

Bioestadística

Sesión 8: REPASO

José Aurelio Pina Romero

Ja.pina@ua.es

Bioestadística - Grado Enfermería

UA- Departamento de Enfermería

¿Qué vamos a ver?

- Repaso

Significación: p (Nivel de significación a posteriori)

Se define como la **probabilidad de observar**, suponiendo que H_0 es cierta, un valor del EC igual o mayor que el observado en la muestra.

P no se fija a priori, sino que es calculado a partir de los datos de la muestra

Si contraste bilateral $\rightarrow p = P(t \leq -|EC|) + P(t \geq |EC|)$

No se rechaza
 $H_0: \mu=70$

Se conoce también como el nivel de significación observado

α

$H_0: \mu=70$

$\bar{X} = 72$

Significación: p (Nivel de significación a posteriori)

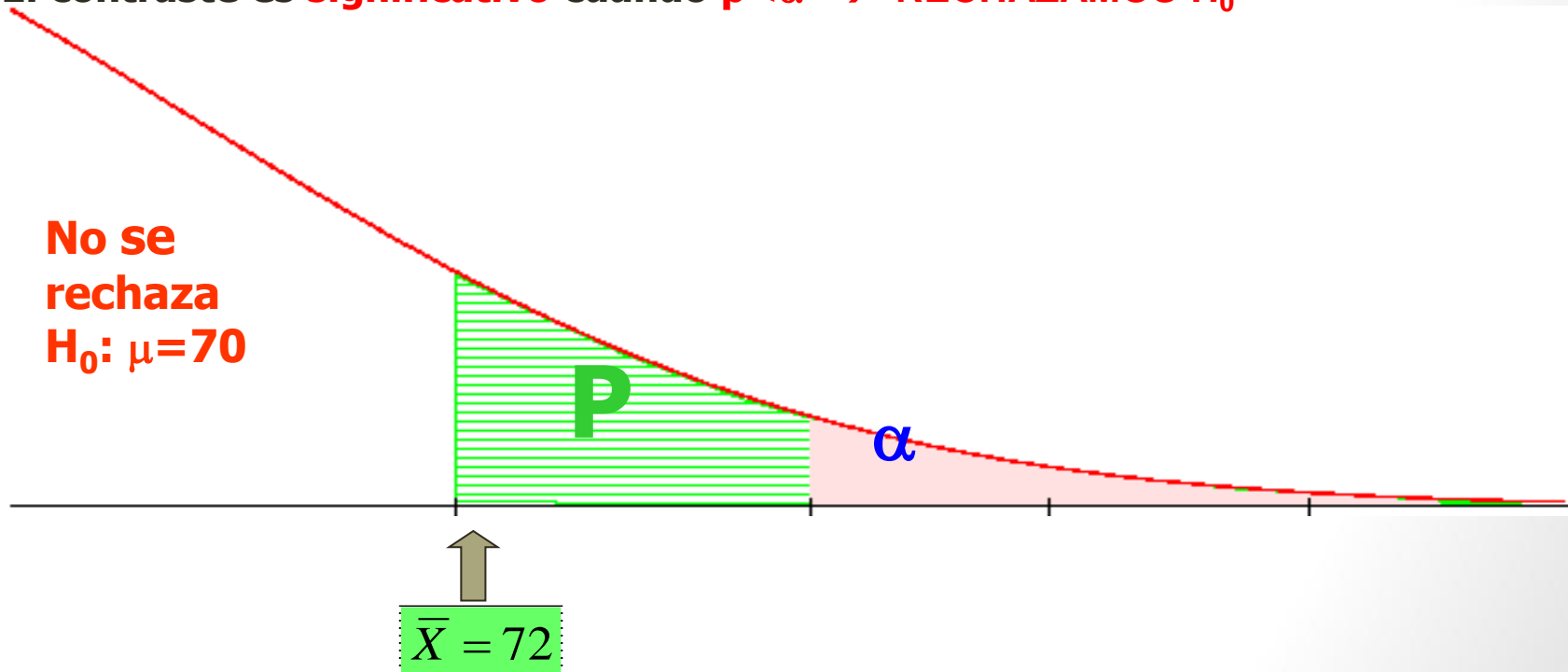
Es la probabilidad que tendría una región crítica que comenzase exactamente en el valor del estadístico obtenido de la muestra.

Es la probabilidad de tener una muestra que **discrepe aún más** que la nuestra de H_0 .

p es conocido **después de** realizar el experimento aleatorio

El contraste es **no significativo** cuando $p > \alpha \rightarrow$ NO RECHAZAMOS H_0

El contraste es **significativo** cuando $p < \alpha \rightarrow$ RECHAZAMOS H_0



Significación : p

RESUMEN

$$p = P(t \leq -|EC|) + P(t \geq |EC|)$$

Se conoce también como el nivel de significación observado

El contraste es **no significativo** cuando $p > \alpha \rightarrow$ NO RECHAZAMOS H_0

El contraste es **significativo** cuando $p < \alpha \rightarrow$ RECHAZAMOS H_0

p-valores "pequeños" proporcionan evidencia en contra de H_0

p-valores "grandes" proporcionan evidencia a favor de H_0

REPASO

La **inferencia estadística** es un proceso mediante el cual se pretende extraer conclusiones sobre determinadas características de la población objeto de estudio (**parámetros**) a partir de los datos de una muestra de dicha población.

