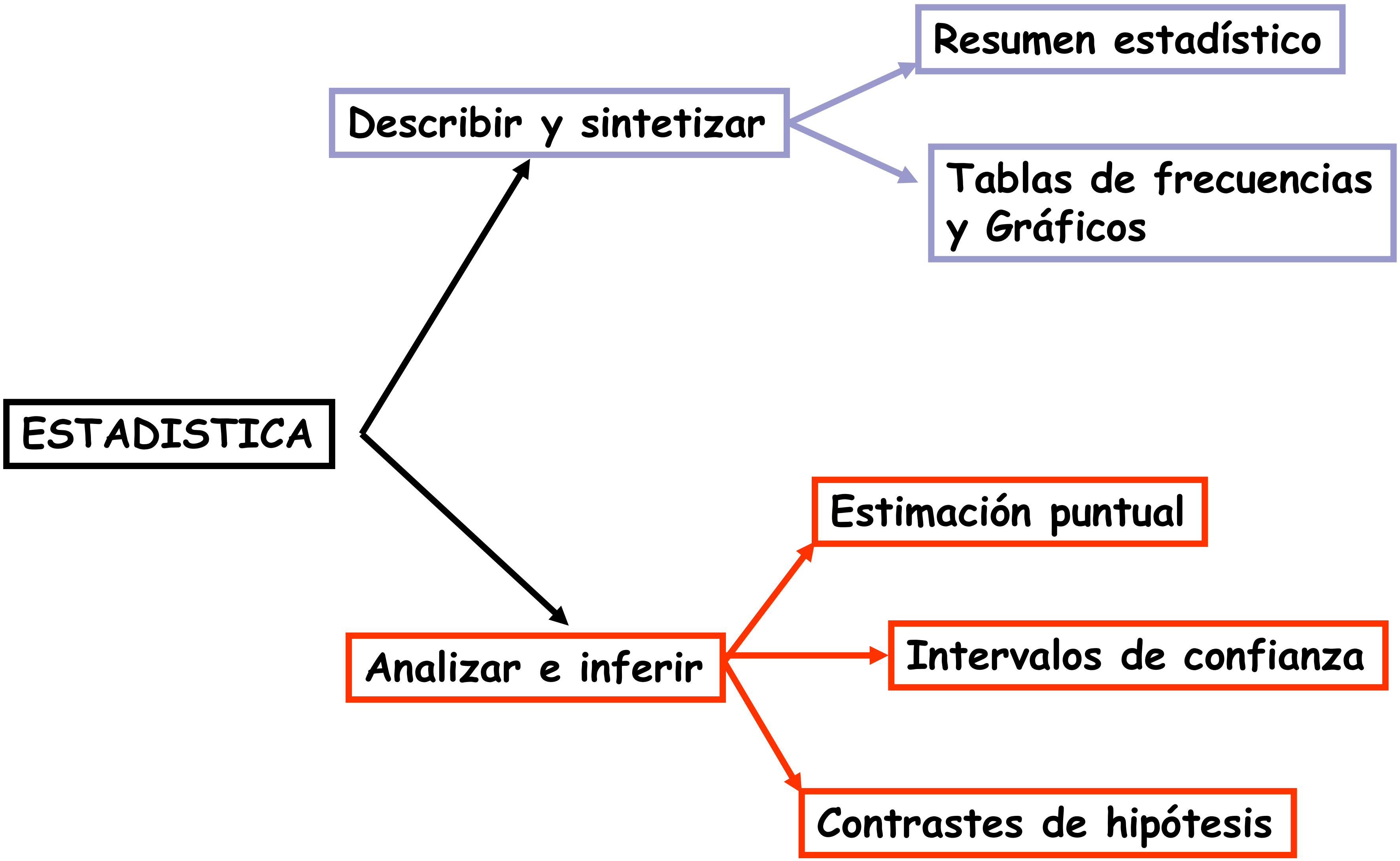
ANÁLISIS DE VARIABLES RELACIONADAS CON LA DOCENCIA E INVESTIGACIÓN EN LA FAC. FARMACIA

Estimación y contraste de hipótesis

Francisco M. Ocaña Peinado @ocanapaco
http://www.ugr.es/local/fmocan
Departamento de Estadística e Investigación Operativa. UGR


ESTADISTICA
Describir y sintetizar
Resumen estadístico
Tablas de frecuencias y Gráficos
Analizar e inferir
Estimación puntual
Intervalos de confianza
Contrastes de hipótesis

Técnicas de Inferencia Estadística

***Estimación puntual*: Pronosticar el valor de un parámetro desconocido a partir del valor que nos proporciona un estadístico de la muestra.

“*Se estima que el porcentaje de personas con obesidad en una población es del 14.3%*”

Estimación por intervalo: Consiste en atribuir al parámetro desconocido un intervalo de valores entre los que se espera que pueda encontrarse el valor del mismo con una probabilidad alta y conocida.

“*Se estima que el peso medio de una población infantil está comprendido entre 21.3 kg y 23.3 kg con una confianza del 99%.*”

Contraste de Hipótesis: Dada una afirmación acerca de una población, determinar si puede rechazarse o no dicha afirmación a partir de los datos de la muestra.

“*La obesidad en poblaciones infantiles afecta en mayor proporción a niños que a niñas*”

I - ESTIMACIÓN

Dado un parámetro poblacional de interés
¿En qué consiste un problema de estimación estadística?

1. Determinar el *mejor modo* de resumir la información muestral para inferir el valor del parámetro

⟶ Concepto de **estimador (puntual)**

2. Dar una medida de la capacidad de variación del estimador de muestra a muestra

⟶ Concepto de **error típico de estimación**

3. Dar la mayor información que proporciona nuestra muestra acerca del parámetro de interés

⟶ Concepto de **Intervalo de confianza** para el parámetro estudiado

Ejemplos de Parámetros de interés

- Medias
- Proporciones
- Medidas de asociación como Riesgos, Coeficientes de correlación

Estimación puntual

Estimación puntual: Hacer un pronóstico numérico de un parámetro de la población con un único valor.

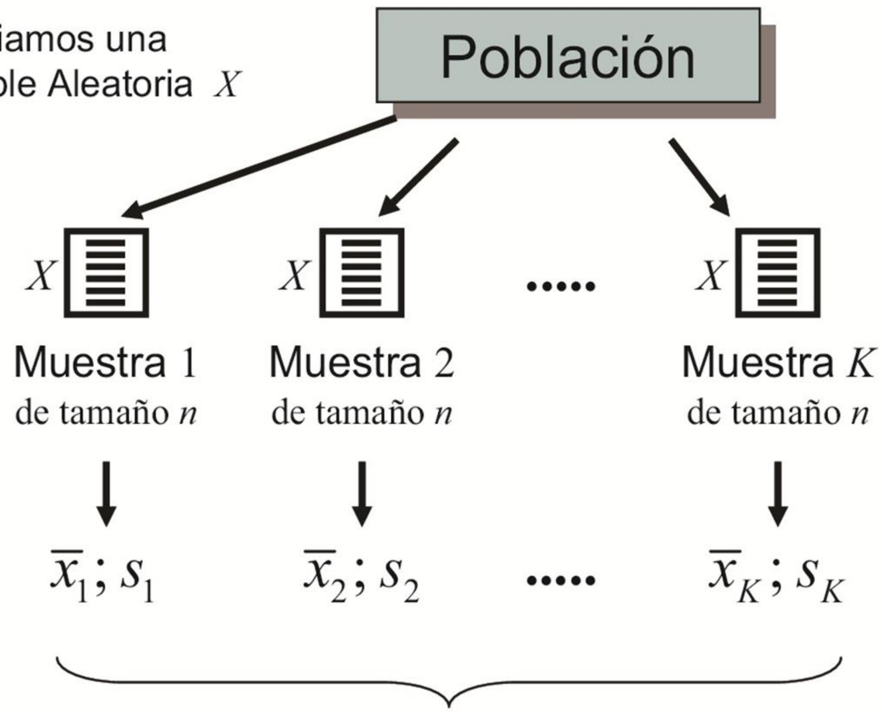
Parámetros	***Estimador Puntual***
μ	$\overline{x}$ (media muestral)
σ	$\hat{s}$ (cuasidesviación típica muestral)
p	$\hat{p}$ (proporción muestral)

Inconvenientes:

- Depende de la muestra particular que se obtenga.

- Existe una incertidumbre total de la proximidad o lejanía del valor puntual obtenido con respecto de la medida teórica.

