

# Inferencia estadística: Prueba de Hipótesis

Jhon Jairo Padilla A., PhD.

# Justificación

- Es una etapa de análisis de datos de un **experimento comparativo**:
  - Se compara un parámetro de una v.a. con un valor dado.
- Estos **experimentos comparativos simples** se encuentran muy a menudo en la práctica.
- Los experimentos comparativos proporcionan bases sólidas para los **problemas de diseño experimental** (se obtiene un modelo de una población a partir de unas muestras).
- Nuestro interés: Experimentos comparativos en los que interviene sólo una población.

# Ejemplo

- Se tienen interés en el tiempo que dura el procesamiento de un producto.
- La velocidad con que se procesa el producto es una variable aleatoria que puede describirse con una distribución de probabilidad.
- Específicamente, se requiere decidir si: *la media de la velocidad de procesamiento es de 50cm/s.*
- Formalmente, esto se expresa como:
  - $H_0: \mu=50\text{cm/s}$
  - $H_1: \mu\neq 50\text{cm/s}$

# Definiciones

- Hipótesis Nula:
  - Es el enunciado  $H_0$ :  
 $\mu=50\text{cm/s}$
- Hipótesis alternativa:
  - Es el enunciado  $H_1$ :  
 $\mu\neq 50\text{cm/s}$
- Hay dos tipos de hipótesis alternativas:
  - De una cola: Ejemplos:  $H_1$ :  
 $\mu < 50\text{cm/s}$  ó  $H_1$ :  $\mu > 50\text{cm/s}$
  - De dos colas: Ejemplo:  $H_1$ :  
 $\mu \neq 50\text{cm/s}$

Conclusión de rechazo de $H_0$	Tipo de hipótesis alternativa
“Mayor que”, “Menor que”, “superior a”, “que exceda a”, “al menos”	Hipótesis de una cola
No implica dirección, ó la afirmación es “diferente de”	Hipótesis de dos colas

# Prueba de hipótesis

- A un proceso que lleva a una decisión acerca de una hipótesis particular se le llama **prueba de una hipótesis**.
- Las hipótesis son siempre **enunciados acerca de la población o distribución bajo estudio**, no son enunciados acerca de la muestra.

# Objetivos de una prueba de hipótesis

- El valor del parámetro utilizado en la hipótesis nula (ej: 50cm/s) puede ser resultado de alguna de tres situaciones. Según la situación, la prueba de hipótesis tendrá un objetivo diferente.

Fuente del parámetro	Objetivo de la prueba
Experiencia pasada, conocimiento previo del proceso, pruebas o experimentos anteriores	Verificar si el parámetro ha sufrido cambios
Teoría o modelo matemático del proceso bajo estudio	Verificar la teoría o modelo
Especificaciones de diseño u obligaciones contractuales	Prueba de conformidad

