Estadística: estadística
Grado en Relaciones Laborales | Curso 2019-2020
Tema 7.B. Introducción a la inferencia estadística

Alejandro Saavedra Nieves

Contraste de hipótesis

- Los procedimientos de inferencia que hemos realizado hasta ahora son:
  - La estimación puntual
  - Los intervalos de confianza

- En esta parte de tema vamos a ver otro procedimiento de inferencia basado en contrastes de hipótesis en el que el objetivo de la experimentación está orientado a corroborar una hipótesis inicial sobre la población de estudio.

Contraste de hipótesis

- Cuando un investigador trata de entender o explicar algo, generalmente formula su problema de investigación por medio de una hipótesis
- Ejemplo: Me planteo si la edad media que tienen las mujeres gallegas cuando deciden tener su primer hijo es igual que en el resto de España (29.3 años)

Contraste de hipótesis

- Cuando un investigador trata de entender o explicar algo, generalmente formula su problema de investigación por medio de una hipótesis
- Ejemplo: Me planteo si la edad media que tienen las mujeres gallegas cuando deciden tener su primer hijo es igual que en el resto de España (29.3 años)

Hipótesis nula

$$H_0 \; : \; \mu \; = \; 29.3$$

- Cuando un investigador trata de entender o explicar algo, generalmente formula su problema de investigación por medio de una hipótesis
- Ejemplo: Me planteo si la edad media que tienen las mujeres gallegas cuando deciden tener su primer hijo es igual que en el resto de España (29.3 años)

Hipótesis nula

$$H_0 \ : \ \mu \ = \ 29.3$$

- Tomo una muestra de 6 mujeres gallegas embarazadas primerizas

Contraste de hipótesis

- Cuando un investigador trata de entender o explicar algo, generalmente formula su problema de investigación por medio de una hipótesis
- Ejemplo: Me planteo si la edad media que tienen las mujeres gallegas cuando deciden tener su primer hijo es igual que en el resto de España (29.3 años)

Hipótesis nula

$$H_0 \;:\; \mu \;=\; 29.3$$

- Tomo una muestra de 6 mujeres gallegas embarazadas primerizas

- $\bar{X} = 30.5$ años

Contraste de hipótesis

- Cuando un investigador trata de entender o explicar algo, generalmente formula su problema de investigación por medio de una hipótesis
- Ejemplo: Me planteo si la edad media que tienen las mujeres gallegas cuando deciden tener su primer hijo es igual que en el resto de España (29.3 años)

Hipótesis nula

$$H_0 \;:\; \mu \;=\; 29.3$$

- Tomo una muestra de 6 mujeres gallegas embarazadas primerizas

- $\bar{X} = 30.5$ años
- ¿Existe suficiente evidencia en los datos para rechazar H_0?
- ¿O la diferencia entre $\bar{X}$ y el valor hipotético de μ puede ser debido al azar?

- Cuando un investigador trata de entender o explicar algo, generalmente formula su problema de investigación por medio de una hipótesis
- Ejemplo:No sé si la edad media que tienen las mujeres gallegas cuando deciden tener su primer hijo es igual que en el resto de España (29.3 años)

Hipótesis nula

$$H_0 \ : \ \mu \ = \ 29.3$$

- Tomo una muestra de 36 mujeres gallegas embarazadas primerizas

- $\bar{X} = 30.5$ años
- ¿Existe suficiente evidencia en los datos para rechazar H_0?
- ¿O la diferencia entre $\bar{X}$ y el valor hipotético de μ puede ser debido al azar?

Contraste de hipótesis

- Llamaremos **hipótesis nula**, y la denotamos por H_0, a la que se da por cierta antes de obtener la muestra. Goza de presunción de inocencia.
- Llamaremos **hipótesis alternativa**, y la denotamos por H_1 (o H_a) a lo que sucede cuando no es cierta la hipótesis nula.

Contraste de hipótesis

- Llamaremos **hipótesis nula**, y la denotamos por H_0, a la que se da por cierta antes de obtener la muestra. Goza de presunción de inocencia.
- Llamaremos **hipótesis alternativa**, y la denotamos por H_1 (o H_a) a lo que sucede cuando no es cierta la hipótesis nula.
- Por gozar la hipótesis nula de presunción de inocencia, sobre la hipótesis alternativa recae la carga de la prueba. Por tanto, cuando rechazamos H_0 en favor de H_1 es porque hemos encontrado pruebas significativas a partir de la muestra.

Representamos este problema de decisión mediante el siguiente gráfico:

		Decisión	
		No se rechaza H_0	Se rechaza H_0
Realidad	H_0 es verdadera	Decisión correcta	Error tipo I
	H_0 es falsa	Error tipo II	Decisión correcta

Representamos este problema de decisión mediante el siguiente gráfico:

		Decisión	
		No se rechaza H_0	Se rechaza H_0
Realidad	H_0 es verdadera	Decisión correcta	Error tipo I
	H_0 es falsa	Error tipo II	Decisión correcta

Observamos que se puede tomar una decisión correcta o errónea.