Estudiamos una
Variable Aleatoria X

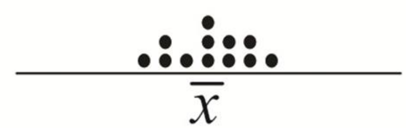


(1) Las muestras se eligen al azar

(2) Las medidas descriptivas son variables aleatorias (continuas)

Muestra	tamaño	media
M_1	n	$\overline{x}_1$
M_2	n	$\overline{x}_2$
...	...	...
M_K	n	$\overline{x}_K$

(3) ¿Se puede decir algo de la distribución que tienen las medias muestrales?

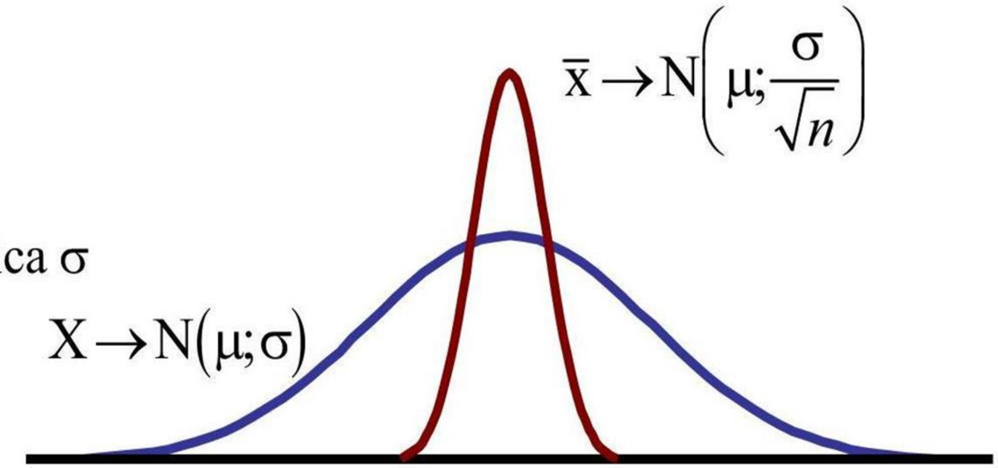


Teorema del Límite Central: para un número de observaciones suficientemente grande (digamos n>60) las medias de v.a. No normales (incluso de v.a. Discretas) tienen distribución aproximadamente normal. La aproximación mejora en la medida en que aumenta n

Distribución muestral de la media

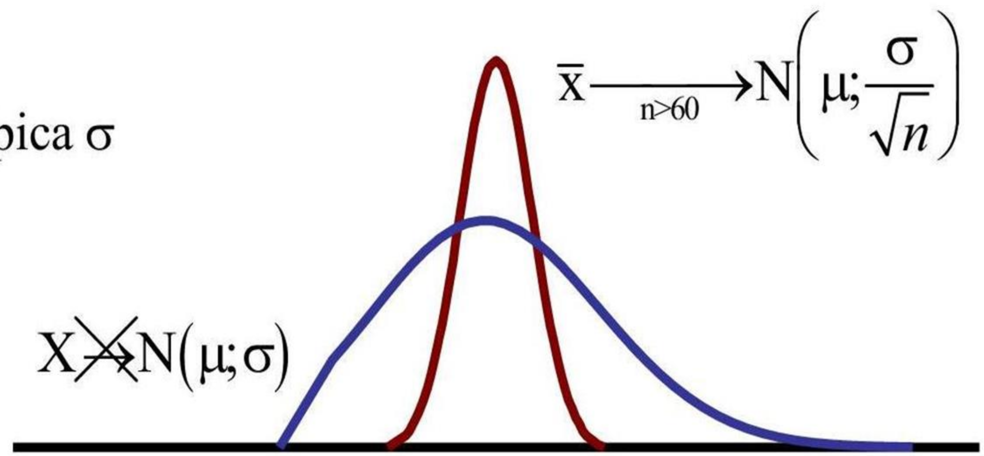
- X Normal con media μ y desviación típica σ

$$\bar{x} \longrightarrow N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$$



- X no normal con media μ y desviación típica σ

$$\bar{x} \xrightarrow[n>60]{} N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$$



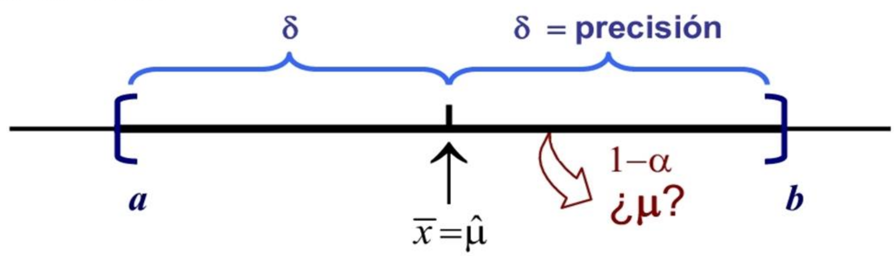
Intervalo de confianza

Concepto: Se trata de asignar al parámetro poblacional desconocido, por ejemplo μ, un intervalo de valores, digamos (*a, b*) entre los cuales está μ con una cierta probabilidad = **confianza** $(1-\alpha)$**.** Es decir, si se cumple que

$$P(a \leq \mu \leq b) = 1 - \alpha$$

diremos entonces que (*a, b*) es un **intervalo de confianza** para el parámetro μ construido al $(1-\alpha)$ % **de confianza** o, lo que es lo mismo, al α **% de error.**

Además de la **cobertura**, un IC proporciona otra magnitud de interés: la **precisión** **(** δ **)**



Niveles de serotonina en sangre en una población (conocida) de 20 pacientes con depresión.
5 muestras aleatorias de tamaño n=5 y los IC a que dan lugar

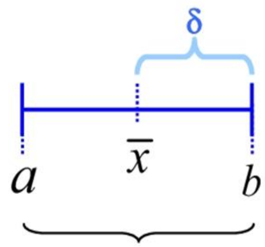
Población

108	118
112	120
123	133
109	127
121	125
136	115
124	117
113	125
118	117
129	110

$\mu = 120$

Muestra	Datos					$\bar{x}$	s	Intervalo (95% conf.)
M1	123	125	118	125	113	120.80	5.215	**(114.325 ; 127.275)**
M2	124	110	115	133	112	118.80	9.576	**(106.912 ; 130.688)**
M3	125	113	117	123	124	120.40	5.177	**(113.973 ; 126.827)**
M4	133	110	136	125	110	122.80	12.357	**(107.459 ; 138.141)**
M5	118	113	117	110	112	114.00	3.391	**(109.790 ; 118.210)** !

Intervalos construidos a nivel de confianza del 95% $\longrightarrow$ $P(a \le \mu \le b) = 1-\alpha = 0.95$



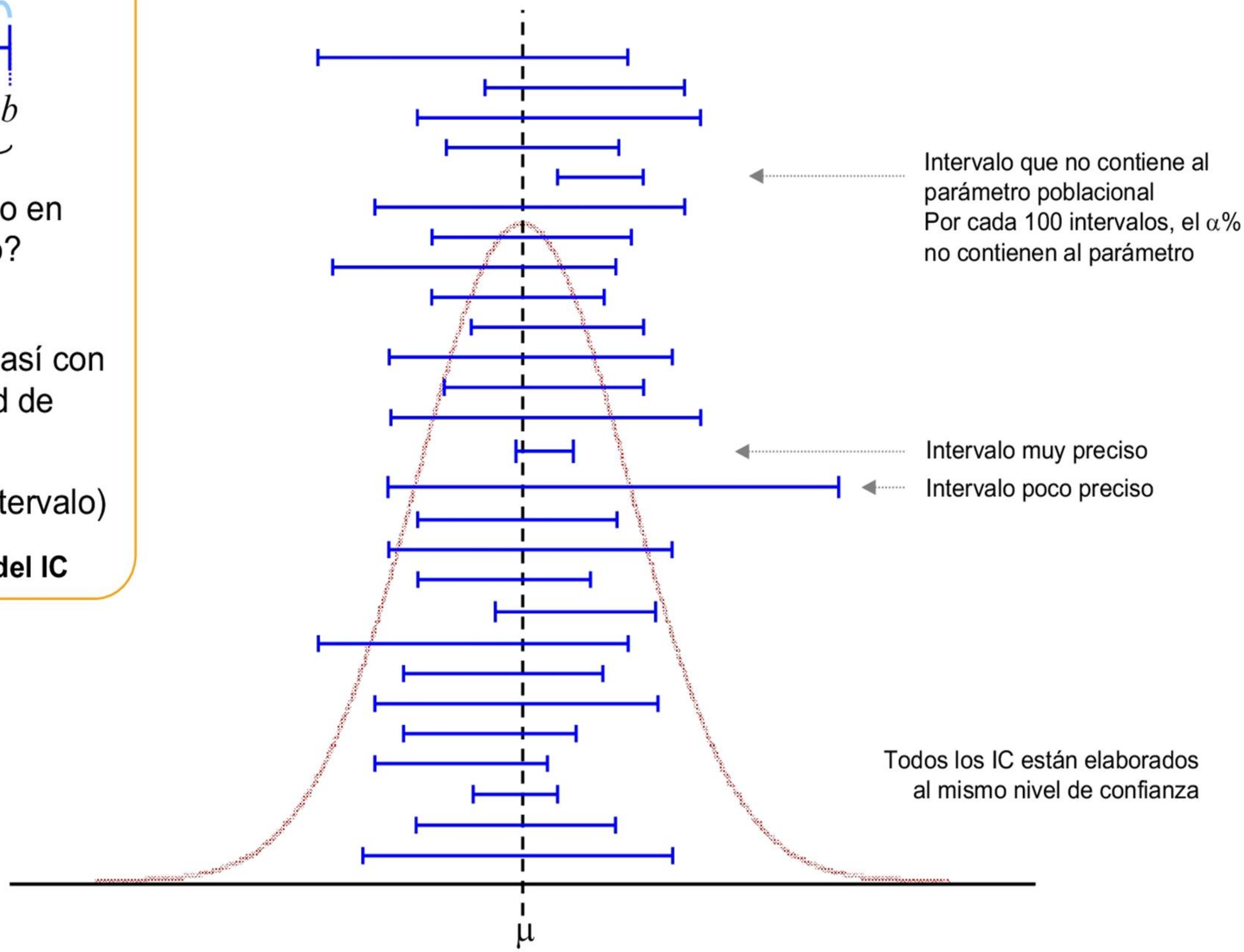
¿está μ contenido en este intervalo?

