruebas de Hipótesis

## Hipótesis Unilaterales

En el ejemplo supongamos que si la rapidez media de combustión es menor que 50 cm/seg se desea demostrar esto con una conclusión fuerte. ¿cómo deben plantearse las hipótesis?

$$H_0: \mu = 50 \text{ cm/seg}$$

$$H_1: \mu < 50 \text{ cm/seg}$$

Nótese que aunque  $H_0$  está planteada como una igualdad, se sobreentiende que incluye cualquier valor de  $\mu$  no especificado por  $H_1$ , es decir, la incapacidad de rechazar  $H_0$  no significa que  $\mu = 50$ , sino que no se tiene evidencia fuerte que apoye a  $H_1$ , es decir, pudiera ser que  $\mu = 50$  o que  $\mu > 50$

# Pruebas de Hipótesis

## Hipótesis Unilaterales

Ejemplo: Un embotellador de refresco desea estar seguro de que las botellas que usa tienen en promedio un valor que supera el mínimo de presión de estallido de 200 psi. El embotellador puede formular una prueba de hipótesis de dos maneras:

$$(1) \begin{array}{l} H_0: \mu = 200 \text{ psi} \\ H_1: \mu < 200 \text{ psi} \end{array}$$

$$(2) \begin{array}{l} H_0: \mu = 200 \text{ psi} \\ H_1: \mu > 200 \text{ psi} \end{array}$$

Con el planteamiento (1), como el rechazo de  $H_0$  es una conclusión fuerte, se obliga al fabricante a demostrar (aportar evidencia) de que las botellas soportan mayor presión que 200 psi.

Con el planteamiento (2), si se rechaza  $H_0$  se concluye que el estallido no ocurre a los 200 psi, es decir, se concluye que las botellas son satisfactorias a menos que halla evidencia fuerte en sentido contrario.

¿Cuál planteamiento es el correcto?

# Pruebas de Hipótesis

## Hipótesis Unilaterales

Es decir, en la Hipótesis alternativa se debe poner la proposición sobre la cual es importante llegar a una conclusión fuerte:

$$(1) \begin{array}{l} H_0: \mu = 200 \text{ psi} \\ H_1: \mu < 200 \text{ psi} \end{array}$$

$$(2) \begin{array}{l} H_0: \mu = 200 \text{ psi} \\ H_1: \mu > 200 \text{ psi} \end{array}$$

# Pruebas de Hipótesis

## Procedimiento general para la prueba de Hipótesis

### Antes de examinar los datos muestrales:

1. Identificar el parámetro de interés
2. Establecer la Hipótesis Nula  $H_0$
3. Especificar una Hipótesis alternativa adecuada  $H_1$
4. Seleccionar un nivel de significancia  $\alpha$

### Usando los datos muestrales:

5. Establecer un estadístico de prueba adecuado
6. Establecer una región de rechazo
7. Calcular todas las cantidades muestrales necesarias para el estadístico
8. Decidir si debe o no rechazar  $H_0$

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la media, varianza conocida

Si se desea probar la Hipótesis:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \leq \mu_0$$

Se puede usar el estadístico de prueba  $Z$  siguiente

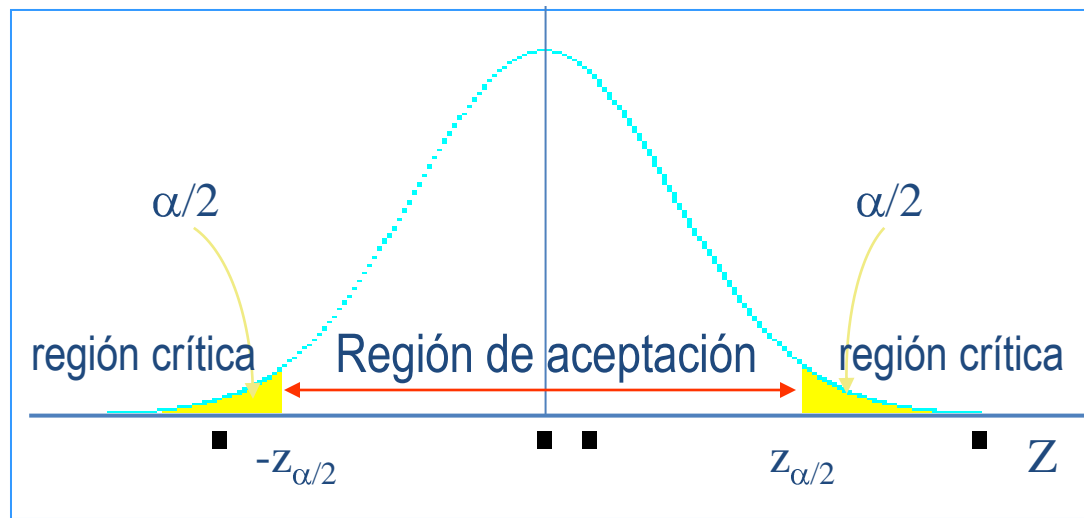
$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{N}}$$

El cual tiene una distribución normal con media cero y varianza 1 (si se cumplen las suposiciones del teorema central del límite)

