# XS3310 Teoría Estadística

I Semestre 2021

Escuela de Estadística

2021-05-24

class: center, middle

## ¿Qué hemos visto hasta ahora?

Todo sobre estimadores puntuales + pivotes e intervalos de confianza. Bootstrap.

## ¿Qué vamos a discutir hoy?

Contrastes de hipótesis - introducción

---

### Contrastes de hipótesis

Tenemos un parámetro $\theta$ que es desconocido pero podemos decir que pertenece a un espacio paramétrico $\Omega$; este espacio incluye todos los posibles valores que $\theta$ podría tomar. Suponga que dividimos $\Omega$ en dos subconjuntos disjuntos $\Omega_0$ y $\Omega_1$ tales que:

$$\Omega_0 \cap \Omega_1 = \emptyset \land \Omega_0 \cup \Omega_1 = \Omega.$$

Si dividimos $\Omega$ de esta manera entonces es de esperar que $\theta$ se encuentre en $\Omega_0$ o en $\Omega_1$ (no puede estar en los dos al mismo tiempo).

Así, podemos tener:

- una **hipótesis nula** tal que $H_0 : \theta \in \Omega_0$
- y una **hipótesis alternativa** tal que $H_1 : \theta \in \Omega_1$.

El procedimiento que sigue consiste en *contrastar* estas hipótesis mediante una *regla de decisión* que favorece una hipótesis sobre la otra al cumplirse cierta condición y viceversa si la condición no se cumple.

---

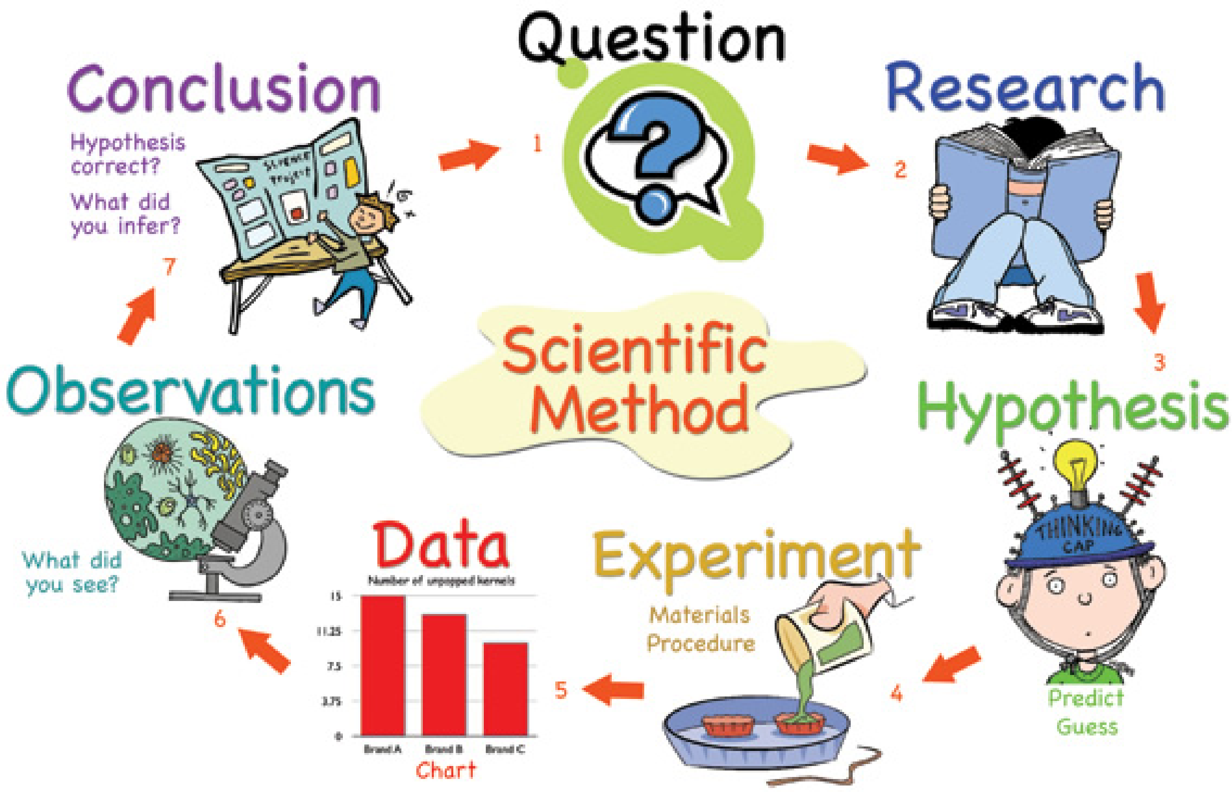


Figure 1: Método científico

## Contrastes de hipótesis

https://www.youtube.com/watch?v=9jW9G8MO4PQ

---

---

## Contrastes de hipótesis

Dos posibles resultados:

- Rechazar la hipótesis nula o
- No rechazar la hipótesis nula.

