



UNC

R Users Group - Ecuador[®]

“Taller de Introducción a la Inferencia Estadística con R”

Unidad 1: La Estadística una Orquesta hecha Instrumento



Andrés Peña M.

a.pena@rusersgroup.com

Julio 2018



Tabla de contenidos

- 1 Estructura del taller
- 2 Introducción
 - Jaime Llopis Pérez
 - Dos Mundos
- 3 Estadística: una orquesta hecha instrumento
 - Comprensión de objetos matemáticos como pedestales
 - Orquesta hecha instrumento
 - Contrastes de hipótesis
 - Modelos probabilísticos compuestos
- 4 Conclusiones y Recomendaciones



UNC

1. Estructura del taller



Estructura del taller

- 1 La estadística una orquesta hecha instrumento. Visión general del método estadístico. Objetos matemáticos como pedestales. Repaso de funciones.

Estructura del taller

- 1 La estadística una orquesta hecha instrumento. Visión general del método estadístico. Objetos matemáticos como pedestales. Repaso de funciones.
- 2 Data management en R. Análisis descriptivo. Medidas de tendencia central, dispersión, posición y forma. Repaso de técnicas de conteo.

Estructura del taller

- 1 La estadística una orquesta hecha instrumento. Visión general del método estadístico. Objetos matemáticos como pedestales. Repaso de funciones.
- 2 Data management en R. Análisis descriptivo. Medidas de tendencia central, dispersión, posición y forma. Repaso de técnicas de conteo.
- 3 Probabilidad y variables aleatorias. Función de densidad y distribución. Principales Distribuciones Continuas y Discretas. Distribuciones asociadas a la normal.

Estructura del taller

- 1 La estadística una orquesta hecha instrumento. Visión general del método estadístico. Objetos matemáticos como pedestales. Repaso de funciones.
- 2 Data management en R. Análisis descriptivo. Medidas de tendencia central, dispersión, posición y forma. Repaso de técnicas de conteo.
- 3 Probabilidad y variables aleatorias. Función de densidad y distribución. Principales Distribuciones Continuas y Discretas. Distribuciones asociadas a la normal.
- 4 Distribuciones en el muestreo. Teorema del Límite Central.

Estructura del taller

- 1 La estadística una orquesta hecha instrumento. Visión general del método estadístico. Objetos matemáticos como pedestales. Repaso de funciones.
- 2 Data management en R. Análisis descriptivo. Medidas de tendencia central, dispersión, posición y forma. Repaso de técnicas de conteo.
- 3 Probabilidad y variables aleatorias. Función de densidad y distribución. Principales Distribuciones Continuas y Discretas. Distribuciones asociadas a la normal.
- 4 Distribuciones en el muestreo. Teorema del Límite Central.
- 5 Inferencia estadística. Estimación puntual: propiedades de los estimadores y principales métodos de estimación. Estimación por intervalos. Pruebas de hipótesis estadísticas. Tipos de errores, pruebas clásicas para una, dos y más muestras. Pruebas no paramétricas más útiles.



UNC

2. Introducción



En una ocasión le preguntaron a un famoso director de orquesta cuál era para él el mejor instrumento de la orquesta. Y contestó que, para él, el mejor instrumento de la orquesta era la **orquesta**.

En una ocasión le preguntaron a un famoso director de orquesta cuál era para él el mejor instrumento de la orquesta. Y contestó que, para él, el mejor instrumento de la orquesta era la **orquesta**.

Jaime Llopis Pérez

- Profesor de Estadística de la Universitat de Barcelona, Universitat Internacional de Catalunya, en la Universitat Abat Oliba y en la Universitat Pompeu Fabra.
- Licenciado en Matemáticas, Medicina, Biología, Filosofía, Filología hispánica y Derecho.
- Doctor en Lingüística.

En una ocasión le preguntaron a un famoso director de orquesta cuál era para él el mejor instrumento de la orquesta. Y contestó que, para él, el mejor instrumento de la orquesta era la **orquesta**.