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la media, varianza conocida

Entonces, para una  $\alpha$  dada podemos establecer las siguientes regiones de aceptación y crítica:



Conclusiones:

Rechazar $H_0$ si:	$z < -z_{\alpha/2} \text{ o } z > z_{\alpha/2}$
No rechazar $H_0$ si:	$-z_{\alpha/2} \leq z \leq z_{\alpha/2}$

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la media, varianza conocida

Ejemplo: Se ilustran los 8 pasos del procedimiento general para el ejemplo del combustible sólido para escape de aeronaves. En este caso se conoce  $\sigma=2$  cm/seg, se desea probar si la media  $\mu$  es de 50 cm/seg. Se selecciona una muestra aleatoria de tamaño  $N=25$ , obteniendo  $\bar{x}=51.3$  cm/seg. Se especifica un nivel de significación de  $\alpha=0.05$ , ¿a qué conclusiones se debe llegar?

- 1) El parámetro de interés es  $\mu$  (rapidez promedio de combustión)
- 2)  $H_0: \mu = 50$  cm/seg
- 3)  $H_1: \mu \neq 50$  cm/seg
- 4)  $\alpha = 0.05$

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la media, varianza conocida

5) La estadística de prueba es

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{N}}$$

6) Rechazar  $H_0$  si  $z > 1.96$  o si  $z < -1.96$  (consecuencia del paso 4)

7) cálculos

$$Z = \frac{51.3 - 50}{2 / \sqrt{25}} = 3.25$$

8) Conclusión como  $z = 3.25 > 1.96$ , se rechaza  $H_0: \mu = 50$  cm/seg con un nivel de significancia  $\alpha = 0.05$

8) Es decir, Se concluye que en base a una muestra de 25 mediciones la rapidez promedio de combustión es diferente de 50 cm/seg, de hecho, existe evidencia fuerte de que ésta es mayor.

# Pruebas de Hipótesis

## Valores P

Una manera de notificar los resultados de una prueba de hipótesis es establecer si la hipótesis nula fue o no rechazada con un nivel especificado  $\alpha$  de significancia

Una alternativa es especificar el nivel de significancia  $\alpha$  más pequeño que conduce al rechazo de la hipótesis nula. A este se le llama el *P-valor*

Este *P-valor* sólo depende de la muestra tomada, es decir, para una muestra y un estadístico calculado se puede obtener su *P-valor* y comparar con un  $\alpha$  especificado. Entonces, si  $P < \alpha$ ,  $H_0$  se rechaza.

# Pruebas de Hipótesis

## Valores P

En el caso de la distribución normal para la prueba sobre la media es fácil calcular el valor P. Si  $z_0$  fue el valor calculado del estadístico de prueba, entonces:

$P = \left\{ \begin{array}{l} \\ \\ \end{array} \right.$	$2 [ 1 - \Phi( z_0 ) ]$	Prueba de dos colas: $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$
	$1 - \Phi(z_0)$	Prueba de cola superior: $H_0: \mu = \mu_0, H_1: \mu > \mu_0$
	$\Phi(z_0)$	Prueba de cola inferior: $H_0: \mu = \mu_0, H_1: \mu < \mu_0$

Donde  $\Phi(z) = P(Z \leq z)$  (Función de distribución normal  $N(0,1)$ )

Para el ejemplo  $z_0 = 3.25$ , entonces  $P = 2(1 - \Phi(3.25)) = 0.0012$ . Es decir,  $H_0$  será rechazada con cualquier nivel de significancia  $\alpha \geq 0.0012$

Si se usa el enfoque del valor P el paso 6 del procedimiento general de prueba de hipótesis ya no es necesario.

# Pruebas de Hipótesis

## Error Tipo II y tamaño de la muestra

Consideremos la hipótesis bilateral  $H_0: \mu = \mu_0$ ,  $H_1: \mu \neq \mu_0$ .

Si  $H_0$  es falsa y la media verdadera es  $\mu = \mu_0 + \delta$  (con  $\delta > 0$ ). El estadístico de prueba

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{N}}$$

se puede escribir como

$$Z = \frac{\bar{X} - (\mu_0 + \delta)}{\sigma / \sqrt{N}} + \frac{\delta \sqrt{N}}{\sigma}$$

Es decir, Si  $H_1$  es verdadera  $Z$  tiene distribución Normal con media y varianza 1.

$$\frac{\delta \sqrt{N}}{\sigma}$$

Por lo tanto, el error Tipo 1 ( $\beta$ ) se puede calcular como

$$\beta \approx \Phi\left(z_{\alpha/2} + \frac{\delta \sqrt{N}}{\sigma}\right)$$

Y si definimos  $\beta = \Phi(-z_\beta)$ , obtenemos

$$\sqrt{N} \approx \frac{(z_{\alpha/2} + z_\beta) \sigma}{\delta}$$

# Pruebas de Hipótesis

## Error Tipo II y tamaño de la muestra

Para el ejemplo del combustible sólido. Si al analista le interesa diseñar la prueba de hipótesis de manera que si el valor verdadero de  $\mu$  es 51 cm/seg se rechace  $H_0$  con una probabilidad alta (por ejemplo 90%) y con el mismo valor anterior de  $\alpha=0.05$