Se espera que sea así con una probabilidad de

$$1-\alpha$$

(la **confianza** del intervalo)

δ es la precisión del IC



Estimación por I. C.

$$IC_{1-\alpha}(\mu) \equiv \left[\bar{x} \pm t_{n-1;\alpha/2} \frac{\hat{s}}{\sqrt{n}} \right] \qquad \bar{x} \to N\left(\mu, \frac{\sigma}{\sqrt{n}} \right)$$

$$IC_{1-\alpha}(p) \equiv \left[\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} \right] \quad \hat{p} \to N\left(p, \sqrt{\frac{pq}{n}} \right) (n \geq 30)$$

Forma general de un IC

$$\begin{array}{c}\text{estimador}\\ \text{centrado}\end{array} \overbrace{\pm \ \textit{valor probabilistico} \times SE(\textit{estimador})}^{\textit{precisión}(\delta)}$$

[nivel de confianza]

[variabilidad]

[tamaño de muestra]

La amplitud de un IC **no** dice "entre cuanto puede variar el parámetro buscado". Es una medida de nuestro grado de imprecisión

Amplitud y precisión aluden a lo mismo, aunque son términos contrapuestos, su valor numérico es el mismo: δ

Tamaño muestral

Objetivo: Calcular *n* para poder estimar el parámetro requerido con una cierta precisión.

Un cálculo muestral inadecuado puede llevar a:

- Conclusiones no válidas, para *n* excesivamente pequeño.
- Desperdiciar recursos si se ha tomado *n* excesivamente alto.

Para la determinación efectiva del tamaño de muestra lo primero que se debe **fijar** es **la precisión deseada** en la estimación **y el nivel de confianza**.

Forma general de un IC

$$\begin{matrix}\textit{estimador} \\ \textit{centrado}\end{matrix} \pm \overbrace{\textit{valor probabilistico} \times SE(\textit{estimador})}^{\textit{precisión}\,(\delta)}$$

[nivel de confianza]

[variabilidad]

[tamaño de muestra]

Tamaño muestral

Objetivo principal del estudio: Estimar una media poblacional, μ, con nivel de confianza (1- α), y precisión δ:

$$n \geq \frac{z_{\alpha/2}^{2}\,\hat{s}_{p}^{2}}{\delta^{2}} \qquad \hat{s}_{p}^{2} \text{ se estima a partir de muestra piloto}$$

Objetivo principal del estudio: Estimar una proporción poblacional, p, con nivel de confianza (1- α), y precisión δ:

$$n \geq \frac{z_{\alpha/2}^{2}\,\hat{p}^{*}\hat{q}^{*}}{\delta^{2}} \qquad \hat{p}^{*} \text{ se estima a partir de muestra piloto}$$

Tomar p = q = 0.5, en caso de duda (criterio de máxima indeterminación)

NOTA: Si hay información previa, la determinación de *n* es aproximada dado que se basa en la información de una muestra aleatoria; que puede no ser *«afortunada»* como representación de la población

Tamaño muestral: ficha técnica encuestas

GAD3

Ficha técnica

- **Ámbito:** Nacional
- **Universo:** Población general mayor de 18 años (incluidas Ceuta y Melilla)
- **Tamaño de la muestra:** 1.000 entrevistas
- **Cuotas:** de sexo, edad en cuatro grupos (18-29, 30-44, 45-64, 65 y más) según el censo.
- **Error muestral:** ± 3,2 % para el total de la muestra y para un grado de confianza del 95,5% (dos sigmas) y en la hipótesis más desfavorable de P=Q=0,5 en el supuesto de muestreo aleatorio simple
- **Procedimiento de recogida de la información:** Entrevista telefónica asistida por ordenador (CATI)
- **Duración media de la entrevista:** 7 minutos de media
- **Fecha de los trabajos de campo:** del 5 al 13 de Febrero de 2014
- **Trabajo de campo:** GAD3

ESOMAR

GAD3 es miembro de **ESOMAR** (*European Society for Opinion and Marketing Research*) y de **AEDEMO** (*Asociación Española de Estudios de Mercado, Marketing y Opinión*) y posee la clasificación LO 3C del Ministerio de Industria: para la realización de encuestas, toma de datos y servicios análogos.

Contraste (Test) de Hipótesis

Conjunto de reglas que se usan para determinar si una afirmación acerca de una población puede rechazarse o no a partir de los resultados obtenidos en una muestra.

En todo contraste de hipótesis, hay 2 enunciados excluyentes:

- **Hipótesis nula o H_0**: Hipótesis que se somete a comprobación para ver si se rechaza o no.
- **Hipótesis alternativa o H_1**: Hipótesis que se acepta cuando se rechaza H_0, es la negación de H_0

Ambas hipótesis **NO** tienen igual fuerza, salvo que se demuestre lo contrario, **hay que asumir que H_0 es cierta**.

Contraste de hipótesis: Filosofía

Contraste de hipótesis $\longrightarrow$ Hipótesis nula (H_0) *versus* alternativa (H_1)

Un contraste de hipótesis tienen una presunción a favor de H_0, de forma similar a como ocurre en los tribunales de justicia, donde hay una presunción de inocencia. Dado que uno es inocente hasta que se demuestre lo contrario, la evidencia aportada debe ser muy consistente para admitir la culpabilidad

Reglas de inferencia negativa $\longrightarrow$ Se da por supuesto que la hipótesis nula es verdadera

$\downarrow$

La decisión por H_1, viene dada a partir del rechazo de H_0

¿Quién es H_0?

- Ejemplo: «*En la ejecución de una prueba diagnóstico del SIDA hay 2 posibilidades: que se obtenga resultado positivo, o que se obtenga resultado negativo. No obstante, la prueba no es 100% exacta, ya que existen los falsos negativos y falsos positivos. Determine quién sería H_0*»

- Hipótesis:
  - **El individuo está enfermo (el virus está en el individuo)**
  - **El individuo está sano (no existe el virus)**
  - Seleccionar la hipótesis nula, **H_0 : El individuo está sano**

¿Quién es H_0?

- Ejemplo: «*Se está estudiando un nuevo fármaco para poder utilizarlo contra el cáncer de piel. Se espera que sea eficaz en la mayoría de los pacientes sobre los que se aplica. La empresa que produce el fármaco desea alguna prueba estadística que apoye tal afirmación*»

- Hipótesis:
  - **El nuevo fármaco es efectivo en la mayoría de pacientes**
  - **El nuevo fármaco no es efectivo en la mayoría de los pacientes**
  - Seleccionar la hipótesis nula, **H_0 : El nuevo fármaco no es efectivo**

Elementos a establecer en Contraste Hipótesis

- **Hipótesis nula H_o**
  - Afirmación en términos de igualdad $= , \leq , \geq$
  - Hipótesis científicamente más simple
  - Los datos pueden refutarla
  - No debería ser rechazada sin una buena razón.
- **Hipótesis Alternativa H_1**
  - Niega a H_0 $\neq , > , <$
  - Hipótesis que se quiere demostrar fuera de toda duda
  - Si nos decantamos por ella se quiere estar prácticamente seguro de que es cierta
  - Los datos pueden mostrar evidencia a favor
  - No debería ser aceptada sin una gran evidencia a favor.