# Resultado de la prueba de hipótesis

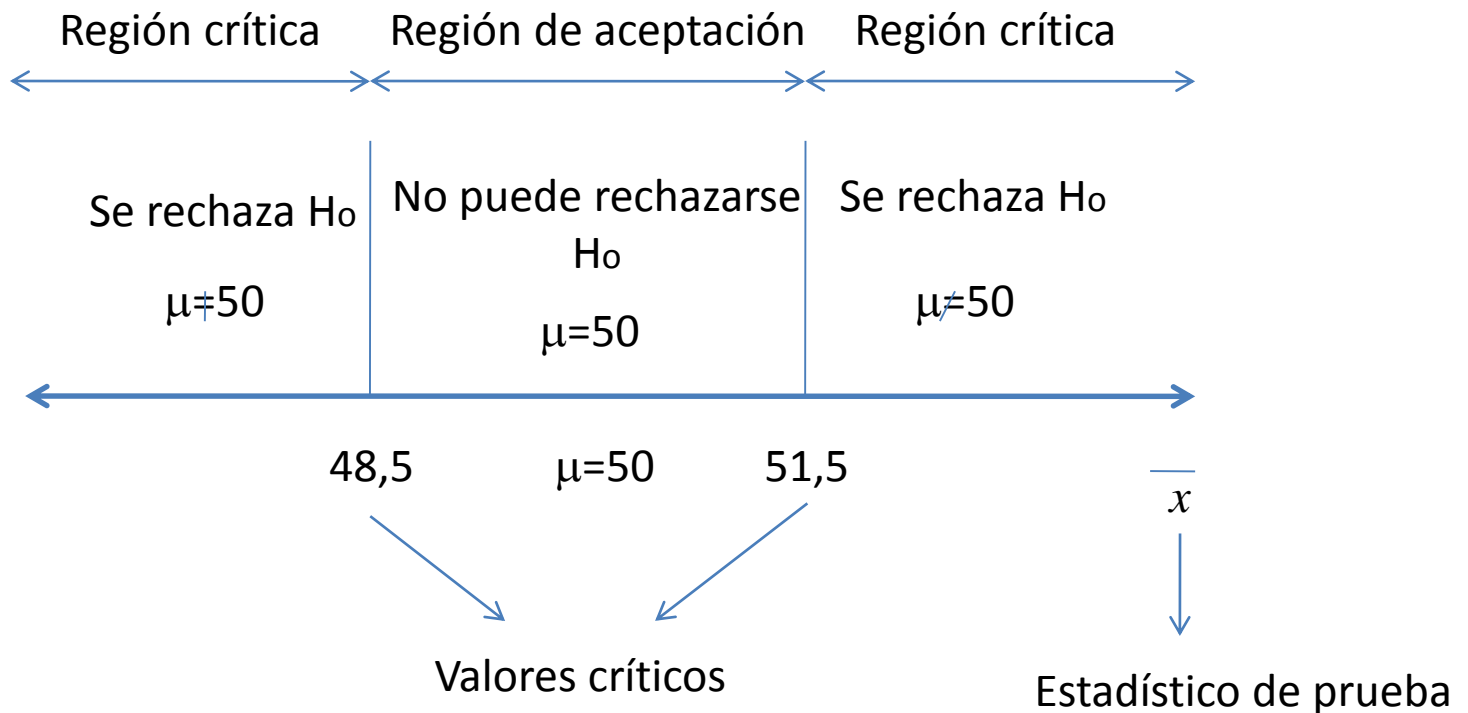
- El resultado obtenido es que:
  - La hipótesis nula es verdadera
  - La hipótesis nula es falsa
- No existe certeza total de que el resultado sea correcto (tendría que examinarse la población completa).
- Existe una probabilidad de llegar a una conclusión incorrecta.
- El rechazo de la hipótesis nula, lleva siempre a aceptar la hipótesis alternativa.

# Procedimiento general

- Tomar una muestra aleatoria
- Calcular un **estadístico de prueba** de los datos muestrales
- Se toma una decisión acerca de la hipótesis nula con base en el estadístico de prueba



# Criterio de aceptación de hipótesis nula



# Conclusiones incorrectas

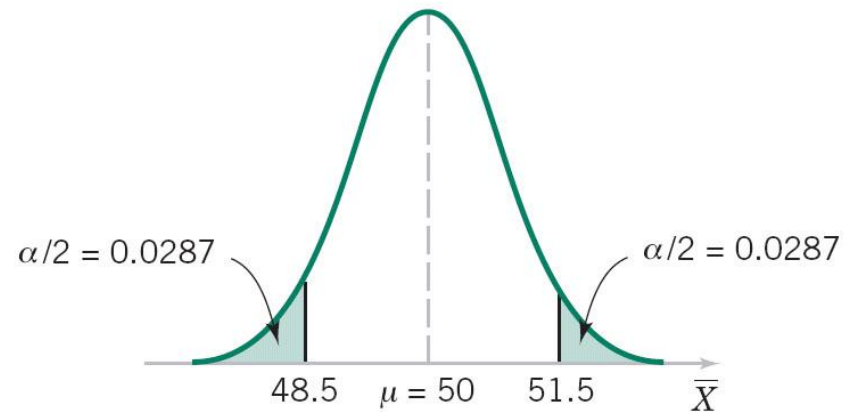
- Podría llegarse a una conclusión incorrecta en varios casos:
  - **Error Tipo I:** Rechazo de la hipótesis nula, cuando en realidad es verdadera  
Puede deberse a que la muestra particular tomada arroja un valor del estadístico en la región crítica cuando en realidad está en la región de aceptación.
  - **Error Tipo II:** Aceptación de la hipótesis nula cuando en realidad es falsa  
Puede deberse a que la media muestral cae en la región de aceptación, cuando en realidad está en la región crítica