En este caso  $\delta=1$ ,  $\sigma=2$ ,  $\alpha=0.05$  por lo tanto:

$$N = 4 * (\Phi^{-1}(0.025) + \Phi^{-1}(0.1))^2 \approx 42$$

Observación: Debe tenerse cuidado cuando se interpretan los resultados basados en una muestra muy grande, ya que es muy probable que se detecte cualquier alejamiento (muy pequeño) respecto al valor hipotético  $\mu_0$ . *Esta diferencia podría no tener ninguna importancia práctica pero conducir al rechazo de  $H_0$*

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la igualdad de dos medias (varianzas conocidas)

Se tienen dos poblaciones de interés. La primera con media  $\mu_1$  y varianza  $\sigma_1^2$  conocidas y la segunda con media  $\mu_2$  y varianza  $\sigma_2^2$  conocidas. Interesa saber si las dos medias son iguales. Se plantean las hipótesis

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Suposiciones: Las dos poblaciones son normales o se cumplen las condiciones del teorema del límite central. Entonces el estadístico  $\bar{X}_1 - \bar{X}_2$  es una variable Normal con media  $\mu_1 - \mu_2$  y varianza  $\sigma_1^2/N_1 + \sigma_2^2/N_2$

Por lo tanto el siguiente estadístico de prueba

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}}$$

Que sigue una  $N(0,1)$  si  $H_0$  es cierta.

Por lo tanto, se rechazará  $H_0$  si  $z_0 > z_{\alpha/2}$  o  $z < z_{-\alpha/2}$

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la igualdad de dos medias (varianzas conocidas)

Ejemplo: Un diseñador quiere reducir el tiempo de secado de una pintura. Se prueban dos fórmulas de pintura. La fórmula 1 es la normal y la fórmula 2 posee un ingrediente secante que se espera reduzca el tiempo de secado. Se sabe que el tiempo de secado tiene una desviación estándar de 8 min y que ésta no se afecta con la adición del nuevo ingrediente. Se pintan 10 artículos con la fórmula 1, y 10 con la fórmula 2, obteniéndose tiempos promedio de secado de  $\bar{x}_1=121$  min, y  $\bar{x}_2=112$  min. respectivamente. ¿A qué conclusión se llega sobre la eficacia del nuevo ingrediente utilizando  $\alpha=0.05$ ?

- 1) Cantidad de interés:  $\mu_1 - \mu_2$
- 2)  $H_0: \mu_1 = \mu_2$
- 3)  $H_1: \mu_1 > \mu_2$  (se busca evidencia fuerte que indique que el tiempo de secado promedio de la muestra 2 es menor)

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la igualdad de dos medias (varianzas conocidas)

4)  $\alpha=0.05$

5) El estadístico de prueba es

$$Z = (\bar{X}_1 - \bar{X}_2) / \sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}$$

6)  $H_0$  se rechazará si  $z > z_{0.05} = 1.645$

7) Sustituyendo los datos, obtenemos  $z = (121 - 112) / (12.8)^{1/2} = 2.52$

8) Conclusión: Puesto que  $z = 2.52 > 1.645$  se rechaza  $H_0$  con un nivel de significancia  $\alpha=0.05$  concluyéndose el nuevo ingrediente sí disminuye el tiempo de secado.

Alternativamente puede calcularse un valor  $P = 1 - \Phi(2.52) = 0.0059$ , es decir, se rechazará  $H_0$  para cualquier nivel de significancia  $\alpha \geq 0.0059$

# **Pruebas de Hipótesis**

## **Identificación Causa - Efecto**

En el ejemplo anterior se supone que fueron asignados de manera aleatoria 10 especímenes a una fórmula (tratamiento) y 10 especímenes a la otra luego se aplicó la pintura en un orden aleatorio a cada espécimen hasta pintar los 20. Este es un Experimento Completamente Aleatorizado.

En un estudio estadístico sobre la incidencia del cáncer pulmonar entre personas que fuman normalmente se hace un seguimiento en el tiempo de los individuos a prueba. Este es un Experimento Observacional.

En este caso no se puede asignar de manera aleatoria un tratamiento u otro (fumar o no fumar) a una porción de los individuos. Por otro lado, el hábito de fumar no es el único factor que influye en el desarrollo de cáncer pulmonar.

# Pruebas de Hipótesis

## Prueba de Hipótesis sobre la media, varianza desconocida

Si la población tiene una distribución Normal con media  $\mu$  y varianza  $\sigma^2$  desconocidas pudiera utilizarse el estadístico  $S^2$  y el procedimiento descrito anteriormente para varianza conocida (esto es válido para  $N$  grande), pero si la muestra es pequeña, tendremos que usar el estadístico siguiente,

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{N}}$$

el cual tiene una distribución  $t$  con  $N-1$  grados de libertad,

Así, para la prueba de Hipótesis bilateral

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Se rechazará  $H_0$  si  $t > t_{\alpha/2, N-1}$  o si  $t < t_{-\alpha/2, N-1}$

# Pruebas de Hipótesis

## Prueba de Hipótesis sobre la media, varianza desconocida

Ejercicio: Los siguientes son datos de pruebas de resistencia a la adhesión, los siguientes datos presentan la carga (en Mpa) a la cual 22 especímenes fallaron

19.8	18.5	17.6	16.7	15.8	15.4
14.1	13.6	11.9	11.4	11.4	8.8
7.5	15.4	15.4	19.5	14.9	12.7
11.9	11.4	10.1	7.9		

¿Sugieren los datos que la carga promedio de falla es mayor que 10Mpa? Supóngase que la carga de falla tiene una distribución Normal y utilice  $\alpha=0.05$ . Desarrolle los 8 pasos del procedimiento general y encuentre un valor P para la prueba.

