eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea
The University of the Basque Country

E.U.I.T.I. Bilbao

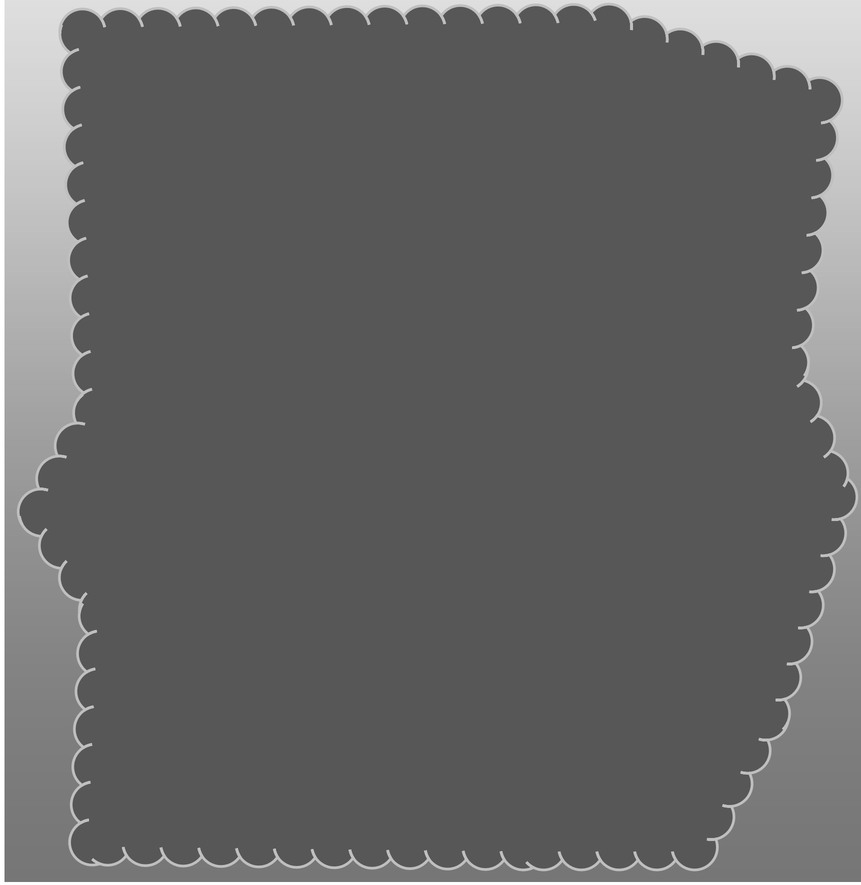
Asignatura: MÉTODOS ESTADÍSTICOS DE LA INGENIERÍA

E.U.I.T.I. Bilbao

Asignatura:
MÉTODOS ESTADÍSTICOS DE LA INGENIERÍA

TEMA 7: CONTRASTES DE HIPÓTESIS

Todos aprendemos de la experiencia y la lección, en esta ocasión, es que nunca se debe perder de vista la alternativa



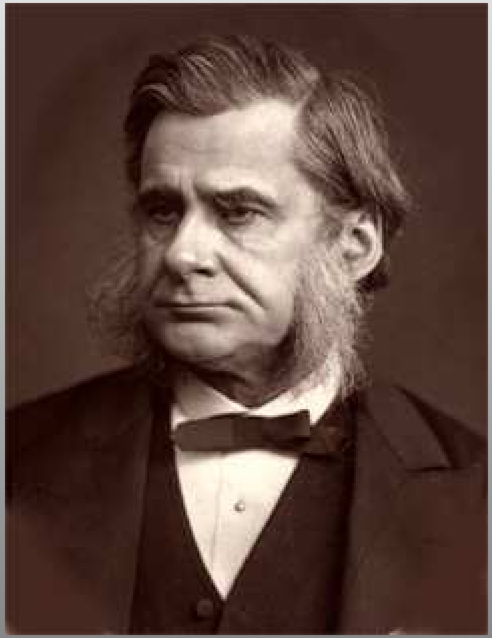
Sherlock Holmes

Las aventuras de Black Peter

Sir Arthur Conan Doyle

Actor: Basil Rathbone

La gran tragedia de la ciencia: la destrucción de una bella hipótesis por un antiestético conjunto de datos



http://es.wikipedia.org/wiki/Thomas_Henry_Huxley

Thomas Henry Huxley

Biólogo inglés

(1825-1895)

1. RESUMEN

Se explica qué se entiende por contraste de hipótesis estadística y se aprende a realizar contrastes de este tipo, a partir de datos muestrales referidos a algún parámetro poblacional desconocido

Palabras clave:

- contraste de hipótesis
- errores tipo I y tipo II
- hipótesis nula e hipótesis alternativa
- estadístico de contraste y *p*-valor
- nivel de confianza y nivel de significación

2. ÍNDICE DEL TEMA

2. ÍNDICE DEL TEMA

3. INTRODUCCIÓN

- como ya se indicó, los métodos más utilizados para realizar inferencias sobre el valor de parámetros de una población son:
  - **estimación** de dichos parámetros: puntual o con intervalos de confianza
  - **contraste de hipótesis**: tomando una decisión sobre un valor hipotético del parámetro
- en ambos casos se usa la información proporcionada por una muestra significativa de la población

Hipótesis estadística: sentencia sobre la naturaleza de una población

- generalmente, se formula en términos de un cierto parámetro de la población

3. INTRODUCCIÓN

<u>Contraste de hipótesis</u>: es una técnica de Estadística inferencial que permite comprobar si la información que proporciona una muestra observada concuerda (o no) con la hipótesis estadística formulada

- por lo general, las hipótesis establecen que un cierto parámetro poblacional tiene un valor que cae dentro de una determinada región
- se trata de decidir si la hipótesis es consistente con los datos observados en la muestra
- intervalos de confianza y contrastes de hipótesis se relacionan entre sí y pueden servir para decidir sobre los parámetros poblacionales

3. INTRODUCCIÓN

Contraste de hipótesis

- es obvio que para decidir si una hipótesis estadística es cierta ó no se debería estudiar toda la población
- como puede ser imposible ó poco práctico se toma una muestra aleatoria representativa de la población para decidir si se acepta o no la hipótesis estadística planteada
- si la información muestral es consistente con la hipótesis estadística entonces se acepta dicha hipótesis; en caso contrario se rechaza (no se indica si es falsa ó verdadera sino si se acepta o se rechaza)

4. CONCEPTOS BÁSICOS

Hipótesis nula: sentencia acerca de un parámetro de la población

- notación: H_0
- es la hipótesis que se somete a contraste
- las observaciones son resultado de la casualidad

Hipótesis alternativa: hipótesis contraria a la nula

- notación: $H_a = H_1$
- hay un efecto real, es decir, las observaciones son el resultado de ese efecto real

4. CONCEPTOS BÁSICOS

Estadístico del contraste: estadístico cuyo valor se determina a partir de los datos de la muestra

► dependiendo del valor que tome, la hipótesis nula es rechazada ó no

Región crítica (ó región de rechazo): es el conjunto de valores del estadístico de contraste para los que se rechaza la hipótesis nula

► el contraste estadístico de H_0 queda completamente especificado al determinar el estadístico de contraste (*EC*) y la región crítica (*RC*)

$$\begin{cases} EC \in RC \;\Rightarrow\; \text{rechazar } H_0 \\ EC \notin RC \;\Rightarrow\; \text{no rechazar } H_0 \end{cases}$$

4. CONCEPTOS BÁSICOS

Hipótesis nula vs. hipótesis alternativa

- hipótesis nula: H_0
  - es la hipótesis que se contrasta
  - los datos pueden refutarla, nunca probarla
  - no debería ser rechazada sin una buena razón
  - el rechazo de H_0 es una sentencia fuerte en el sentido de que la hipótesis no parece ser consistente con los datos observados
  - no rechazar la hipótesis nula es una sentencia débil que se debería interpretar en el sentido de que H_0 es consistente con los datos
  - si se intenta desacreditar una hipótesis ésta deberá ser la nula

4. CONCEPTOS BÁSICOS

Hipótesis nula vs. hipótesis alternativa

- hipótesis alternativa: H_a
  - es la hipótesis complementaria de la nula
  - los datos pueden mostrar evidencias a favor
  - no debería ser aceptada sin una gran evidencia a favor
  - si se intenta probar una cierta hipótesis deberá designarse como alternativa

4. CONCEPTOS BÁSICOS

Hipótesis nula vs. hipótesis alternativa

- hipótesis nula: H_0
  - suele expresarse como una igualdad del tipo, $\boxed{H_0 : \theta = \theta_0}$
  - θ es un parámetro de una población
  - θ_0 es un valor hipotético para ese parámetro
- hipótesis alternativa: H_a
  - $\boxed{H_a : \theta \neq \theta_0}$: contraste bilateral ó de dos colas ó de dos extremos
  - $\boxed{H_a : \theta > \theta_0}$: contraste unilateral a la derecha ó de una cola a la derecha ó de un extremo a la derecha
  - $\boxed{H_a : \theta < \theta_0}$: contraste unilateral a la izquierda ó de una cola a la izquierda ó de un extremo a la izquierda

4. CONCEPTOS BÁSICOS

Hipótesis nula vs. hipótesis alternativa

- hipótesis nula: H_0
  - también, es frecuente expresar H_0 como negación exacta de H_a en cuyo caso sí puede ser una desigualdad no estricta
  - el nombre de nula proviene de que H_0 representa la hipótesis que se mantiene a no ser que los datos muestrales indiquen su falsedad; debe entenderse en el sentido de "neutra"

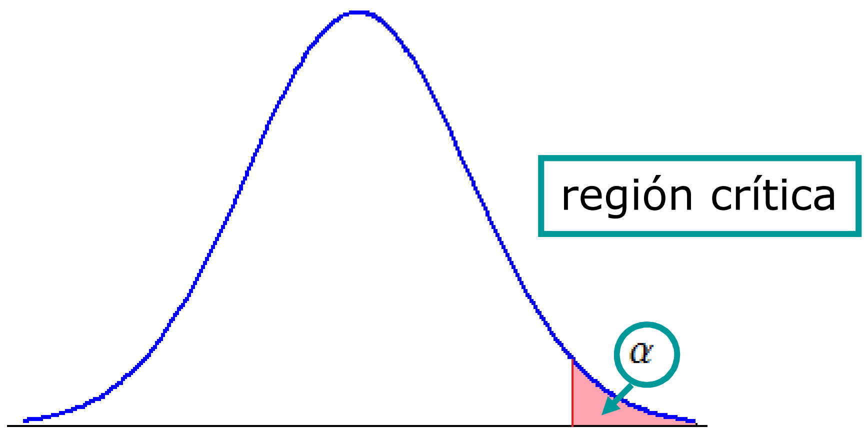
- Nota 1. Ambas hipótesis deben formularse antes de recoger y analizar la muestra que se va a utilizar para realizar el contraste de hipótesis

- Nota 2. La decisión de realizar un tipo de contraste unilateral o bilateral depende de la información que se posee teniendo en cuenta que la hipótesis que se cree cierta es H_a

4. CONCEPTOS BÁSICOS

<u>Nivel de significación (α)</u>: valor que se fija para el contraste de una hipótesis nula

- procedimiento clásico para contrastar una hipótesis nula:
  - primero, fijar un nivel de significación α
  - si H_0 es cierta, obligar a que la probabilidad de rechazar H_0 sea menor o igual que α