# Nivel de Significación de la prueba

- Es la **probabilidad de que ocurra un error de tipo I**.
- Se denota por la letra  $\alpha$ .
- También se le llama **amplitud** de la prueba.
- Suponiendo que se cumplen las condiciones del teorema del límite central, la **probabilidad de error tipo I** será el área de la región sombreada, con:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2,5}{\sqrt{10}} = 0,79$$

Valor dado para el problema



**Figure 9-2** The critical region for  $H_0: \mu = 50$  versus  $H_1: \mu \neq 50$  and  $n = 10$ .

Debe hacerse la transformación de  $X$  a una v.a normal estándar:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Luego,  $\alpha = 0,0287 + 0,0287 = 0,0574$

# Interpretación del resultado

- Para el ejemplo anterior, el 5,74% de las posibles muestras aleatorias que se tomen llevarían a un error tipo I (Rechazo de la hipótesis cuando en realidad es verdadera).
- Cómo se puede reducir la probabilidad de error tipo I?
  - Ampliando la región de aceptación
  - Ó aumentando el número de muestras

# Probabilidad de error de tipo II

- A la probabilidad de que ocurra un error de tipo II se le denotará por  $\beta$ .
- Ejemplo:
  - Se tiene una hipótesis nula con  $\mu=50$ , pero en la realidad  $\mu=52$ .
  - Los valores críticos que limitan la región de aceptación son 48.5 y 51.5
  - Según la muestra tomada, existirá una probabilidad  $\beta$  de que la media esté en la región crítica:  $48.5 \leq \bar{x} \leq 51.5$  (área azul en la figura)

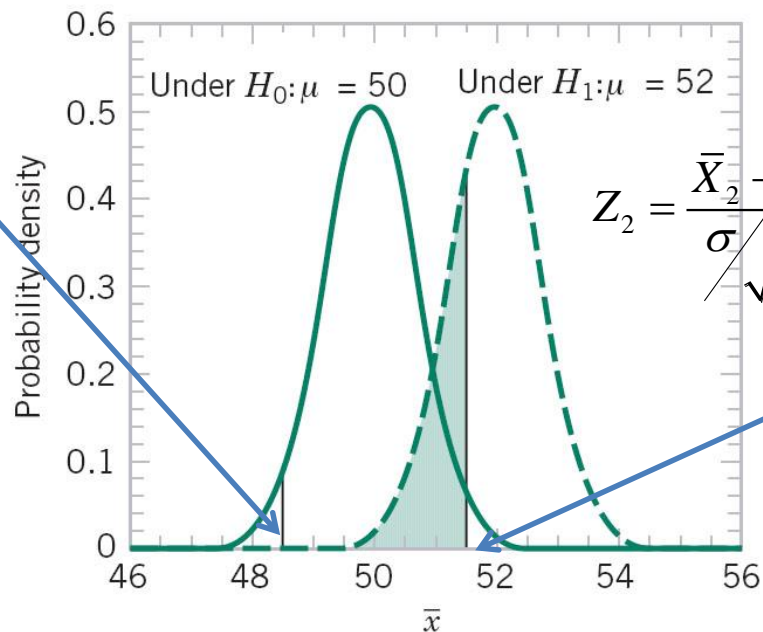
$$Z_1 = \frac{\bar{X}_1 - \mu}{\sigma / \sqrt{n}} = \frac{48,5 - 52}{0,79} = -4,43$$