# Pruebas de Hipótesis

## P-valor de una prueba t

El P-valor es el más pequeño nivel de significancia para el que  $H_0$  debe rechazarse, esto es el área de la cola (de la curva de densidad de probabilidad) que está más allá del valor del estadístico (en este caso t). o el doble de esta área en pruebas bilaterales.

## Selección del Tamaño de la Muestra

En todas las pruebas de hipótesis estadísticas se puede calcular el tamaño de la muestra (N) adecuada en función de la magnitud del error de tipo I que se permite. En cada tipo de prueba se encuentran fórmulas diferentes para N.

# Pruebas de Hipótesis

## Otras pruebas de Hipótesis

En forma similar a como se describió el caso de la media y la diferencia de medias, se pueden realizar diferentes pruebas de hipótesis para estos mismos u otros parámetros, lo único que cambia en cada caso es:

- Las suposiciones sobre la distribución de la población
- El estadístico elegido y por consiguiente
- La distribución del estadístico.

En la siguiente tabla se resumen algunas de las pruebas de hipótesis más utilizadas

# Pruebas de Hipótesis

## Otras pruebas paramétricas de Hipótesis

Prueba sobre	Hipótesis Nula	Suposiciones	Estadístico de Prueba
La media	$\mu = \mu_0$	$\sigma^2$ conocida	Normal
	$\mu = \mu_0$	$\sigma^2$ desconocida	T
Igualdad de medias	$\mu_1 = \mu_2$	$\sigma_1^2 = \sigma_2^2$ conocidas	Normal
	$\mu_1 = \mu_2$	$\sigma_1^2 = \sigma_2^2$ desconocidas	T
	$\mu_1 = \mu_2$	$\sigma_1^2 \neq \sigma_2^2$ conocidas	T
La varianza	$\sigma^2 = \sigma_0^2$	Dist. Normal, N pequeña	Ji <sup>2</sup>
	$\sigma^2 = \sigma_0^2$	N grande	Normal
Igualdad de dos varianzas	$\sigma_1^2 = \sigma_2^2$		F
Una proporción	$p = p_0$		Normal
Igualdad de dos proporciones	$p_1 = p_2$		Normal

# Pruebas de Hipótesis

## Pruebas de Hipótesis No Paramétricas

Las pruebas de hipótesis anteriores se llaman paramétricas porque suponen conocida la distribución de la población y la hipótesis es acerca de los parámetros de dicha distribución.

Otra clase de hipótesis es: No se sabe cual es la distribución de la población y se desea probar la hipótesis de que cierta distribución en particular será un modelo satisfactorio. Por ejemplo, tal vez se requiera probar si la distribución es Normal

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

- Se parte de una muestra aleatoria de tamaño  $N$ , proveniente de una población cuya distribución de probabilidad es desconocida.
- Las  $N$  observaciones se acomodan en un histograma de frecuencia con  $k$  intervalos de clase. Sea  $O_i$  la  $i$ -ésima frecuencia de clase
- De la distribución de probabilidad propuesta se calcula la frecuencia esperada  $E_i$  en el  $i$ -ésimo intervalo de clase

- El estadístico de prueba es

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

El cual tiene una distribución  $\chi^2$  con  $k-p-1$  grados de libertad si la población sigue la distribución propuesta. (donde  $p$  es el número de parámetros de la población)

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

- La aproximación mejora a medida que N es más grande
- La hipótesis debe rechazarse si el valor del estadístico de prueba es

$$\chi^2 > \chi^2_{1-\alpha, k-p-1}$$

- Precaución: Si las frecuencias esperadas son muy pequeñas el estadístico  $\chi^2$  no reflejará el alejamiento entre lo observado y lo esperado. (Se considera que valores menores de 5 son pequeños)
- Si en una prueba resultan frecuencias esperadas pequeñas, se pueden combinar intervalos de clase adyacentes para aumentar estos valores, ya que no es necesario que los anchos de clase sean del mismo tamaño

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

Ejemplo 1.- Un algoritmo para generar enteros pseudoealeatorios de 0 a 9 'se prueba para determinar si tiene una distribución uniforme, para ello se generan 1000 números, obteniendo la siguiente tabla de frecuencia. ¿Existe evidencia de que el generador funciona de manera correcta?. Utilice  $\alpha=0.05$

	0	1	2	3	4	5	6	7	8	9
$O_i$	94	93	112	101	104	95	100	99	108	94
$E_i$	100	100	100	100	100	100	100	100	100	100

Como  $E_i$  se puede calcular sin estimar ningún parámetro a partir de la muestra, entonces  $p=0$  y el estadístico será  $\sum j_i^2$  con  $k-p-1=10-0-1=9$  grados de libertad.