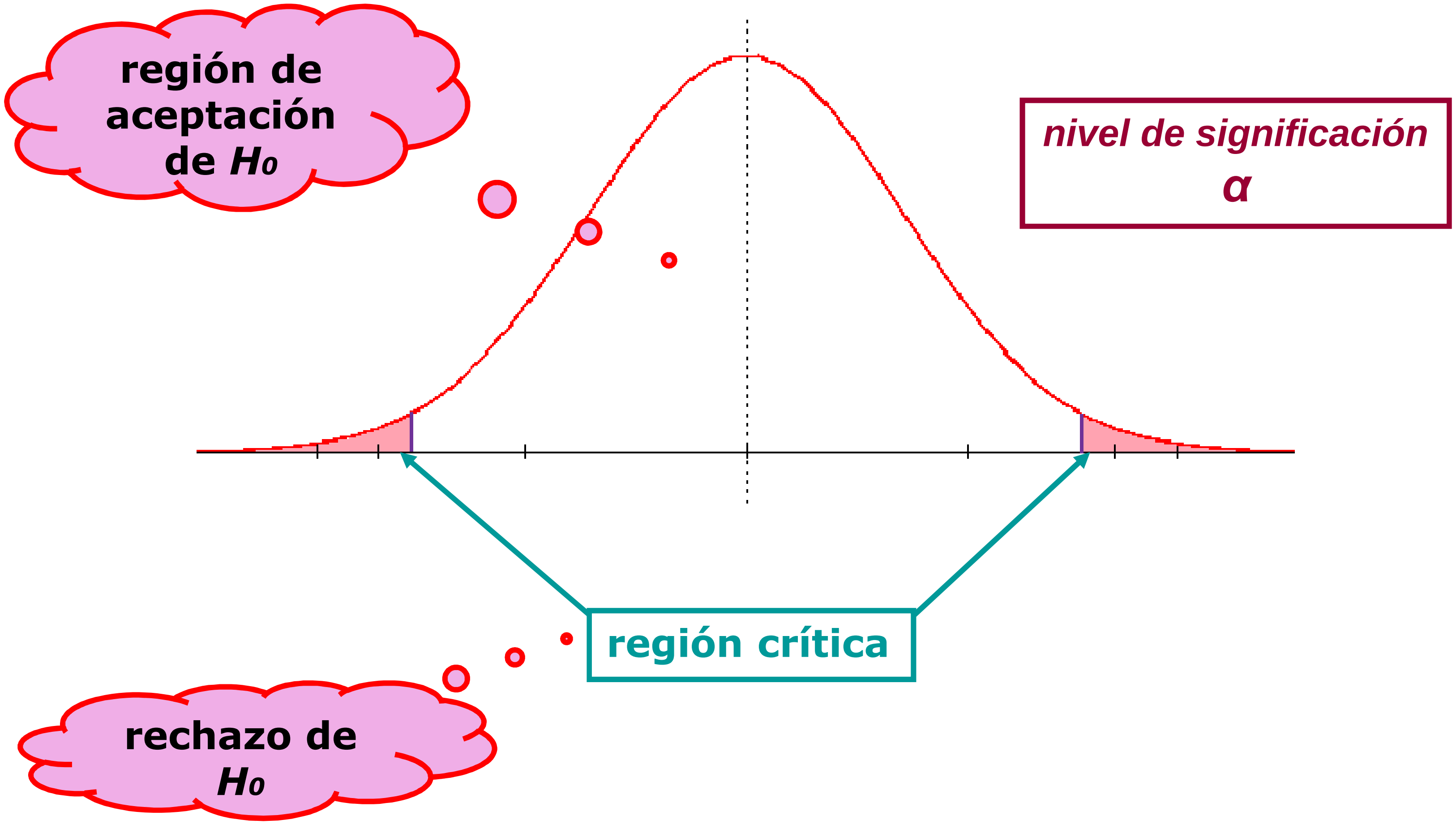
4. CONCEPTOS BÁSICOS

Región crítica y nivel de significación

- región crítica, *RC*
  - conocida antes de realizar el experimento
  - valores experimentales que refutarían H_0 (valores improbables del estimador según las condiciones dadas)
  - su amplitud depende del nivel de significación
  - su posición depende de H_a
- nivel de significación, α
  - valor pequeño: 1%, 5%
  - fijado de antemano por el investigador
  - es la probabilidad de rechazar H_0 cuando es cierta

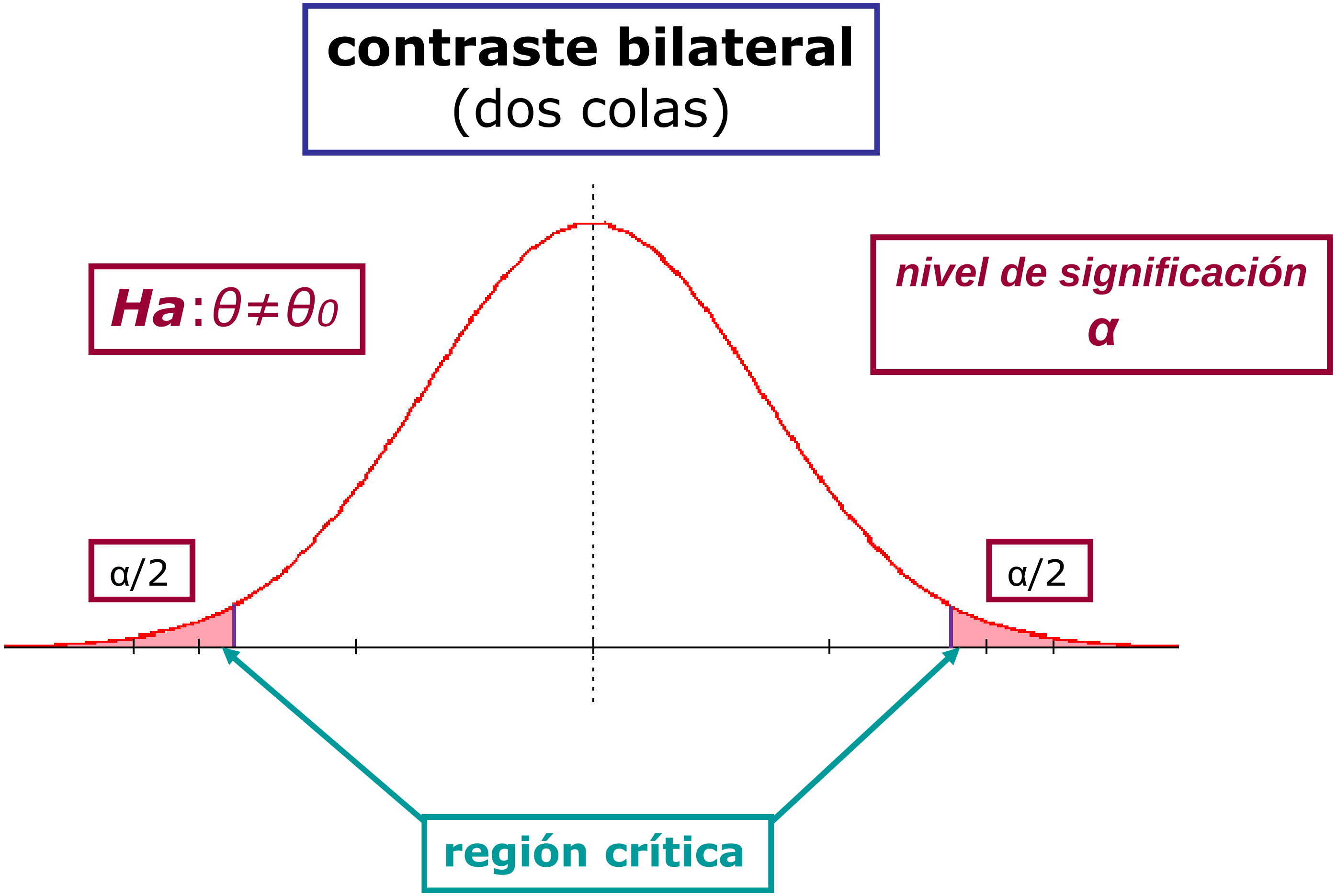
4. CONCEPTOS BÁSICOS

Región crítica y nivel de significación



4. CONCEPTOS BÁSICOS

Región crítica y nivel de significación



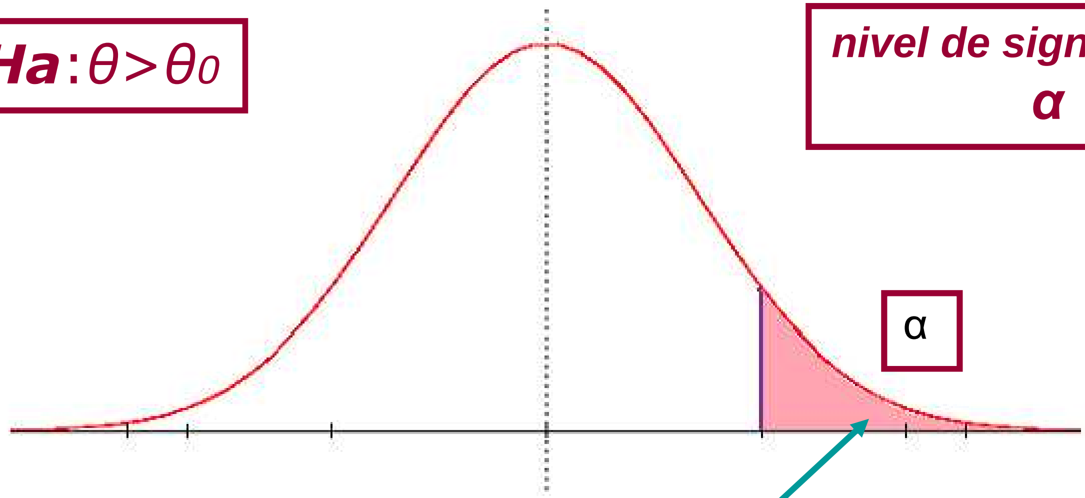
4. CONCEPTOS BÁSICOS

Región crítica y nivel de significación

contraste unilateral a la derecha
(una cola a la derecha)

Ha: $\theta > \theta_0$

nivel de significación
α



región crítica

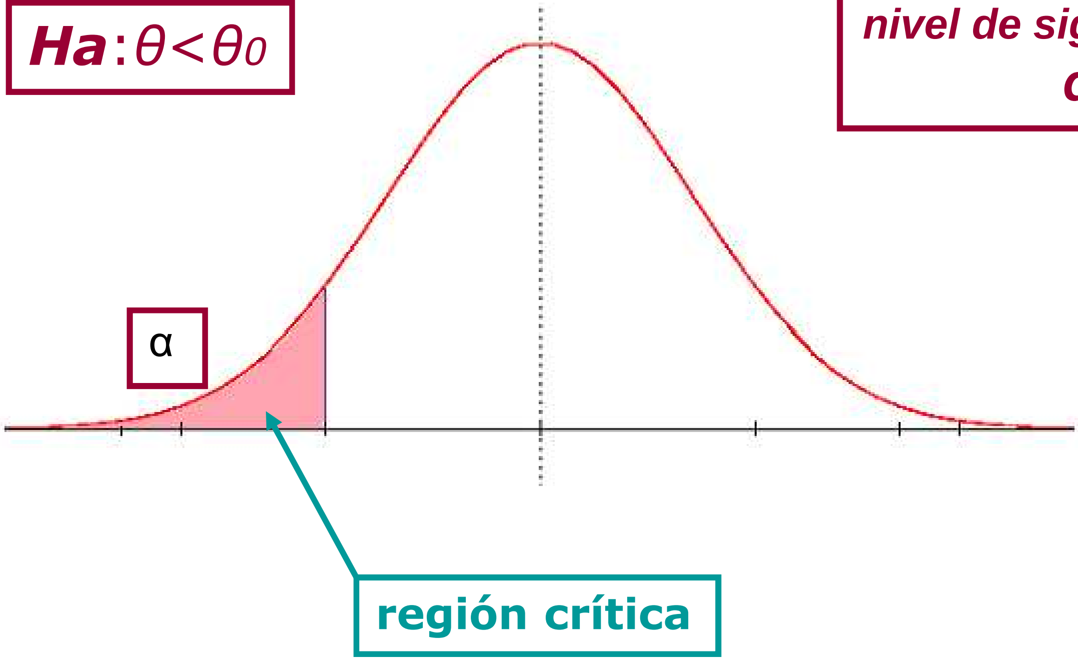
4. CONCEPTOS BÁSICOS

Región crítica y nivel de significación

contraste unilateral a la izquierda
(una cola a la izquierda)

Ha: $\theta < \theta_0$

nivel de significación
α



4. CONCEPTOS BÁSICOS

Decisiones en un contraste de hipótesis

- tipos de decisiones que proporciona un contraste:
  - si el valor del estadístico de contraste, para la muestra dada, cae en la región de rechazo se puede afirmar con **cierto nivel de confianza** que los datos de la muestra permiten rechazar la hipótesis nula a favor de la alternativa
  - si el valor del estadístico de contraste, para la muestra dada, no cae en la región de rechazo no se puede afirmar con el **nivel de confianza exigido** que los datos de la muestra permiten rechazar la hipótesis nula a favor de la alternativa
- Nota. La hipótesis nula carece de confianza, se asume como un punto de partida pero se abandona cuando los datos empíricos muestran evidencias claras en su contra y a favor de la hipótesis alternativa

4. CONCEPTOS BÁSICOS

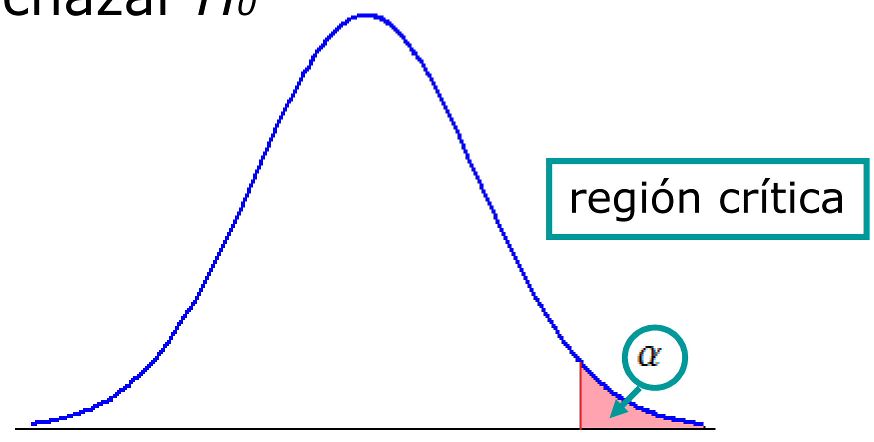
Ejemplo 1. El estadístico sir Ronald A. Fisher explicó el procedimiento de un contraste de hipótesis usando una anécdota sobre una señora y el té.
Una señora afirma que si prueba una taza de té con leche sabe distinguir si se ha añadido primero el té ó la leche.
El experimento consiste en preparar ocho tazas de té con leche (cuatro de un tipo y cuatro de otro) y darlas a probar a la señora de forma aleatoria.