Elementos a establecer en Contraste Hipótesis

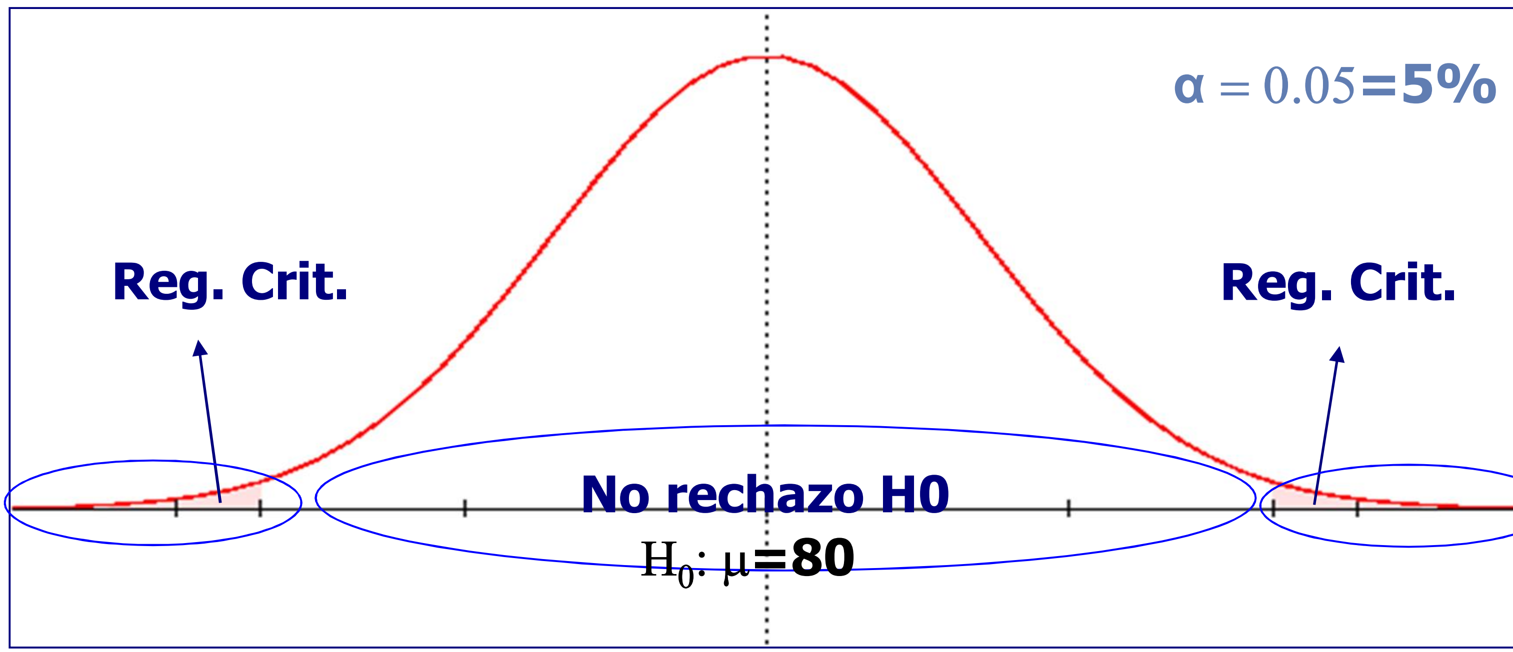
- **Nivel de significación, α**
  - Probabilidad [Rechazar H_0 / H_0 es cierta] = α
  - Es un valor tanto más pequeño cuántas más garantías se precisen de que una decisión a favor de H_1 sea correcta. Usualmente se toma α = 0.05

- **Estadístico de Contraste**
  - Variable aleatoria que se usa para realizar el test
  - Es una función de los valores de la muestra que resume la información relevante de ella.
  - **Si se cumple H_0**, su comportamiento es conocido

Región crítica y nivel de significación

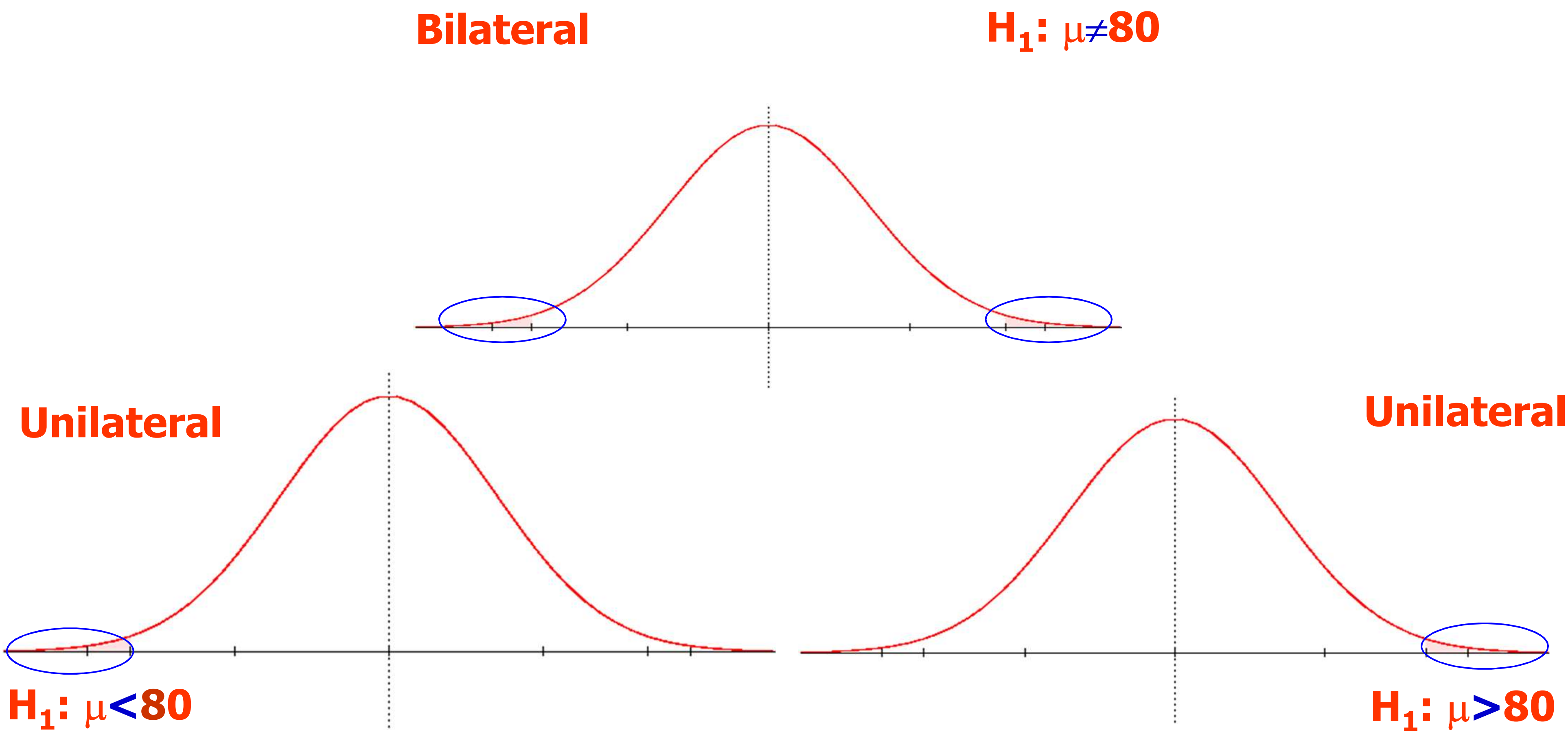
Región crítica

- Valores '*improbables*' **si Ho cierta**
- Es conocida antes de tomar la muestra: son resultados experimentales que harían rechazar H_0



