Capítulo 1

Contrastes de Hipótesis paramétricos y no-paramétricos.

Estadística Inductiva o Inferencia Estadística: Conjunto de métodos que se fundamentan en la Teoría de la Probabilidad y que tienen por finalidad generalizar los resultados, obtenidos mediante una muestra, a toda una población.

CONTRASTES DE HIPÓTESIS

Procedimientos para aceptar o rechazar una hipótesis que se emite acerca de un parámetro u otra característica de la población.

ETAPAS DEL PROCESO

1) El investigador formula una hipótesis sobre un parámetro poblacional, por ejemplo que toma un determinado valor
2) Selecciona una muestra de la población
3) Comprueba si los datos están o no de acuerdo con la hipótesis planteada, es decir compara la observación con la teoría
    a) Si lo observado es incompatible con lo teórico entonces el experimentador puede rechazar la hipótesis planteada y proponer una nueva teoría.
    b) Si lo observado es compatible con lo teórico entonces el experimentador puede continuar como si la hipótesis fuera cierta.

TIPOS DE HIPÓTESIS

- H_0 : **Hipótesis Nula** es la hipótesis sobre la que se desea decidir
- H_1 : **Hipótesis Alternativa** es la hipótesis que se acepta, si se rechaza la hipótesis nula. Generalmente la hipótesis alternativa es la negación de la hipótesis nula .

★ Un Contraste o Test de Hipótesis es un procedimiento mediante el cual nos decidimos por H_0 o por H_1.

TIPOS DE ERRORES

- **Error de Tipo I o error α** : Rechazar la hipótesis H_0 cuando es cierta.

$$P\left[\text{rechazar } H_0 \;/\; H_0 \text{ es cierta}\right] = \alpha; \quad 0 \leq \alpha \leq 1$$

- **Error de Tipo II o error β**: Aceptar la hipótesis H_0 cuando es falsa

$$P\left[\text{decidir } H_0 \;/\; H_0 \text{ es falsa}\right] = \beta; \quad 0 \leq \beta \leq 1$$

TIPOS DE REGIONES

- **Región Crítica** o **Región de Rechazo**: Los valores del estadístico de contraste que nos conducen a rechazar la hipótesis H_0 forman la Región Crítica o Región de Rechazo del contraste
- **Región de Aceptación**: Los valores del estadístico de contraste que nos conducen a decidir H_0 forman la Región de Aceptación.

★ **Nivel de significación:** Es el error α, es decir la probabilidad de que el estadístico de contraste caiga en la región de rechazo.

POTENCIA DE UN CONTRASTE

$$\boxed{P(\theta) = 1 - \beta(\theta) = P\left[\text{rechazar } H_0 \;/\; H_0 \text{ es falsa}\right] = P\left[\text{decidir } H_1 \;/\; H_1 \text{ es cierta}\right]}$$

		Decisión	
		Rechazar H_0	Aceptar H_0
Hipótesis cierta	H_0	α	Decisión correcta
Hipótesis falsa	H_0	Decisión correcta Potencia	β

RESULTADO DE UN CONTRASTE DE HIPÓTESIS

★ **Estadísticamente Significativo:** cuando se rechaza H_0

★ **Estadísticamente No-Significativo**: cuando se acepta H_0.

CONTRASTES PARAMÉTRICOS

Se conoce la forma de la distribución y los parámetros son desconocidos

TIPOS DE REGIONES CRÍTICAS

$$\boxed{\begin{array}{l} H_0 \equiv \theta \leqslant \theta_0 \\ H_1 \equiv \theta > \theta_0 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv \theta \geqslant \theta_0 \\ H_1 \equiv \theta < \theta_0 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv \theta = \theta_0 \\ H_1 \equiv \theta \neq \theta_0 \end{array}}$$

- $H_1 \equiv \theta > \theta_0$ ó $H_1 \equiv \theta < \theta_0$: **Hipótesis Alternativa es Unilateral. (Región Crítica Unilateral)**
- $H_1 \equiv \theta \neq \theta_0$: **Hipótesis Alternativa es Bilateral (R. C. Bilateral.)**

★ $H_0 \equiv \theta = \theta_0$: **Hipótesis nula Sencilla o Simple**

★ $H_0 \equiv \theta \leqslant \theta_0$ ó $H_0 \equiv \theta \geqslant \theta_0$: **Hipótesis nula Compuesta**