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

- 1) Variable de interés: distribución de los números pseudoaleatorios
- 2)  $H_0$ : La distribución es uniforme en el intervalo de 0 a 9
- 3)  $H_1$ : La distribución No es uniforme en ese intervalo
- 4)  $\alpha = 0.05$
- 5) El estadístico de prueba es 
$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$
- 6) Se rechazará  $H_0$  si  $\chi^2 > \chi^2_{0.05,9} = 16.92$
- 7) Cálculos  
$$\chi^2 = 0.01 * ((94-100)^2 + (93-100)^2 + \dots + (94-100)^2) = 3.72$$
- 8) Conclusiones: como  $3.72 < 16.92$  No es posible rechazar la hipótesis. Por lo tanto parece ser que el generador de números aleatorios trabaja bien.

¿Cual es el P-valor de la prueba ?

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

Ejemplo 2.- Se propone que el número de defectos en tarjetas de circuito impreso sigue una distribución de Poisson. Se obtiene una muestra de 60 tarjetas y se observa el número de defectos, con los siguientes resultados:

<i>Defectos</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4 o más</i>
<i>O<sub>i</sub></i>	32	15	9	4	0

Distribución de Poisson. Es una distribución discreta cuya función de probabilidad es

$$f(x) = \frac{e^{-\mu} \mu^x}{x!}$$

Definida para  $x=0,1,2,3,\dots\infty$ . Donde  $\mu$  es la media de  $X$

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

Cálculo de las frecuencias Esperadas  $E_i$ :

Un estimador para la media  $\mu$  de la distribución de Poisson es la media muestral, es decir,  $(32 \times 0 + 15 \times 1 + 9 \times 2 + 4 \times 3) / 60 = 0.75$  fallas/tarjeta. Usando este valor de  $m$  obtenemos la siguiente tabla de frecuencias esperadas:

$x$	0	1	2	3	4 o más
$F(x)$	0.472	0.354	0.133	0.033	0.0073
$E_i$	28.32	21.24	7.98	1.98	0.44

Para evitar que las últimas dos frecuencias esperadas sean menores que 5 combinamos las últimas tres celdas para obtener:

$x$	0	1	2 o más
$E_i$	28.32	21.24	10.44
$O_i$	32	15	13

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

- 1) Variable de interés: La forma de distribución de los defectos en tarjetas de circuito impreso
- 2)  $H_0$ : La distribución es de Poisson
- 3)  $H_1$ : La distribución No es Poisson
- 4)  $\alpha = 0.05$
- 5) El estadístico de prueba es  $\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$ , el cual tiene una distribución  $\chi^2$  con  $k-p-1=3-1-1=1$  grado de libertad
- 6) Se rechazará  $H_0$  si  $\chi^2 > \chi^2_{0.05,1}=3.84$
- 7) Cálculos  
 $\chi^2 = (94-100)^2/28.32 + (93-100)^2/21.24 + (94-100)^2/10.44 = 2.94$
- 8) Conclusiones: como  $2.94 < 3.84$ . No es posible rechazar la hipótesis. Por lo tanto parece ser que la distribución de defectos en las placas de circuito impreso es Poisson

El P-valor de la prueba es  $P=0.9861$

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

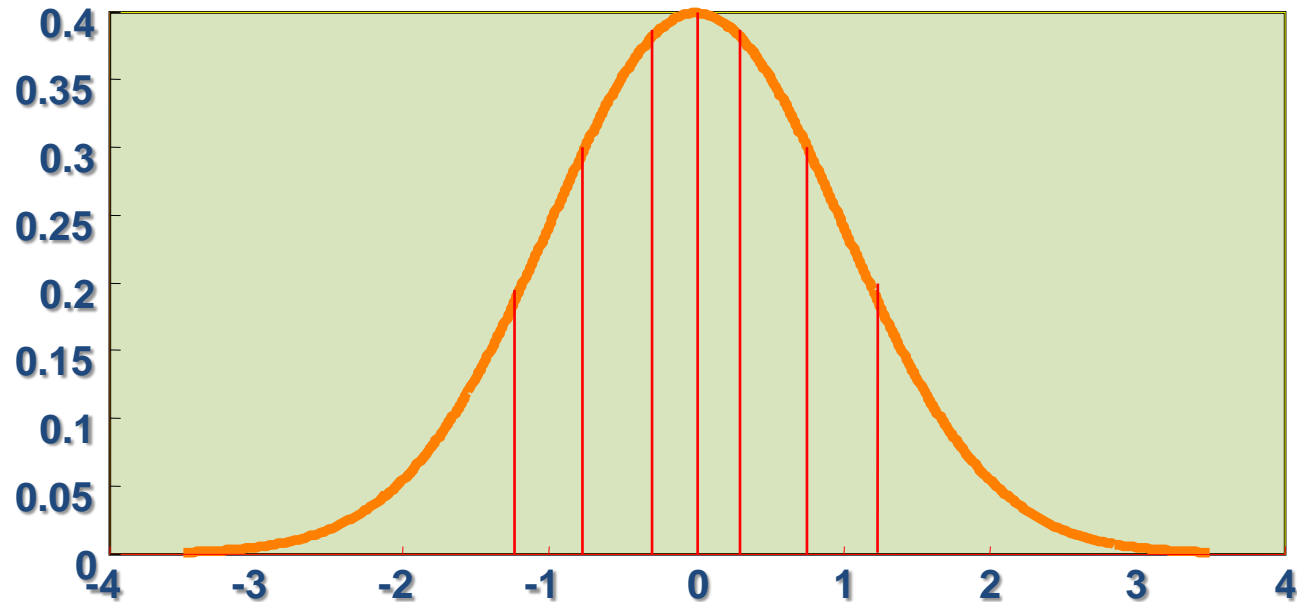
Ejemplo 3.- Se desea determinar con  $\alpha=0.05$  si el voltaje de salida de una fuente de alimentación está descrito por una distribución Normal. Se toma una muestra aleatoria de  $N=100$  fuentes, determinándose los siguientes valores muestrales  $\bar{x} = 5.04$ ,  $s = 0.08$ .