- Error de tipo I: cuando rechazamos la hipótesis nula, siendo cierta.
- Error de tipo II: cuando aceptamos la hipótesis nula, siendo falsa.

Contraste de hipótesis. Analogía con un juicio

Supongamos un juicio en el que se trata de decidir la culpabilidad o inocencia de un acusado.

Supongamos un juicio en el que se trata de decidir la culpabilidad o inocencia de un acusado.

- Hipótesis nula: el acusado es inocente (todo acusado es inocente hasta que se demuestre lo contrario).

Contraste de hipótesis. Analogía con un juicio

Supongamos un juicio en el que se trata de decidir la culpabilidad o inocencia de un acusado.

- Hipótesis nula: el acusado es inocente (todo acusado es inocente hasta que se demuestre lo contrario).
- Hipótesis alternativa: el acusado es culpable.

Contraste de hipótesis. Analogía con un juicio

Supongamos un juicio en el que se trata de decidir la culpabilidad o inocencia de un acusado.

- Hipótesis nula: el acusado es inocente (todo acusado es inocente hasta que se demuestre lo contrario).
- Hipótesis alternativa: el acusado es culpable.
- Juicio: es el procedimiento en el cual se trata de probar la culpabilidad del acusado y la evidencia debe ser muy fuerte para que se rechace la inocencia (H_0) en favor de la culpabilidad (H_a).

Contraste de hipótesis. Analogía con un juicio

Supongamos un juicio en el que se trata de decidir la culpabilidad o inocencia de un acusado.

- Hipótesis nula: el acusado es inocente (todo acusado es inocente hasta que se demuestre lo contrario).
- Hipótesis alternativa: el acusado es culpable.
- Juicio: es el procedimiento en el cual se trata de probar la culpabilidad del acusado y la evidencia debe ser muy fuerte para que se rechace la inocencia (H_0) en favor de la culpabilidad (H_a).
- Decisión: el veredicto.

Contraste de hipótesis. Analogía con un juicio

Supongamos un juicio en el que se trata de decidir la culpabilidad o inocencia de un acusado.

- Hipótesis nula: el acusado es inocente (todo acusado es inocente hasta que se demuestre lo contrario).
- Hipótesis alternativa: el acusado es culpable.
- Juicio: es el procedimiento en el cual se trata de probar la culpabilidad del acusado y la evidencia debe ser muy fuerte para que se rechace la inocencia (H_0) en favor de la culpabilidad (H_a).
- Decisión: el veredicto.
- Error de tipo I: condenar a un inocente.

Supongamos un juicio en el que se trata de decidir la culpabilidad o inocencia de un acusado.

- Hipótesis nula: el acusado es inocente (todo acusado es inocente hasta que se demuestre lo contrario).
- Hipótesis alternativa: el acusado es culpable.
- Juicio: es el procedimiento en el cual se trata de probar la culpabilidad del acusado y la evidencia debe ser muy fuerte para que se rechace la inocencia (H_0) en favor de la culpabilidad (H_a).
- Decisión: el veredicto.
- Error de tipo I: condenar a un inocente.
- Error de tipo II: absolver a un culpable.

- La probabilidad del error de tipo I se denota por α y se denomina **nivel de significación**.

Nivel de significación

$$\alpha \quad = \quad P(\text{Rechazar } H_0/H_0 \text{ es cierta})$$

- La probabilidad del error de tipo II se denota por β

$$\beta = P(\text{No rechazar } H_0/H_0 \text{ es falsa})$$

- **Potencia:** Es la probabilidad de detectar que una hipótesis es falsa.

Potencia

$$\text{Potencia} \quad = \quad P(\text{Rechazar } H_0/H_0 \text{ es falsa}) \quad = \quad 1 - \beta$$

Región crítica. Contrastes bilaterales y unilaterales

- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0

- Ejemplo: ¿Difiere la edad media de las madres primerizas en Galicia de la edad media de las madres primerizas en el resto de España (29.3 años)?

- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0
- Ejemplo: ¿Difiere la edad media de las madres primerizas en Galicia de la edad media de las madres primerizas en el resto de España (29.3 años)?

Contraste bilateral

$$H_0 : \mu = 29.3$$
$$H_1 : \mu \neq 29.3$$

Región crítica. Contrastes bilaterales y unilaterales

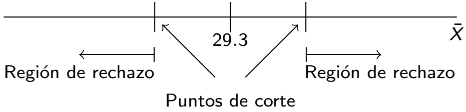
- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0

- Ejemplo: ¿Difiere la edad media de las madres primerizas en Galicia de la edad media de las madres primerizas en el resto de España (29.3 años)?