CRITERIOS GENERALES PARA LOS CONTRASTES

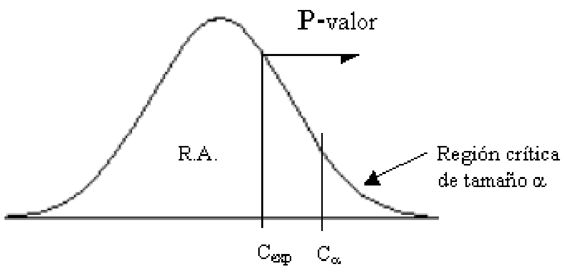
- Calcular una cantidad experimental ($C_{\exp}$) a partir de los datos
- Calcular una cantidad teórica (C_α) a partir de las tablas

$$\boxed{\text{Si } C_{\exp} < C_\alpha \Rightarrow \text{Aceptar } H_0 \quad ; \quad \text{Si } C_{\exp} \geq C_\alpha \Rightarrow \text{Rechazar } H_0}$$

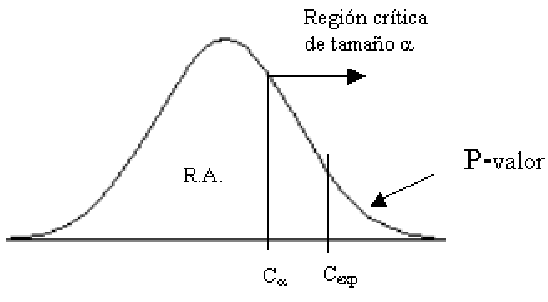
NIVEL MÍNIMO DE SIGNIFICACIÓN

Nivel crítico (valor P o **$P-$value o nivel mínimo de significación**): Es el error de la primera región crítica de rechazo. Es el área que deja a la derecha la $C_{\exp}$

$$\boxed{\text{Si } \alpha < P \Rightarrow \text{Aceptar } H_0 \quad ; \quad \text{Si } \alpha \geq P \Rightarrow \text{Rechazar } H_0}$$



Aceptar H_0

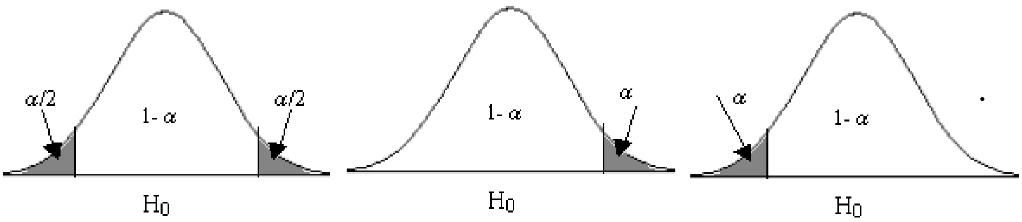


Rechazar H_0

CONTRASTES DE HIPÓTESIS DE UNA POBLACIÓN NORMAL

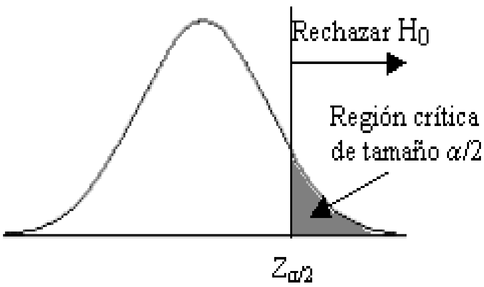
CONTRASTES DE HIPÓTESIS PARA LA MEDIA DE UNA POBLACIÓN NORMAL

$$\boxed{\begin{array}{l} H_0 \equiv \mu = \mu_0 \\ H_1 \equiv \mu \neq \mu_0 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv \mu \leqslant \mu_0 \\ H_1 \equiv \mu > \mu_0 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv \mu \geqslant \mu_0 \\ H_1 \equiv \mu < \mu_0 \end{array}}$$



a) Varianza poblacional σ^2 conocida.

a1) $\boxed{\begin{array}{l} H_0 \equiv \mu = \mu_0 \\ H_1 \equiv \mu \neq \mu_0 \end{array}}$