Para evitar valores de frecuencias esperadas muy pequeños, de antemano se elige el ancho de los intervalos de clase de manera que la frecuencia esperada sea constante  $F_i = N / k$ .

Así, si  $k=8$  clases, se buscarán 8 intervalos de clase que dividan la curva de densidad normal en 8 áreas iguales, como se muestra en la siguiente figura para media 0 y varianza 1.

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste



Para la distribución  $N(0,1)$  los límites de los 8 intervalos son

$-\infty, -1.15, -0.675, -0.32, 0, 0.32, 0.675, 1.15, +\infty$ ,

por lo tanto para el ejemplo, los límites son

$-\infty, 4.948, 4.986, 5.014, 5.040, 5.066, 5.094, 5.132, +\infty$

Con esta elección se obtiene la siguiente tabla de frecuencias para la muestra

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

Intervalo de Clase	$O_i$	$E_i$
De $-\infty$ a 4.948	12	12.5
De 4.948 a 4.986	14	12.5
De 4.986 a 5.014	12	12.5
De 5.014 a 5.040	13	12.5
De 5.040 a 5.066	12	12.5
De 5.066 a 5.094	11	12.5
De 5.094 a 5.132	12	12.5
De 5.132 a $+\infty$	14	12.5
Suma:	100	100

# Pruebas de Hipótesis

## Prueba $\chi^2$ de la Bondad de Ajuste

- 1) La variable de interés es el tipo de distribución del voltaje dado por una fuente de alimentación
- 2)  $H_0$ : El tipo de distribución es Normal
- 3)  $H_1$ : El tipo de distribución no es Normal
- 4)  $\alpha = 0.05$
- 5) El estadístico de prueba es 
$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$
- 6) Para determinar los intervalos de clase se requirió estimar  $\mu$  y  $\sigma$ , por lo tanto los grados de libertad son  $k-p-1=8-2-1=5$ , por lo tanto se rechazará  $H_0$  si  $\chi^2 > \chi^2_{0.05,5} = 11.07$
- 7) Cálculos:  
$$\chi^2 = (1/12.5)[(12-12.5)^2 + (14-12.5)^2 + \dots + (14-12.5)^2] = 0.64$$
- 8) Conclusiones: como  $0.64 < 11.07$ , no es posible rechazar  $H_0$ , por lo tanto no hay evidencia fuerte de que la distribución no sea Normal.  
El valor P de la prueba (para  $\chi^2 = 0.64$ ) es  $P=0.9861$ .

# Pruebas de Hipótesis

## Gráfica de Probabilidad

La gráfica de probabilidad es un método gráfico que permite determinar si una muestra de datos se ajusta a una distribución propuesta en base a un análisis visual subjetivo.

Originalmente esta gráfica se realizaba sobre un papel especial llamado papel de probabilidad diseñado con las escalas adecuadas para las diferentes distribuciones.

### Procedimiento:

- Se ordena la muestra de menor a mayor:  $x_1, x_2, \dots, x_N$
- Se grafica sobre el papel de probabilidad la frecuencia acumulada observada  $(i-0.5)/N$  contra el valor de los datos ordenados
- Si los puntos obtenidos se desvían significativamente de una línea recta, el modelo propuesto no será el apropiado.