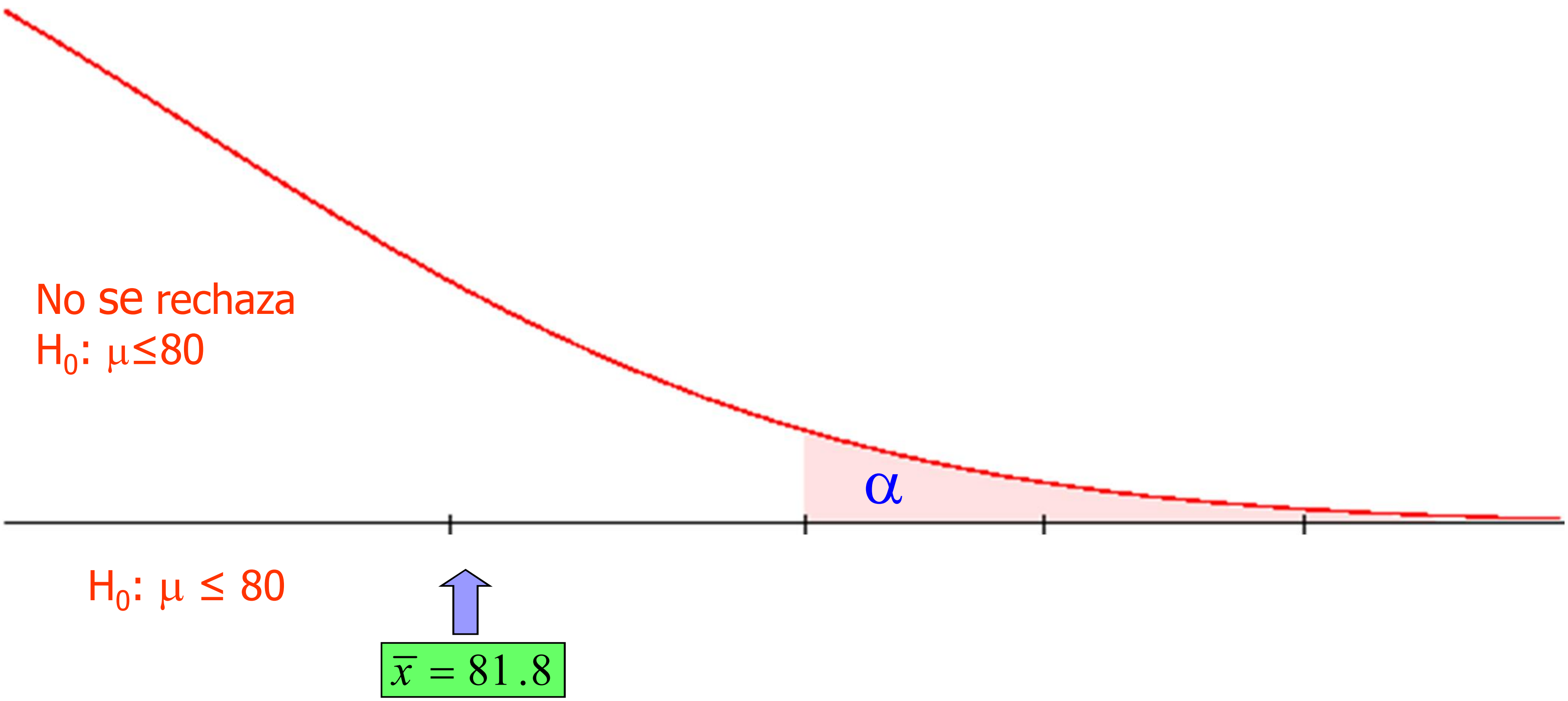
Contrastes: unilateral y bilateral

La posición de la región crítica depende de H_1



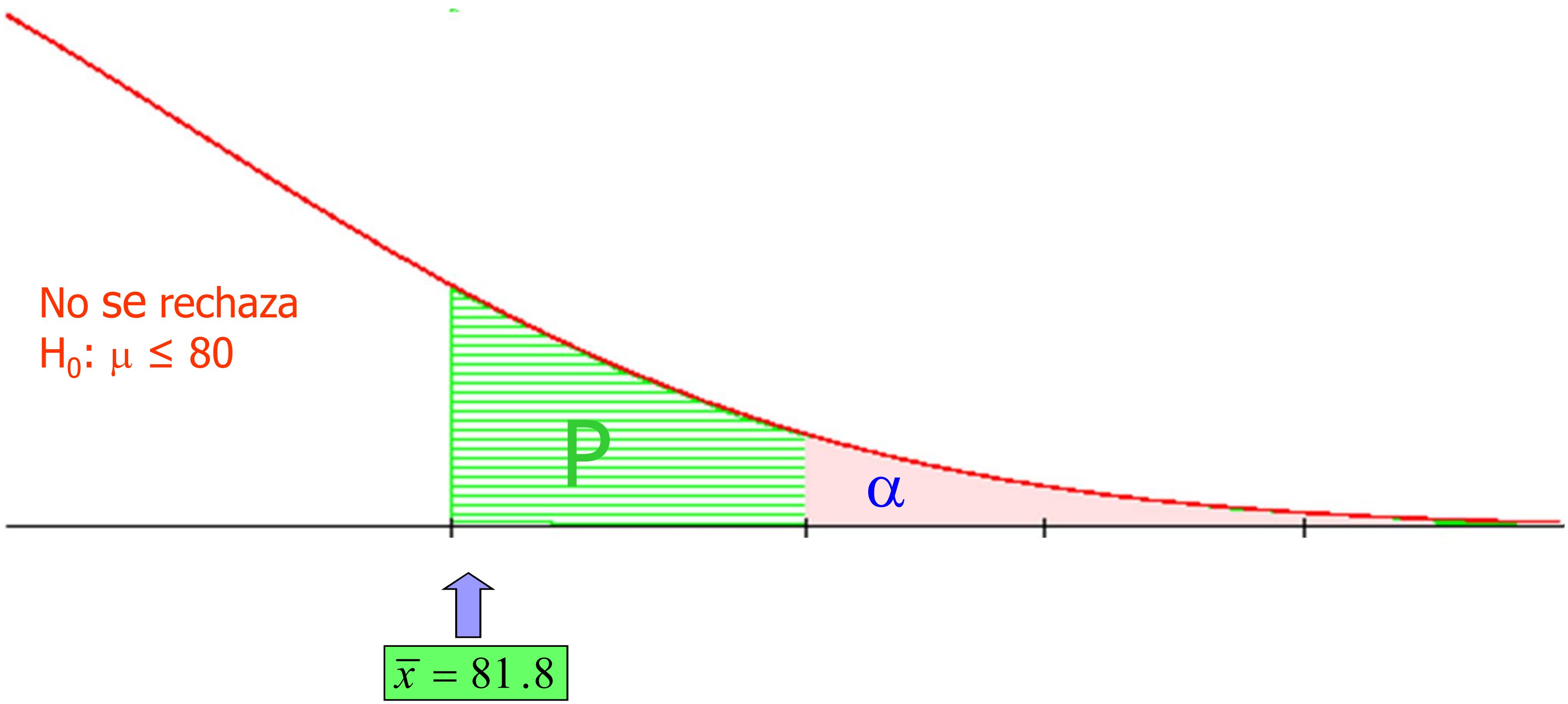
Significación (p-valor): interpretación

- Probabilidad de equivocarse al rechazar H_0
- Grado de credibilidad de H_0 según esa muestra
- Conocido una vez extraída la muestra.
- La probabilidad de que el rechazo de H_0 pueda explicarse simplemente por el azar o la casualidad.



