

## Tema 12: Contrastes Paramétricos

### Presentación y Objetivos.

Se comienza este tema introduciendo la terminología y conceptos característicos de los contrastes de hipótesis, típicamente a través de un ejemplo sencillo procediendo posteriormente a su definición formal. Se distinguirán los tipos de hipótesis, las etapas en la formulación de un contraste, los tipos de errores a los que estamos sujetos, potencia, nivel de significación y el método del nivel crítico o p-valor. Estos dos últimos servirán para tomar la decisión final. Hay que hacer hincapié en que el rechazo de la hipótesis nula supone una fuerte evidencia de que ésta no está soportada por los datos. Sin embargo, su aceptación no supone que sea fuerte la evidencia a su favor, sino que la evidencia en su contra no ha sido determinante. Por ello, antes que aceptar la hipótesis a contrastar (hipótesis nula), se preferirá no rechazarla. Una vez introducidos en el lenguaje de los contrastes, se comienza con el estudio del catálogo, similar al de intervalos de confianza. Se comienza con el contraste para la media de una población normal que se puede ilustrar con un ejemplo. Un contraste permitirá decidir, entre otras cosas, si un determinado tratamiento o procedimiento ha tenido algún efecto o el efecto deseado. La relación entre intervalos de confianza y contrastes de hipótesis debe quedar ilustrada de forma clara y concisa ya que las variables pivote utilizadas en los primeros son análogas a los estadísticos de contraste utilizados en los segundos, facilitando de esta forma la comprensión y estudio de la materia.

Los objetivos de este tema son:

- Conocer la motivación del contraste paramétrico para simplificar un modelo.
- Manejar la terminología básica del contraste de hipótesis.
- Saber aplicar en diferentes casos un contraste de hipótesis clásico y entender su relación con los intervalos de confianza.

### Esquema Inicial

1. Introducción.
2. Tipos de hipótesis.

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

---

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70**

Confianza en poblaciones normales.

8. Contraste para la comparación de medias en poblaciones normales.
9. Contraste para la comparación de varianzas en poblaciones normales.
10. Contrastes de hipótesis para proporciones.
11. Relación entre intervalo de confianza y contraste de hipótesis.

## Desarrollo del Tema

### 1. Introducción

Una **hipótesis estadística** es una afirmación sobre alguna característica (parámetro o forma), desconocida de la población de interés. El hecho de probar esa hipótesis estadística es esencialmente decidir si la afirmación está apoyada por la evidencia muestral.

Contrastaremos una hipótesis comparando lo que debería pasar si la hipótesis es cierta con lo que realmente pasa (lo observado en la muestra). Si coinciden, dentro de un margen de error admisible, mantendremos la hipótesis. En caso contrario, la rechazaremos y buscaremos nuevas hipótesis que puedan explicar los datos observados.

#### Ejemplo:

1. El departamento de programación de una empresa lleva mucho tiempo construyendo un software muy complicado. Su director ha observado durante este tiempo que el tiempo de compilación y montaje de los distintos módulos sigue una distribución  $N(5, 0.1)$ . Le ofrecen una herramienta que aceleraría este tiempo, pero como es muy cara, antes de decidir comprarla o no, pide una versión de evaluación y recoge los tiempos de compilación de 6 módulos genéricos: 4.71, 4.82, 5.01, 4.75, 4.82 y 4.95. Con estos datos tiene que decidir si existe evidencia para suponer que la nueva herramienta altera el tiempo medio de compilación.

### 2. Tipos de Hipótesis

Llamamos hipótesis estadística a una suposición que determina parcial o totalmente la distribución de una o varias variables aleatorias. Puede:

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

---

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70**

En un contraste distinguimos los dos tipos básicos de hipótesis:

1. Hipótesis nula  $H_0$ : Es la hipótesis que se contrasta, es decir, la que mantendremos y consideraremos como verdadera a no ser que los datos indiquen su falsedad.
2. Hipótesis alternativa  $H_1$ : Si rechazamos  $H_0$  estamos aceptando implícitamente una hipótesis alternativa  $H_1$  (que reflejará situaciones de interés si  $H_0$  es falsa).

Además, si el contraste es paramétrico, las hipótesis pueden ser:

1. Hipótesis simple: Aquélla que especifica un único valor para el o los parámetros en cuestión. Ej:  $\theta = \theta_0$ .
2. Hipótesis compuesta: Aquélla que da un rango de valores posibles. Ej:  $\theta > \theta_0$ ,  $a \leq \theta \leq b$ .