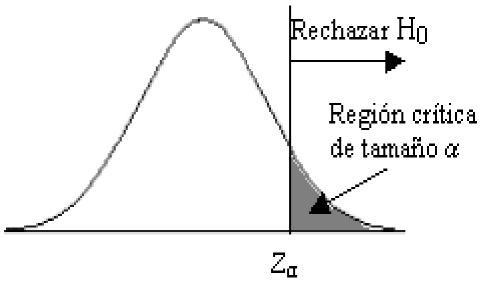


Si $|Z_{\text{exp}}| < z_{\alpha/2} \Rightarrow$ Se acepta H_0

Si $|Z_{\text{exp}}| \geq z_{\alpha/2} \Rightarrow$ Se rechaza H_0

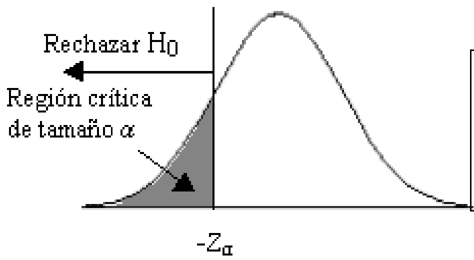
.

. a2) $\boxed{\begin{array}{l} H_0 \equiv \mu \leqslant \mu_0 \\ H_1 \equiv \mu > \mu_0 \end{array}}$



Si $Z_{\exp} < z_\alpha \Rightarrow$ Se acepta H_0
Si $Z_{\exp} \geq z_\alpha \Rightarrow$ Se rechaza H_0

. a3) $\boxed{\begin{array}{l} H_0 \equiv \mu \geqslant \mu_0 \\ H_1 \equiv \mu < \mu_0 \end{array}}$



Si $Z_{\exp} > -z_\alpha \Rightarrow$ Se acepta H_0
Si $Z_{\exp} \leqslant -z_\alpha \Rightarrow$ Se rechaza H_0

b) Varianza poblacional σ^2 desconocida

Hipótesis alternativa	Regla de decisión
$H_1 \equiv \mu \neq \mu_0$	Rechazar H_0 cuando $\|t_{\exp}\| \geq t_{\alpha/2}$
$H_1 \equiv \mu > \mu_0$	Rechazar H_0 cuando $t_{\exp} \geq t_\alpha$
$H_1 \equiv \mu < \mu_0$	Rechazar H_0 cuando $t_{\exp} \leq -t_\alpha$

CONTRASTES DE HIPÓTESIS PARA LA VARIANZA DE UNA POBLACIÓN NORMAL

$$\boxed{\begin{array}{l} H_0 \equiv \sigma^2 = \sigma_0^2 \\ H_1 \equiv \sigma^2 \neq \sigma_0^2 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv \sigma^2 \leqslant \sigma_0^2 \\ H_1 \equiv \sigma^2 > \sigma_0^2 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv \sigma^2 \geqslant \sigma_0^2 \\ H_1 \equiv \sigma^2 < \sigma_0^2 \end{array}}$$

Hipótesis alternativa	Regla de decisión
$H_1 \equiv \sigma^2 \neq \sigma_0^2$	Rechazar H_0 cuando $\chi^2_{\text{exp}} \geq \chi^2_{\alpha/2}$ ó $\chi^2_{\text{exp}} \leq \chi^2_{1-\alpha/2}$
$H_1 \equiv \sigma^2 > \sigma_0^2$	Rechazar H_0 cuando $\chi^2_{\text{exp}} \geq \chi^2_{\alpha}$
$H_1 \equiv \sigma^2 < \sigma_0^2$	Rechazar H_0 cuando $\chi^2_{\text{exp}} \leq \chi^2_{1-\alpha}$

a) Media poblacional conocida

b) Media poblacional desconocida.

CONTRASTES DE HIPÓTESIS PARA EL PARÁMETRO P DE UNA DISTRIBUCIÓN BINOMIAL

$$\boxed{\begin{array}{l} H_0 \equiv p = p_0 \\ H_1 \equiv p \neq p_0 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv p \leqslant p_0 \\ H_1 \equiv p > p_0 \end{array}} \quad \text{ó} \quad \boxed{\begin{array}{l} H_0 \equiv p \geqslant p_0 \\ H_1 \equiv p < p_0 \end{array}}$$