Nunca decimos que se 'acepta' la hipótesis nula ya que es imposible que un experimento confirme una hipótesis. Sin embargo, si encontramos suficiente evidencia que sugiere que una hipótesis sea falsa podemos 'rechazar' la hipótesis. No obstante, por simplicidad, en el resto de este tema cuando mencione que 'se acepta $H_0$' se entenderá como 'no se rechaza $H_0$".

Supongamos que tenemos datos que provienen de una población Normal con media desconocida $\mu$ y variancia 1, además, que $\mu$ puede tomar solamente dos valores: $\mu = 1$ y $\mu = -1$. De este modo tenemos las hipótesis $H_0 : \mu = 1$ y $H_1 : \mu = -1$. Como con el método científico nuestro siguiente paso consiste en obtener datos y observar cuál de estas dos hipótesis es la más factible para los datos observados. En un ejemplo como este esperaríamos que si $H_0$ es cierto entonces hayan más datos positivos que negativos y que si $H_1$ es cierto entonces hayan más datos negativos que positivos.

---

## Contrastes de hipótesis

Supongamos que resumimos los datos de la muestra en el estadístico $\overline{X}$ y comparamos este valor contra un valor $k$ de manera que si $\overline{X}$ es mayor a $k$ entonces preferimos $H_0$ mientras que si $\overline{X}$ es menor a $k$ entonces preferimos a $H_1$, o como decimos usualmente, rechazamos $H_0$. Nótese que existe toda una región de posibles valores de $\overline{X}$ donde podemos rechazar $H_0$; esta región se llama la *región crítica* del contraste.

> Definición: **Región crítica**. Consideremos las hipótesis $H_0 : \theta \in \Omega_0$ y $H_1 : \theta \in \Omega_1$. Sea $X_1, X_2, ..., X_n$ una muestra aleatoria seleccionada sobre una distribución con parámetros desconocido $\theta$. Sea $S$ el espacio muestral para el vector aleatorio n-dimensional $\mathbf{X} = \{X_1, X_2, ..., X_n\}$ es decir, $S$ es el resultado de todas las posibles muestras aleatorias de tamaño $n$. Sea $\delta$ un contraste que consiste en dividir el espacio muestral $S$ en dos subconjuntos, el primero de ellos contiene los valores de $\mathbf{X}$ para los cuales se rechaza $H_0$ y se denomina como la **región crítica del contraste**, denotada como $RC_\delta$.

---

## Contrastes de hipótesis

En los contrastes de hipótesis con los que acostumbramos trabajar solemos definir la región crítica en términos del estadístico de prueba. En el ejemplo que veníamos discutiendo anteriormente el *estadístico de prueba* es el que se usa en el contraste para resumir la información de la muestra es decir, $\overline{X}$. La región crítica se define como todos los valores de $\overline{X}$ que sean menores a $k$. Por otro lado, el valor $k$ es el valor que separa la región crítica de la región donde no rechazamos $H_0$, llamada la *región crítica complemento* (denotada como $RC_\delta^c$); por su parte el valor de $k$ se llama el *valor crítico* del contraste. Pueden existir contrastes con más de un valor crítico.

---

## Contrastes de hipótesis

Dos características que tienen los contrastes: la potencia y el tamaño. El enfoque está en encontrar un "buen" contraste según estas características. A continuación se define la potencia de la prueba:

> Definición. **Función potencia:** Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de una población con parámetro desconocido $\theta$ y sea $RC_\delta$ la región crítica de un contraste $\delta$ respecto a $\theta$. La **potencia del contraste**, denotada Potencia($\theta$), es la probabilidad que el contraste indique rechazar $H_0$ para un valor dado de $\theta$ es decir, Potencia($\theta$) $= P(\mathbf{X} \in RC|\theta)$

---

## Contrastes de hipótesis

La función potencia nos da la probabilidad de que el estadístico de prueba se encuentre en la región crítica para un valor dado de $\theta$. Un contraste ideal, pero imposible en la práctica, sería tal que siempre rechace $H_0$ cuando este es falso y que no lo rechace cuando es cierto. Es decir, queremos que la Potencia para valores de $\theta$ en $\Omega_0$ sea cero mientras que para valores de $\theta$ en $\Omega_1$ queremos que la Potencia sea uno. De este modo estamos tomando la decisión correcta el 100% de las veces. En la práctica queremos tener un contraste que tenga una función potencia lo más parecida a la función de potencia ideal.