En los contrastes paramétricos, si  $H_0$  es simple, del tipo  $\theta = \theta_0$ , los casos más importantes de hipótesis alternativa son:

1.  $H_1 : \theta \neq \theta_0$ . El contraste es BILATERAL .
2.  $H_1 : \theta > \theta_0$ , ( ó  $\theta < \theta_0$ ). El contraste es UNILATERAL.

### 3. Tipos de errores

Al decidir rechazar  $H_0$  o no rechazarla se pueden cometer dos tipos de errores, error de tipo I o error de tipo II.

El error de tipo I se comete cuando rechazamos  $H_0$  siendo cierta. Su probabilidad se denota con  $\alpha$  y se llama **nivel de significación**.

$$\alpha = P(\text{rechazar } H_0 \mid H_0 \text{ cierta}) = P(\text{cometer ET I})$$

El error de tipo II se comete cuando se decide aceptar  $H_0$  siendo falsa. Su probabilidad se denota con  $\beta$  (cuando la hipótesis alternativa es simple).

$$\beta = P(\text{no rechazar } H_0 \mid H_0 \text{ falsa}) = P(\text{cometer ET II})$$

Lo ideal sería reducir  $\alpha$  y  $\beta$  a la vez, pero si disminuimos uno el otro aumentará. La única forma de disminuir los dos a la vez sería aumentando el tamaño muestral.

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

---

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70

{  $H_1 : \theta \neq \theta_0$  } {  $H_1 : \theta > \theta_0$  }

Nosotros supondremos  $H_0 : \theta = \theta_0$  y, según sea  $H_1$ , consideraremos contrastes Unilaterales y Bilaterales.

2. Determinación de  $d$ . Definir una medida de discrepancia,  $d(\theta_0, \hat{\theta})$ , entre el valor  $\theta_0$  que propone  $H_0$  para el parámetro y el valor estimado del parámetro a partir de la muestra,  $\hat{\theta}$ . Esta variable aleatoria debe tener una distribución conocida si  $H_0$  es cierta.

En general y mientras no se diga lo contrario, las medidas de discrepancia que usaremos serán las variables pivote del tema anterior en las que sustituiremos el parámetro genérico de interés  $\theta$  ( $\mu, \sigma^2, \mu_1 - \mu_2$ , etc.) por el valor concreto  $\theta_0$  propuesto por  $H_0$ . Así, si  $H_0$  es cierta, la medida  $d$  tendrá una distribución conocida. Sólo faltará determinar los valores de  $d$  que nos harán rechazar  $H_0$ .

3. Determinación de la región de rechazo. La *región de rechazo* o *región crítica* se define como el conjunto de valores de la medida de discrepancia  $d$  que nos hacen rechazar la hipótesis nula  $H_0$ .
4. Cálculo de  $d$  y toma de la decisión. Con todo lo anterior establecido, se toma la muestra  $(x_1, \dots, x_n)$ , se calcula  $\hat{\theta}$  y  $\hat{d} = d(\theta_0, \hat{\theta})$ .

## 5. Región de Rechazo

### 5.1. Nivel de Significación

Fijado el nivel de significación  $\alpha$ , la región de rechazo se determina a partir de la distribución de  $d(\theta_0, \hat{\theta})$ , cuando  $H_0$  sea cierta.

1.  $\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta > \theta_0 \end{cases}$  Unilateral por la derecha

Elegimos  $d_c$  tal que  $P(d > d_c | H_0 \text{ cierta}) = \alpha$ . La región de rechazo será  $d > d_c$  y la de no rechazo  $d \leq d_c$ .

2.  $\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{cases}$  Unilateral por la izquierda

Elegimos  $d_c$  tal que  $P(d < d_c | H_0 \text{ cierta}) = \alpha$ . La región de rechazo será  $d < d_c$  y la de no rechazo  $d \geq d_c$ .



**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

---

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70**

razo es  $|\alpha| > \alpha_c$ .