Contraste bilateral

$$H_0 : \mu = 29.3$$
$$H_1 : \mu \neq 29.3$$

- Si estamos interesados en determinar si μ **difiere significativamente** de 29.3, deberíamos rechazar H_0 si $\bar{X}$ está "lejos" de 29.3 **en ambas direcciones**.



Región crítica. Contrastes bilaterales y unilaterales

- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0
- Ejemplo: ¿Es la edad media de las madres primerizas en Galicia mayor que la edad media de las madres primerizas en el resto de España (29.3 años)?

Región crítica. Contrastes bilaterales y unilaterales

- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0
- Ejemplo: ¿Es la edad media de las madres primerizas en Galicia mayor que la edad media de las madres primerizas en el resto de España (29.3 años)?

Contraste unilateral

$$H_0 : \mu \leq 29.3$$
$$H_1 : \mu > 29.3$$

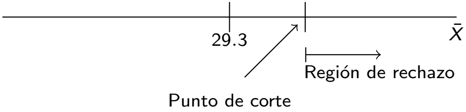
- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0

- Ejemplo: ¿Es la edad media de las madres primerizas en Galicia mayor que la edad media de las madres primerizas en el resto de España (29.3 años)?

Contraste unilateral

$$H_0 : \mu \leq 29.3$$
$$H_1 : \mu > 29.3$$

- Si estamos interesados en determinar si μ **es significativamente mayor** que 29.3, deberíamos rechazar H_0 si $\bar{X}$ está "lejos" de 29.3 **en una sola dirección**.



Región crítica. Contrastes bilaterales y unilaterales

- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0
- Ejemplo: ¿Es la edad media de las madres primerizas en Galicia menor que la edad media de las madres primerizas en el resto de España (29.3 años)?

Región crítica. Contrastes bilaterales y unilaterales

- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0

- Ejemplo: ¿Es la edad media de las madres primerizas en Galicia menor que la edad media de las madres primerizas en el resto de España (29.3 años)?

Contraste unilateral

$$H_0 : \mu \geq 29.3$$
$$H_1 : \mu < 29.3$$

Región crítica. Contrastes bilaterales y unilaterales

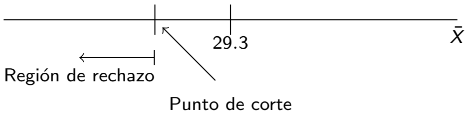
- Debemos establecer una regla de decisión para determinar cuando rechazamos o no la hipótesis nula H_0

- Ejemplo: ¿Es la edad media de las madres primerizas en Galicia menor que la edad media de las madres primerizas en el resto de España (29.3 años)?

Contraste unilateral

$$H_0 : \mu \geq 29.3$$
$$H_1 : \mu < 29.3$$

- Si estamos interesados en determinar si μ **es significativamente menor** que 29.3, deberíamos rechazar H_0 si $\bar{X}$ está "lejos" de 29.3 **en una sola dirección**.



Etapas en la resolución de un contraste de hipótesis

- Especificar las hipótesis nula H_0 y alternativa H_1.
- Elegir un estadístico de contraste apropiado para el parámetro.
- Fijar el nivel de significación α en base a cómo de importante se considere rechazar H_0 cuando realmente es cierta.
- Al fijar un nivel de significación, α, se obtiene implícitamente una división en dos regiones del conjunto de posibles valores del estadístico de contraste:
  - La **región de rechazo** o región crítica que tiene probabilidad α (bajo H_0).
  - La **región de aceptación** que tiene probabilidad $1-\alpha$ (bajo H_0).
- Si el valor del estadístico cae en la región de rechazo, los datos no son compatibles con H_0 y la rechazamos. Entonces se dice que el contraste es **estadísticamente significativo**, es decir existe evidencia estadísticamente significativa a favor de H_1.
- Si el valor del estadístico cae en la región de aceptación, no existen razones suficientes para rechazar la hipótesis nula con un nivel de significación α, y el contraste se dice **estadísticamente no significativo**, es decir no existe evidencia a favor de H_1.

Contraste sobre la media μ de una población normal con σ^2 conocida

Sea $X_1, X_2, \ldots, X_n$ una muestra formada por n variables independientes y con la misma distribución $N(\mu, \sigma)$.

- Supongamos que la varianza σ^2 es conocida
- Se desea contrastar una hipótesis relativa a la media, μ.

Contraste bilateral (hipótesis nula simple)

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

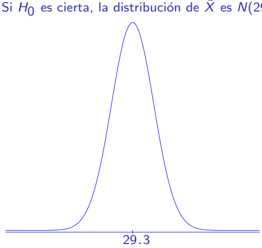
- El sentido común nos aconseja rechazar la hipótesis nula de que la media poblacional es μ_0 cuando la media muestral $\bar{X}$ sea muy distinta de μ_0.