4. CONCEPTOS BÁSICOS

- establecer hipótesis: se asume que la señora no sabe distinguir, realmente, si la leche se añade antes ó después del té, es decir, decide de forma aleatoria
  - ***H_0***: la señora no sabe distinguir realmente si se ha añadido primero la leche ó el té
  - ***H_a***: la señora sabe distinguir realmente si se ha añadido primero la leche ó el té
- toma de una muestra:
  - la señora prueba ocho tazas de té con leche
  - se contabiliza el número de veces que acierta
- nivel de significancia:
  - si la señora distingue correctamente, al menos, siete de las 8 tazas de té con leche se rechaza la hipótesis nula

4. CONCEPTOS BÁSICOS

- estadístico del contraste: si la señora decide de forma aleatoria
  - ***X***: "número de aciertos"
  - $X \sim B(8;0.5)$
- análisis de datos:
  - región de rechazo (ó crítica): $X \geq 7$
  - región de aceptación: conjunto de valores del estadístico del contraste que implican no rechazar H_0



4. CONCEPTOS BÁSICOS

Ejemplo 2. Supóngase que un investigador de la EPA (Agencia de Protección Ambiental) quiere determinar si el nivel medio, μ, de cierto tipo de contaminante liberado a la atmósfera por una compañía química cumple con las pautas de la EPA.

Si el límite superior establecido por la EPA es de tres partes de millón el investigador querrá utilizar una muestra (mediciones diarias de la contaminación) para decidir si la compañía está violando la ley ó no, es decir, decidir si $\mu > 3$

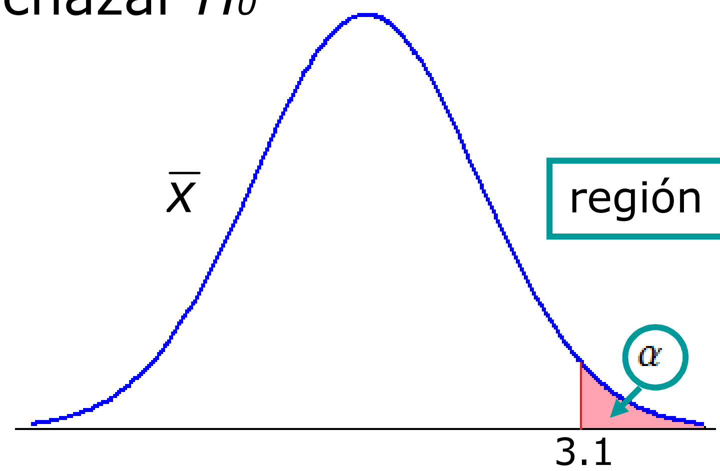
4. CONCEPTOS BÁSICOS

- establecer hipótesis: se asume que la compañía no viola la ley
  - H_0: $\mu=3$
  - H_a: $\mu>3$
- toma de una muestra :
  - la decisión se basará en una muestra de tamaño $n=30$ lecturas de contaminación diarias
  - se calcula la media muestral del contaminante vertido
- nivel de significancia:
  - si la EPA observa, en una muestra de tamaño $n=30$, que $\bar{x} \geq 3.1$ entonces se rechaza la hipótesis nula

4. CONCEPTOS BÁSICOS

- estadístico del contraste: la distribución muestral de la media muestral

- análisis de datos:
  - región de rechazo (ó crítica): $\bar{x} \geq 3.1$
  - región de aceptación: conjunto de valores del estadístico del contraste que implican no rechazar H_0



4. CONCEPTOS BÁSICOS

Ejemplo 3. Similitud con un juicio penal.

- establecer hipótesis: se asume la inocencia del acusado
  - H_0: el acusado es no culpable
  - H_a: el acusado es culpable
- toma de una muestra y análisis de datos:
  - se recogen pruebas o evidencias: huellas digitales, testimonios de testigos, muestras de ADN, etc.
- nivel de significancia:
  - si hay evidencias, "más allá de toda duda razonable", se rechaza la hipótesis nula y se declara culpable al acusado
  - en caso contrario, se acepta la hipótesis nula (no culpable)

5. ETAPAS DE UN CONTRASTE

1. suposiciones elementales:
   - establecer el parámetro poblacional que se va a considerar
   - concretar el modelo de probabilidad que se va a usar en la descripción
2. formular una hipótesis nula, H_0, que se desea probar (contrastar)
3. formular una hipótesis alternativa, H_a, contra la cual se va a probar la hipótesis nula (generalmente, la complementaria)
4. escoger el estadístico para la prueba
5. elegir un nivel de significación, α, apropiado

5. ETAPAS DE UN CONTRASTE

6. determinar la distribución de probabilidad de la prueba estadística

7. calcular el valor del estadístico a partir de una muestra aleatoria de datos (proceso de estimación)

8. establecer la región crítica, *RC*, para el contraste de hipótesis (zona de rechazo/aceptación)

9. aceptar ó rechazar la hipótesis nula comparando el valor calculado para el estadístico de la prueba con los valores que definen la *RC* (normalmente implica una estimación por intervalo)

6. ERRORES EN UN CONTRASTE

- el contraste de una hipótesis estadística implica una toma de decisión
  - a favor de H_0
  - en contra de H_0 (a favor de H_a)
- como el contraste de una hipótesis sólo puede tener dos resultados entonces, al tomar la decisión, sólo puede haber dos tipos de error
- las probabilidades de cometer estos dos tipos de error miden el riesgo de tomar una decisión incorrecta al realizar un contraste de hipótesis y, por tanto, permiten medir la bondad de dicho contraste

6. ERRORES EN UN CONTRASTE

Tipos de error

- **error tipo I** (ó **falso negativo**): se rechaza la hipó-tesis nula, H_0, siendo cierta
  - nivel de significación, α: probabilidad de cometer un error del tipo I (rechazar H_0 cuando es H_0 es cierta)
  - nivel de confianza, 1-α: probabilidad de aceptar la hipótesis nula, H_0, cuando es cierta
- **error tipo II** (ó **falso positivo**): se acepta la hipó-tesis nula, H_0, siendo falsa
  - β: probabilidad de cometer un error del tipo II
  - potencia, 1-β: probabilidad de rechazar la hipótesis nula, H_0, cuando es falsa

6. ERRORES EN UN CONTRASTE

Tipos de error

- por ejemplo, en un juicio:
  - error de tipo I: implica declarar culpable a un inocente
  - error de tipo II: implica declarar inocente a un culpable

- ¿cuál de los dos errores es más grave?
  - depende de cada contraste
  - en general, se pretende acotar el error de tipo I y tratar de minimizar el error de tipo II
  - entonces, se trata de elegir contrastes lo más potentes posible garantizando que la probabilidad del error de tipo I es inferior a un determinado nivel

6. ERRORES EN UN CONTRASTE

Tipos de error

- al realizar un contraste de hipótesis se pueden tomar dos decisiones correctas:
  - rechazar la hipótesis nula cuando es falsa
  - no rechazar la hipótesis nula cuando es verdadera

- y se pueden cometer dos tipos de error
  - rechazar la hipótesis nula cuando es verdadera
  - no rechazar la hipótesis nula cuando es falsa

6. ERRORES EN UN CONTRASTE

Tipos de error

- cuadro resumen

	H_0 cierta	H_0 falsa
rechazar H_0	error de tipo I (α)	decisión correcta
aceptar H_0	decisión correcta	error de tipo II (β)

6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 11

Un examinador plantea 10 cuestiones del tipo verdadero-falso. Para comprobar la hipótesis de que los estudiantes contestan al azar se adopta la siguiente regla de decisión:

- si, al menos, 7 respuestas son correctas el estudiante no ha contestado al azar

Calcular la probabilidad de cometer un error de tipo I

Solución

- error de tipo I: rechazar H_0 cuando es cierta
- hipótesis nula, H_0: el estudiante contesta al azar
- hipótesis alternativa, H_a: el estudiante NO contesta al azar

6. ERRORES EN UN CONTRASTE

<u>Ejemplo</u>: ejercicio 11

Solución

- *X*: "número de respuestas correctas (si H_0 es cierta)"
- $X \sim B(10;0.5)$
- si $X \geq 7$: se asume que el estudiante NO contesta al azar y, por tanto, se rechaza la hipótesis nula, H_0

$$P\left(X \geq 7 \,|\, X \approx B\left(10;0.5\right)\right) = 1 - P\left(X \leq 6 \,|\, X \approx B\left(10;0.5\right)\right) = \boxed{0.171875}$$

=1-DISTR.BINOM(6;10;0,5;1)

nivel de significación
α

6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 11

Solución

- X: "número de respuestas correctas (si H_0 es cierta)"
- $X \sim B(10;0.5)$ → $X' \approx N\left(5; \sqrt{\frac{5}{2}}\right)$

corrección por continuidad

$$P\left(X \geq 7 \,|\, X \approx B(10; 0.5) \right) = 1 - P\left(X \leq 6 \,|\, X \approx B(10; 0.5) \right) =$$

$$= 1 - P\left(X' \leq 6.5 \,|\, X' \approx N\left(5; \sqrt{2.5}\right) \right) =$$

tipificación

$$= 1 - P\left(Z \leq \frac{6.5-5}{\sqrt{2.5}} \,|\, Z \approx N(0;1) \right) = \boxed{0.17139}$$

=1-DISTR.NORM.ESTAND(1,5/RAIZ(5/2))

6. ERRORES EN UN CONTRASTE

<u>Ejemplo</u>: ejercicio 1

Para contrastar la hipótesis de que una moneda es buena (probabilidad de cara p=0.5) se adopta la siguiente regla de decisión:

- aceptarla si el número de caras obtenidas en una muestra de 100 tiradas está entre 40 y 60, ambas inclusive
- rechazarla en caso contrario

1. Hallar la probabilidad de rechazar la hipótesis p=0.5 cuando sea correcta
2. Representar gráficamente la regla de decisión y el resultado del apartado anterior, es decir, α
3. Probabilidad de cometer un error del tipo II, es decir, de aceptar la hipótesis de que la moneda es buena cuando, realmente, la probabilidad de cara es p=0.7
4. Probabilidad de obtener al menos 55 caras en 100 tiradas si p=0.5

6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 1

Apartado 1: solución

- hipótesis: $\begin{cases} H_0 : p = 0.5 \\ H_a : p \neq 0.5 \end{cases}$

- estadístico del contraste: si la moneda es buena (H_0 cierta)

X: "número de caras obtenidas en 100 lanzamientos"