## 5.2. Nivel Crítico o $p$ -valor

Otra forma de determinar la Región de Rechazo es mediante el nivel crítico o  $p$ -valor. Éste se define como *la probabilidad de observar una discrepancia peor o igual que la observada cuando  $H_0$  es cierta*, entendiéndose por peor que rechace la hipótesis nula con más evidencia. El  $p$ -valor se calcula a partir del valor obtenido  $\hat{d}$  utilizando la distribución de la medida de discrepancia  $d$ .

$$1. \begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta > \theta_0 \end{cases}$$

El nivel crítico es  $p = P(d \geq \hat{d} | H_0 \text{ cierta})$ .

$$2. \begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{cases}$$

El nivel crítico es  $p = P(d \leq \hat{d} | H_0 \text{ cierta})$ .

$$3. \begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$$

Si  $d$  tiene una distribución simétrica, el nivel crítico es  $p = P(|d| \geq |\hat{d}| | H_0 \text{ cierta})$ .

Cuanto menor sea  $p$ , menor será la probabilidad de obtener una discrepancia como la observada, y menor la credibilidad de  $H_0$ .

- Si  $p > 0.2$  diremos que no existe evidencia muestral para rechazar  $H_0$ .
- Si  $0.01 \leq p \leq 0.2$  estará en la región de duda. Rechazaremos  $H_0$  dependiendo de nuestra opinión a priori y de las consecuencias prácticas de aceptar y rechazar  $H_0$ .
- Si  $p < 0.01$ , rechazamos  $H_0$ .

## 6. Contraste para la media en poblaciones normales

Queremos contrastar que la media de una v.a. normal con parámetros desconocidos es  $\mu_0$ .

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

---

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70

Cartagena99

Para el contraste unilateral por la derecha, con  $H_1 : \mu > \mu_0$ , la región de aceptación de  $H_0$  es:  $(-\infty, t_{n-1, \alpha})$  y el p-valor es  $p = P(d \geq \hat{d} | H_0)$ .

Si  $\sigma$  es conocida utilizamos

$$d = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

y no rechazamos  $H_0$  si  $\hat{d} \in (-z_{\frac{\alpha}{2}}, z_{\frac{\alpha}{2}})$ .

Para poblaciones generales, no normales, y con tamaño muestral grande ( $n > 30$ ), usamos que  $\bar{X}$  es asintóticamente normal. No rechazamos  $H_0$ , en contrastes bilaterales, si  $\hat{d} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \in (-z_{\frac{\alpha}{2}}, z_{\frac{\alpha}{2}})$

**Ejemplo:**

- La longitud de las alas de los grillos de una región puede aproximarse a una distribución normal. Cuando la media se mantiene en 1 cm. apenas se percibe su sonido por las noches. Las alas de 15 grillos elegidos al azar midieron (en cm.):

0.7 1.4 1.3 1 1.8 1.2 0.8 1.7  
1.6 0.9 1.4 1.5 0.7 1.8 0.9

Contrastar la hipótesis de que en esa región pueda dormirse plácidamente, sin ruido, con un nivel de significación del 5 %.

## 7. Contraste para la varianza en poblaciones normales

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$$

La medida de discrepancia es

$$d = \frac{n-1}{\sigma_0^2} s^2 \sim \chi_{n-1}^2 \quad \text{si } H_0 \text{ es cierta}$$

Para  $\alpha$  fijo, la región de aceptación es:  $(\chi_{n-1, 1-\frac{\alpha}{2}}^2, \chi_{n-1, \frac{\alpha}{2}}^2)$

Para el contraste unilateral por la derecha, con  $H_1 : \sigma^2 > \sigma_0^2$ , la región de aceptación de  $H_0$

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

---

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70**



## 8. Contraste para la comparación de medias en poblaciones normales independientes

Tenemos dos poblaciones normales independientes y queremos contrastar si tienen la misma media.

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

### 8.1. Varianzas conocidas

La medida de discrepancia es:

$$d = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1) \quad \text{si } H_0 \text{ es cierta}$$

y para  $\alpha$  fijo, la región de aceptación es  $(-z_{\frac{\alpha}{2}}, z_{\frac{\alpha}{2}})$  y el nivel crítico o p-valor es  $p = P(|d| \geq |\hat{d}| | H_0)$ .

### 8.2. Varianzas desconocidas pero iguales

Consideremos como medida de discrepancia:

$$d = \frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2} \quad \text{si } H_0 \text{ es cierta}$$

Fijado  $\alpha$ , la región de aceptación es  $(-t_{n_1+n_2-2, \frac{\alpha}{2}}, t_{n_1+n_2-2, \frac{\alpha}{2}})$ .