Contraste sobre la media μ de una población normal con σ^2 conocida

- La edad de las madres primerizas en Galicia sigue una distribución normal con $\sigma = 2$ años.
- Tomamos una muestra de $n = 36$ madres primerizas gallegas. Queremos contrastar si la edad media de las madres primerizas en Galicia difiere de la media de madres primerizas en el resto de España (29.3 años).

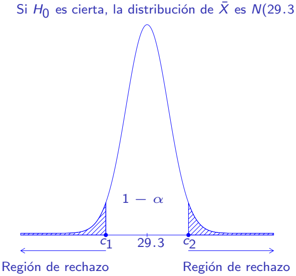
Contraste sobre la media μ de una población normal con σ^2 conocida

- La edad de las madres primerizas en Galicia sigue una distribución normal con $\sigma = 2$ años.
- Tomamos una muestra de $n = 36$ madres primerizas gallegas. Queremos contrastar si la edad media de las madres primerizas en Galicia difiere de la media de madres primerizas en el resto de España (29.3 años).



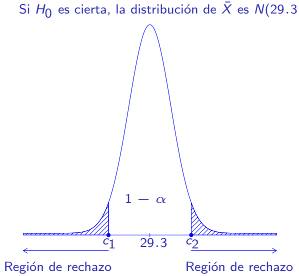
Contraste sobre la media μ de una población normal con σ^2 conocida

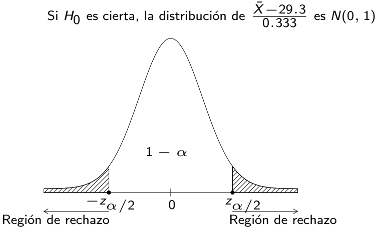
- La edad de las madres primerizas en Galicia sigue una distribución normal con $\sigma = 2$ años.
- Tomamos una muestra de $n = 36$ madres primerizas gallegas. Queremos contrastar si la edad media de las madres primerizas en Galicia difiere de la media de madres primerizas en el resto de España (29.3 años).



Contraste sobre la media μ de una población normal con σ^2 conocida

- La edad de las madres primerizas en Galicia sigue una distribución normal con $\sigma = 2$ años.
- Tomamos una muestra de $n = 36$ madres primerizas gallegas. Queremos contrastar si la edad media de las madres primerizas en Galicia difiere de la media de madres primerizas en el resto de España (29.3 años).

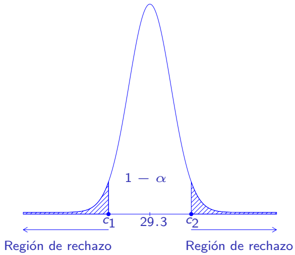




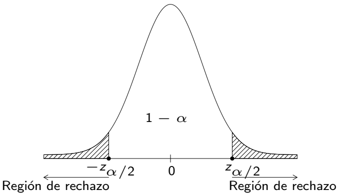
Contraste sobre la media μ de una población normal con σ^2 conocida

- La edad de las madres primerizas en Galicia sigue una distribución normal con $\sigma = 2$ años.
- Tomamos una muestra de $n = 36$ madres primerizas gallegas. Queremos contrastar si la edad media de las madres primerizas en Galicia difiere de la media de madres primerizas en el resto de España (29.3 años).
- Observamos que $\bar{X} = 30.5$ años. En base a la muestra, ¿podrías concluir que la edad media de las madres primerizas en Galicia difiere de la edad media de las madres primerizas en el resto de España?

Si H_0 es cierta, la distribución de $\bar{X}$ es $N(29.3, 2/6)$

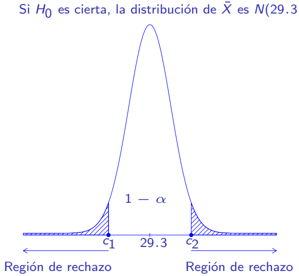


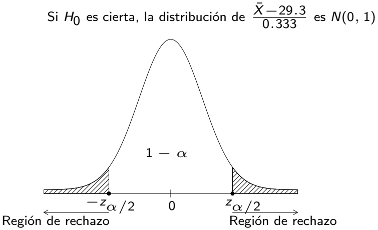
Si H_0 es cierta, la distribución de $\frac{\bar{X}-29.3}{0.333}$ es $N(0, 1)$



Contraste sobre la media μ de una población normal con σ^2 conocida

- La edad de las madres primerizas en Galicia sigue una distribución normal con $\sigma = 2$ años.
- Tomamos una muestra de $n = 36$ madres primerizas gallegas. Queremos contrastar si la edad media de las madres primerizas en Galicia difiere de la media de madres primerizas en el resto de España (29.3 años).
- Observamos que $\bar{X} = 30.5$ años. En base a la muestra, ¿podrías concluir que la edad media de las madres primerizas en Galicia difiere de la edad media de las madres primerizas en el resto de España?