$X \sim B(100;0.5) \longrightarrow X' \sim N(50;5)$

$$\begin{cases} n \cdot p = 50 \\ \sqrt{n \cdot p \cdot q} = 5 \end{cases}$$

6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 1

Apartado 1: solución

- nivel de significación, α:

$$\boxed{\alpha = P\left(error\ tipo\ I\right) = P\left(rechazar\ H_0 \mid H_0\ cierta\right)}$$

$$\alpha = P\left(X \notin [40, 60] \mid X \approx B(100; 0.5)\right)$$

$$\alpha = 1 - P\left(40 \leq X \leq 60 \mid X \approx B(100; 0.5)\right)$$

corrección por continuidad

$$\alpha = 1 - P\left(39.5 \leq X' \leq 60.5 \mid X' \approx N(50; 5)\right)$$

tipificación

$$\boxed{\alpha = 1 - P\left(-2.1 \leq Z \leq 2.1 \mid Z \approx N(0; 1)\right) = 0.03572}$$

6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 1

Apartado 1: solución

- nivel de significación, α: $\boxed{\alpha = 1 - P\left(-2.1 \le Z \le 2.1\right)}$

$$\alpha = 1 - \left[\Phi(2.1) - \Phi(-2.1)\right] = 1 - \left(2\Phi(2.1) - 1\right) = 2 - 2\Phi(2.1)$$

Excel: =2-2*DISTR.NORM.ESTAND(2,1)=0.03572844

◇ sin realizar la aproximación mediante la normal

Excel: =1−(DISTR.BINOM(60;100;0,5;1)−
−DISTR.BINOM(40;100;0,5;1))=0.04604407

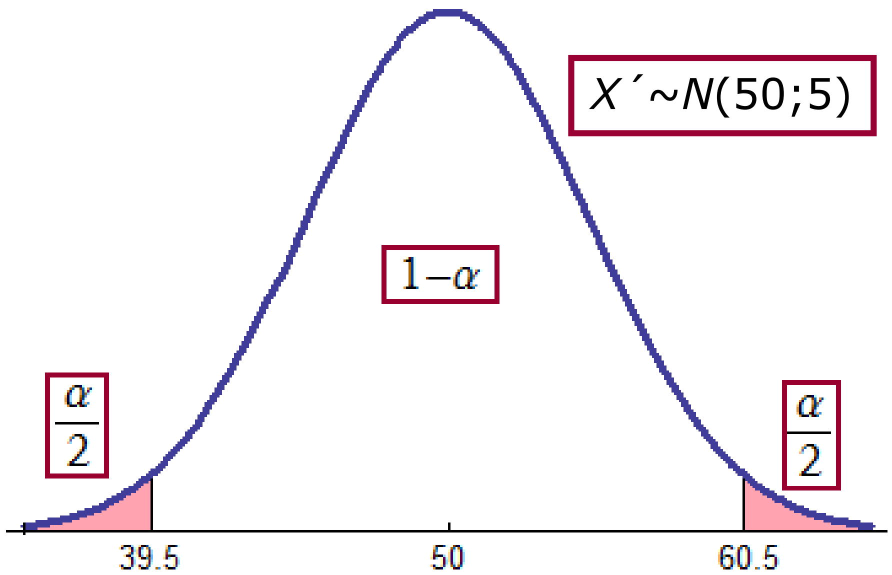
6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 1

Apartado 2: solución

X: "número de caras obtenidas en 100 lanzamientos"

$X \sim B(100;0.5) \longrightarrow X' \sim N(50;5)$



6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 1

Apartado 3: solución

X: "número de caras obtenidas en 100 lanzamientos (si p=0.7)"

$X \sim B(100;0.7) \longrightarrow X' \sim N(70;4.5826)$

$$\begin{cases} n \cdot p = 70 \\ \sqrt{n \cdot p \cdot q} = \sqrt{21} \end{cases}$$

$$\beta = P\left(error\ tipo\ II\right) = P\left(aceptar\ H_0 \mid H_0\ falsa\left(p = 0.7\right)\right)$$

$$\beta = P\left(X \in [40, 60] \mid X \approx B\left(100; 0.7\right)\right)$$

6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 1

Apartado 3: solución

$$\beta = P\left(X \in [40, 60] \mid X \approx B(100; 0.7)\right)$$

◇ sin realizar la aproximación mediante la normal

Excel: =DISTR.BINOM(60;100;0,7;1)−
DISTR.BINOM(40;100;0,7;1)=0.020988575

◇ aproximación mediante la normal

$$\beta = P\left(X \in [40, 60] \mid X \approx B(100; 0.7)\right) = P\left(40 \leq X \leq 60 \mid X \approx B(100; 0.7)\right)$$

$$\beta = P\left(39.5 \leq X' \leq 60.5 \mid X' \approx N\left(70; \sqrt{21}\right)\right)$$

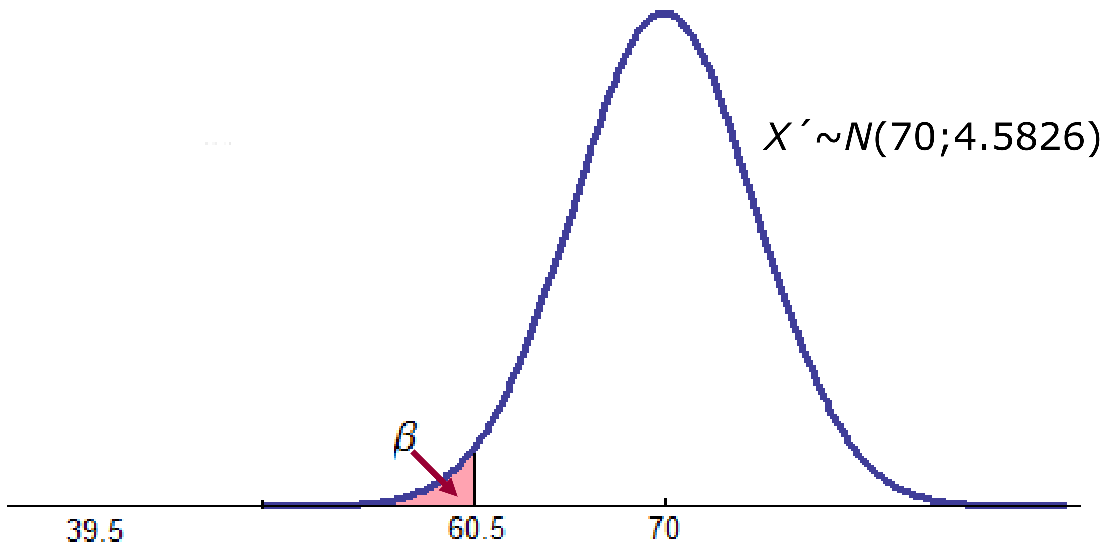
$$\beta = P\left(-6.66 \leq Z \leq -2.07\right) = 0.01922617$$

Excel: =DISTR.NORM.ESTAND(-2,07)−
−DISTR.NORM.ESTAND(-6,66) =0.01922617

6. ERRORES EN UN CONTRASTE

<u>Ejemplo</u>: ejercicio 1

<u>Apartado 3</u>: solución



- <u>potencia del contraste</u>: 1-β=0.98077

6. ERRORES EN UN CONTRASTE

Ejemplo: ejercicio 1

Apartado 4: solución

$$P\left(X \geq 55 \mid X \approx B(100; 0.5)\right)$$

◇ sin realizar la aproximación mediante la normal

Excel: =1−DISTR.BINOM(54;100;0,5;1)=0.18410081

◇ aproximación mediante la normal

$$P\left(X \geq 55 \mid X \approx B(100; 0.5)\right) = P\left(X' > 54.5 \mid X' \approx N(50; 5)\right) = P\left(Z > 0.9\right)$$

$$P\left(Z > 0.9\right) = 1 - \Phi(0.9) = 0{,}18406013$$

Excel: =1−DISTR.NORM.ESTAND(0,9) =0.18406013

7. *p-valor* DE UN CONTRASTE

- históricamente, la forma más común de realizar un contraste de hipótesis consiste en elegir un nivel bajo de significación que determina un límite para el error de tipo I que se está dispuesto a asumir

- el nivel de significación determina toda la región de rechazo

- examinando si el valor del estadístico del contraste cae en la región de rechazo se concluye si se acepta ó se rechaza la hipótesis nula a favor de la alternativa con el nivel de confianza requerido

7. *p-valor* DE UN CONTRASTE

- sin embargo, debido a la generalización del uso de los ordenadores, surge otra forma de actuar ante un contraste de hipótesis

- se calcula el valor del estadístico del contraste y se valora lo extremo que resulta bajo la distribución en el muestreo de la hipótesis nula

- si el valor es más extremo que el nivel de significación deseado se rechaza la hipótesis nula a favor de la alternativa

- la medida de cuán extremo es el valor del estadístico se llama ***p-valor***

7. *p-valor* DE UN CONTRASTE

- supóngase que se quiere contrastar una hipótesis estadística simple del tipo $H_0\!:\theta=\theta_0$ frente a alguna hipótesis alternativa ($H_a\!:\theta\neq\theta_0$, $H_a\!:\theta>\theta_0$ ó $H_a\!:\theta<\theta_0$)

> El ***p-valor*** ó **nivel crítico** asociado al contraste se define como el mínimo nivel de significación con el que la hipótesis nula será rechazada a favor de la alternativa.

- también, el *p-valor* es la probabilidad de obtener una discrepancia entre los datos muestrales y la hipótesis nula, H_0, mayor o igual que la observada en la muestra analizada cuando H_0 es cierta

7. *p-valor* DE UN CONTRASTE

- como, normalmente, se elige como nivel de significación máximo $\alpha=0.05$ la regla de decisión en un contraste, dado el *p-valor* sería:
  - si $p<0.05$, se rechaza H_0 con más de un 95% de confianza
  - si $p\geq 0.05$, no se puede rechazar H_0 a favor de H_a con, al menos, un 95% de confianza

- esta es la regla de decisión más habitual pero resulta simplista ya que no se proporciona el valor exacto del *p-valor*
  - no es lo mismo rechazar una hipótesis con, *al menos*, un 95% de confianza para diferentes valores del *p-valor*
  - debe proporcionarse siempre el *p-valor* de un contraste para que el interesado tome la decisión correspondiente

7. *p-valor* DE UN CONTRASTE

- en resumen, el *p-valor* permite utilizar cualquier otro nivel de significación ya que si se considera un cierto nivel α:
  - si $p < \alpha$, se rechaza H_0 con más de un $(1-\alpha)\%$ de confianza
  - si $p \geq \alpha$, no se puede rechazar H_0 en favor de H_a con, al menos, un $(1-\alpha)\%$ de confianza
- como conclusión, siempre que se realice un contraste de hipótesis debe facilitarse el *p-valor* asociado

7. *p-valor* DE UN CONTRASTE

- cuanto menor es el *p-valor* mayores son las evidencias para rechazar la hipótesis nula:

  - si $p < 0.01$, hay muchísimas evidencias para rechazar H_0
  - si $0.01 \leq p < 0.05$, hay fuertes evidencias para rechazar H_0
  - si $0.05 \leq p < 0.1$, hay pocas evidencias para rechazar H_0
  - si $p \geq 0.1$, no hay evidencias para rechazar H_0

7. *p-valor* DE UN CONTRASTE

Cálculo

- debe distinguirse entre contrastes unilaterales (una cola) y bilaterales (dos colas)
- **contrastes bilaterales**:
  - $H_0: \theta = \theta_0$ frente a $H_a: \theta \neq \theta_0$
  - dos colas: el rechazo de θ_0 puede producirse debido a que el estadístico del contraste toma valores muy altos o muy bajos
- **contrastes unilaterales**:
  - $H_0: \theta = \theta_0$
  - a la derecha: el rechazo de θ_0 se produce porque el estadístico del contraste toma valores muy altos $H_a: \theta > \theta_0$
  - a la izquierda: el rechazo de θ_0 se produce porque el estadístico del contraste toma valores muy bajos $H_a: \theta < \theta_0$