Hipótesis alternativa	Regla de decisión
$H_1 \equiv p \neq p_0$	Rechazar H_0 cuando $\lvert Z_{\text{exp}} \rvert \geq z_{\alpha/2}$
$H_1 \equiv p > p_0$	Rechazar H_0 cuando $Z_{\text{exp}} \geq z_{\alpha}$
$H_1 \equiv p < p_0$	Rechazar H_0 cuando $Z_{\text{exp}} \leq -z_{\alpha}$

CONTRASTES DE HIPÓTESIS EN DOS POBLACIONES NORMALES INDEPENDIENTES

CONTRASTES DE COMPARACIÓN DE DOS MEDIAS

$$H_0 \equiv \mu_1 = \mu_2 \quad H_1 \equiv \mu_1 \neq \mu_2$$ ó $$H_0 \equiv \mu_1 \leqslant \mu_2 \quad H_1 \equiv \mu_1 > \mu_2$$ ó $$H_0 \equiv \mu_1 \geqslant \mu_2 \quad H_1 \equiv \mu_1 < \mu_2$$

a) Varianzas poblaciones conocidas

b) Varianzas poblaciones desconocidas

CONTRASTES DE COMPARACIÓN DE DOS VARIANZAS

$$H_0 \equiv \sigma_1^2 = \sigma_2^2 \quad H_1 \equiv \sigma_1^2 \neq \sigma_2^2$$ ó $$H_0 \equiv \sigma_1^2 \leqslant \sigma_2^2 \quad H_1 \equiv \sigma_1^2 > \sigma_2^2$$ ó $$H_0 \equiv \sigma_1^2 \geqslant \sigma_2^2 \quad H_1 \equiv \sigma_1^2 < \sigma_2^2$$

Hip. alternativa	Regla de decisión
$H_1 \equiv \sigma_1^2 \neq \sigma_2^2$	Rechazar H_0 si $F_{\exp} \geq F_{\alpha/2}$ o $F_{\exp} \leq F_{1-\alpha/2}$
$H_1 \equiv \sigma_1^2 > \sigma_2^2$	Rechazar H_0 si $F_{\exp} \geq F_{\alpha}$
$H_1 \equiv \sigma_1^2 < \sigma_2^2$	Rechazar H_0 si $F_{\exp} \leq F_{1-\alpha}$

a) Media poblacional conocida

b) Media poblacional desconocida

CONTRASTES DE COMPARACIÓN DE DOS PROPORCIONES

$$H_0 \equiv p_1 = p_2 \quad H_1 \equiv p_1 \neq p_2$$ ó $$H_0 \equiv p_1 \leqslant p_2 \quad H_1 \equiv p_1 > p_2$$ ó $$H_0 \equiv p_1 \geqslant p_2 \quad H_1 \equiv p_1 < p_2$$

CONTRASTES DE HIPÓTESIS PARA COMPARAR DOS MEDIAS DE VARIABLES NORMALES: MUESTRAS APAREADAS

NOTA: Dos muestras se dicen **independientes** cuando las obsevaciones de una de ellas no condicionan para nada a las observaciones de la otra, siendo **dependientes** en caso contrario.

El tipo de dependencia que se considera a estos efectos es muy especial: cada dato de una muestra tiene un homónimo en la otra con el que está relacionado, de ahí el nombre alternativo de **muestras apareadas.**

Ejemplo: Consideremos que se desea estudiar el efecto de un fármaco presuntamente antihipertensivo. El experimento podría planificarse:

a) Se toman 20 hipertensos al azar, se le aplica el fármaco a 10 de ellos dejando sin tratamiento a los otros 10. Transcurrido un tiempo se miden las presiones sanguíneas de ambos grupos y se contrasta la hipótesis $H_0 \equiv \mu_1 = \mu_2$ para evaluar si las medias son iguales o no. Las dos muestras están formadas por individuos distintos, sin relación entre sí $\Longrightarrow$ **muestras independientes**

b) Se administra el fármaco a los 20 hipertensos disponibles y se anota su presión sanguínea antes y después de la administración del mismo. En este caso los datos vienen dados por parejas (presión antes y después) y parece lógico que tales datos se encuentren relacionados entre sí $\Longrightarrow$ **muestras apareadas**

$\boxed{\begin{array}{l} H_0 \equiv \mu_1 = \mu_2 \\ H_1 \equiv \mu_1 \neq \mu_2 \end{array}}$ ó $\boxed{\begin{array}{l} H_0 \equiv \mu_1 \leqslant \mu_2 \\ H_1 \equiv \mu_1 > \mu_2 \end{array}}$ ó $\boxed{\begin{array}{l} H_0 \equiv \mu_1 \geqslant \mu_2 \\ H_1 \equiv \mu_1 < \mu_2 \end{array}}$