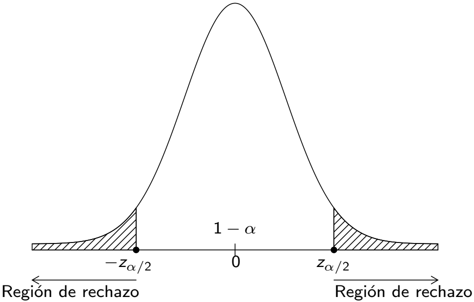


Rechazamos la hipótesis nula $H_0 : \mu = 29.3$ frente a $H_1 : \mu \neq 29.3$ si

$$\frac{30.5-29.3}{0.333} \leq -z_{\alpha/2} \quad \text{ó} \quad \frac{30.5-29.3}{0.333} \geq z_{\alpha/2}$$

Contraste sobre la media μ de una población normal con σ^2 conocida

Si H_0 es cierta, la distribución de $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ es $N(0,1)$

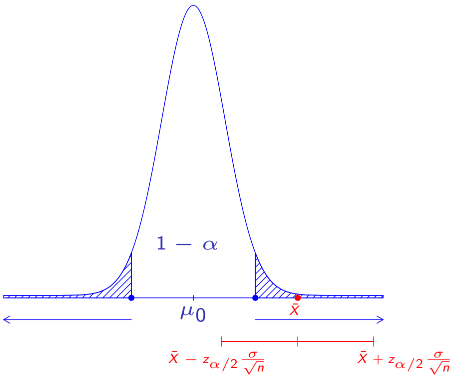


Rechazamos la hipótesis nula $H_0 : \mu = \mu_0$ frente a $H_1 : \mu \neq \mu_0$ si

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq -z_{\alpha/2} \quad \text{ó} \quad \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \geq z_{\alpha/2}$$

Relación entre el contraste bilateral y los Intervalos de confianza

- $H_0 : \mu = \mu_0$
- Si H_0 es cierta, la distribución de $\bar{X}$ es $N\left(\mu_0, \frac{\sigma}{\sqrt{n}}\right)$



- Rechazamos $H_0 : \mu = \mu_0$ con una significación α si μ_0 no pertenece al intervalo de confianza para μ de nivel $1 - \alpha$

Contraste sobre la media μ de una población normal con σ^2 conocida

Sea $X_1, X_2, \ldots, X_n$ una muestra formada por n variables independientes y con la misma distribución $N(\mu, \sigma)$.

- Supongamos que la varianza σ^2 es conocida
- Se desea contrastar una hipótesis relativa a la media, μ.

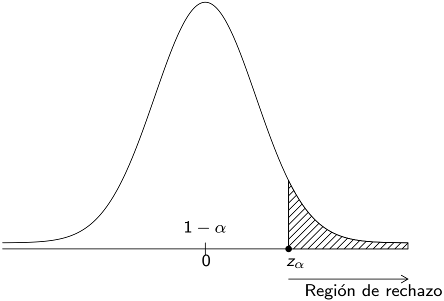
Contraste unilateral (hipótesis nula compuesta)

$$H_0 : \mu \leq \mu_0$$
$$H_1 : \mu > \mu_0$$

- El sentido común nos aconseja rechazar la hipótesis nula de que la media poblacional es μ_0 cuando la media muestral $\bar{X}$ sea "considerablemente mayor" que μ_0.

Contraste sobre la media μ de una población normal con σ^2 conocida

Si H_0 es cierta, la distribución de $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ es $N(0,1)$



Rechazamos la hipótesis nula $H_0 : \mu \leq \mu_0$ frente a $H_1 : \mu > \mu_0$ si

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha$$

Contraste sobre la media μ de una población normal con σ^2 conocida

Sea $X_1, X_2, \ldots, X_n$ una muestra formada por n variables independientes y con la misma distribución $N(\mu, \sigma)$.

- Supongamos que la varianza σ^2 es conocida
- Se desea contrastar una hipótesis relativa a la media, μ.

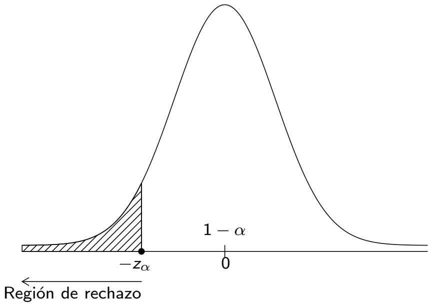
Contraste unilateral
(hipótesis nula compuesta)

$$H_0 : \mu \geq \mu_0$$
$$H_1 : \mu < \mu_0$$

- El sentido común nos aconseja rechazar la hipótesis nula de que la media poblacional es μ_0 cuando la media muestral $\bar{X}$ sea "considerablemente menor" que μ_0.