Jaime Llopis Pérez

- Profesor de Estadística de la Universitat de Barcelona, Universitat Internacional de Catalunya, en la Universitat Abat Oliba y en la Universitat Pompeu Fabra.
- Licenciado en Matemáticas, Medicina, Biología, Filosofía, Filología hispánica y Derecho.
- Doctor en Lingüística.



En una ocasión le preguntaron a un famoso director de orquesta cuál era para él el mejor instrumento de la orquesta. Y contestó que, para él, el mejor instrumento de la orquesta era la **orquesta**.

Jaime Llopis Pérez

- Profesor de Estadística de la Universitat de Barcelona, Universitat Internacional de Catalunya, en la Universitat Abat Oliba y en la Universitat Pompeu Fabra.
- Licenciado en Matemáticas, Medicina, Biología, Filosofía, Filología hispánica y Derecho.
- Doctor en Lingüística.



Ámbitos muy diversos:
“[...] entiendo necesario
para alguien que quiere
dedicarse a la Estadística”.

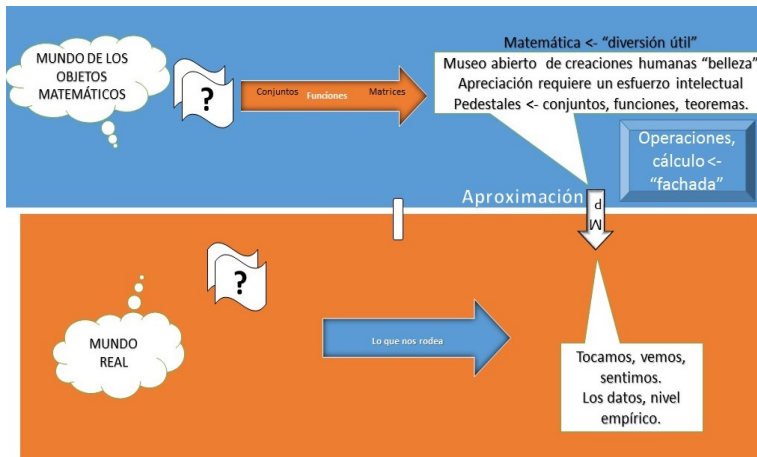
Jaime Llopis Pérez



Por qué?

Dos Mundos

- Filosofía de Platón “Dos mundos”: mundo real vs el mundo de las ideas.





UNC

3. Estadística: una orquesta hecha instrumento

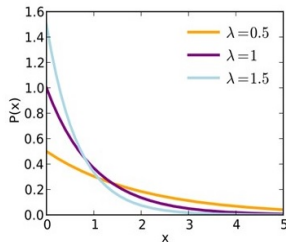


Comprensión de objetos matemáticos como pedestales

- Función de densidad de la distribución exponencial:

$$f(x) = P(x) = \begin{cases} \lambda e^{-\lambda x} & \text{para } x \geq 0 \\ 0 & \text{en caso contrario} \end{cases}$$

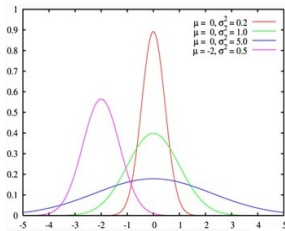
$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{para } x < 0 \\ 1 - e^{-\lambda x} & \text{para } x \geq 0 \end{cases}$$