---

## Contrastes de hipótesis

Ejemplo de una función de potencia para un caso práctico.

---

## Contrastes de hipótesis

Ejemplo de una función de potencia para el caso ideal.

---

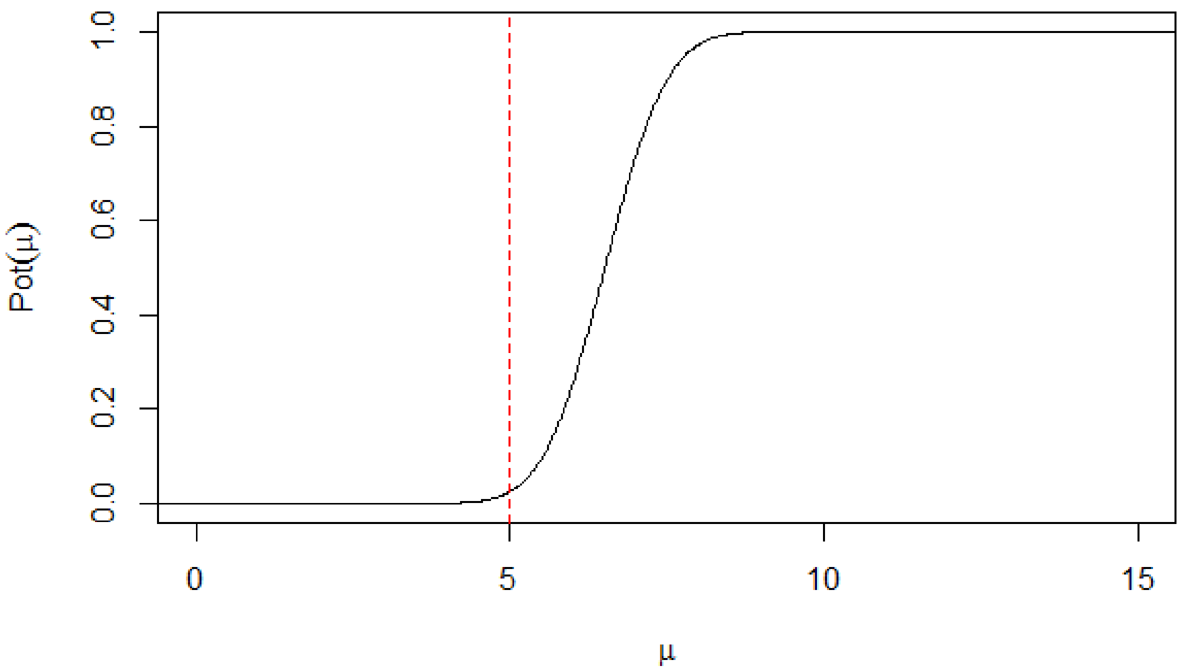


Figure 2: Función de potencia real

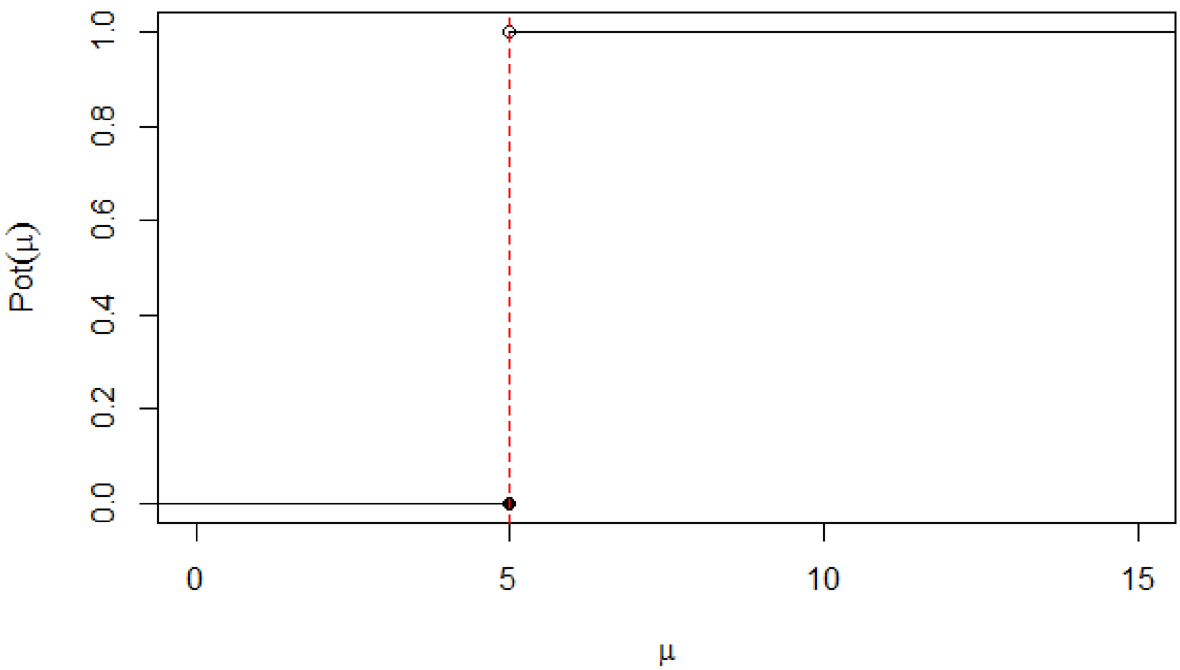


Figure 3: Función de potencia ideal

## Contrastes de hipótesis

Ejemplo: Sea $X_1, X_2, ..., X_4$ una muestra aleatoria de una población Normal con media desconocida $\mu$ y variancia 1. Suponga que se tienen las hipótesis $H_0 : \mu = 1$ y $H_1 : \mu = -1$. Para contrastar estas hipótesis se va a rechazar $H_0$ si $\overline{X}$ es menor a 0. Obtenga la función potencia para todos los valores de $\mu$ dentro de $\Omega$.

Solución: En este caso todo el espacio paramétrico $\Omega$ se compone solo por dos valores de $\mu$: $\mu = 1$ y $\mu = -1$. Por lo tanto debemos calcular la Potencia para cada uno de estos valores. Empezaremos por la Potencia en $\mu = 1$:

$$\text{Potencia}(1) = P(\overline{X} < 0|\mu = 1) = P\left(\frac{\overline{X}-1}{\frac{1}{\sqrt{4}}} < \frac{0-1}{\frac{1}{\sqrt{4}}}\right) = P(Z < -2) \approx 0.023$$

Seguimos con la Potencia para $\mu = -1$:

$$\text{Potencia}(-1) = P(\overline{X} < 0|\mu = -1) = P\left(\frac{\overline{X}+1}{\frac{1}{\sqrt{4}}} < \frac{0+1}{\frac{1}{\sqrt{4}}}\right) = P(Z < 2) \approx 0.977$$

---

## Contrastes de hipótesis

Podríamos interpretar el primer valor diciendo que para este contraste la probabilidad de rechazar la hipótesis nula cuando esta es cierta es de 0.023, mientras que el segundo valor se puede interpretar diciendo que la probabilidad de rechazar la hipótesis nula cuando esta es falsa es de 0.977. Podemos ver que tenemos un buen contraste, ya que la probabilidad de cometer una equivocación al tomar nuestra decisión es muy baja.

Nótese que en cualquier contraste de hipótesis podemos tener cuatro posibles resultados. Podemos rechazar la hipótesis nula cuando esta es falsa, podemos rechazar la hipótesis nula cuando es verdadera, podemos no rechazar la hipótesis nula cuando esta es falsa o podemos no rechazar la hipótesis nula cuando esta es verdadera. Dos de estos resultados constituyen cometer una equivocación: rechazar la hipótesis nula cuando es cierta y no rechazar la hipótesis nula cuando es falsa. Estos errores reciben el nombre de *Error Tipo I* y *Error Tipo II*, respectivamente.