Contraste sobre la media μ de una población normal con σ^2 conocida

Si H_0 es cierta, la distribución de $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ es $N(0,1)$



Rechazamos la hipótesis nula $H_0 : \mu \geq \mu_0$ frente a $H_1 : \mu < \mu_0$ si

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha$$

El p-valor

- A medida que el nivel de significación α disminuye es más difícil rechazar la hipótesis nula (manteniendo los mismos datos).
- Hay un valor de α a partir del cual ya no podemos rechazar H_0. A dicho valor se le se le llama el *p-valor del contraste* y se denota por p.
- Es decir, si el nivel de significación es menor que p ya no se rechaza H_0.

> - Si $\alpha < p$ no podemos rechazar H_0 a nivel α.
> - Si $\alpha \geq p$ podemos rechazar H_0 a nivel α.

Cálculo del p-valor

Fijado el parámetro bajo estudio, el estadístico de contraste y su distribución D_c, E_c denota el valor del estadístico para la muestra dada. Entonces,

- Caso 1. Contraste bilateral.

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

El p-valor es igual a $P(D_c \geq |E_c|)$.

- Caso 2. Unilateral derecho.

$$\begin{array}{l} H_0 : \mu \leq \mu_0 \\ H_1 : \mu > \mu_0 \end{array}$$

El p-valor es igual a $P(D_c \geq E_c)$.

- Caso 3. Unilateral izquierdo.

$$\begin{array}{l} H_0 : \mu \geq \mu_0 \\ H_1 : \mu < \mu_0 \end{array}$$

El p-valor es igual a $P(D_c \leq E_c)$.

El p-valor: un ejemplo

Ejemplo. Queremos probar que la edad media de los habitantes de Estados Unidos es superior a 70 años, sabiendo que la varianza de las alturas es igual a 8.9^2. Para ello contamos con una muestra de $n = 100$ individuos cuya edad media es 71.8 años. Calcula el p-valor en el contraste:

$$H_0 : \mu = 70$$
$$H_1 : \mu \neq 70$$

Solución. Tenemos que $\bar{x} = 71.8$ e $\sigma = 8.9$, de forma que el estadístico de contraste (E_c) es

$$E_c = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{71.8 - 70}{8.9/\sqrt{100}} = 2.02$$

Entonces, $D_c = \frac{\bar{x}-\mu_0}{S/\sqrt{n}} \sim N(0,1)$ y

p-valor $= P(D_c > |2.02|) = P(D_c > 2.02) + P(D_c < -2.02) = 2P(D_c > 2.02)$.

Resulta que el p-valor del contraste es $2 \times 0.0216 = 0.0434$

Contraste sobre la media μ de una población normal con σ^2 desconocida

Sea $X_1, X_2, \ldots, X_n$ una muestra formada por n variables independientes y con la misma distribución $N(\mu, \sigma)$.

- Supongamos que σ^2 es desconocida
- Se desea contrastar una hipótesis relativa a la media, μ.
- Si H_0 es cierta,

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \in t_{n-1}$$

- Recuerda que:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}$$

Contraste sobre la media de una población normal con varianza desconocida

Rechazamos la hipótesis nula $H_0 : \mu = \mu_0$ frente a $H_1 : \mu \neq \mu_0$ si

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -t_{n-1,\alpha/2} \quad \text{ó} \quad \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_{n-1,\alpha/2}$$

Rechazamos la hipótesis nula $H_0 : \mu \leq \mu_0$ frente a $H_1 : \mu > \mu_0$ si

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_{n-1,\alpha}$$

Rechazamos la hipótesis nula $H_0 : \mu \geq \mu_0$ frente a $H_1 : \mu < \mu_0$ si

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -t_{n-1,\alpha}$$

t con $n-1$ g.l.

Contraste sobre una proporción (muestras grandes)

Rechazamos la hipótesis nula $H_0 : p = p_0$ frente a $H_1 : p \neq p_0$ si

$$\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \leq -z_{\alpha/2} \quad \text{ó} \quad \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq z_{\alpha/2}$$

Rechazamos la hipótesis nula $H_0 : p \leq p_0$ frente a $H_1 : p > p_0$ si

$$\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq z_{\alpha}$$

Rechazamos la hipótesis nula $H_0 : p \geq p_0$ frente a $H_1 : p < p_0$ si

$$\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \leq -z_{\alpha}$$

Contraste sobre la varianza σ de una población normal con μ conocida

Sea $X_1, X_2, \ldots, X_n$ una muestra formada por n variables independientes y con la misma distribución $N(\mu, \sigma)$.