# Pruebas de Hipótesis

## Gráfica de Probabilidad

Ejemplo: Las siguientes son diez observaciones sobre la duración en minutos de las baterías de computadoras portátiles:

176, 183, 185, 190, 191, 192, 201, 205, 214, 220

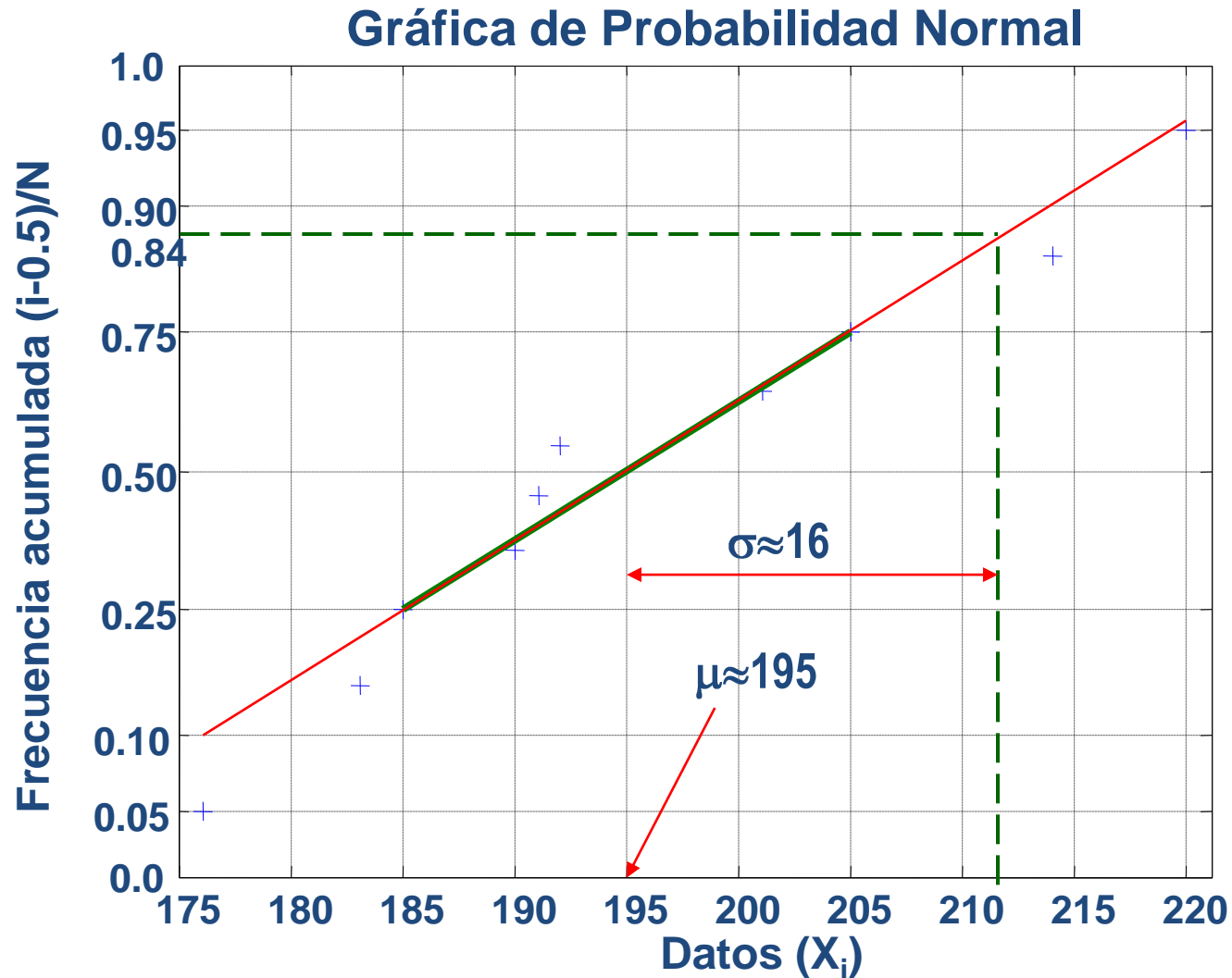
Utilizar la gráfica de probabilidad para determinar si la muestra corresponde a una distribución Normal.

Procedimiento: Formamos la tabla de los datos ordenados y las frecuencias acumuladas  $(i-0.5)/N$  siguiente:

$i$	1	2	3	4	5	6	7	8	9	10
$x_i$	176	183	185	190	191	192	201	205	214	220
$(i-0.5)/10$	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95

# Pruebas de Hipótesis

## Gráfica de Probabilidad



# Pruebas de Hipótesis

## Gráfica de Probabilidad

Observaciones:

- Al analizar la gráfica debe recordarse que el eje vertical está graduado en percentiles, por ello la media se encuentra en el percentil 50.
- Los puntos más confiables son los que están entre el percentil 25 y el 75, de hecho, la línea trazada debe unir estos percentiles
- Se puede obtener una gráfica sobre papel normal ajustando la escala vertical de acuerdo a  $z_i$ , donde  $\Phi(z_i) = (i-0.5)/N$ , para el ejemplo:

$i$	1	2	3	4	5	6	7	8	9	10
$(i-0.5)/10$	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95
$z_i$	-1.64	-1.04	-0.67	-0.39	-0.13	0.13	0.39	0.67	1.04	1.64

4) En SPSS, ver “Gráficos P-P” en Estadísticos Descriptivos.

# Pruebas de Hipótesis

## Tablas de Contingencia

Una tabla de contingencia es una herramienta que nos permite poner a prueba si *dos criterios de clasificación de una misma muestra* son independientes o no, por ejemplo:

Población	Criterio 1	Criterio 2
Ingenieros recién egresados	Salario inicial	Institución de origen
Estudiantes	Nivel Socioeconómico	Promedio académico
Número de fallas en un proceso	Maquinaria utilizada	Turno
Estudiantes	Calif. en Materia 1	Calif. en Materia 2
Fallas en un transformador	Tipo de falla	Ubicación
Etc...		