---

## Contrastes de hipótesis

**Error Tipo I y Error Tipo II**

| | $H_0$ cierto | $H_0$ falso |
|---|---|---|
| Aceptar $H_0$ | Acierto | Error Tipo II |
| Rechazar $H_0$ | Error Tipo I | Acierto |

---

## Contrastes de hipótesis

---

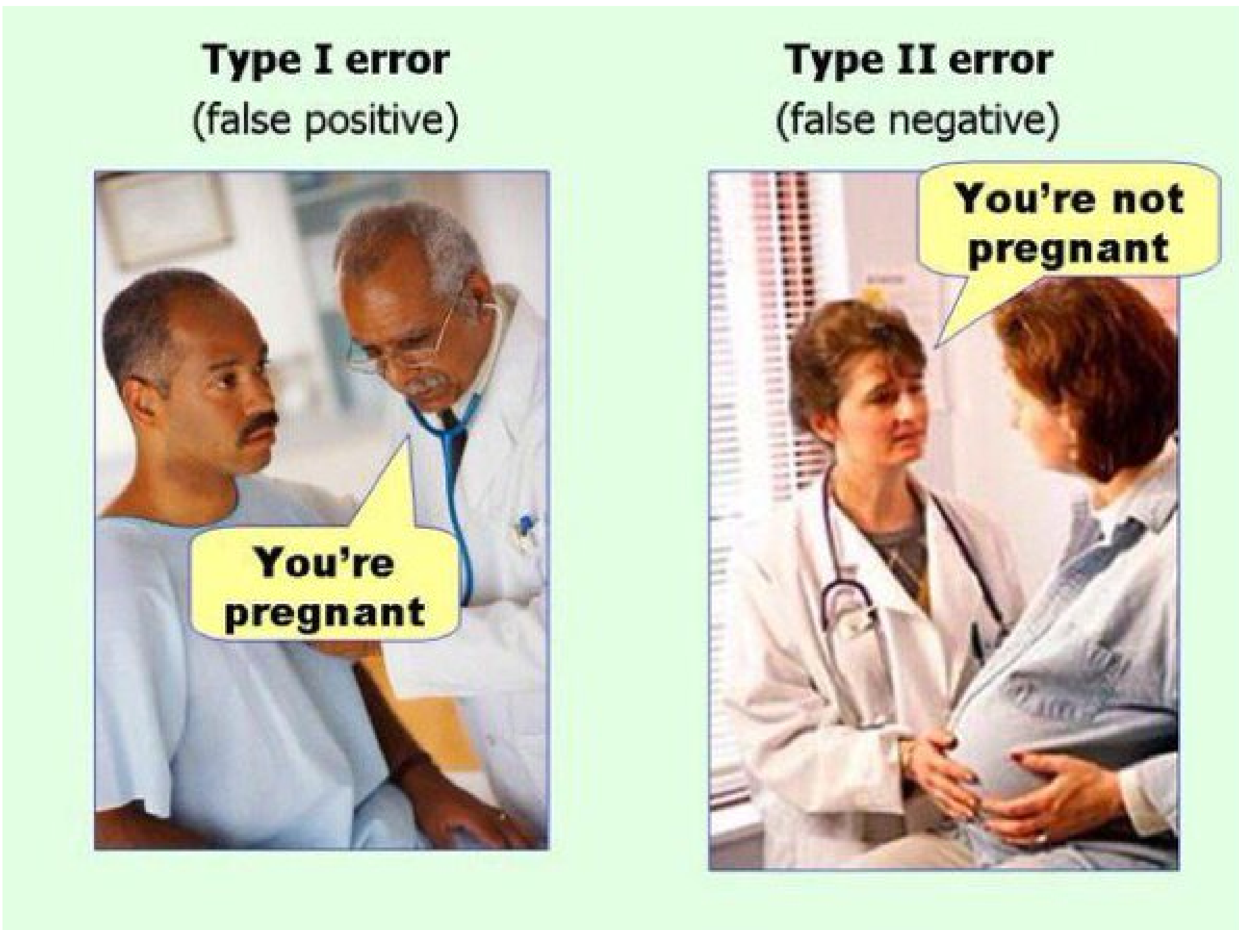


Figure 4: Error Tipo I y Error Tipo II

## Contrastes de hipótesis

Podemos definir la **probabilidad** de cometer estos dos tipos de errores en términos de la función potencia:

> Definición. **Probabilidad de Error Tipo I y Error Tipo II:** Para un espacio paramétrico $\Omega$, dividido en dos subconjuntos $\Omega_0$ y $\Omega_1$, de un parámetro $\theta$, se define la probabilidad de cometer Error Tipo I, denotada como $\alpha(\delta)$ como:
>
> $$\begin{aligned}\alpha(\delta) &= P(\text{rechazar } H_0 \text{ cuando es cierta}) \\ &= P\left(\mathbf{X} \in RC_\delta | \theta \in \Omega_0\right) = \text{Potencia}\left(\theta | \theta \in \Omega_0\right)\end{aligned}$$
>
> Por otra parte la probabilidad de cometer Error Tipo II, denotada como $\beta(\delta)$, se define como:
>
> $$\begin{aligned}\beta(\delta) &= P(\text{no rechazar } H_0 \text{ cuando es falsa}) = P\left(\mathbf{X} \in RC_\delta^c | \theta \in \Omega_1\right) \\ &= 1 - P\left(\mathbf{X} \in RC_\delta | \theta \in \Omega_1\right) = 1 - \text{Potencia}(\theta | \theta \in \Omega_1)\end{aligned}$$

---

## Contrastes de hipótesis

De la definición quedan claro algunos puntos sobre la probabilidad de Error Tipo I y la probabilidad de Error Tipo II. Primero, podemos ver que estos son funciones de $\theta$; vamos a tener valores distintos de estas probabilidades dependiendo del valor de $\theta$ que estemos evaluando. Nótese en la Figura 1 como en la gráfica de la potencia real todo aquello antes de la línea punteada corresponde a $\Omega_0$, por lo tanto la curva que está en ese dominio pertenece a $\alpha(\mu)$. Segundo, el $\alpha$ y $\beta$ no son probabilidades complementarias ya que están definidos en dos regiones distintas de $\Omega$. No obstante, estos presentan una relación inversamente proporcional, de forma que a medida de que aumenta la probabilidad de cometer un error, disminuye la probabilidad de cometer el otro.

---

## Contrastes de hipótesis

En el ejemplo anterior podemos observar que solo tenemos un valor para la probabilidad de cada tipo de error. En ese caso:

$\alpha(1) = \text{Potencia}(1) = 0.023$

$\beta(-1) = 1 - \text{Potencia}(-1) = 1 - 0.977 = 0.023$

Por pura coincidencia ambos errores tienen la misma probabilidad en el ejemplo anterior. A continuación veremos otra forma de describir un contraste.

> Definición. **Tamaño del contraste**. Para un contraste de hipótesis $H_0 : \theta \in \Omega_0$ y $H_1 : \theta \in \Omega_1$ se denomina tamaño del contraste al valor $\alpha_c$ dado por:
>
> $$\alpha_c = \text{Sup}(\text{Potencia}(\theta | \theta \in \Omega_0)) = \text{Sup}(\alpha(\delta))$$
>
> Es decir, el tamaño de un contraste es la máxima probabilidad de cometer un Error Tipo I.

---

## Contrastes de hipótesis

Para la mayoría de los contrastes su tamaño se va a encontrar en los límites de $\Omega_0$. De no ser ese el caso podríamos derivar $\alpha(\delta)$ y ver si alcanza algún máximo dentro de su dominio.

**Ejemplo:** Suponga que se selecciona una muestra aleatoria $X_1, X_2, ..., X_n$ sobre una distribución Uniforme en $(0, \theta)$, con $\theta$ desconocido y mayor a cero. Con esta muestra se desea contrastar las hipótesis:

$$H_0 : 3 \leq \theta \leq 4 \qquad \Omega_0 = [3, 4]$$

$$H_1 : \theta < 3 \quad o \quad \theta > 4 \qquad \Omega_1 = ]0, 3[ \cup ]4, +\infty[$$

En este caso utilizaremos a $X_{(n)}$ para crear la región crítica. Sabemos que $X_{(n)}$ es el estimador de máxima verosimilitud de $\theta$ y que siempre será menor a $\theta$ (aunque con muestras grandes se va a aproximar mucho). Supongamos que por estas razones vamos a aceptar $H_0$ si $2.9 \leq X_{(n)} \leq 3.9$ y se rechaza en caso contrario. Entonces la región crítica del contraste incluye todas las muestras $\mathbf{X}$ para las cuales $X_{(n)} < 2.9$ o $X_{(n)} > 3.9$, es decir, $RC_\delta = \left\{\mathbf{X} | X_{(n)} \in \mathbf{X} \text{ y } X_{(n)} < 2.9 \text{ o } X_{(n)} > 3.9\right\}$.

---

## Contrastes de hipótesis

Determine una expresión para la función potencia de este contraste.

