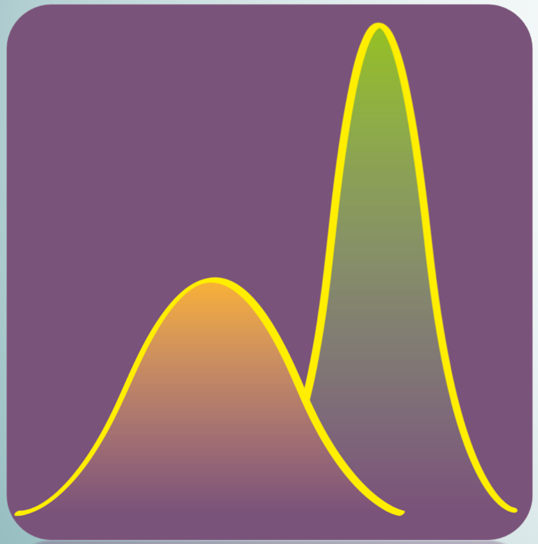
Estadística II

Tema 6. Contrastación de hipótesis paramétricas



Marta Guijarro Garvi

David Gutiérrez Sobrao

Departamento de Economía

Tema 6: contrastación de hipótesis paramétricas

Índice de contenidos

Tema 6: contrastación de hipótesis paramétricas

Resultados de aprendizaje

- Saber formular las hipótesis nula y alternativa.
- Saber aplicar las distintas etapas del proceso de realización de un contraste.
- Saber resolver contrastes para la media y la varianza en el muestreo de una población normal.
- Saber resolver contrastes para la proporción en el muestreo de una población de Bernoulli.
- Saber resolver contrastes de igualdad de medias y varianzas en el muestreo de dos poblaciones normales independientes.
- Saber resolver contrastes de igualdad de proporciones en el muestreo de poblaciones de Bernoulli independientes.

Principales conceptos

Sea X una población con distribución de probabilidades dada por $f(x;\theta)$. Una **hipótesis estadística** es una afirmación sobre el parámetro θ .

Una hipótesis **simple** asigna un valor al parámetro y una hipótesis **compuesta** un rango de valores.

La hipótesis **nula**, H_0, es la hipótesis que se contrasta y es cierta mientras los datos no evidencien lo contrario. Suponiendo cierta la hipótesis nula (*bajo* H_0), la distribución de probabilidades de X está completamente determinada. La hipótesis **alternativa**, H_1, asigna al parámetro valores diferentes a los de la hipótesis nula.

Principales conceptos

El procedimiento estadístico se aplica en los tipos de contrastes que se recogen en la siguiente tabla.

Contraste **bilateral** (contraste de dos colas)	Contraste **unilateral izquierdo** (contraste de una cola)*	Contraste **unilateral derecho** (contraste de una cola)**
$H_0: \theta = \theta_0$ $H_1: \theta \neq \theta_0$	$H_0: \theta = \theta_0$ $H_1: \theta < \theta_0$	$H_0: \theta = \theta_0$ $H_1: \theta > \theta_0$

*Este contraste se plantea también para resolver: $H_0: \theta \geq \theta_0$ frente a $H_1: \theta < \theta_0$.

**Este contraste se plantea también para resolver: $H_0: \theta \leq \theta_0$ frente a $H_1: \theta > \theta_0$.

Procedimiento estadístico de contraste de hipótesis

Un **contraste de hipótesis** es una regla de decisión que permite rechazar o no la hipótesis nula a partir de la realización de una muestra.

El **error de tipo I** consiste en rechazar la hipótesis nula cuando es cierta y el **error de tipo II** en no rechazarla cuando es falsa.

El **nivel de significación** de un contraste, α, es la probabilidad de cometer el error de tipo I, mientras que β es la probabilidad de cometer el error de tipo II; la **potencia** de un contraste es la probabilidad de rechazar la hipótesis nula cuando es falsa y, por tanto, es igual a $1-\alpha$.