El **error muestral o aleatorio** es el que se comete por el hecho de seleccionar una muestra y no toda la población. Está relacionado con el tamaño de la muestra.

Se controla con el tamaño muestral y con un análisis estadístico eficiente.

↑ tamaño ↓ error

El **error sistemático o sesgo**, aparece cuando se introduce un error en el diseño del trabajo, ya sea en la selección de los individuos, en la información recogida o en su análisis.

Debilidades diseño metodológico.
Debilidades en la recolección de datos.
Afectan a su validez.

REPASO

- **Contraste de Hipótesis:** Procedimiento estadístico mediante el cual se investiga la aceptación o rechazo de una afirmación acerca de una o varias características de una o más variables en una o más poblaciones.
- **¿Qué es una hipótesis?**
Una creencia o afirmación sobre la **población**, principalmente sus parámetros:
 - Media
 - Varianza
 - Proporción/Tasa

$$a = P(\text{cometer error tipo I}) = P(\text{Rechazar } H_0/H_0 \text{ cierta})$$

$$b = P(\text{cometer error tipo II}) = P(\text{Aceptar } H_0/H_0 \text{ falsa})$$

V/F

- La **inferencia estadística** es un proceso mediante el cual se pretende extraer conclusiones sobre determinadas características de la población objeto de estudio (parámetros) a partir de los datos de una muestra de dicha población.
- El **error muestral** es el que se comete por el hecho de seleccionar una muestra y no toda la población.
- El **error sistemático o sesgo** es el que se comete por falta de representatividad de la muestra.
- El **parámetro** es una característica o medida obtenida mediante el uso de todos los valores de datos de una **población** específica.
- Un **estimador** es una cantidad numérica **calculada sobre una muestra** y que esperamos que sea una buena **aproximación** de cierta cantidad con el mismo significado en la población (**parámetro**).

V/F

- El **muestreo aleatorio o probabilístico** requiere que todos los individuos de la población tengan una probabilidad conocida y distinta de cero de ser incluidos en la muestra.
- Cualquier característica calculada a partir de los datos de la muestra (cualquier función de los datos de la muestra) es un **Estadístico**.
- Un **parámetro** es una característica cuantitativa conocida referida a la muestra.

V/F

- En un contraste de hipótesis el error tipo I es el que se comete al rechazar la hipótesis nula cuando es cierta.

$$a = P(\text{cometer error tipo I}) = P(\text{Rechazar } H_0/H_0 \text{ cierta})$$

$$b = P(\text{cometer error tipo II}) = P(\text{Aceptar } H_0/H_0 \text{ falsa})$$

V/F

- En un estudio nutricional realizado entre 60 individuos de la población de estudiantes de la UA, se sabe que el intervalo de confianza al 95% para la media poblacional de colesterol (en mg/100 ml) es de (195, 205). A partir de estos resultados conteste a las siguientes preguntas:
 - El **estimador puntual** de la media poblacional de colesterol es de 200 mg/100 ml.
 - **Podemos afirmar con seguridad** que la media poblacional se encontrará entre 195 y 205.
 - **Sería admisible** un valor de la media poblacional de colesterol superior a 206 mg/100ml al 95% de confianza.
 - **Sería admisible** un valor de la media poblacional de colesterol superior a 200 mg/100ml al 95% de confianza.
 - Si construimos un intervalo al 99% la amplitud del IC aumentaría (sería menos preciso).

V/F

- En un estudio sobre litiasis biliar realizado en el hospital de Orihuela se obtuvo información sobre una muestra de 384 pacientes, en el que se pretendía valorar las diferencias en el nivel promedio de colesterol según su género. Los resultados se muestran en la siguiente tabla:

	Promedio de colesterol (mg/100 ml)	Desviación típica de colesterol (mg/100 ml)	N.º pacientes
Hombres	226,58	49,39	92
Mujeres	228,06	51,87	108

- Se puede concluir, **sin necesidad** de cálculos adicionales, que existe evidencia de que las medias poblacionales de colesterol entre los hombres y mujeres son significativamente distintas, ya que las medias muestrales son distintas.
- Para responder a la **cuestión anterior** será necesario realizar un contraste de comparación de medias entre las medias poblacionales de colesterol entre hombres y mujeres.

V/F

- La **hipótesis nula** deberá ser que las medias poblacionales de colesterol de los dos grupos de pacientes son iguales
- Es necesario realizar, en primer lugar, un **contraste de igualdad de varianzas** para determinar la distribución muestral y el estadístico de contraste indicado para comparar las medias poblacionales.
- Si el **valor de la p** del contraste finalmente es de 0,06 se rechazaría la hipótesis de igualdad de medias al nivel $\alpha=0,05$.