7. *p-valor* DE UN CONTRASTE

Cálculo

- supóngase que se quiere contrastar una hipótesis estadística simple del tipo $H_0:\theta=\theta_0$ frente a alguna hipótesis alternativa ($H_a:\theta\neq\theta_0$, $H_a:\theta>\theta_0$ ó $H_a:\theta<\theta_0$)
- supóngase, además, que el estadístico del contraste es S y que el valor que toma para la muestra es s
- con la definición dada de *p-valor* su cálculo es:
  - contraste unilateral a la izquierda ($H_a:\theta<\theta_0$): $\boxed{p=P\left(S\leq s\,|\,H_0\right)}$
  - contraste unilateral a la derecha ($H_a:\theta>\theta_0$): $\boxed{p=P\left(S>s\,|\,H_0\right)}$
  - contraste bilateral ($H_a:\theta\neq\theta_0$):

$$\boxed{p=2\times min\left\{P\left(S\leq s\,|\,H_0\right),P\left(S>s\,|\,H_0\right)\right\}}$$

DIFERENTES CONTRASTES DE HIPÓTESIS

8. INTRODUCCIÓN

- se van a enunciar diferentes contrastes de hipótesis para la media, la varianza ó la proporción de una población y para comparar las medias, las varianzas y las proporciones en dos poblaciones distintas
- no interesa cómo se deducen sino cómo se utilizan en la práctica
- aclaración:
  - cuando los datos proceden de una distribución normal es muy sencillo obtener la distribución del estadístico del contraste
  - si los datos no proceden de variables normales esta cuestión resulta mucho más difícil aunque, si el tamaño de la muestra es grande, el teorema central del límite garantiza que los parámetros obtenidos a partir de sumas basadas en las muestras siguen, aproximadamente, una distribución normal

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- población normal
- desviación típica conocida: σ

1. Formular las hipótesis nula y alternativa:

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

Contraste bilateral

2. Fijar un nivel de significación: $\boxed{\alpha}$

3. Calcular el estadístico del contraste. Bajo los supuestos adicionales y suponiendo cierta la hipótesis nula:

$$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \approx N(0;1)$$

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- población normal
- desviación típica conocida: σ

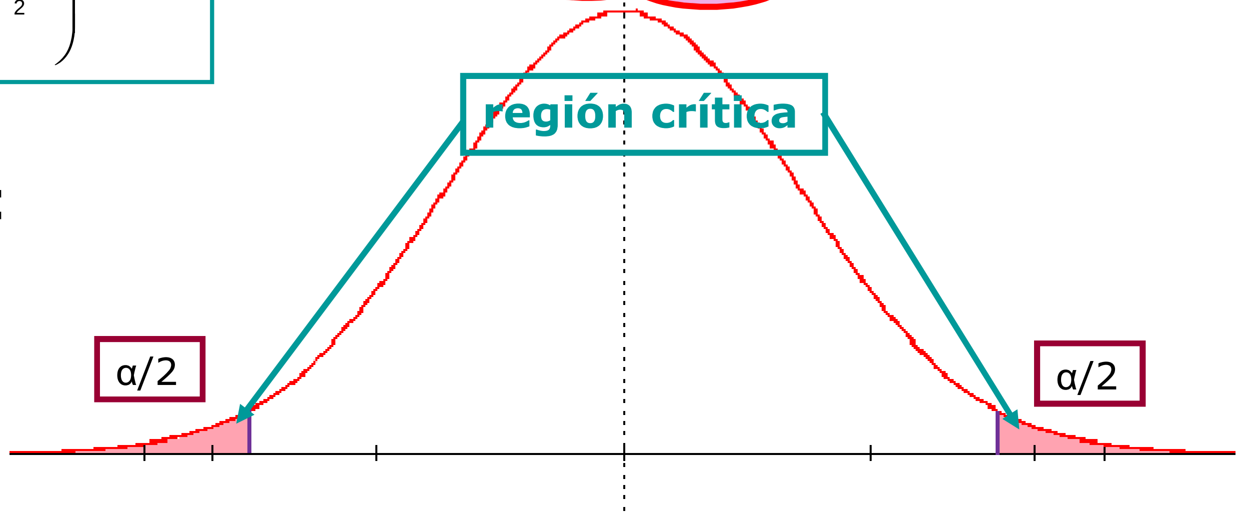
4. Hallar la región crítica. Se sabe que:

$$P\left(\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| > z_{\frac{\alpha}{2}}\right)=\alpha$$

$\alpha = P\,(error\ tipo\ I\,)$

La región crítica es:

$$\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| > z_{\frac{\alpha}{2}}$$



9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

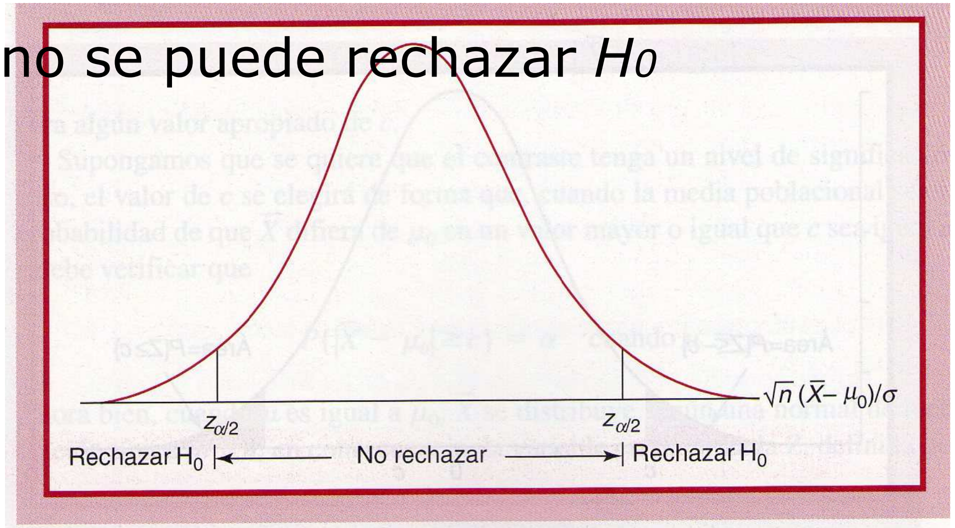
- población normal
- desviación típica conocida: σ

5. Se analiza el valor del estadístico del contraste:

- si cae en la región crítica se rechaza H_0
- si no cae en la región crítica no se puede rechazar H_0

6. Se establece la conclusión

Ross, M.S.; Introducción a la Estadística; Ed. Reverté S.A. (2005)



Contraste de H_0: $\mu = \mu_0$ frente a H_1: $\mu \neq \mu_0$.

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- población normal
- desviación típica conocida: σ

1. Formular las hipótesis nula y alternativa:

$$\boxed{\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu > \mu_0 \end{cases}}$$

Contraste unilateral

2. Fijar un nivel de significación: $\boxed{\alpha}$

3. Calcular el estadístico del contraste. Bajo los supuestos adicionales y suponiendo cierta la hipótesis nula:

$$\boxed{\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \approx N(0;1)}$$

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- población normal
- desviación típica conocida: σ

4. Hallar la región crítica. Se sabe que:

$$P\left(\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| > z_{\frac{\alpha}{2}}\right) = \alpha$$

$\alpha = P\,(error\ tipo\ I\,)$

Resulta razonable rechazar la hipótesis nula, H_0, en favor de la alternativa, H_a, cuando el valor de la media muestral sea mayor que μ_0; por tanto, la región crítica es:

$$\boxed{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} > z_{\alpha}}$$

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

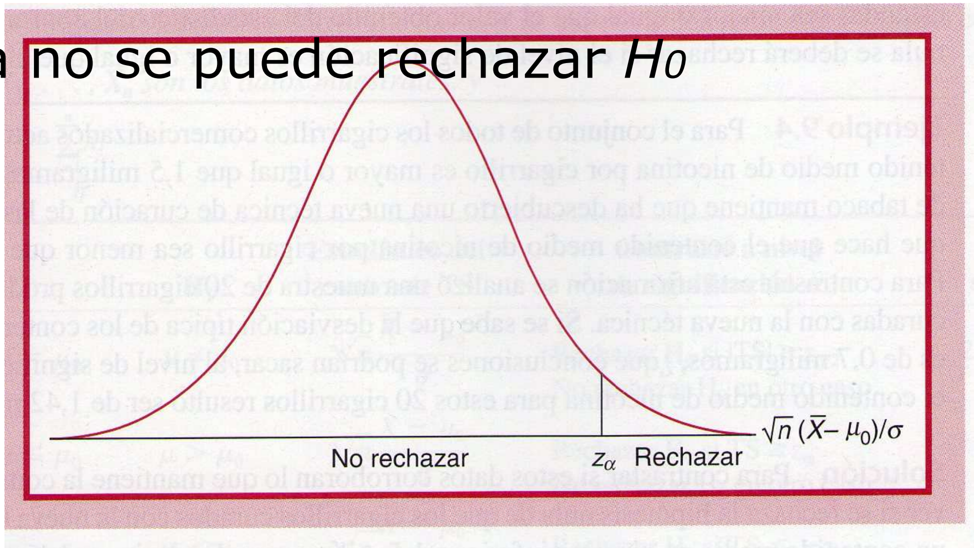
Supuestos adicionales:

- población normal
- desviación típica conocida: σ

5. Se analiza el valor del estadístico del contraste:

- si cae en la región crítica se rechaza H_0
- si no cae en la región crítica no se puede rechazar H_0

6. Se establece la conclusión



Contrastando H_0: $\mu \leq \mu_0$ frente a H_1: $\mu > \mu_0$.

Ross, M.S.; Introducción a la Estadística; Ed. Reverté S.A. (2005)

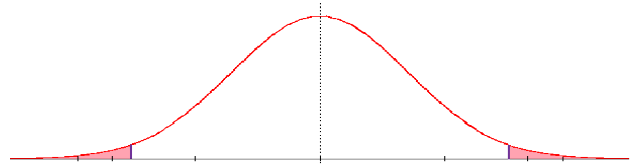
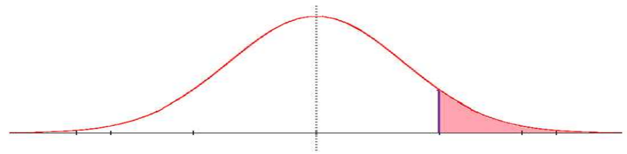
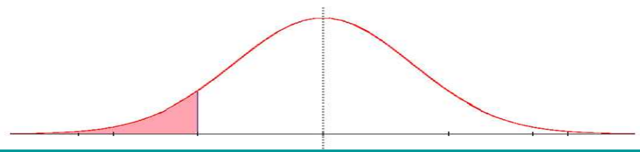
9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- población normal
- desviación típica conocida: σ

Cálculo del *p*-valor

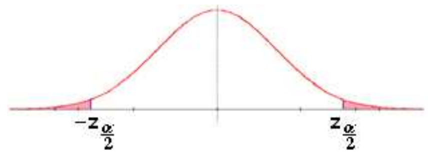
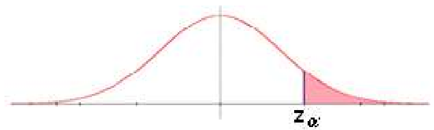
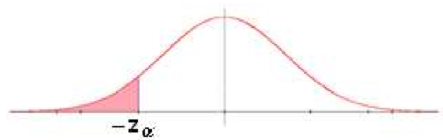
Test		***p*-valor**	
bilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{cases}$	$p-valor = 2 \cdot P\left(Z > \left\lvert \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \right\rvert \right)$	
unilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu > \mu_0 \end{cases}$	$p-valor = P\left(Z > \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \right)$	
	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu < \mu_0 \end{cases}$	$p-valor = P\left(Z < \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \right)$	

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- población normal **ó** n≥30
- desviación típica poblacional conocida: σ

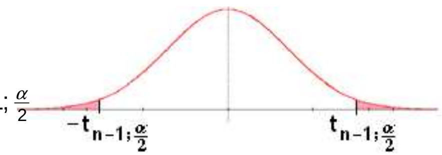
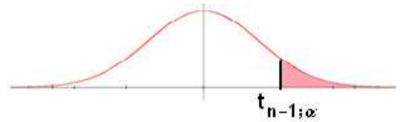
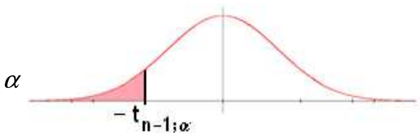
	Hipótesis	Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{cases}$	$\boxed{\overline{X} \approx N\left(\mu_0 ; \frac{\sigma}{\sqrt{n}}\right)}$	$\left\lvert \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \right\rvert > z_{\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu > \mu_0 \end{cases}$		$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} > z_{\alpha}$ 
	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu < \mu_0 \end{cases}$		$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -z_{\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- población normal con **muestras pequeñas** ($n \leq 30$)
- desviación típica poblacional desconocida