Procedimiento estadístico de contraste de hipótesis

Para determinar la regla de decisión, se fija el nivel de significación α , esto es, la probabilidad del error de tipo I que se está en disposición de asumir.

El **valor crítico** del contraste es el valor a partir del cual se rechaza H_0. Se determina gracias a la distribución conocida del estadístico de contraste bajo la hipótesis nula.

La **región crítica** de un contraste está formada por todas las realizaciones de la muestra para las cuales se rechaza la hipótesis nula y la **región de aceptación** por aquéllas que conducen al no rechazo de dicha hipótesis.

El ***p*-valor** de un contraste es el mínimo nivel de significación para el que se rechaza la hipótesis nula a partir de la realización de la muestra.

Procedimiento estadístico de contraste de hipótesis

El procedimiento de contraste de hipótesis tiene las siguientes etapas:

1. Descripción del modelo de distribución de probabilidades de la población.
2. Formulación de las hipótesis nula y alternativa.
3. Establecimiento del nivel de significación.
4. Selección del estadístico de contraste y determinación de su distribución de probabilidades bajo H_0
5. Identificación de la región crítica del contraste.
6. Realización de la muestra y cálculo del valor del estadístico de contraste.
7. Toma de decisión estadística: rechazo o no rechazo de la hipótesis nula.

La resolución de los contrastes bilaterales[1] puede realizarse mediante intervalos de confianza. Así, una vez determinado el nivel de significación α, el contraste podría resolverse mediante el intervalo de confianza calculado para la realización de la muestra con nivel de confianza igual a $1-\alpha$. En tal caso, la regla de decisión consiste en rechazar la hipótesis nula cuando el parámetro postulado por ella no pertenece al intervalo de confianza y en no rechazarla, cuando ocurre lo contrario.

[1]En realidad, de todos los contrastes, aunque este procedimiento es más utilizado en este caso.

Contrastes para los parámetros de una población normal: media (desviación típica conocida)

La tabla presenta los contrastes para μ en el muestreo de una población normal (σ conocida), así como la expresión y la distribución del estadístico de contraste bajo H_0, la obtención del valor crítico y la regla de decisión en cada caso.

Hipótesis del contraste	**Expresión y distribución del estadístico bajo H_0**	**Obtención del valor crítico**	**Regla de decisión**
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \to N(0,1)$	$p\left[Z \leq z_{\alpha/2}\right] = 1-\dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\left\lvert \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \right\rvert > z_{\alpha/2}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$		$p\left[Z \geq z_{\alpha}\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > z_{\alpha}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$		$p\left[Z \leq -z_{\alpha}\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < -z_{\alpha}$

Contrastes para los parámetros de una población normal: media

La tabla presenta los contrastes para μ en el muestreo de una población normal, así como la expresión y la distribución del estadístico de contraste bajo H_0, la obtención del valor crítico y la regla de decisión en cada caso.

Hipótesis del contraste	Expresión y distribución del estadístico bajo H_0	Obtención del valor crítico	Regla de decisión
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$\dfrac{\bar{X}-\mu_0}{S/\sqrt{n-1}} \to t_{n-1}$	$p\left[t_{n-1} \leq t_{\alpha/2}\right] = 1 - \dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\left\lvert\dfrac{\bar{x}-\mu_0}{s/\sqrt{n-1}}\right\rvert > t_{\alpha/2}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$		$p\left[t_{n-1} \geq t_{\alpha}\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{\bar{x}-\mu_0}{s/\sqrt{n-1}} > t_{\alpha}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$		$p\left[t_{n-1} \leq -t_{\alpha}\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{\bar{x}-\mu_0}{s/\sqrt{n-1}} < -t_{\alpha}$

Contrastes para los parámetros de una población normal: varianza

La tabla presenta los contrastes para σ^2 en el muestreo de una población normal, así como la expresión y la distribución del estadístico de contraste bajo H_0, la obtención del valor crítico y la regla de decisión en cada caso.