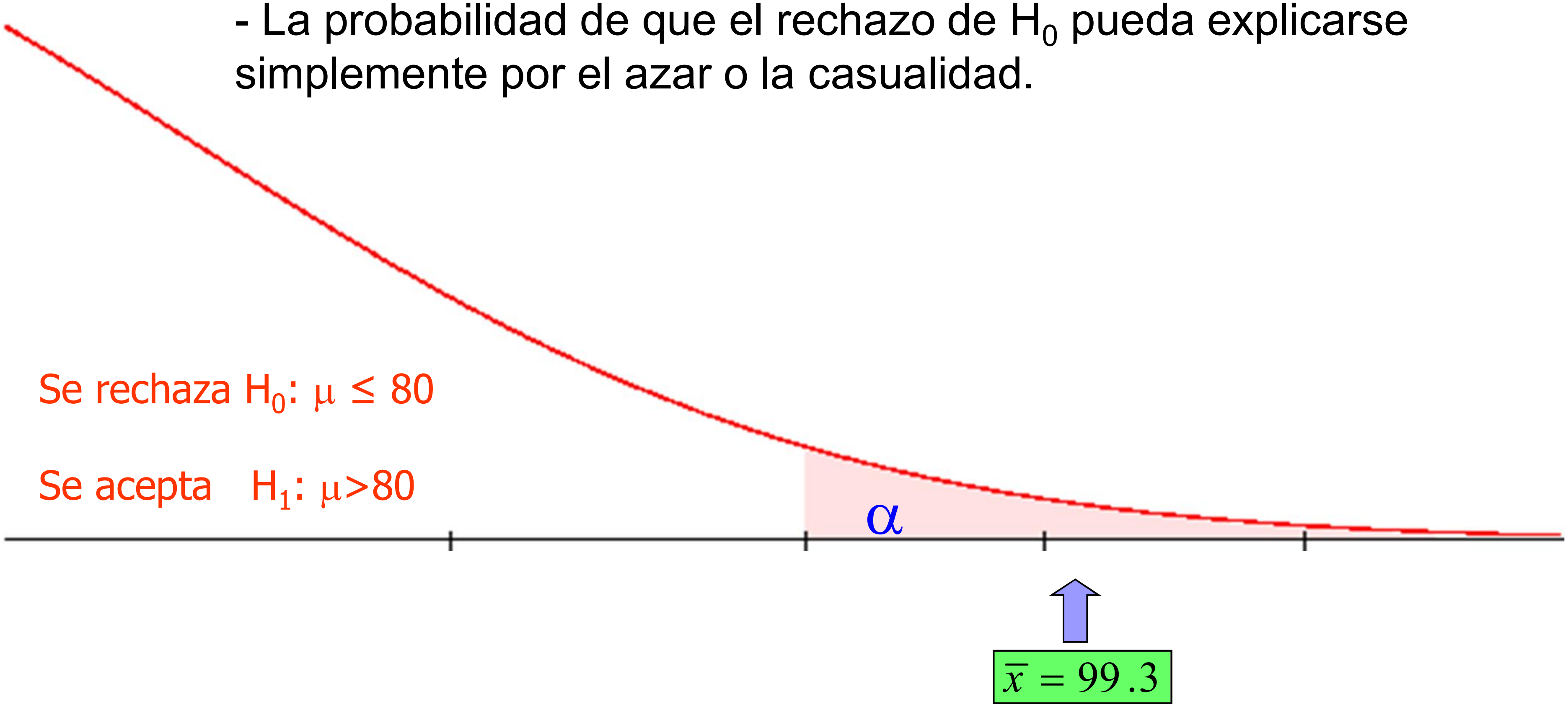
Significación (p-valor): interpretación

- Probabilidad de equivocarse al rechazar H_0
- Grado de credibilidad de H_0 según esa muestra
- Conocido una vez extraída la muestra.
- La probabilidad de que el rechazo de H_0 pueda explicarse simplemente por el azar o la casualidad.



Significación (p-valor): interpretación

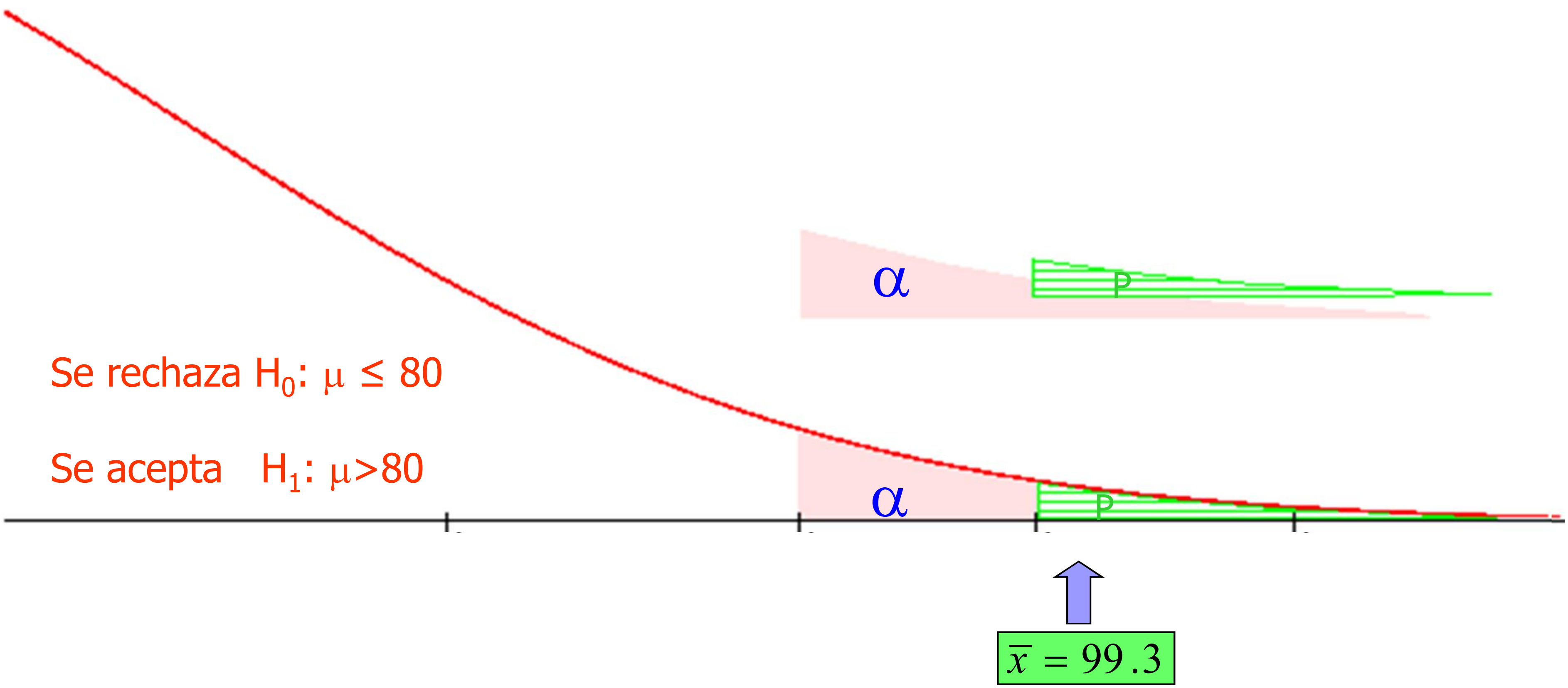
- Probabilidad de equivocarse al rechazar H_0
- Grado de credibilidad de H_0 según esa muestra
- Conocido una vez extraída la muestra.
- La probabilidad de que el rechazo de H_0 pueda explicarse simplemente por el azar o la casualidad.



Significación (p-valor): interpretación

El contraste es **estadísticamente significativo** **si** **p-valor ≤ α**

Es decir, si el resultado experimental discrepa más de "lo tolerado" *a priori*.



Riesgos al tomar decisiones

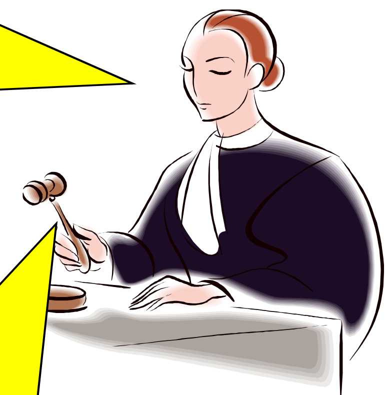
Ejemplo 1: Se juzga a un individuo por la *presunta* comisión de un delito

- H_0: Hipótesis nula
  - Es inocente

Los datos pueden refutarla

La que se acepta si las pruebas no indican lo contrario

Rechazarla por error tiene graves consecuencias



- H_1: Hipótesis alternativa
  - Es culpable

No debería ser aceptada sin una gran evidencia a favor

Rechazarla por error tiene consecuencias menos graves

Tipos de error al tomar una decisión

		Realidad	
		Inocente	Culpable
Veredicto	Inocente	OK	Error Menos grave
	Culpable	Error Muy grave	OK

Riesgos al contrastar hipótesis

Ejemplo 2: Se cree que un nueva dieta ofrece buenos resultados

- H_0: Hipótesis nula
  - (Ej.1) Es inocente
  - (Ej.2) El tratamiento no tiene efecto

- H_1: Hipótesis alternativa
  - (Ej.1) Es culpable
  - (Ej.2) El tratamiento es útil

Tipos de error al contrastar hipótesis

	Realidad	
Decisión	H_0 Cierta	H_0 Falsa
No Rechazo H_0 (*Retener* H_0)	**Correcto** *El tratamiento no tiene efecto y así se decide* Probabilidad **1 - α** «Especificidad del test»	**Error de tipo II** *El tratamiento sí tiene efecto pero no se percibe.* Probabilidad **β** «Falso negativo»
Rechazo H_0 Acepto H_1	**Error de tipo I** *El tratamiento no tiene efecto pero se decide que sí.* Probabilidad **α** «Falso positivo»	**Correcto** *El tratamiento tiene efecto y el experimento lo confirma.* Probabilidad **1- β** «Sensibilidad del test»