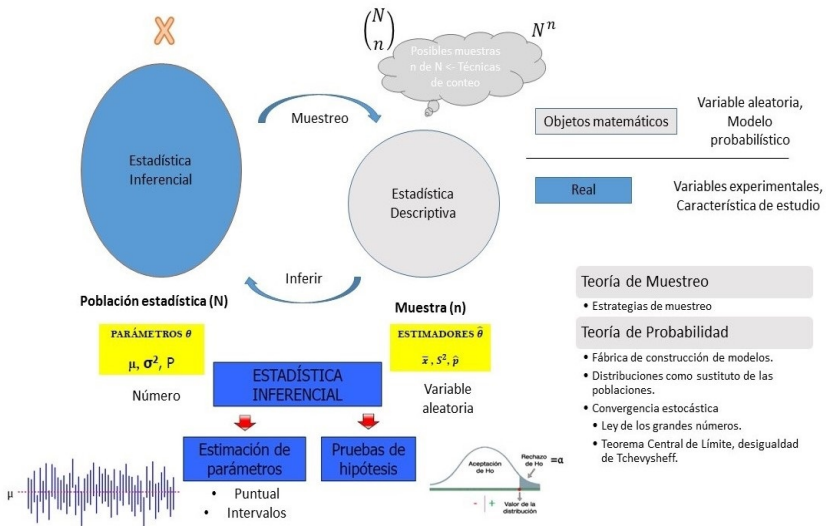
- Función de densidad de la normal:

$$\phi_{\mu, \sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

$$\Phi_{\mu, \sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(u-\mu)^2}{2\sigma^2}} du, \quad x \in \mathbb{R}.$$



Orquesta hecha instrumento



Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Las afirmaciones en H_0 son siempre **conservadoras**:

- Un mundo indiferenciado;

Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Las afirmaciones en H_0 son siempre **conservadoras**:

- Un mundo indiferenciado;
- Las cosas son iguales;

Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Las afirmaciones en H_0 son siempre **conservadoras**:

- Un mundo indiferenciado;
- Las cosas son iguales;
- Variables no relacionadas;

Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Las afirmaciones en H_0 son siempre **conservadoras**:

- Un mundo indiferenciado;
- Las cosas son iguales;
- Variables no relacionadas;
- Estado primitivo en donde todo está por demostrarse.

Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Las afirmaciones en H_0 son siempre **conservadoras**:

- Un mundo indiferenciado;
- Las cosas son iguales;
- Variables no relacionadas;
- Estado primitivo en donde todo está por demostrarse.

Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Las afirmaciones en H_0 son siempre **conservadoras**:

- Un mundo indiferenciado;
- Las cosas son iguales;
- Variables no relacionadas;
- Estado primitivo en donde todo está por demostrarse.

Probabilidades	Rechazar H_0	no Rechazar H_0
H_0 verdadera	α	$1 - \alpha$
H_0 falsa	$1 - \beta$	β

$$p\text{-value} = \mathbb{P}(\text{resultado tan extremo o más} \mid H_0)$$

Contrastes de hipótesis

Lo que afirmamos **versus** lo que vemos. La H_0 la tomaremos como cierta hasta que no se demuestre lo contrario. Ej: juicio.

Las afirmaciones en H_0 son siempre **conservadoras**:

- Un mundo indiferenciado;
- Las cosas son iguales;
- Variables no relacionadas;
- Estado primitivo en donde todo está por demostrarse.

Probabilidades	Rechazar H_0	no Rechazar H_0
H_0 verdadera	α	$1 - \alpha$
H_0 falsa	$1 - \beta$	β

$$p\text{-value} = \mathbb{P}(\text{resultado tan extremo o más} \mid H_0)$$

La distribución del estadístico bajo la hipótesis nula es el valor del estadístico calculado a todas las muestras posibles de tamaño n , si fuera cierta la hipótesis nula. Imagen teórica de como deberían ser las cosas si fuera cierta la hipótesis.