El nivel crítico o p-valor es

$$p = P(|d| > |\hat{d}| | H_0) = P\left(|t_{n_1+n_2-2}| \geq \left| \frac{\bar{x} - \bar{y}}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right|\right)$$

### 8.3. Varianzas desconocidas y distintas

En este caso utilizamos como medida de discrepancia:

$$d = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{n_1+n_2-2-\Delta} \quad \text{si } H_0 \text{ es cierta}$$

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

---

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70**

$$\begin{cases} H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

Cartagena99

Tomamos como medida de discrepancia:

$$d = \frac{S_2^2}{S_1^2} \sim F_{n_2-1, n_1-1} \quad \text{si } H_0 \text{ es cierta}$$

Para  $\alpha$  fijo, la región de aceptación es

$$(F_{n_2-1, n_1-1, 1-\frac{\alpha}{2}}, F_{n_2-1, n_1-1, \frac{\alpha}{2}})$$

**Ejemplo:**

- Una fábrica tiene dos máquinas que producen arandelas de acero de un grosor de 1.3 mm. según los planos. Se han medido los espesores de muestras independientes de cada máquina con precisión de décimas de mm. y se han obtenido los siguientes resultados:

Máquina 1:	1.17	1.30	1.22	1.39	1.32	
Máquina 2:	1.20	1.27	1.33	1.15	1.37	1.26

Supuesta normalidad, ¿podemos decir que ambas máquinas están trabajando en condiciones semejantes? Utilizar un nivel de significación  $\alpha = 0.05$ .

## 10. Contrastes de Hipótesis para proporciones

### 10.1. Para una proporción

Sea  $(X_1, \dots, X_n)$  una m.a.s. de una distribución de  $Bern(p)$ , queremos realizar el siguiente contraste:

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

Si  $n$  es lo suficientemente grande, aplicamos el teorema central del límite y la medida de discrepancia a considerar es:

$$d = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0}} \sim N(0, 1) \quad \text{si } H_0 \text{ es cierta}$$



**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE**  
**LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

---

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS**  
**CALL OR WHATSAPP:689 45 44 70**

¿CÓMO CAMBIO EN EL NÚMERO DE VOTOS?

## 10.2. Comparación de dos proporciones

Ahora tenemos dos muestras independientes  $(X_1, \dots, X_{n_1})$  m.a.s. de  $Bern(p_1)$ ,  $(Y_1, \dots, Y_{n_2})$  m.a.s. de  $Bern(p_2)$  y queremos contrastar:

$$\begin{cases} H_0 : p_1 = p_2 = p_0 \\ H_1 : p_1 \neq p_2 \end{cases}$$

Si  $n_1$  y  $n_2$  son grandes, utilizamos como medida de discrepancia:

$$d = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_0 \hat{q}_0}{n_1} + \frac{\hat{p}_0 \hat{q}_0}{n_2}}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\hat{p}_0 \hat{q}_0}{n_1} + \frac{\hat{p}_0 \hat{q}_0}{n_2}}} \sim N(0, 1) \quad \text{si } H_0 \text{ es cierta}$$

Donde  $\hat{p}_0 = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} = \frac{n_1 \bar{X} + n_2 \bar{Y}}{n_1 + n_2}$

### Ejemplo:

6. La proporción de unidades defectuosas en un lote de  $n_1 = 100$  unidades del proveedor  $A$  es 0.04 mientras que en un lote de  $n_2 = 150$  unidades del  $B$  la proporción es 0.07. ¿Hay evidencia suficiente para admitir diferencias entre los proveedores?

## 11. Relación entre Intervalo de Confianza y Contraste de Hipótesis

En un contraste bilateral:

Se acepta al nivel  $\alpha$  la hipótesis  $H_0 : \theta = \theta_0 \Leftrightarrow$  El IC- $(1 - \alpha)100\%$  construido para  $\theta$  incluye el valor  $\theta_0$ .

Por ejemplo, para  $\mu$  en una  $N(\mu, \sigma)$  y fijado  $\alpha$ :

IC	CH
$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$	$d = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$
$\mu \in \left( \bar{X} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right)$	$\bar{X} \in \left( \mu_0 \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right)$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE  
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

---

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS  
CALL OR WHATSAPP:689 45 44 70

Cartagena99