$$\beta = P(-4,43 \leq Z \leq -0,63)$$

$$\beta = P(Z \leq -0,63) - P(Z \leq -4,43)$$

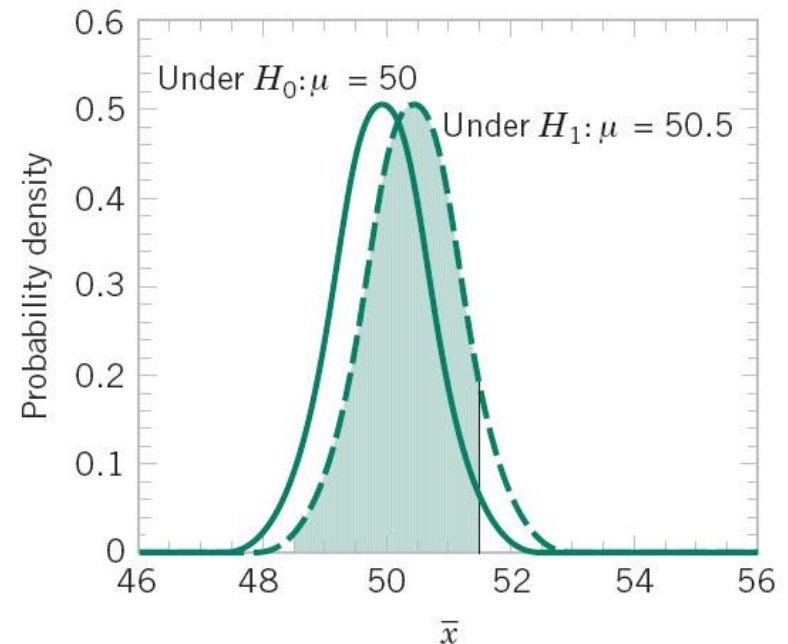
$$\beta = 0,2643 - 0,000 = 0,2643$$

**Figure 9-3** The probability of type II error when  $\mu = 52$  and  $n = 10$ .



# Probabilidad de error de tipo II

- La probabilidad de error de tipo II aumenta conforme el verdadero valor de  $\mu$  se acerca al valor propuesto en la hipótesis.
- Para el ejemplo de la figura, el área azul ( $\beta$ ) aumenta con respecto al ejemplo anterior.
- Si están muy cerca los valores hipotético y real, no preocupa mucho equivocarse y aceptar la hipótesis.



**Figure 9-4** The probability of type II error when  $\mu = 50.5$  and  $n = 10$ .

# Factores que afectan $\alpha$ y $\beta$

Región de Aceptación	Tamaño de la muestra (n)	$\alpha$	$\beta$ con $\mu=52$	$\beta$ con $\mu=50,5$
$48,5 < x < 51,5$	10	0,0576	0,2643	0,8923
$48 < x < 52$	10	0,0114	0,5000	0,9705
$48,5 < x < 51,5$	16	0,0164	0,2119	0,9445
$48 < x < 52$	16	0,0014	0,5000	0,9918

- El tamaño de la región crítica y por consiguiente la probabilidad  $\alpha$  de un error de tipo I, siempre puede reducirse mediante la selección apropiada de los valores críticos.
- Los errores de tipo I y II están relacionados. Una reducción de la probabilidad de un tipo de error resulta en un incremento en la probabilidad del otro (si no cambia el tamaño de la muestra, n)
- Un incremento del tamaño de la muestra (n) generalmente dará como resultado una reducción tanto de  $\alpha$  como de  $\beta$  (si los valores críticos permanecen constantes).
- Cuando una hipótesis nula es falsa,  $\beta$  se incrementa conforme el verdadero valor del parámetro se aproxima al valor hipotético propuesto en la hipótesis nula.

# Características de las conclusiones

- El analista puede controlar los valores críticos, luego puede controlar la probabilidad de rechazar incorrectamente a  $H_0$ .
- Por tanto, al rechazo de la hipótesis nula  $H_0$  siempre se le considera una **conclusión robusta**.
- La probabilidad de error de tipo II depende del tamaño de la muestra ( $n$ ) y del valor verdadero del parámetro a probar.
- Se considera que la decisión de aceptar  $H_0$  es una **conclusión débil** (a menos que  $\beta$  sea muy pequeña).
- **Terminología:**
  - En lugar de decir que “se acepta  $H_0$ ”, se prefiere decir “No puede rechazarse  $H_0$ ”
  - Por tanto “No puede rechazarse  $H_0$ ” significa que no se ha encontrado evidencia suficiente para rechazar  $H_0$ .
  - El no poder rechazar  $H_0$  no significa necesariamente que exista una alta probabilidad de que sea verdadera. Puede significar simplemente que se requieren más datos para llegar a una conclusión robusta.



# Potencia de una prueba estadística

- La potencia de una prueba estadística es la probabilidad de rechazar la hipótesis nula  $H_0$  cuando la hipótesis alternativa es verdadera.
- $\text{Potencia} = 1 - \beta$
- La potencia de una prueba estadística es la probabilidad de rechazar correctamente una hipótesis nula.