Hipótesis del contraste	**Expresión y distribución del estadístico bajo H_0**	**Obtención del valor crítico**	**Regla de decisión**
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$\dfrac{n \cdot S^2}{\sigma_0^2} \to \chi^2_{n-1}$	$p\left[\chi^2_{n-1} \leq a\right] = p\left[\chi^2_{n-1} \geq b\right] = \dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{n \cdot s^2}{\sigma_0^2} < a$ o bien $\dfrac{n \cdot s^2}{\sigma_0^2} > b$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$		$p\left[\chi^2_{n-1} \leq c\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{n \cdot s^2}{\sigma_0^2} < c$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$		$p\left[\chi^2_{n-1} \geq d\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{n \cdot s^2}{\sigma_0^2} > d$

Contrastes para los parámetros de una población normal

Ejemplo 6.1. Sea X una población con distribución $N(\mu,1)$. Sea $n = 50, \bar{x} = 10,5$ y α=0,05 , entonces, la regla de decisión para el contraste $H_0{:}\mu=10$ frente a $H_1{:}\mu>10$ es:

Fijado un nivel de significación α=0,05, se rechaza H_0 si, para la realización de la muestra, se cumple:

$$\frac{\bar{x}-10}{1/\sqrt{50}} > 1,645$$

donde $p[Z \leq 1,645] = 0,95$. Como

$$\frac{10,5-10}{1/\sqrt{50}} = 3,53 > 1,645$$

se rechaza la hipótesis nula.

Ejemplo 6.2. Sea X una población con distribución $N(\mu,\sigma)$. Sea $n = 80$, $s = 1,5$ y α=0,05, entonces, la regla de decisión para el contraste $H_0{:}\sigma^2=5$ frente a $H_1{:}\sigma^2<5$ es:

Fijado un nivel de significación α=0,05 , se rechaza H_0 si, para la realización de la muestra, se cumple:

$$\frac{80 \cdot s^2}{5} < 59,52$$

donde $p[\chi_{80-1} \leq 59,52] = 0,05$. Como

$$\frac{80 \cdot 1,5^2}{5} = 36 < 59,52$$

se rechaza la hipótesis nula.

Contrastes para el parámetro de una población de Bernoulli: proporción

La tabla presenta los contrastes para p en el muestreo de una población de Bernoulli, así como la expresión y la distribución del estadístico de contraste bajo H_0, la obtención del valor crítico y la regla de decisión en cada caso.

Hipótesis del contraste	Expresión y distribución del estadístico bajo H_0	Obtención del valor crítico	Regla de decisión
$H_0: p = p_0$ $H_1: p \neq p_0$	$\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 \cdot (1 - p_0)}{n}}} \xrightarrow{\text{aprox.}} N(0,1)$	$p\left[Z \leq z_{\alpha/2}\right] = 1 - \dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\left\lvert \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 \cdot (1 - p_0)}{n}}} \right\rvert > z_{\alpha/2}$
$H_0: p = p_0$ $H_1: p > p_0$		$p\left[Z \geq z_{\alpha}\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 \cdot (1 - p_0)}{n}}} > z_{\alpha}$
$H_0: p = p_0$ $H_1: p < p_0$		$p\left[Z \leq -z_{\alpha}\right] = \alpha$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 \cdot (1 - p_0)}{n}}} < -z_{\alpha}$

Contrastes para el parámetro de una población de Bernoulli: proporción

Ejemplo 6.3. Sea X una población con distribución $b(p)$. Si $n = 200$, $\hat{p} = 0,75$ y α=0,05 , entonces, la regla de decisión para el contraste H_0:p=0,8 frente a H_1:$p \neq 0,8$ es:

Fijado un nivel de significación α=0,05, se rechaza H_0 si, para la realización de la muestra, se cumple:

$$\left| \frac{\hat{p} - 0,8}{\sqrt{\frac{0,8 \cdot (1 - 0,8)}{100}}} \right| > 1,96$$

donde $p[Z \leq 1,96] = 0,975.$ Como

$$\left| \frac{0,75 - 0,8}{\sqrt{\frac{0,8 \cdot (1 - 0,8)}{100}}} \right| = 1,25 < 1,96$$

no se rechaza la hipótesis nula

Contrastes para dos poblaciones normales independientes: igualdad de medias

La tabla presenta el contraste bilateral de igualdad de medias en el muestreo de dos poblaciones normales independientes, así como la expresión y la distribución del estadístico de contraste bajo H_0, la obtención del valor crítico y la regla de decisión en cada caso.