Tipos de error al contrastar hipótesis

- Toda decisión por H_1 lleva implícita un posible error, llamado error α, o **error Tipo I**:

$$\alpha = P\,[\text{Decidir } H_1 \,/\, H_0 \text{ es cierta}] \equiv \text{Error Tipo I}$$

- Toda decisión por H_0 lleva implícita un posible error, llamado error β, o **error Tipo II**:

$$\beta = P\,[\text{Decidir } H_0 \,/\, H_1 \text{ es cierta}] \equiv \text{Error Tipo II}$$

- **Consideraciones:**
  - α está controlado. Por eso **una decisión por H_1 es siempre fiable** (es posible que sea errónea, pero es muy improbable)
  - β no está controlado. Por eso **una decisión por H_0 NO es fiable** (lo único que indica es que no se ha podido demostrar que H_0 es falsa)

Potencia de un contraste de hipótesis

Se llama ***potencia de un contraste*** a la capacidad del test para detectar hipótesis alternativas verdaderas, es decir 1- β

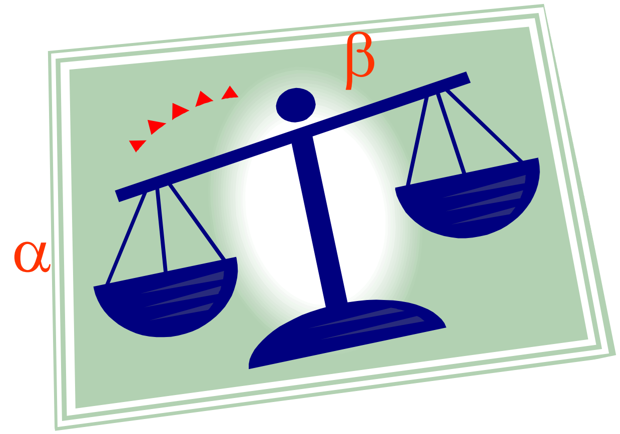
Potencia = 1- β = P [Decidir H_1 / H_1 es cierta]

Un contraste o test de hipótesis es mejor **cuanto más potente** sea

Test Paramétricos vs No Paramétricos

Las **pruebas no paramétricas son menos potentes**: son más exigentes al rechazar H_0, y por tanto tienen menos posibilidades de acertar cuando no la rechazan (más posibilidad de cometer un error de tipo II).

No se puede tener todo



- Cuando α disminuye, β aumenta (la potencia decrece).
- Para un tamaño muestral fijo, no se pueden reducir a la vez ambos tipos de error.
- En general, las muestras pequeñas producen potencia pequeña (difícil rechazar H_0).
- Para *n* escasa, es muy difícil detectar diferencias significativas.
- Para reducir β, y por tanto aumentar la potencia, hay que aumentar el tamaño muestral.

Pasos para resolver un contraste de hipótesis con SPSS

1. **Fijar H_0 y H_1** de acuerdo al problema.
2. Seleccionar el test adecuado para resolver el contraste: existirá una variable aleatoria (estadístico de contraste) asociada a dicho contraste, tal que **si H_0 es cierta, tendrá un comportamiento conocido**.
3. **Verificar** que se cumplen las **condiciones** para poder aplicar el test del paso 2.
4. Fijar el nivel de significación, **α**
5. Seleccionar una **muestra (aleatoria) de tamaño *n***, para la cual el estadístico de contraste tomará un valor numérico (se denomina valor experimental).
6. Adoptar una **decisión** sobre el rechazo o no de H_0: **Se rechaza H_0** si el **p-valor** asociado al valor experimental del estadístico de contraste es **$\leq \alpha$** . En caso contrario, **no se rechaza H_0**

Nota: Si $0.05 < \text{p-valor} \leq 0.15$, conviene $\uparrow n$ y repetir el test

Conclusiones

- En Ciencia, H_0 y H_1 no tienen el mismo peso:
  - H_0 : Hipótesis afirmativa o positiva (científicamente más simple).
  - H_1 : Generalmente la hipótesis que queremos demostrar.

- **Sobre el criterio de rechazo**

Se rechaza H_0 (contraste significativo) si p-valor $\leq \alpha$

En caso contrario, no se rechaza H_0

- Cualquier decisión tiene riesgo:
  - Rechazar H_0 (Aceptar H_1): Podemos cometer error de tipo I (controlado)
  - No rechazar (*Retener*) H_0: Podemos cometer error de tipo II (no controlado), por eso **H_0 no se acepta** en ningún caso

Escribir los resultados de un contraste

- Sea cual se la decisión tomada hay que indicar **siempre** el nivel de significación al que se trabaja, así como el p-valor asociado al contraste.

"En todos los análisis realizados, se fijó el nivel de significación, α, en 0.05."

"Los resultados anteriores muestran que hay diferencias estadísticamente significativas entre el grupo control y el grupo de tratamiento (p-valor = 0.003)"

"En la comparativa por edades, los resultados muestran diferencias estadísticamente significativas para las variables analizadas, excepto para el nivel de actividad física (p-valor = 0.329), el índice de alimentación saludable (p-valor = 0.356), el colesterol total (p-valor = 0.425), y en el colesterol LDL (p-valor = 0.115)"

- A la hora de escribir un p-valor, se deben **utilizar al menos 3 cifras decimales.** Cuando el p-valor es muy pequeño es posible indicarlo de una forma más breve escribiendo *p-valor < 0.001*. **NO escribir *p-valor = 0.000***

Sobre cálculo del tamaño de muestra fijada la potencia

- En algunos casos concretos, es posible fijando previamente 1-β y α, calcular el tamaño de muestra adecuado (véase por ejemplo la casuística que desarrolla *Wittes J. "Sample size calculations for randomized controlled trials". Epidemiologic Reviews 2002;24:39-53).*
- Factores que influyen en el tamaño de muestra:
  - α y 1-β seleccionados
  - **La magnitud que se considera relevante en la diferencia entre la hipótesis nula y alternativa**: «*Es más fácil detectar una diferencia grande que una diferencia pequeña*». «*Un experimento para detectar grandes diferencias no requerirá un n tan grande como otro con el que se desee detectar diferencias pequeñas*»
  - **La variabilidad de la variable respuesta en población que se muestrea**: «*cuanto mayor sea la variabilidad, más difícil es el proceso de estimación (más n se requiere)*»

Ejemplo cálculo de tamaño muestral fijada la potencia

- Se desea contrastar diferencia de medias entre 2 poblaciones (caso del tipo control/tratamiento).
- **Hipótesis**: la respuesta en ambas poblaciones tiene distribución Normal y variabilidad igual y conocida (σ), y se toman **grupos equilibrados en tamaño**.
- La diferencia promedio entre grupos es la medida de eficacia del tratamiento:
  - $H_0 : \mu_1 = \mu_2$; (Tratamiento no efectivo)
  - $H_1 : \mu_1 \neq \mu_2$; (Medias signif. distintas). **¿Tratamiento efectivo?**
- Hay que **fijar previamente d,** que es la mínima diferencia para la cual se considera que el tratamiento es efectivo.
- Es decir, el investigador supone que:
  - Grupo 1: Respuesta media es μ_1
  - Grupo 2*:* Respuesta media es $\mu_2 = \mu_1 \pm$ **d**

Ejemplo cálculo de tamaño muestral fijada la potencia

- Fijando α y β, y suponiendo conocida la variabilidad de la respuesta en la población σ (desviación típica poblacional), se tiene que el **tamaño muestral total del estudio** es:

$$n = 2 \cdot \frac{\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 \sigma^2}{d^2}$$

donde z_t es el percentil de orden t de una Normal estándar N(0,1)

Interpretación: Para nivel de significación α, el *n* obtenido garantiza, que el test realizado con tal tamaño de muestra será significativo el (1-β)x100% de las veces en que la verdadera hipótesis H_1 se diferencie de H_0 en la cantidad *d* especificada

Ejemplo cálculo de tamaño muestral fijada la potencia

Ejemplo: Para potencia del 80% y nivel de significación del 5%, se desea detectar una diferencia mínima de 6 unidades, y suponiendo σ=10 unidades, se tendría que:

$$n = 2 \cdot \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 \sigma^2}{d^2}$$

donde z_t es el percentil de orden t de una Normal estándar N(0,1)

$$\alpha = 0.05 \Rightarrow z_{1-\alpha/2} = z_{0.975} = 1.96$$

$$1-\beta = 0.8 \Rightarrow z_{1-\beta} = z_{0.8} = 0.84$$

$$d = 6 \;;\; \sigma = 10; \text{ se tiene } \; n = 43.6048 \Rightarrow n = 44$$

Interpretación: Para nivel de significación del 5%, con una asignación de al menos n = 22 en ambos grupos, se tendría una potencia del 80% para tratar de detectar si el efecto del tratamiento difiere con el del control en al menos 6 unidades