# Potencia de una prueba estadística

- La potencia es una medida de la **sensibilidad** de una prueba estadística.
- **Sensibilidad:** capacidad de la prueba para detectar diferencias.
- **Ejemplo:**
  - $n=10$
  - $\mu=52$
  - $H_0: \mu=50$

Potencia =  $1 - \beta = 1 - 0,2643 = 0,7357$
- **Significado:**
  - Si la media verdadera es 52, esta prueba rechazará correctamente la hipótesis  $H_0: \mu=50$  y detectará esta diferencia el 73,57% de las veces.
  - Si el valor de la sensibilidad se considera muy bajo, el analista puede incrementar  $\alpha$  o el tamaño de la muestra  $n$ .

# Procedimiento general de las pruebas de hipótesis

1. Por el contexto del problema, identificar el parámetro de interés.
2. Establecer la hipótesis nula  $H_0$ .
3. Especificar una hipótesis alternativa apropiada,  $H_1$ .
4. Elegir un nivel de significación  $\alpha$ .
5. Establecer un estadístico de la prueba apropiado.
6. Establecer la región de rechazo del estadístico.
7. Calcular las cantidades muestrales necesarias, sustituirlas en la ecuación del estadístico de la prueba, y calcular ese valor.
8. Decidir si deberá rechazarse o no  $H_0$  y contextualizar la decisión del problema.

# Intervalo de confianza para la media, varianza conocida

- Si  $x$  es la media muestral de una muestra aleatoria de tamaño  $n$  de una población con varianza conocida  $\sigma^2$ , un intervalo de confianza del  $100(1-\alpha)$  por ciento para  $\mu$  está dado por

$$x - z_{\alpha/2} \sigma / \sqrt{n} \leq \mu \leq x + z_{\alpha/2} \sigma / \sqrt{n}$$

- Donde  $z_{\alpha/2}$  es el punto porcentual  $100\alpha/2$  superior de la distribución normal estándar.
- El intervalo de confianza que se obtiene con esta aproximación dará buenos resultados si el número de muestras  $n \geq 30$  (o si la distribución de la población es normal).

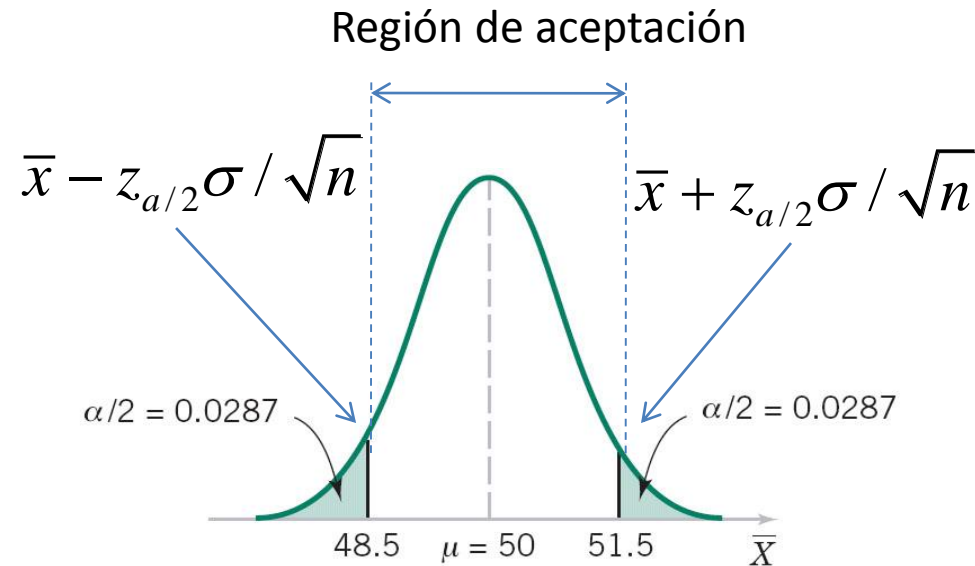
# Relación entre intervalos de confianza y pruebas de hipótesis

- Si  $[l, u]$  es un intervalo del  $100(1-\alpha)$  por ciento para el parámetro  $\theta$ , entonces la prueba de amplitud  $\alpha$  de la hipótesis

$$H_0: \theta = \theta_0$$

$$H_1: \theta \neq \theta_0$$

Llevará al rechazo de  $H_0$  si y sólo si  $\theta_0$  no está en el intervalo de confianza  $[l, u]$  del  $100(1-\alpha)$  por ciento.