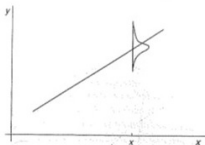
Contrastes de hipótesis

Cuando...		
$H_1: \neq$	$H_1: <$	$H_1: >$
<p>Rechazo de $H_0 = \frac{\alpha}{2}$</p> <p>Aceptación de H_0</p> <p>Rechazo de $H_0 = \frac{\alpha}{2}$</p> <p>Estadístico de prueba</p> <p>Valor de la distribución - + Valor de la distribución</p>	<p>Rechazo de $H_0 = \alpha$</p> <p>Aceptación de H_0</p> <p>Estadístico de prueba</p> <p>Valor de la distribución - +</p>	<p>Aceptación de H_0</p> <p>Rechazo de $H_0 = \alpha$</p> <p>Estadístico de prueba</p> <p>- + Valor de la distribución</p>

Modelos probabilísticos compuestos

- Regresión lineal:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$



$$Y_i = \beta_0 + \sum \beta_p X_{pi} + \epsilon_i$$

- Tablas de contingencia:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}$$

- Análisis de la varianza:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

Tratamiento		
Nivel 1	Nivel 2	Nivel 3
4	5	2
5	4	1
2	7	2
6	8	3

- Análisis de covarianza:

$$y_{ij} = \mu + \tau_i + B(x_{ij} - \bar{x}) + \epsilon_{ij}$$



Tratamiento					
Nivel 1		Nivel 2		Nivel 3	
x	y	x	y	x	y
2	5	3	5	6	10
3	4	4	6	7	12
1	6	2	7	5	10
5	10	5	8	8	18

$$F = \frac{\text{variance between treatments}}{\text{variance within treatments}}$$



UNC

4. Conclusiones y Recomendaciones



Conclusiones y Recomendaciones

- La Estadística es el estudio de la **variabilidad** realizado mediante las herramientas aportadas por las **matemáticas**. Lo que se pretende es modelar la variabilidad y la incertidumbre.

Conclusiones y Recomendaciones

- La Estadística es el estudio de la **variabilidad** realizado mediante las herramientas aportadas por las **matemáticas**. Lo que se pretende es modelar la variabilidad y la incertidumbre.
- Es posible aproximarse al mundo real a través de los modelos probabilísticos que son **maquetas** de las poblaciones.

Conclusiones y Recomendaciones

- La Estadística es el estudio de la **variabilidad** realizado mediante las herramientas aportadas por las **matemáticas**. Lo que se pretende es modelar la variabilidad y la incertidumbre.
- Es posible aproximarse al mundo real a través de los modelos probabilísticos que son **maquetas** de las poblaciones.
- La Estadística es una orquesta formada por la suma de muchos instrumentos, de muchas técnicas que pretenden ayudarnos a **entender mejor la realidad**.

Conclusiones y Recomendaciones

- La Estadística es el estudio de la **variabilidad** realizado mediante las herramientas aportadas por las **matemáticas**. Lo que se pretende es modelar la variabilidad y la incertidumbre.
- Es posible aproximarse al mundo real a través de los modelos probabilísticos que son **maquetas** de las poblaciones.
- La Estadística es una orquesta formada por la suma de muchos instrumentos, de muchas técnicas que pretenden ayudarnos a **entender mejor la realidad**.
- De estas técnicas se obtiene **mayor provecho** si se **unifican**, si se **comprenden los fundamentos**, si se **llega a las ideas que hay detrás** de todas ellas como conjunto.



Conclusiones y Recomendaciones

- La Estadística es el estudio de la **variabilidad** realizado mediante las herramientas aportadas por las **matemáticas**. Lo que se pretende es modelar la variabilidad y la incertidumbre.
- Es posible aproximarse al mundo real a través de los modelos probabilísticos que son **maquetas** de las poblaciones.
- La Estadística es una orquesta formada por la suma de muchos instrumentos, de muchas técnicas que pretenden ayudarnos a **entender mejor la realidad**.
- De estas técnicas se obtiene **mayor provecho** si se **unifican**, si se **comprenden los fundamentos**, si se **llega a las ideas que hay detrás** de todas ellas como conjunto.
- La Estadística es mejor instrumento, se aplica con más eficacia, cuando todas sus técnicas se ponen en juego ante un problema real concreto, cuando es una **orquesta** que suena como un **único instrumento!!!**





UNC

Gracias!!!