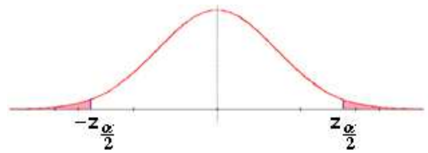
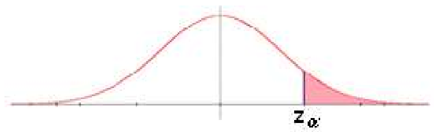
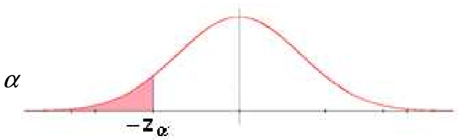
Hipótesis		Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{cases}$	$\boxed{\dfrac{\overline{X} - \mu_0}{\hat{s}/\sqrt{n}} \approx t_{n-1}}$	$\left\lvert \dfrac{\overline{X} - \mu_0}{\hat{s}/\sqrt{n}} \right\rvert > t_{n-1;\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu > \mu_0 \end{cases}$		$\dfrac{\overline{X} - \mu_0}{\hat{s}/\sqrt{n}} > t_{n-1;\alpha}$ 
	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu < \mu_0 \end{cases}$		$\dfrac{\overline{X} - \mu_0}{\hat{s}/\sqrt{n}} < -t_{n-1;\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la media de una población

Supuestos adicionales:

- tamaño muestral: $n \geq 30$
- desviación típica poblacional desconocida

Hipótesis		Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{cases}$	$\boxed{\overline{X} \approx N\left(\mu_0 ; \frac{\hat{s}}{\sqrt{n}}\right)}$	$\left\lvert \frac{\overline{X} - \mu_0}{\hat{s}/\sqrt{n}} \right\rvert > z_{\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu > \mu_0 \end{cases}$		$\frac{\overline{X} - \mu_0}{\hat{s}/\sqrt{n}} > z_{\alpha}$ 
	$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu < \mu_0 \end{cases}$		$\frac{\overline{X} - \mu_0}{\hat{s}/\sqrt{n}} < -z_{\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la media de una población: ejemplo

Ejemplo: ejercicio 4

Las tensiones de rotura de los cables fabricados en un proceso A se distribuyen normalmente con media 1800 y desviación típica 100. Se desea conocer si con un nuevo proceso B se aumenta la tensión media para lo que se toman 50 muestras observándose una tensión media de 1850.

1. ¿Se puede afirmar la mejoría del nuevo proceso B al nivel de significación 0.01?
2. Calcular el nivel crítico para $\bar{x} = 1850$

9. CONTRASTES DE HIPÓTESIS

Para la media de una población: ejercicio 4

Apartado 1: solución

- nivel de significación: $\alpha = 0.01$
- hipótesis: $\begin{cases} H_0 : \mu = 1800 \\ H_a : \mu > 1800 \end{cases}$

  H_0: el proceso B no mejora el A (su media no es mayor que 1800)
- estadístico del contraste: $\overline{X} \approx N\left(\mu_0 ; \dfrac{\sigma}{\sqrt{n}}\right)$
- región crítica: $\dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} > z_\alpha \Rightarrow \overline{x} > \mu_0 + z_\alpha \cdot \dfrac{\sigma}{\sqrt{n}}$

9. CONTRASTES DE HIPÓTESIS

Para la media de una población: ejercicio 4

Apartado 1: solución

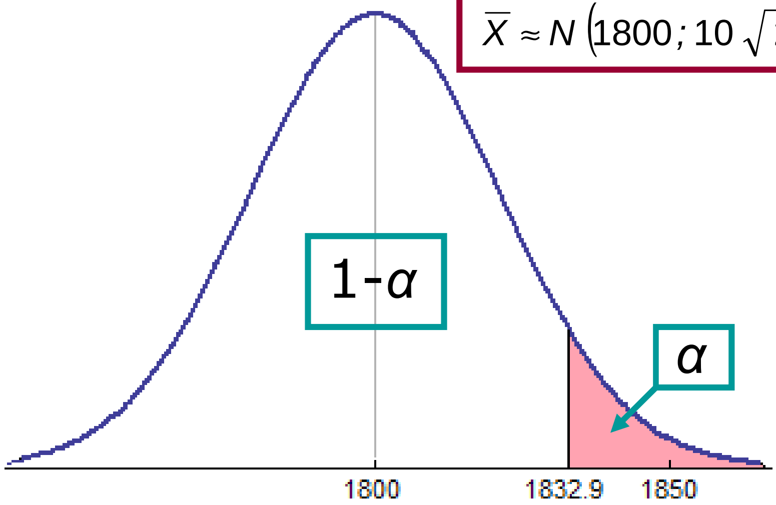
$$\bar{x} = 1850 \in RC \equiv \{\bar{x} \in R \;/\; \bar{x} > 1832.89956\}$$

$$\left.\begin{array}{l} \mu_0 = 1800 \\ \frac{\sigma}{\sqrt{n}} = 10 \cdot \sqrt{2} \\ \boxed{z_\alpha = 2.32635} \end{array}\right\} \Rightarrow \boxed{\bar{x} > 1832.89956}$$

rechazar H_0

=DISTR.NORM.ESTAND.INV(0,99)

$$\bar{X} \approx N\left(1800; 10\sqrt{2}\right)$$



9. CONTRASTES DE HIPÓTESIS

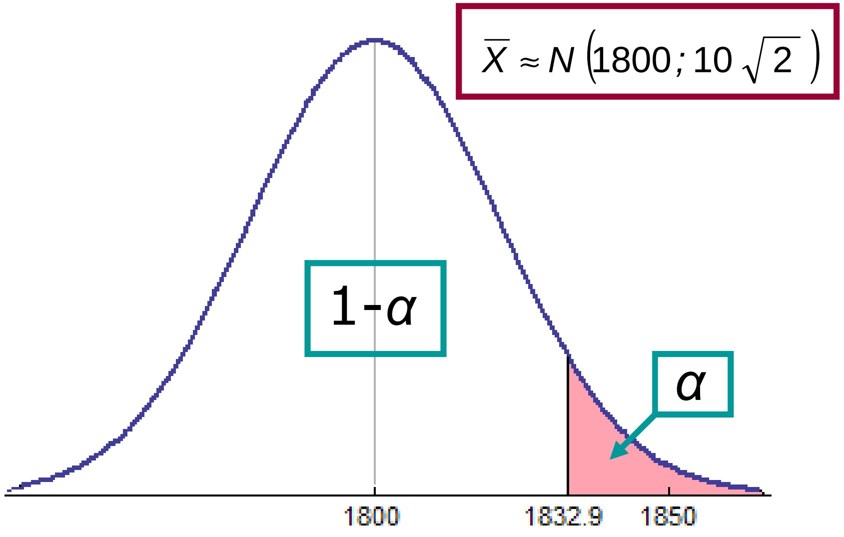
Para la media de una población: ejercicio 4

Apartado 2: solución

- contraste unilateral:

$$p = P\left(\overline{X} \geq 1850 \mid \overline{X} \approx N\left(1800; 10\sqrt{2}\right)\right) = P\left(Z \geq 3.5355\right) = 0.0002035$$

- valor muy pequeño, lo que da mucha confianza (~99.978%) para rechazar la hipótesis nula, H_0

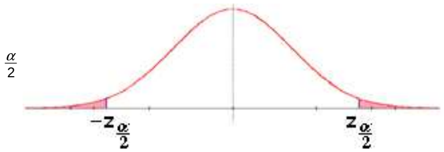
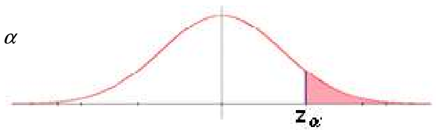
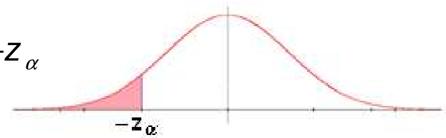


9. CONTRASTES DE HIPÓTESIS

Para la diferencia de medias de dos poblaciones

Supuestos adicionales:

- poblaciones normales **ó** $n_1, n_2 \geq 30$ (muestras independientes)
- desviaciones típicas poblacionales conocidas: σ_1, σ_2

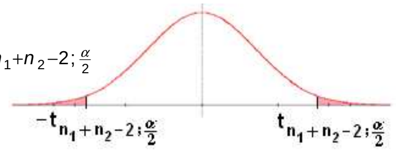
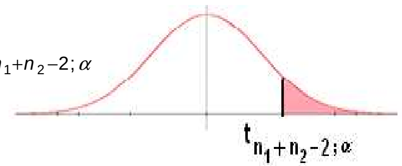
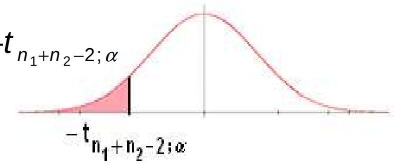
	Hipótesis	Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 \neq 0 \end{cases}$		$\dfrac{\lvert \bar{x}_1 - \bar{x}_2 \rvert}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} > z_{\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 > 0 \end{cases}$	$\boxed{\bar{X}_1 - \bar{X}_2 \approx N\left(0\,;\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}\right)}$	$\dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} > z_{\alpha}$ 
	$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 < 0 \end{cases}$		$\dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < -z_{\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la diferencia de medias de dos poblaciones

Supuestos adicionales:

- poblaciones normales (muestras independientes)
- desviaciones típicas poblacionales desconocidas pero iguales

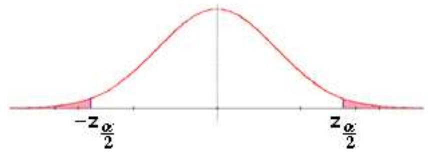
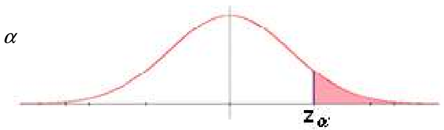
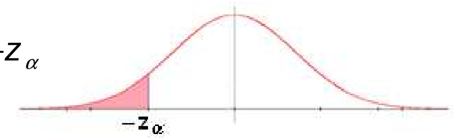
	Hipótesis	Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 \neq 0 \end{cases}$	$\boxed{\dfrac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \approx t_{n_1+n_2-2}}$ $\boxed{S_P^2 = \dfrac{(n_1-1)\cdot \hat{s}_1^2 + (n_2-1)\cdot \hat{s}_2^2}{n_1+n_2-2}}$	$\dfrac{\lvert \overline{x}_1 - \overline{x}_2 \rvert}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} > t_{n_1+n_2-2;\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 > 0 \end{cases}$		$\dfrac{\overline{x}_1 - \overline{x}_2}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} > t_{n_1+n_2-2;\alpha}$ 
	$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 < 0 \end{cases}$		$\dfrac{\overline{x}_1 - \overline{x}_2}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} < -t_{n_1+n_2-2;\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la diferencia de medias de dos poblaciones

Supuestos adicionales:

- $n_1, n_2 \geq 30$ (muestras independientes)
- desviaciones típicas poblacionales desconocidas

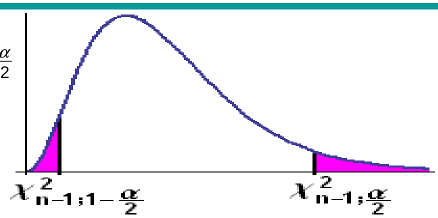
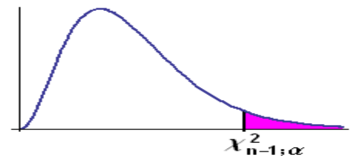
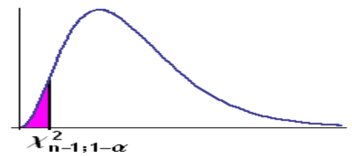
	Hipótesis	Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_a: \mu_1 - \mu_2 \neq 0 \end{cases}$		$\dfrac{\lvert \bar{x}_1 - \bar{x}_2 \rvert}{\sqrt{\dfrac{\hat{s}_1^2}{n_1} + \dfrac{\hat{s}_2^2}{n_2}}} > z_{\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_a: \mu_1 - \mu_2 > 0 \end{cases}$	$\boxed{\bar{X}_1 - \bar{X}_2 \approx N\left(0 ; \sqrt{\dfrac{\hat{s}_1^2}{n_1} + \dfrac{\hat{s}_2^2}{n_2}}\right)}$	$\dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\hat{s}_1^2}{n_1} + \dfrac{\hat{s}_2^2}{n_2}}} > z_{\alpha}$ 
	$\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_a: \mu_1 - \mu_2 < 0 \end{cases}$		$\dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\hat{s}_1^2}{n_1} + \dfrac{\hat{s}_2^2}{n_2}}} < -z_{\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la varianza de una población