- Supongamos que μ es conocida
- Se desea contrastar una hipótesis relativa a la varianza, σ.
- Si H_0 es cierta,

$$\frac{nS_X^2}{\sigma_0^2} \in \chi_n^2$$

- Recuerda que:

$$S_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Contraste sobre la varianza σ de una población normal con μ conocida

Rechazamos la hipótesis nula $H_0 : \sigma^2 = \sigma_0^2$ frente a $H_1 : \sigma^2 \neq \sigma_0^2$ si

$$\frac{nS_X^2}{\sigma_0^2} \leq \chi^2_{n,1-\alpha/2} \quad \text{ó} \quad \frac{nS_X^2}{\sigma_0^2} \geq \chi^2_{n,\alpha/2}$$

Rechazamos la hipótesis nula $H_0 : \sigma^2 \leq \sigma_0^2$ frente a $H_1 : \sigma^2 > \sigma_0^2$ si

$$\frac{nS_X^2}{\sigma_0^2} \geq \chi^2_{n,\alpha}$$

Rechazamos la hipótesis nula $H_0 : \sigma^2 \geq \sigma_0^2$ frente a $H_1 : \sigma^2 < \sigma_0^2$ si

$$\frac{nS_X^2}{\sigma_0^2} \leq \chi^2_{n,1-\alpha}$$

χ^2 con n g.l.

Contraste sobre la varianza σ de una población normal con μ desconocida

Sea $X_1, X_2, \ldots, X_n$ una muestra formada por n variables independientes y con la misma distribución $N(\mu, \sigma)$.

- Supongamos que μ es desconocida
- Se desea contrastar una hipótesis relativa a la varianza, σ.
- Si H_0 es cierta,

$$\frac{(n-1)S_c^2}{\sigma_0^2} \in \chi^2_{n-1}$$

- Recuerda que:

$$S_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Contraste sobre la varianza σ de una población normal con μ desconocida

Rechazamos la hipótesis nula $H_0 : \sigma^2 = \sigma_0^2$ frente a $H_1 : \sigma^2 \neq \sigma_0^2$ si

$$\frac{(n-1)S_c^2}{\sigma_0^2} \leq \chi^2_{n-1,1-\alpha/2} \quad \text{ó} \quad \frac{(n-1)S_c^2}{\sigma_0^2} \geq \chi^2_{n-1,\alpha/2}$$

Rechazamos la hipótesis nula $H_0 : \sigma^2 \leq \sigma_0^2$ frente a $H_1 : \sigma^2 > \sigma_0^2$ si

$$\frac{(n-1)S_c^2}{\sigma_0^2} \geq \chi^2_{n-1,\alpha}$$

Rechazamos la hipótesis nula $H_0 : \sigma^2 \geq \sigma_0^2$ frente a $H_1 : \sigma^2 < \sigma_0^2$ si

$$\frac{(n-1)S_c^2}{\sigma_0^2} \leq \chi^2_{n-1,1-\alpha}$$

χ^2 con $n-1$ g.l.

Contrastes referidos a las medias de dos poblaciones normales

- En algunas ocasiones estamos interesados en contrastes sobre la diferencia de medias $\mu_1 - \mu_2$ de dos poblaciones.
- Tenemos dos muestras:
  - Una muestra formada por n_1 variables independientes y con la misma distribución $N(\mu_1, \sigma_1)$
  - Una muestra formada por n_2 variables independientes y con la misma distribución $N(\mu_2, \sigma_2)$
- Suponemos que las muestras son **independientes** (los individuos donde se han obtenido las mediciones de la población 1 son distintos de los individuos donde se han obtenido las mediciones de la población 2).
- Suponemos que las **varianzas σ_1^2 y σ_2^2 son conocidas**.

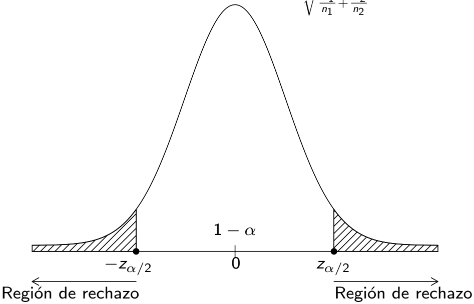
Contraste bilateral (hipótesis nula simple)

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

- El sentido común nos aconseja rechazar la hipótesis nula de que $\mu_1 = \mu_2$ cuando $\bar{X}_1 - \bar{X}_2$ sea muy distinta de 0.