**Figure 9-2** The critical region for  $H_0: \mu = 50$  versus  $H_1: \mu \neq 50$  and  $n = 10$ .

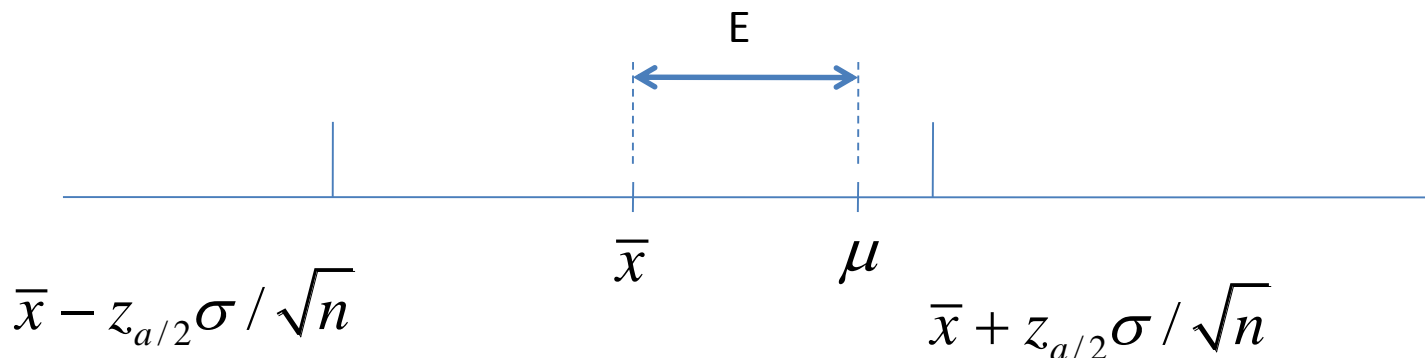
# Elección del nivel de confianza

- Cuál nivel de confianza escoger? ...95%...99%?
- Es razonable un nivel de confianza alto?
- Intervalo confianza con  $\alpha=0.01$ :
  - $Z_{\alpha/2}=z_{0,005}=2,58$
  - Longitud del intervalo:  $2(2,58\sigma / \sqrt{n}) = 5,16\sigma / \sqrt{n}$
- Intervalo confianza con  $\alpha=0.05$ :
  - $Z_{\alpha/2}=z_{0,025}=1,96$
  - Longitud del intervalo:  $2(1.96\sigma / \sqrt{n}) = 3,92\sigma / \sqrt{n}$
- El intervalo de confianza es más largo para una confianza del 99%.
- Es deseable un intervalo de confianza que sea lo suficientemente corto para fines de toma de decisiones, pero que ofrezca a su vez una confianza adecuada. Esto se puede manejar eligiendo un tamaño de muestra adecuado

# Elección del tamaño de la muestra

- Si se usa la media muestral ( $\bar{x}$ ) como estimación de  $\mu$ , puede tenerse una confianza del  $100(1-\alpha)$  por ciento de que el error  $|\bar{x}-\mu|$  no excederá una cantidad especificada  $E$  cuando el tamaño de la muestra es

$$n = \left( \frac{z_{\alpha/2} \sigma}{E} \right)^2$$



# Ejemplo

- Suponga que se quisiera que el error al estimar la velocidad media de procesamiento de un producto fuera menor que 1.5cm/s, con una confianza del 95%. Suponga una  $\sigma=2$ .
- Solución:
  - El valor de  $\alpha$  es  $\alpha=0,05$
  - Luego  $z_{\alpha/2}=z_{0,025}=1,96$
  - Por tanto, el tamaño de la muestra será

$$n = \left( \frac{z_{\alpha/2} \sigma}{E} \right)^2 = \left[ \frac{(1.96) \cdot 2}{1.5} \right]^2 = 6,83 \approx 7$$



# Prueba de la bondad del ajuste

- Es un tipo de prueba de hipótesis donde se desconoce la distribución de probabilidad de la población.
- El **objetivo** es determinar si una distribución particular será satisfactoria como modelo de la población.
- **Algunos métodos:** gráficas de probabilidad, prueba ji-cuadrada, prueba de cuantiles.