# Pruebas de Hipótesis

## Tablas de Contingencia

Procedimiento:

Se forma una tabla de frecuencias observadas  $O_{ij}$ , donde:

$i$ =No. de renglón= nivel de clasificación  $i$  del criterio 1 ( $i=1,2,3,\dots,r$ )

$j$ =No. de columna= nivel de clasificación  $j$  del Criterio 2 ( $j=1,2,3,\dots,c$ )

<b><i>Criterio1 \ Criterio2</i></b>	<b><i>Nivel 1</i></b>	<b><i>Nivel 2</i></b>	<b><i>...</i></b>	<b><i>Nivel c</i></b>
<b><i>Nivel 1</i></b>	$O_{11}$	$O_{12}$		$O_{1c}$
<b><i>Nivel 2</i></b>	$O_{21}$	$O_{22}$		$O_{2c}$
<b><i>...</i></b>			<b><i>...</i></b>	
<b><i>Nivel r</i></b>	$O_{r1}$	$O_{r1}$	<b><i>...</i></b>	$O_{rc}$

# Pruebas de Hipótesis

## Tablas de Contingencia

Consideraciones: Si los criterios son independientes (Hipótesis Nula): La probabilidad de que un elemento elegido al azar caiga en la  $ij$ -ésima celda es  $p_{ij}=u_i v_j$ ,

donde  $u_i$ = probabilidad de que caiga en el renglón  $i$

$v_j$ = probabilidad de que caiga en la columna  $j$

Son estimadores para  $u_i, v_j$  :

$$\hat{u}_i = \frac{1}{N} \sum_{j=1}^c O_{ij}$$

$$\hat{v}_j = \frac{1}{N} \sum_{i=1}^r O_{ij}$$

Por lo tanto, la frecuencia esperada en cada celda es  $E_{ij} = Np_{ij} = Nu_i v_j$ , es decir

$$E_{ij} = \frac{1}{N} \sum_{j=1}^c O_{ij} \sum_{i=1}^r O_{ij}$$

# Pruebas de Hipótesis

## Tablas de Contingencia

Para N grande el siguiente estadístico

$$\chi^2 = \sum_{j=1}^c \sum_{i=1}^r \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Tiene una distribución  $\chi^2$  con  $(r-1)(c-1)$  grados de libertad siempre que la Hipótesis nula sea verdadera.

Por lo tanto, la Hipótesis de independencia se deberá rechazar si el estadístico  $\chi^2 > \chi^2_{\alpha, (r-1)(c-1)}$ .

# Pruebas de Hipótesis

## Tablas de Contingencia

Ejemplo: Los empleados de una compañía eligen uno de tres posibles planes de pensión. La gerencia desea saber con  $\alpha=0.05$  si la preferencia en la elección es independiente de la clasificación del contrato (asalariados y por horas). De una muestra aleatoria de 500 empleados se obtiene la siguiente tabla de contingencia

<b>Tipo de contrato</b>	<b>Plan 1</b>	<b>Plan 2</b>	<b>Plan 3</b>	<b>Total</b>
<b>Asalariados</b>	160	140	40	340
<b>Por Horas</b>	40	60	60	160
<b>Total</b>	200	200	100	500

# Pruebas de Hipótesis

## Tablas de Contingencia

Solución: Necesitaremos las frecuencias esperadas, para ello calculamos estimados de  $u_i$ ,  $v_j$  para  $i=1,2$ ,  $j=1,2,3$ :

$$\begin{array}{lll} u_1=0.68, & u_2=0.32, & \\ v_1=0.4, & v_2=0.4, & v_3=0.2 \end{array}$$

Con esto calculamos las frecuencias esperadas, por ejemplo

$$E_{11} = Nu_1v_1 = 500(0.68)(0.4) = 136$$

El resto se muestran en la siguiente tabla

Tipo de contrato	Plan 1	Plan 2	Plan 3	Total
Asalariados	136	136	68	340
Por Horas	64	64	32	160
Total	200	200	100	500

# Pruebas de Hipótesis

## Tablas de Contingencia

- 1) La variable de interés es la preferencia de los empleados por los planes de pensión
  - 2)  $H_0$ : La preferencia es independiente del tipo de contrato
  - 3)  $H_1$ : La preferencia no es independiente del tipo de contrato
  - 4)  $\alpha=0.05$
  - 5) El estadístico de prueba es
- $$\chi^2 = \sum_{j=1}^c \sum_{i=1}^r \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$
- 6) Como  $r=2$ ,  $c=1$ ,  $\chi^2$  tiene 2 grados de libertad, por lo tanto  $H_0$  debe rechazarse si  $\chi^2 > \chi^2_{0.05,2} = 5.99$
  - 7) Cálculos:  $\chi^2 = 49.63$
  - 8) Como  $49.63 > 5.99$ , se rechaza la hipótesis de independencia. El P-valor para  $\chi^2 = 49.63$  es  $P = 1.671 \times 10^{-11}$