Contrastes referidos a las medias de dos poblaciones normales

Si H_0 es cierta, la distribución de $\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ es $N(0,1)$



Rechazamos la hipótesis nula $H_0 : \mu_1 = \mu_2$ frente a $H_1 : \mu_1 \neq \mu_2$ si

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \leq -z_{\alpha/2} \quad \text{ó} \quad \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \geq z_{\alpha/2}$$

Contrastes referidos a las medias de dos poblaciones normales

Rechazamos la hipótesis nula $H_0 : \mu_1 \leq \mu_2$ frente a $H_1 : \mu_1 > \mu_2$ si

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq z_\alpha$$

Rechazamos la hipótesis nula $H_0 : \mu_1 \geq \mu_2$ frente a $H_1 : \mu_1 < \mu_2$ si

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq -z_\alpha$$

Contrastes referidos a las medias de dos poblaciones normales (var. desconocidas)

- En algunas ocasiones estamos interesados en contrastes sobre la diferencia de medias $\mu_1 - \mu_2$ de dos poblaciones.
- Tenemos dos muestras:
  - Una muestra formada por n_1 variables independientes y con la misma distribución $N(\mu_1, \sigma_1)$
  - Una muestra formada por n_2 variables independientes y con la misma distribución $N(\mu_2, \sigma_2)$
- Suponemos que las muestras son **independientes** (los individuos donde se han obtenido las mediciones de la población 1 son distintos de los individuos donde se han obtenido las mediciones de la población 2).
- Suponemos que las **varianzas σ_1^2 y σ_2^2 son desconocidas pero iguales**.
- Recuerda que si suponemos que las varianzas de las dos poblaciones son iguales el mejor estimador de la varianza será:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

Contrastes referidos a las medias de dos poblaciones normales (var. desconocidas)

Rechazamos la hipótesis nula $H_0 : \mu_1 = \mu_2$ frente a $H_1 : \mu_1 \neq \mu_2$ si

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} \leq -t_{n_1+n_2-2,\alpha/2} \quad \text{ó} \quad \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} \geq t_{n_1+n_2-2,\alpha/2}$$

Rechazamos la hipótesis nula $H_0 : \mu_1 \leq \mu_2$ frente a $H_1 : \mu_1 > \mu_2$ si

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} \geq t_{n_1+n_2-2,\alpha}$$

Rechazamos la hipótesis nula $H_0 : \mu_1 \geq \mu_2$ frente a $H_1 : \mu_1 < \mu_2$ si

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} \leq -t_{n_1+n_2-2,\alpha}$$

t con $n_1 + n_2 - 2$ g.l.

Contrastes referidos a las varianzas de dos poblaciones normales

- Tenemos dos muestras:
  - Una muestra formada por n_1 variables independientes y con la misma distribución $N(\mu_1, \sigma_1)$
  - Una muestra formada por n_2 variables independientes y con la misma distribución $N(\mu_2, \sigma_2)$
- En algunas ocasiones estamos interesados en contrastes sobre la ratio de varianzas $\frac{\sigma_1^2}{\sigma_2^2}$ de dos poblaciones.
- Si H_0 es cierta, $\frac{\sigma_1^2}{\sigma_2^2} = 1$. Equivalentemente, tenemos que $\sigma_1^2 = \sigma_2^2$.
- Suponemos que las muestras son **independientes** (los individuos donde se han obtenido las mediciones de la población 1 son distintos de los individuos donde se han obtenido las mediciones de la población 2).
- En la comparación de dos varianzas se tiene que, si H_0 es cierta,

$$\frac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1}.$$

Contrastes referidos a las varianzas de dos poblaciones normales

Rechazamos la hipótesis nula $H_0 : \sigma_1^2 = \sigma_2^2$ frente a $H_1 : \sigma_1^2 \neq \sigma_2^2$ si

$$\frac{S_1^2}{S_2^2} \leq F_{n_1-1,n_2-1,1-\alpha/2} \quad \text{ó} \quad \frac{S_1^2}{S_2^2} \geq F_{n_1-1,n_2-1,\alpha/2}$$

Rechazamos la hipótesis nula $H_0 : \sigma_1^2 \leq \sigma_2^2$ frente a $H_1 : \sigma_1^2 > \sigma_2^2$ si

$$\frac{S_1^2}{S_2^2} \geq F_{n_1-1,n_2-1,\alpha}$$

Rechazamos la hipótesis nula $H_0 : \sigma_1^2 \geq \sigma_2^2$ frente a $H_1 : \sigma_1^2 < \sigma_2^2$ si

$$\frac{S_1^2}{S_2^2} \leq F_{n_1-1,n_2-1,1-\alpha}$$

t con $n_1 + n_2 - 2$ g.l.

Consejos prácticos

1. ¿Qué queremos comparar?:
    - media,
    - varianza,
    - igualdad de medias o varianzas,
    - proporciones,...
2. Elegir el estadístico adecuado y su distribución.
3. Comparación con el cuantil correspondiente dependiendo si el contraste es o no bilateral.
4. Toma de decisión: ¿acepto o rechazo H_0?