# Procedimiento

- Se toma una muestra aleatoria de tamaño  $n$  de la población cuya distribución de probabilidad es desconocida.
- Las  $n$  observaciones se ordenan en un histograma con  $k$  intervalos de clase:
  - Sea  $O_i$  la frecuencia observada en el intervalo de clase  $i$ .
- Se calcula la frecuencia esperada a partir de la distribución de probabilidad hipotética.
- Para el intervalo de clase  $i$ -ésimo, denotado por  $E_i$ , el estadístico de prueba es

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

# Interpretación de resultados

- Si la población sigue la distribución hipotética propuesta,  $\chi_0^2$  tiene aproximadamente una distribución ji-cuadrada con  $k-p-1$  grados de libertad.
- Donde  $p$  representa el número de parámetros de la distribución hipotética, estimados por los estadísticos muestrales.
- Esta aproximación mejora conforme  $n$  se incrementa.
- La hipótesis de que la distribución de la población es la distribución hipotética se rechazaría si el valor calculado del estadístico de prueba  $\chi_0^2 > \chi_{\alpha, k-p-1}^2$  .

# Condiciones de la prueba Ji-cuadrado

- Si las frecuencias esperadas son muy pequeñas (tanto como 1 ó 2), entonces el estadístico de prueba  $\chi_0^2$  no reflejará la desviación de las frecuencias observadas y las esperadas, sino únicamente la pequeña magnitud de las frecuencias esperadas.
- El mínimo de las frecuencias esperadas debe estar entre 3 y 5.
- Algunos autores proponen que una frecuencia esperada podría ser tan pequeña como 1 ó 2 siempre que la mayoría de las frecuencias excedan de 5.
- Si alguna frecuencia es muy pequeña, es mejor combinar su clase con una clase adyacente, tanto en los valores observados como en los valores obtenidos de la distribución hipotética. Esto reduce  $k$ . (No es necesario que los intervalos de clase tengan la misma anchura).

# Ejemplo: Una distribución de Poisson

- Se propone hipotéticamente que el número de defectos en tarjetas de circuitos impresos sigue una distribución de Poisson. Se ha colectado una muestra aleatoria de  $n=60$  tarjetas de circuitos impresos y se observó el número de defectos. Los datos de la tabla son el resultado.

Número de defectos	Frecuencia observada
0	32
1	15
2	9
3	4

# Solución

- La media es desconocida y debe estimarse de los datos muestrales.
- La media muestral será entonces:

$$\bar{x} = \frac{(32 \cdot 0) + (15 \cdot 1) + (9 \cdot 2) + (4 \cdot 3)}{60} = 0,75$$

- Por tanto, la tasa media de la distribución de Poisson será  $\lambda=0,75$ .
- Se calcula la probabilidad para cada valor de  $x$  con la expresión de la distribución de Poisson (con  $t=1$ ):

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

- Obteniéndose los valores de la tabla para la frecuencia esperada como  $E_i=np_i$


# Solución

## Frecuencia esperada

Número de defectos	Probabilidad	Frecuencia esperada
0	0,472	28,32
1	0,354	21,24
2	0,133	7,98
3 ó más	0,041	2,46

## Frecuencias corregidas

Número de defectos	Frecuencia observada	Frecuencia esperada
0	32	28,32
1	15	21,24
2 ó más	13	10,44



Se unen en una sola clase,  
por ser la última  
frecuencia menor de 3

# Solución: Prueba de hipótesis

1. La variable de interés es la forma de distribución de los defectos en las tarjetas de circuitos impresos.
2.  $H_0$ : la forma de la distribución de los defectos es de Poisson.
3.  $H_1$ : La forma de la distribución de los defectos no es de Poisson.
4.  $\alpha=0,05$
5. El estadístico de la prueba es  $\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$
6. Se rechaza  $H_0$  si  $\chi_0^2 > \chi_{0,05,3-1-1}^2 = \chi_{0,05,1}^2 = 3,84$
7. Cálculos:  $\chi_0^2 = \frac{(32 - 28,32)^2}{28,32} + \frac{(15 - 21,24)^2}{21,24} + \frac{(13 - 10,44)^2}{10,44} = 2,94$
8. Conclusión: Puesto que el estadístico ji-cuadrado es menor que 3.84, no puede rechazarse la hipótesis nula de que la distribución de los defectos en las tarjetas de circuitos impresos es de Poisson.