Hipótesis del contraste	**Expresión y distribución del estadístico bajo H_0**	**Obtención del valor crítico**	**Regla de decisión**
Desviaciones típicas conocidas			
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	$\dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)}} \rightarrow N(0,1)$	$p\left[Z \leq z_{\alpha/2}\right] = 1-\dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\left\lvert \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)}} \right\rvert > z_{\alpha/2}$
Desviaciones típicas desconocidas e iguales			
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	$\dfrac{\sqrt{\dfrac{n_1 \cdot n_2}{n_1+n_2}}\left(\bar{X}_1 - \bar{X}_2\right)}{\sqrt{\dfrac{n_1 \cdot S_1^2 + n_2 \cdot S_2^2}{n_1+n_2-2}}} \rightarrow t_{n_1+n_2-2}$	$p\left[t_{n_1+n_2-2} \leq t_{\alpha/2}\right] = 1-\dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\left\lvert \dfrac{\sqrt{\dfrac{n_1 \cdot n_2}{n_1+n_2}}\left(\bar{x}_1 - \bar{x}_2\right)}{\sqrt{\dfrac{n_1 \cdot s_1^2 + n_2 \cdot s_2^2}{n_1+n_2-2}}} \right\rvert > t_{\alpha/2}$

Contrastes para dos poblaciones normales independientes: igualdad de varianzas

La tabla presenta el contraste bilateral de igualdad de varianzas en el muestreo de dos poblaciones normales independientes, así como la expresión y la distribución del estadístico de contraste bajo H_0, la obtención del valor crítico y la regla de decisión.

Hipótesis del contraste	**Expresión y distribución del estadístico bajo H_0**	**Obtención del valor crítico**	**Regla de decisión**
$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 \neq \sigma_2^2$	$\dfrac{S_{c_1}^2}{S_{c_2}^2} \rightarrow F_{n_1-1,n_2-1}$	$p\left[F_{n_1-1,n_2-1} \leq a\right] = p\left[F_{n_1-1,n_2-1} \geq b\right] = \dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\dfrac{s_{c_1}^2}{s_{c_2}^2} < a$ o bien $\dfrac{s_{c_1}^2}{s_{c_2}^2} > b$

Contrastes para dos poblaciones normales independientes: igualdad de varianzas

Ejemplo 6.4. Sean X_1 y X_2, poblaciones independientes con distribuciones $N(\mu_1,\sigma_1)$ y $N(\mu_2,\sigma_2)$. Si $n_1 = 25$ y $n_2 = 31$, $\bar{x}_1 = 45$, $\bar{x}_2 = 42$, s_1^2=4,8 y s_2^2=7,7, entonces, la regla de decisión para el contraste $H_0 : \sigma_1^2 = \sigma_2^2$ frente a $H_1 : \sigma_1^2 \neq \sigma_2^2$ es:

Fijado un nivel de significación α=0,1 , se rechaza H_0 si, para la realización de la muestra, se cumple:

$$\frac{s_{c_1}^2}{s_{c_2}^2} < 0,51 \quad \text{o bien} \quad \frac{s_{c_1}^2}{s_{c_2}^2} > 1,89,$$

donde $p\left[F_{25-1,31-1} \leq 0,51\right] = p\left[F_{25-1,31-1} \geq 1,89\right] = 0,05.$ Como

$$\frac{25 \cdot 4,8/24}{31 \cdot 7,7/30} = 0,63, \text{ entonces, } 0,51<0,63<1,89,$$

no se rechaza la hipótesis nula y puede suponerse que las varianzas son iguales.