Hipótesis alternativa	Regla de decisión
$H_1 \equiv \mu \neq \mu_0$	Rechazar H_0 cuando $\lvert t_{\exp}\rvert \geq t_{n-1;\alpha/2}$
$H_1 \equiv \mu > \mu_0$	Rechazar H_0 cuando $t_{\exp} \geq t_{n-1;\alpha}$
$H_1 \equiv \mu < \mu_0$	Rechazar H_0 cuando $t_{\exp} \leq -t_{n-1;\alpha}$

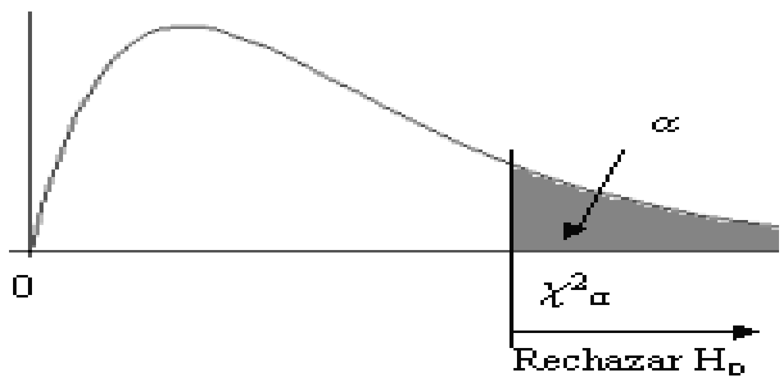
CONTRASTES NO-PARAMÉTRICOS

CONTRASTES DE HIPÓTESIS BASADOS EN LA CHI-CUADRADO DE PEARSON

NO SE CONOCE LA FORMA DE LA DISTRIBUCIÓN DE LA POBLACIÓN

El contraste que planteamos, a diferencia de los estudiados, no se refiere a un valor concreto de un parámatro desconocido

$H_0 \equiv X \rightsquigarrow L(X)$ (sigue una ley)
$H_1 \equiv X \not\rightsquigarrow L(X)$ (no sigue dicha ley)



Si $\chi^2_{\exp} < \chi^2_{\alpha} \Rightarrow$ Se acepta H_0

Si $\chi^2_{\exp} \geq \chi^2_{\alpha} \Rightarrow$Se rechaza H_0

CONTRASTES PARA LA BONDAD DE AJUSTE

El objetivo de los Contrastes de Bondad de Ajuste a Distribuciones es determinar a través de una muestra aleatoria, si una variable aleatoria sigue una cierta distribución teórica dada de antemano (Binomial, Poisson, Normal, Uniforme etc).

$H_0 \equiv$ La distribución teórica está conforme con la distribución empírica
$H_1 \equiv$ La distribución teórica no está conforme con la distribución empírica

CONTRASTES PARA LA INDEPENDENCIA DE DOS CARACTERES

El objetivo de estos contrastes es comprobar si dos características cualitativas están relacionadas entre sí. Por ejemplo, ¿Existe relación entre el color de la piel y el color del pelo? o ¿existe relación entre fumar cigarillos y la predisposición a desarrollar el cáncer de pulmón?

$H_0 \equiv$ Los caracteres A y B son independientes
$H_1 \equiv$ Los caracteres A y B no son independientes

CONTRASTES DE HOMOGENEIDAD

El problema general es determinar si varias muestras cualitativas se pueden considerar procedentes de una misma población en cuyo caso decimos que las muestras son homogéneas. Ejemplos de problemas de homogeneidad se pueden plantear en términos de comprobar si varios tratamientos, que curan una misma enfermedad, aplicados a un cierto tipo de enfermos son homogéneos respecto a los resultados obtenidos.

Bibliografía utilizada:

★ **Lara Porras A.M. (2002).** *"Estadística para Ciencias Biológicas y Ciencias Ambientales. Problemas y Exámenes Resueltos". Ed. Proyecto Sur.*

★ **Martín Andrés, A. y Luna del Castillo, J. de D. (1990).** *"Bioestadística para las Ciencias de la Salud". Ed. Norma*

⧫ **Temporalización:** Dos horas