**Solución:** Recordemos que la función potencia se define como la probabilidad de rechazar la hipótesis nula, es decir, de que la muestra se encuentre en la región crítica:

$\text{Potencia}(\theta) = P(\mathbf{X} \in RC_\delta) = P\left(X_{(n)} < 2.9 \text{ o } X_{(n)} > 3.9\right) =$

$P\left(X_{(n)} < 2.9\right) + P\left(X_{(n)} > 3.9\right) = F_{X_{(n)}}(2.9) + 1 - F_{X_{(n)}}(3.9)$

Recordemos que en este caso $X_{(n)} \sim \text{Potencial}(n, \theta)$, por lo tanto su función de distribución viene dada por:

$$F_{X_{(n)}}(x) = \begin{cases} 0 \quad si \quad x \leq 0 \\ \left(\dfrac{x}{\theta}\right)^n \quad si \quad 0 < x < \theta \\ 1 \quad si \quad x \geq \theta \end{cases}$$

---

## Contrastes de hipótesis

La evaluación de la expresión anterior va a depender del valor de $\theta$, ya que dependiendo de este se va a evaluar la función de distribución en distintos argumentos. En este ejemplo tenemos tres casos distintos:

Si $0 < \theta \leq 2.9$:

$F_{X_{(n)}}(2.9) = 1 \qquad \text{y} \qquad F_{X_{(n)}}(3.9) = 1$

$\Rightarrow \text{Potencia}(\theta) = 1 + 1 - 1 = 1$

Si $2.9 \leq \theta \leq 3.9$:

$F_{X_{(n)}}(2.9) = \left(\dfrac{2.9}{\theta}\right)^n \quad \text{y} \quad F_{X_{(n)}}(3.9) = \left(\dfrac{3.9}{\theta}\right)^n = 1$

$\Rightarrow \text{Potencia}(\theta) = \left(\dfrac{2.9}{\theta}\right)^n + 1 - 1 = \left(\dfrac{2.9}{\theta}\right)^n$

## Contrastes de hipótesis

Finalmente, si $\theta > 4$:

$$F_{X_{(n)}}(2.9) = \left(\frac{2.9}{\theta}\right)^n \quad \text{y} \quad F_{X_{(n)}}(3.9) = \left(\frac{3.9}{\theta}\right)^n$$

$$\Rightarrow \text{Potencia}(\theta) = \left(\frac{2.9}{\theta}\right)^n + 1 - \left(\frac{3.9}{\theta}\right)^n$$

De esta manera podemos resumir la función potencia en la siguiente expresión:

$$\text{Potencia}(\theta) = \begin{cases} 1 \quad si \quad 0 < \theta \leq 2.9 \\ \left(\frac{2.9}{\theta}\right)^n \quad si \quad 2.9 < \theta \leq 3.9 \\ \left(\frac{2.9}{\theta}\right)^n + 1 - \left(\frac{3.9}{\theta}\right)^n \quad si \quad \theta > 3.9 \end{cases}$$

## Contrastes de hipótesis

Podemos graficar la forma de esta función para un tamaño de muestra dado. La siguiente figura muestra la forma de la Potencia con $n = 10$, con $\Omega_0$ delimitado por las líneas rojas punteadas.

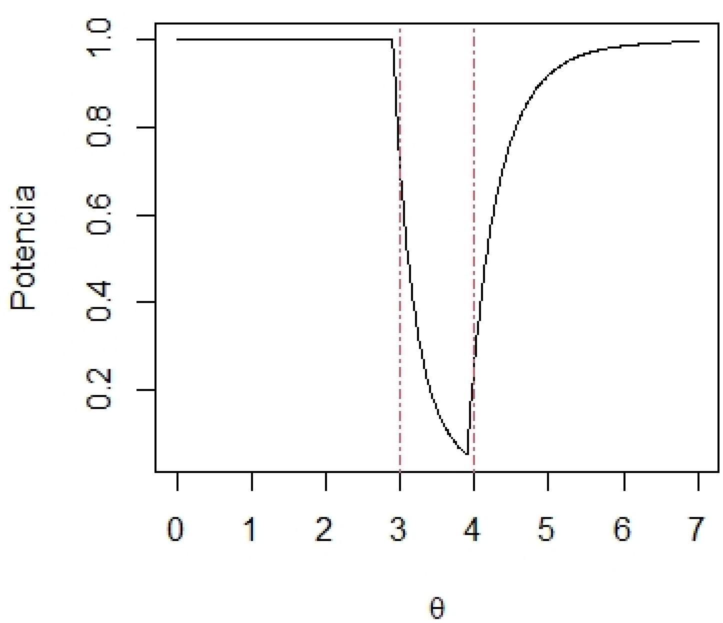


Figure 5: Función potencia con $n = 10$

## Contrastes de hipótesis

- De la Figura anterior podemos detallar que el tamaño de contraste se encuentra cuando $\theta = 3$. Con $n = 10$ tenemos que el tamaño del contraste sería:

$$\alpha_c = \text{Potencia}(3) = \left(\frac{2.9}{3}\right)^{10} = 0.712$$

## Contrastes de hipótesis

Podemos decir que para una muestra de tamaño 10, este contraste tiene una máxima probabilidad de cometer un Error Tipo I de 0.712. Como un ejemplo adicional podemos calcular la probabilidad de cometer Error Tipo 2 para algún valor de $\theta$ en $\Omega_1$. Para propósitos del ejemplo usaremos $\theta = 5$:

$$\beta(5) = 1 - \text{Potencia}(5) = 1 - \left(\left(\frac{2.9}{5}\right)^{10} + 1 - \left(\frac{3.9}{5}\right)^{10}\right) = 0.079$$

Un uso que tiene el tamaño como una característica del contraste es que podemos definir una región crítica para un tamaño de contraste dado.

## Contrastes de hipótesis

**Ejemplo:** Sea $X_1, X_2, ..., X_n$ una muestra aleatoria tal que $X_j \sim N(\mu, 1)$. Se desea contrastar las hipótesis $H_0 : \mu = 5$ contra la alternativa $H_1 : \mu > 5$. La región crítica se define como $RC_\delta = \left\{\mathbf{X} | \overline{X} > c\right\}$. Encuentre el valor de $c$ para que este contraste tenga un tamaño igual a 0.10.

**Solución:** Note que el contraste se rechaza cuando $\overline{X}$ es mayor que $c$, por lo tanto podemos definir la potencia de la siguiente forma:

$\text{Potencia}(\mu) = P(\overline{X} > c|\mu) = P\left(Z > \sqrt{n}(c - \mu)\right)$

Sabemos que el tamaño del contraste se define como la máxima probabilidad de cometer Error Tipo I, pero en este caso solo tenemos una probabilidad de Error Tipo I, que sucede cuando $\mu = 5$. Por lo tanto:

$\alpha_c = \text{Potencia}(5) = P\left(Z > \sqrt{n}(c - 5)\right) = 0.10$

## Contrastes de hipótesis

Podemos utilizar las tablas para buscar el valor de la normal estándar que acumula a su derecha una probabilidad de 0.10 (que equivale a una probabilidad acumulada a su izquierda de 0.90). Por lo tanto tenemos que

$\sqrt{n}(c - 5) = 1.28$

$\Rightarrow c = 5 + \frac{1.28}{\sqrt{n}}$

Una observación importante de resaltar es que si tenemos una distribución de probabilidad discreta no es posible encontrar un tamaño de contraste para cualquier valor de $c$, ya que las variable aleatorias discretas tienen probabilidades acumuladas a solo ciertos valores. Esto se puede solucionar por medio de contrastes aleatorizados, que dependen del lanzamiento de una moneda con cierta probabilidad de caer escudo, o con bootstrap (que veremos al final de esta sección).

---

### Contrastes de hipótesis

Otro punto importante de destacar es que en la práctica usualmente utilizamos la hipótesis nula puntual $H_0 : \theta = \theta_0$ ya que esta no pierde generalidad. Supongamos que se quiere contrastar $H_0 : \theta \leq \theta_0$ contra $H_1 : \theta > \theta_0$. En casi todos los casos la función potencia de un contraste es monótona en todo el dominio de $\theta$; en este caso sería una función creciente. Ya que el tamaño se define como la máxima potencia en $\Omega_0$ esto va a suceder en su extremo derecho, el cual es $\theta_0$. Esto ocasiona que el contraste tenga el mismo tamaño a que si la hipótesis nula hubiera sido $H_0 : \theta = \theta_0$. Por lo tanto en la práctica es más común encontrar la hipótesis nula formulada en $\theta_0$, siendo este el valor de $\Omega_0$ más próximo a $\Omega_1$.

---

class: center, middle

# ¿Qué discutimos hoy?

Contrastes de hipótesis - introducción