Supuestos adicionales:

- población normal
- media poblacional desconocida

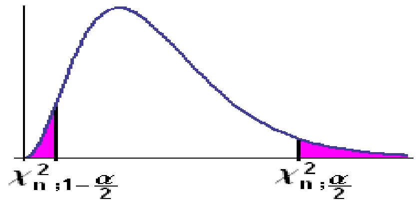
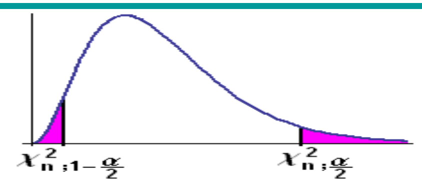
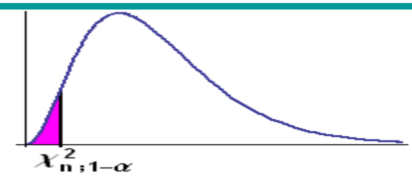
Hipótesis		Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 \neq \sigma_0^2 \end{cases}$	$\boxed{\dfrac{(n-1)\cdot \hat{s}^2}{\sigma_0^2} \approx \chi^2_{n-1}}$	$\chi^2_{n-1} < \chi^2_{n-1;1-\frac{\alpha}{2}}$ $\cup$ $\chi^2_{n-1} > \chi^2_{n-1;\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 > \sigma_0^2 \end{cases}$		$\chi^2_{n-1} > \chi^2_{n-1;\alpha}$ 
	$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 < \sigma_0^2 \end{cases}$		$\chi^2_{n-1} < \chi^2_{n-1;1-\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la varianza de una población

Supuestos adicionales:

- población normal
- media poblacional conocida: μ

	Hipótesis	Estadístico del contraste	Región crítica	
bilateral	$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 \neq \sigma_0^2 \end{cases}$	$$\frac{\sum_{i=1}^{i=n} (x_i - \mu)^2}{\sigma_0^2} \approx \chi_n^2$$	$\chi_n^2 < \chi^2_{n;1-\frac{\alpha}{2}}$ $\cup$ $\chi_n^2 > \chi^2_{n;\frac{\alpha}{2}}$	
unilateral	$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 > \sigma_0^2 \end{cases}$		$\chi_n^2 > \chi^2_{n;\alpha}$	
	$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 < \sigma_0^2 \end{cases}$		$\chi_n^2 < \chi^2_{n;1-\alpha}$	

9. CONTRASTES DE HIPÓTESIS

Para el cociente de dos varianzas

Supuestos adicionales:

- poblaciones normales (muestras independientes)
- medias poblacionales desconocidas

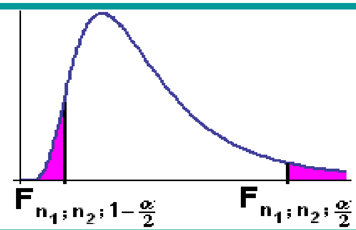
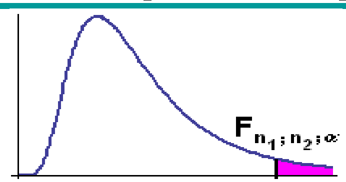
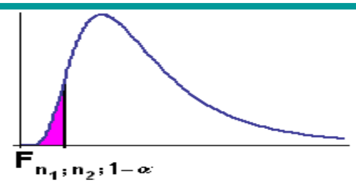
Hipótesis		Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_a: \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{cases}$	$\boxed{\frac{\hat{S}_1^2}{\hat{S}_2^2} \approx F_{n_1-1\,;\,n_2-1}}$	$F_{n_1-1\,;\,n_2-1} < F_{n_1-1\,;\,n_2-1;\,1-\frac{\alpha}{2}}$ $\cup$ $F_{n_1-1\,;\,n_2-1} > F_{n_1-1\,;\,n_2-1;\,\frac{\alpha}{2}}$
unilateral	$\begin{cases} H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_a: \frac{\sigma_1^2}{\sigma_2^2} > 1 \end{cases}$		$F_{n_1-1\,;\,n_2-1} > F_{n_1-1\,;\,n_2-1;\,\alpha}$
	$\begin{cases} H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_a: \frac{\sigma_1^2}{\sigma_2^2} < 1 \end{cases}$		$F_{n_1-1\,;\,n_2-1} < F_{n_1-1\,;\,n_2-1;\,1-\alpha}$

9. CONTRASTES DE HIPÓTESIS

Para el cociente de dos varianzas

Supuestos adicionales:

- poblaciones normales (muestras independientes)
- medias poblacionales conocidas: μ_1, μ_2

	Hipótesis	Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_a : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{cases}$	$\boxed{\frac{suma\,1}{suma\,2} \approx F_{n_1;n_2}}$	$F_{n_1;n_2} < F_{n_1;n_2;1-\frac{\alpha}{2}}$ $\cup$ $F_{n_1;n_2} > F_{n_1;n_2;\frac{\alpha}{2}}$ 
unilateral	$\begin{cases} H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_a : \frac{\sigma_1^2}{\sigma_2^2} > 1 \end{cases}$	$\boxed{suma\,k = \frac{\sum_{i=1}^{i=n_k} (x_{ki} - \mu_k)^2}{n_k} \quad (k = 1, 2)}$	$F_{n_1;n_2} > F_{n_1;n_2;\alpha}$ 
	$\begin{cases} H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_a : \frac{\sigma_1^2}{\sigma_2^2} < 1 \end{cases}$		$F_{n_1;n_2} < F_{n_1;n_2;1-\alpha}$ 

9. CONTRASTES DE HIPÓTESIS

Para la proporción *p* en una población

Supuestos adicionales:

- tamaño muestral grande: $\boldsymbol{n \geq 30\ (np_0 \geq 5,\ n(1-p_0) \geq 5)}$
- para $n<30$ el proceso se basa en la distribución binomial

	Hipótesis	Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0: p = p_0 \\ H_a: p \neq p_0 \end{cases}$	$\boxed{\hat{p} \approx N\left(p_0;\sqrt{\dfrac{p_0(1-p_0)}{n}}\right)}$ $\boxed{\hat{p} = \dfrac{n^o\ \text{éxitos}}{n}}$	$\dfrac{\lvert \hat{p}-p_0 \rvert}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} > z_{\frac{\alpha}{2}}$
unilateral	$\begin{cases} H_0: p = p_0 \\ H_a: p > p_0 \end{cases}$		$\dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} > z_{\alpha}$
	$\begin{cases} H_0: p = p_0 \\ H_a: p < p_0 \end{cases}$		$\dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} < -z_{\alpha}$

9. CONTRASTES DE HIPÓTESIS

Para la diferencia de proporciones, p_1 y p_2, en dos poblaciones

- dos poblaciones en las que hay proporciones p_1 y p_2 de individuos con la característica *éxito*
- se comparan las proporciones mediante la toma de muestras de tamaño n_1 y n_2
- proporciones de *éxito* en las muestras: $\hat{p}_1$ y $\hat{p}_2$

$$\boxed{\hat{p}_i = \frac{n^o \text{ éxitos muestra } i}{n_i} \quad (i = 1, 2)}$$

9. CONTRASTES DE HIPÓTESIS

Para la diferencia de proporciones, *p_1* y *p_2*, en dos poblaciones

Supuestos adicionales (tamaños muestrales grandes):

- ***$n_1 \geq 30$ ($n_1 p_1 \geq 5$, $n_1(1-p_1) \geq 5$)***
- ***$n_2 \geq 30$ ($n_2 p_2 \geq 5$, $n_2(1-p_2) \geq 5$)***

	Hipótesis	Estadístico del contraste	Región crítica
bilateral	$\begin{cases} H_0: p_1 - p_2 = 0 \\ H_a: p_1 - p_2 \neq 0 \end{cases}$	$\boxed{\hat{p}_1 - \hat{p}_2 \approx N\left(0\,;\sqrt{\hat{p}\,(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right)}$	$\dfrac{\lvert \hat{p}_1 - \hat{p}_2 \rvert}{\sqrt{\hat{p}\,(1-\hat{p})\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} > z_{\frac{\alpha}{2}}$
unilateral	$\begin{cases} H_0: p_1 - p_2 = 0 \\ H_a: p_1 - p_2 > 0 \end{cases}$	$\boxed{\hat{p} = \dfrac{n_1\,\hat{p}_1 + n_2\,\hat{p}_2}{n_1 + n_2}}$	$\dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\,(1-\hat{p})\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} > z_{\alpha}$
	$\begin{cases} H_0: p_1 - p_2 = 0 \\ H_a: p_1 - p_2 < 0 \end{cases}$		$\dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\,(1-\hat{p})\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} < -z_{\alpha}$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

La duración media de una muestra de 10 bombillas es de 1250 h. con una cuasidesviación típica muestral de 115 h. Se cambia el material del filamento por otro nuevo y se obtiene una muestra de 13 bombillas en la que la media es de 1340 h. con una cuasi-desviación típica muestral de 106 h.

1. Con un nivel de confianza del 95%, ¿puede aceptarse que las varianzas, antes y después del cambio, son iguales?
2. ¿Ha aumentado la duración media de las bombillas? Hallar el nivel crítico ó *p*-valor del contraste.

Datos:

- muestra 1: $\begin{cases} n_1 = 10 \\ \bar{x}_1 = 1250 \\ \hat{s}_1 = 115 \end{cases}$
- muestra 2: $\begin{cases} n_2 = 13 \\ \bar{x}_2 = 1340 \\ \hat{s}_2 = 106 \end{cases}$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 1: solución

- nivel de significación: $\boxed{\alpha = 0.05}$
- hipótesis: $\left\{ \begin{array}{l} H_0 : \dfrac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_a : \dfrac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{array} \right.$

  H₀: las varianzas son iguales (antes y después del cambio)
- estadístico del contraste: $\boxed{\dfrac{\hat{s}_1^2}{\hat{s}_2^2} \approx F_{n_1-1,\, n_2-1}}$
- región crítica: $\boxed{\dfrac{\hat{s}_1^2}{\hat{s}_2^2} > F_{n_1-1,\, n_2-1;\, \frac{\alpha}{2}} \quad \cup \quad \dfrac{\hat{s}_1^2}{\hat{s}_2^2} < F_{n_1-1,\, n_2-1;\, 1-\frac{\alpha}{2}}}$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 1: solución

- valor del estadístico del contraste: $\dfrac{\hat{s}_1^2}{\hat{s}_2^2} = \dfrac{115^2}{106^2} = 1.1770203$
- región crítica, *RC*:
  - límite superior: $F_{n_1-1,\, n_2-1;\, \frac{\alpha}{2}} = F_{9,\,12;\,0.025} = 3.4358$

    =DISTR.F.INV(0,025;9;12)
  - límite inferior: $F_{n_1-1,\, n_2-1;\, 1-\frac{\alpha}{2}} = F_{9,\,12;\,0.975} = 0.2585$