Ejemplo cálculo de tamaño muestral fijada la potencia

En el caso de **grupos no equilibrados habrá que tomar más muestra** para mantener la potencia fijada, siendo la expresión general la siguiente:

$$n = \frac{r+1}{r} \cdot \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 \sigma^2}{d^2}$$

Siendo *r la razón deseada entre ambos grupos (0 < r ≤ 1)*

- ***r = 1/2, Razón 1 a 2 entre grupos***

$$\alpha = 0.05 \quad 1-\beta = 0.8 \quad d = 6 \quad \sigma = 10$$

$$\text{se tiene } \; n = 65.407 \cong n = 66 \quad n_1 = 22 \quad n_2 = 44$$

- ***r = 1/5, Razón 1 a 5 entre grupos***

$$\alpha = 0.05 \quad 1-\beta = 0.8 \quad d = 6 \quad \sigma = 8$$

$$\text{se tiene } \; n = 83.62 \cong n = 84 \quad n_1 = 14 \quad n_2 = 70$$

Sensibilidad del tamaño de muestra con α y 1-β

TABLE 1. Necessary sample size as a function of power and α level, relative to the sample size required for a study with an α level of 0.05 and 80 percent power*

α	Power			
	70%	80%	90%	95%
0.05	0.8	1.0	1.3	1.7
0.01	1.2	1.5	1.9	2.3
0.001	1.8	2.2	2.7	3.1

* To read the table, choose a power and an α level. Suppose one is interested in a trial with 90 percent power and an α level of 0.01. The entry of 1.9 in the table means that such a trial would require 1.9 times the sample size required for a trial with 80 percent power and an α level of 0.05.

.........::::::::: TEST PARAMÉTRICOS :::::::::..........

	DATOS INDEPENDIENTES	DATOS PAREADOS
2 MEDIAS	Varianzas conocidas: **Normal** Varianzas desconocidas: Homogéneas: **t de Student** No homogéneas: **t de Student + Welch**	**t de Student datos pareados**
2 VARIANZAS	**F-Snedecor**	**t de Student datos pareados**
2 PROPORCIONES	**Normal** (muestras grandes>100) **Test exacto Fisher** (muestras pequeñas<100)	**Test de Mc-Nemar**
+ 2 PROPORCIONES	χ^2	**Test de Cochran**
+ 2 MUESTRAS	**ANOVA** (distribución F de Snedecor)	

NORMALIDAD: D´Agostino, Kolmogorov, Shapiro-Wilks
ALEATORIEDAD: Rachas, Haldane-Hart
HOMOGENEIDAD VARIANZAS (+ de 2): Barlett (χ^2) , Levene (F-Snedecor)

TEST A POSTERIORI (de comparaciones múltiples)
Bonferroni, Tukey, Scheffé, Newman-Keuls, Dunnett

.........::::::::: TEST NO PARAMÉTRICOS :::::::::.........

	DATOS INDEPENDIENTES	DATOS PAREADOS
2 MEDIAS	**U de Mann-Whitney**	**Wilcoxon** (test signos)
+ 2 MEDIAS	**Kruskal-Wallis** (χ^2)	**Friedman** (χ^2)

Variable respuesta		Variable explicativa	Supuestos	Muestreo	Técnica
Cualitativa	Binaria (dicotómica)	Binaria	Se deben verificar las condiciones de validez de los métodos no exactos (χ^2, Mcnemar)	1 sola muestra	Pruebas de asociación en tablas 2x2
				2 muestras independientes	Tests exactos (Fisher, Barnard) y aprox. χ^2
				2 muestras apareadas	Test exacto de Fisher
					Test de McNemar
		Policotómica/ordinal (K categorías)	Condiciones de validez test χ^2	1 muestra	Tablas de contingencia 2xK
				k muestras independientes	
		Cuantitativa		1 muestra	Regresión logística
	Policotómica (R categorías)	Policotómica/ordinal (C categorías)	Condiciones de validez test χ^2	1 o K muestras independientes	Tablas de contingencia RxC
		Cuantitativa		1 muestra	Regresión logística policotómica
	Ordinal (R cats.)	Cualitativa (C categorías)	Condiciones de validez test χ^2		Tablas de contingencia RxC
		Ordinal (C categorías)		1, R o C muestras	Tabla de contingencia
					Métodos no paramétricos basados en la asignación de valores numéricos a las clases
		Cuantitativa		1 muestra	Regresión logística ordinal
Cuantitativa	Discreta*, Continua	Binaria	Respuesta con distribución normal	2 muestras independientes	Test de Student/Welch de comparación de medias
				2 muestras relacionadas	Test de Student de comparación de medias
			Respuesta con libre distribución		
			$n \leq 60$	2 muestras independientes	Test de Mann-Whitney / Wilcoxon
				2 muestras relacionadas	Test de Wilcoxon
			$n > 60$	2 muestras independientes	Test de Student/Welch
				2 muestras relacionadas	Test de Student de comparación de medias
		Policotómica (K categorías)	Respuesta con distribución normal	K muestras independientes	ANOVA de 1 factor
				K muestras relacionadas	ANOVA unifactorial con medidas repetidas
			Respuesta con libre distribución		
			$n \leq 60$	K muestras independientes	Test de Kruskal-Wallis
				K muestras relacionadas	Test de Friedman
			$n > 60$	K muestras independientes	ANOVA de 1 factor
				K muestras relacionadas	ANOVA unifactorial con medidas repetidas
		Ordinal	Asignación de valores numéricos a las categorías		Correlación (Spearman)
		Cuantitativa	Relación lineal	Muestreo I o II	Regresión lineal simple
			resp. Normal	Muestreo I o II	+ Correlación de Pearson
			resp. No normal	Muestreo I o II	+ Correlación de Spearman
			No linealidad	Muestreo I o II	Regresión no lineal
			Relac. Monotónica	Muestreo I o II	+ Correlación de Spearman

* Variable respuesta discreta:

- con distribución conocida (por ejemplo Poisson) → Métodos exactos
- sin distribución conocida
  - $n \leq 60$ → Métodos no paramétricos
  - $n > 60$ → Métodos basados en la distribución normal (son aproximados; requieren cpc)

Test paramétricos vs no paramétricos

- ***Seleccionar** un test no paramétrico si:*
  - Si el objetivo del estudio es una variables cualitativa
  - Si la variable es un rango y claramente no hay distribución normal (escalas de dolor, escala Likert, índice de Apgar, etc...)
  - Presencia de valores *outliers*
  - Hay completa certeza de que la variable no es normal (si bien se puede intentar alguna transformación de los datos)

- **Tamaño de muestra**
  - Para tamaños de muestra grandes, los test no paramétricos son ligeramente menos potentes que los paramétricos.
  - Tamaños pequeños de muestra (difícil detectar la normalidad):
    - ¿Qué sucede al aplicar una prueba paramétrica sin normalidad? No se puede confiar en el TLC, **p-valores pueden ser inexactos**
    - ¿Qué sucede cuando se utiliza una prueba no paramétrica con normalidad? Los **p-valores que tienden a ser demasiado altos**

Test paramétricos vs no paramétricos

Conclusión:

- Si se tiene una variable cuantitativa y se tiende a la normalidad, con homocedasticidad entre grupos, y *n* suficiente ($n \geq 30$), se deben utilizar las pruebas estadísticas paramétricas.
- Usar tests no paramétricos en caso de que no se cumplan los requisitos anteriores; sobre todo cuando la **normalidad y homocedasticidad** de la variable en estudio esté en duda, y además se tenga $n < 30$

Razonamiento:

- Si $n < 30$ en una/varias muestras, la potencia estadística de las pruebas paramétricas y no paramétricas es similar.
- Si *n* aumenta las muestras **las pruebas paramétricas aumentan su potencia respecto a las no paramétricas**.
- Por ello, las pruebas no paramétricas están indicadas cuando $n < 30$, o bien cuando hay una muestra mayor, pero no se cumplen los requisitos de aplicabilidad de las pruebas paramétricas.

ANÁLISIS DE VARIABLES RELACIONADAS CON LA DOCENCIA E INVESTIGACIÓN EN LA FAC. FARMACIA

Estimación y contraste de hipótesis

Francisco M. Ocaña Peinado @ocanapaco
http://www.ugr.es/local/fmocan
Departamento de Estadística e Investigación Operativa. UGR