Contrastes para dos poblaciones normales independientes: igualdad de varianzas

Ejemplo 6.5 (continuación del Ejemplo 6.4) Puesto que puede suponerse que las varianzas son iguales, la regla de decisión para el contraste $H_0 : \mu_1 = \mu_2$ frente a $H_1 : \mu_1 \neq \mu_2$ es:

Fijado un nivel de significación α=0,05, se rechaza H_0 si, para la realización de la muestra, se cumple:

$$\left| \frac{\sqrt{\frac{25 \cdot 31}{25+31}} \left(\bar{x}_1 - \bar{x}_2\right)}{\sqrt{\frac{25 \cdot s_1^2 + 31 \cdot s_2^2}{25+31-2}}} \right| > 1,96,$$

donde $p\left[t_{25+31-2} \leq 1,96\right] = 0,975.$ Como

$$\left| \frac{\sqrt{\frac{25 \cdot 31}{25+31}} \left(45-42\right)}{\sqrt{\frac{25 \cdot 4,8 + 31 \cdot 7,7}{25+31-2}}} \right| = 4,33 > 1,96,$$

se rechaza la hipótesis nula de igualdad de medias.

Contrastes para los parámetros de dos distribuciones de Bernoulli independientes: igualdad de proporciones

La tabla presenta el contraste bilateral de igualdad de proporciones en el muestreo de dos poblaciones de Bernoulli independientes, así como la expresión y la distribución del estadístico de contraste bajo H_0, la obtención del valor crítico y la regla de decisión.

Hipótesis del contraste	**Expresión y distribución del estadístico bajo H_0**	**Obtención del valor crítico**	**Regla de decisión**
$H_0 : p_1 = p_2$ $H_1 : p_1 \neq p_2$	$\dfrac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\dfrac{\widehat{p}_0 \cdot (1-\widehat{p}_0)}{n_1} + \dfrac{\widehat{p}_0 \cdot (1-\widehat{p}_0)}{n_2}}} \xrightarrow{\text{aprox.}} N(0,1)$ Donde $\widehat{p}_0 = \dfrac{n_1 \cdot \widehat{p}_1 + n_2 \cdot \widehat{p}_2}{n_1 + n_2}$	$p\left[Z < z_{\alpha/2}\right] = 1 - \dfrac{\alpha}{2}$	Fijado un nivel de significación α, se rechaza H_0 si, para la realización de la muestra, se cumple que $\left\| \dfrac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\dfrac{\widehat{p}_0 \cdot (1-\widehat{p}_0)}{n_1} + \dfrac{\widehat{p}_0 \cdot (1-\widehat{p}_0)}{n_2}}} \right\| > z_{\alpha/2}$

Contrastes para los parámetros de dos distribuciones de Bernoulli independientes: igualdad de proporciones

Ejemplo 6.5. Sean X_1 y X_2, poblaciones independientes con distribuciones $b(p_1)$ y $b(p_2)$.
Si $n_1 = 350$, $n_2 = 300$, $\hat{p}_1 = 0{,}6$, $\hat{p}_2 = 0{,}5$ y α=0,05, entonces, la regla de decisión para el contraste $H_0 : p_1 = p_2$ frente a $H_1 : p_1 \neq p_2$ es:

Fijado un nivel de significación α=0,05, se rechaza H_0 si

$$\left| \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_0 \cdot (1-\hat{p}_0)}{350} + \dfrac{\hat{p}_0 \cdot (1-\hat{p}_0)}{300}}} \right| > 1{,}96$$

con

$$\hat{p}_0 = \frac{350 \cdot 0{,}6 + 300 \cdot 0{,}5}{350 + 300} = 0{,}55$$

y donde $p[Z \leq 1{,}96] = 0{,}975$. Como

$$\left| \frac{0{,}6 - 0{,}5}{\sqrt{\dfrac{0{,}55 \cdot (1-0{,}55)}{350} + \dfrac{0{,}55 \cdot (1-0{,}55)}{300}}} \right| = 2{,}55 > 1{,}96$$

se rechaza la hipótesis nula de igualdad de proporciones.