    =DISTR.F.INV(0,975;9;12)

$$F_{n_1-1,\, n_2-1;\, 1-\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,\, n_1-1;\, \frac{\alpha}{2}}} = \frac{1}{F_{12,\,9;\,0.025}} = 0.2585$$

tablas

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 1: solución

- decisión:
  - ◇ el estadístico del contraste no pertenece a la región crítica

$$\boxed{\frac{\hat{s}_1^2}{\hat{s}_2^2} \notin RC}$$

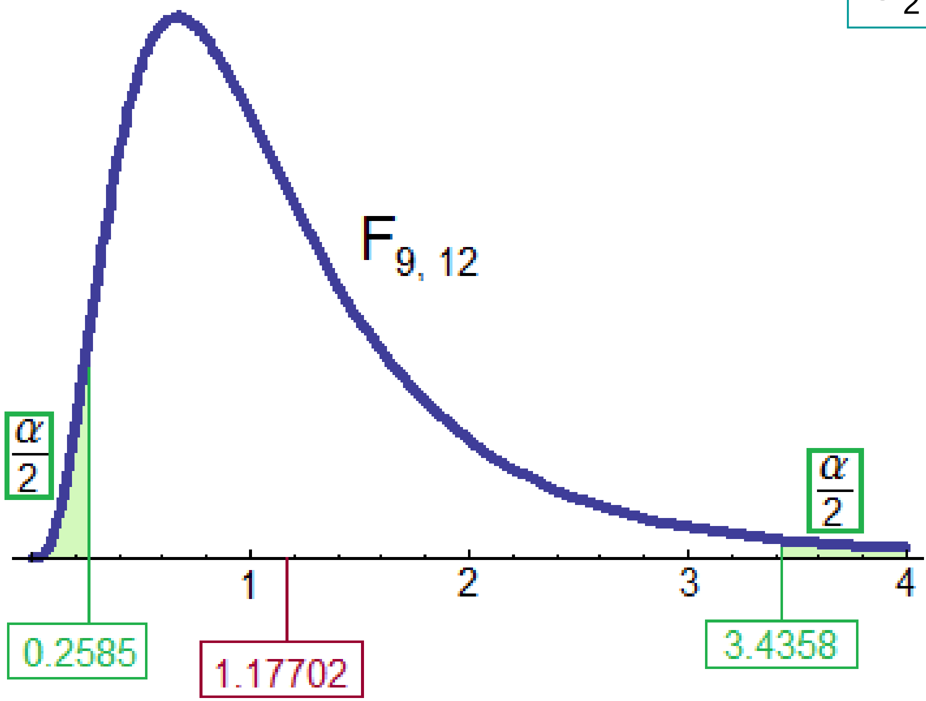
  - ◇ no se puede rechazar la hipótesis nula, por tanto, se asume que las varianzas son iguales

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 1: solución

- decisión:
  - no se puede rechazar la hipótesis nula: $\dfrac{\hat{s}_1^{\,2}}{\hat{s}_2^{\,2}} \notin RC$



9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 2: solución

- nivel de significación: $\alpha = 0.05$
- hipótesis: $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_a : \mu_1 < \mu_2 \end{cases}$

H_0: las medias son iguales (antes y después del cambio)

$$S_P = \sqrt{\frac{(n_1 - 1)\cdot \hat{s}_1^2 + (n_2 - 1)\cdot \hat{s}_2^2}{n_1 + n_2 - 2}}$$

- estadístico del contraste: $\dfrac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \approx t_{n_1+n_2-2}$
- región crítica: $\dfrac{\overline{x}_1 - \overline{x}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < -t_{n_1+n_2-2\,;\,\alpha}$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 2: solución

- valor del estadístico del contraste:

$$\left.\begin{array}{l} \bar{x}_1 - \bar{x}_2 = 1250 - 1340 = -90 \\ S_P = \sqrt{\frac{9 \cdot 115^2 + 12 \cdot 106^2}{21}} = 109.94739 \\ \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.4206222 \end{array}\right\} \Rightarrow \boxed{\frac{\bar{x}_1 - \bar{x}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = -1.946101}$$

- región crítica, *RC*:
  - límite: $\boxed{-t_{n_1+n_2-2\,;\,\alpha} = -t_{21;\,0.05} = -1.72074}$

=DISTR.T.INV(2*0,05;21)

una cola

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 2: solución

- región crítica, *RC*:
  - límite: $$-t_{n_1+n_2-2\,;\,\alpha} = -t_{21\,;\,0.05} = -1.72074$$

TABLAS

$$-t_{21\,;\,0.95} = -1.7207$$

9. CONTRASTES DE HIPÓTESIS

<u>Ejemplo</u>: ejercicio 15

<u>Apartado 2</u>: solución

- <u>decisión</u>:
  - el estadístico del contraste pertenece a la región crítica

$$\boxed{\frac{\bar{x}_1 - \bar{x}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = -1.946101 < -t_{21\,;\,0.05} = -1.72074}$$

  - se rechaza la hipótesis nula; se asume que la duración media de las bombillas es mayor tras el cambio de filamento

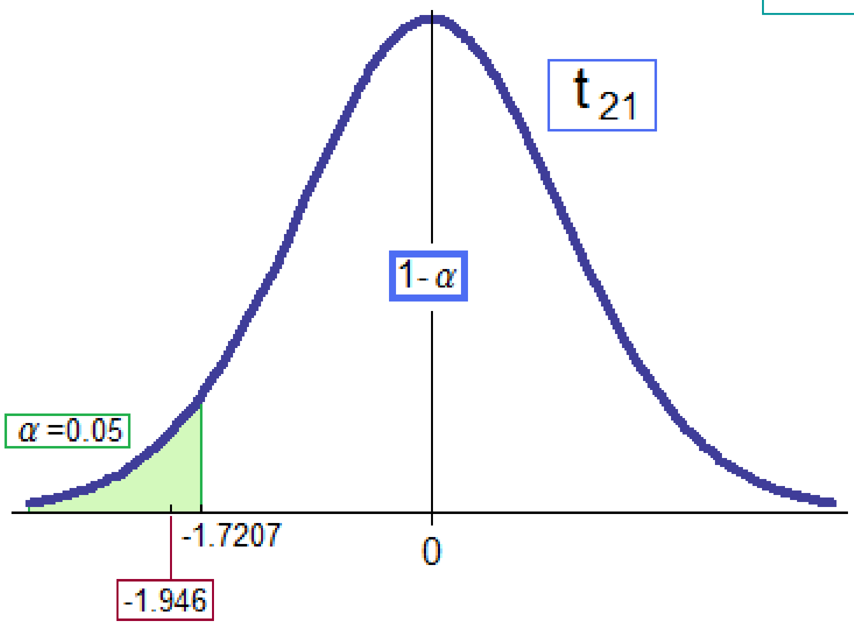
9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 2: solución

- decisión:
  - se rechaza la hipótesis nula:

$$\frac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \in RC$$



9. CONTRASTES DE HIPÓTESIS

<u>Ejemplo</u>: ejercicio 15

<u>Apartado 2</u>: solución

- <u>cálculo del *p*-valor (contraste unilateral a la izquierda)</u>:

$$\boxed{p = P\left(S \leq s \mid H_0\right)}$$

$$p = P\left(\frac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < -1.946101 \;\middle|\; \frac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \approx t_{21} \right) = 0.03257$$

=DISTR.T(1,9461;21;1)

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 2: solución

- cálculo del *p*-valor (contraste unilateral a la izquierda):

$$\boxed{p = P\left(S \leq s \mid H_0\right)}$$

$$p = P\left(\frac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < -1.946101 \;\middle|\; \frac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \approx t_{21} \right)$$

TABLAS

$$p = P\left(\frac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > 1.946101 \;\middle|\; \frac{\overline{X}_1 - \overline{X}_2}{S_P \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \approx t_{21} \right)$$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 15

Apartado 2: solución

- cálculo del *p*-valor (INTERPOLACIÓN):

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 6

Con objeto de estudiar los pesos de un cierto tipo de frutos se ha tomado una muestra de tamaño 10 unidades obteniéndose una cuasivarianza muestral de 402gr. Suponiendo que el peso de los frutos se distribuye normalmente,

1. ¿puede rechazarse, con un nivel de significación α=0.05, la hipótesis de que la varianza poblacional es 1000gr.?

2. calcular el nivel crítico

Datos:

- muestra: $\begin{cases} n = 10 \\ \hat{s}^2 = 402 \end{cases}$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 6

Apartado 1: solución

- nivel de significación: $\alpha = 0.05$

- hipótesis: $\begin{cases} H_0 : \sigma^2 = 1000 \\ H_a : \sigma^2 \neq 1000 \end{cases}$

  H0: la varianza es 1000gr.

- estadístico del contraste: $\dfrac{(n-1)\cdot \hat{s}^2}{\sigma_0^2} \approx \chi^2_{n-1}$

- región crítica: $\dfrac{(n-1)\cdot \hat{s}^2}{\sigma_0^2} < \chi^2_{n-1;\, 1-\frac{\alpha}{2}} \quad \cup \quad \dfrac{(n-1)\cdot \hat{s}^2}{\sigma_0^2} > \chi^2_{n-1;\, \frac{\alpha}{2}}$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 6

Apartado 1: solución

- valor del estadístico del contraste:

$$\frac{(n-1)\cdot \hat{s}^2}{\sigma_0^2} = \frac{9\cdot 402}{1000} = 3.618$$

- región crítica, *RC*:
  - límite superior: $\chi^2_{n-1;\,\frac{\alpha}{2}} = \chi^2_{9\,;\,0.025} = 19.02277$

    =PRUEBA.CHI.INV(0,025;9)

  - límite inferior: $\chi^2_{n-1;\,1-\frac{\alpha}{2}} = \chi^2_{9\,;\,0.975} = 2.70039$

    =PRUEBA.CHI.INV(0,975;9)

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 6

Apartado 1: solución

- región crítica, *RC*:
  - límite superior: $\chi^2_{n-1;\,\frac{\alpha}{2}} = \chi^2_{9\,;\,0.025} = 19.02277$

TABLAS

  - límite inferior: $\chi^2_{n-1;\,1-\frac{\alpha}{2}} = \chi^2_{9\,;\,0.975} = 2.70039$

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 6

Apartado 1: solución

- decisión:
  - el estadístico del contraste no pertenece a la región crítica

$$2.70039 < 3.618 < 19.0228 \Rightarrow \boxed{\frac{(n-1)\cdot\hat{s}^2}{\sigma_0^2} = 3.618 \notin RC}$$

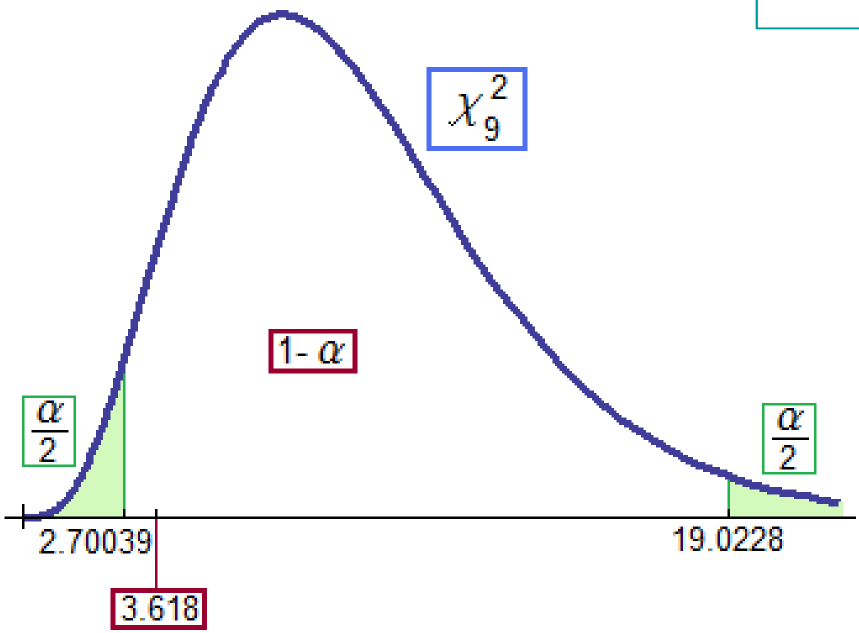
  - no se puede rechazar la hipótesis nula, por tanto, se asume que la varianza es igual a 1000

9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 6

Apartado 1: solución

- decisión:
  - ◇ no se puede rechazar H_0: $$\frac{(n-1)\cdot \hat{s}^2}{\sigma_0^2} = 3.618 \notin RC$$



9. CONTRASTES DE HIPÓTESIS

Ejemplo: ejercicio 6

Apartado 2: solución

- cálculo del *p*-valor (contraste bilateral):

$$\boxed{p = 2 \times min\left\{P\left(S \leq s \mid H_0\right), P\left(S > s \mid H_0\right)\right\}}$$

- $P\left(\frac{(n-1)\cdot\hat{s}^2}{\sigma_0^2} > 3.618 \mid \frac{(n-1)\cdot\hat{s}^2}{\sigma_0^2} \approx \chi_9^2\right) = 0.934711$

=DISTR.CHI(3,618;9)

- $P\left(\frac{(n-1)\cdot\hat{s}^2}{\sigma_0^2} \leq 3.618 \mid \frac{(n-1)\cdot\hat{s}^2}{\sigma_0^2} \approx \chi_9^2\right) = 0.06529$

=1-DISTR.CHI(3,618;9)

$$\boxed{p = 2\cdot 0.06529 = 0.13058}$$