

Estadística General

Tema 1: Introducción a la Estadística

Definición de Estadística, Método Científico, Población y Muestra.



Prof. José G. Páez

¿Para qué sirve la estadística?

- La Ciencia se ocupa en general de fenómenos observables
- La Ciencia se desarrolla observando hechos, formulando leyes que los explican y realizando experimentos para validar o rechazar dichas leyes
- Los modelos que crea la ciencia son de tipo determinista o **aleatorio (estocástico)**
- La **Estadística** se utiliza como **tecnología al servicio** de las ciencias donde la variabilidad y la incertidumbre forman parte de su naturaleza
- “La **Bioestadística** [...] enseña y ayuda a investigar en todas las áreas de las **Ciencias de la Vida donde la variabilidad** no es la excepción sino la **regla**”
Carrasco de la Peña (1982)

Antecedentes



Su origen empieza posiblemente en la isla de Cerdeña, donde existen monumentos prehistóricos con evidencias claves. En China Confucio, en uno de sus clásicos "Shu-King" escrito hacia el año 550 a.C., nos narra cómo el Rey Yao en el año 2238 mandó hacer una estadística agrícola, industrial y comercial. También en los antiguos monumentos egipcios se encontraron interesantes documentos en los que se demuestra la sabia organización y administración de este pueblo. Roma, con su perfecta organización político, jurídica y administrativa; favoreció el desarrollo de la Estadística. La Estadística pasó así a ser la descripción cuantitativa de las cosas notables de un estado.

Definición

La Estadística es la Ciencia de la

- **Sistematización, recogida, ordenación y presentación** de los datos referentes a un fenómeno que presenta variabilidad o incertidumbre para su estudio metódico, con objeto de
- **deducir las leyes** que rigen esos fenómenos,
- y poder de esa forma hacer previsiones sobre los mismos, tomar **decisiones** u obtener **conclusiones**.

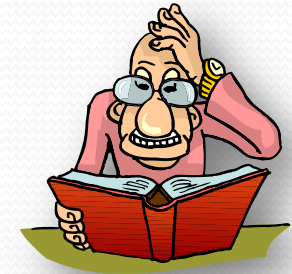
Descriptiva

Probabilidad

Inferencia

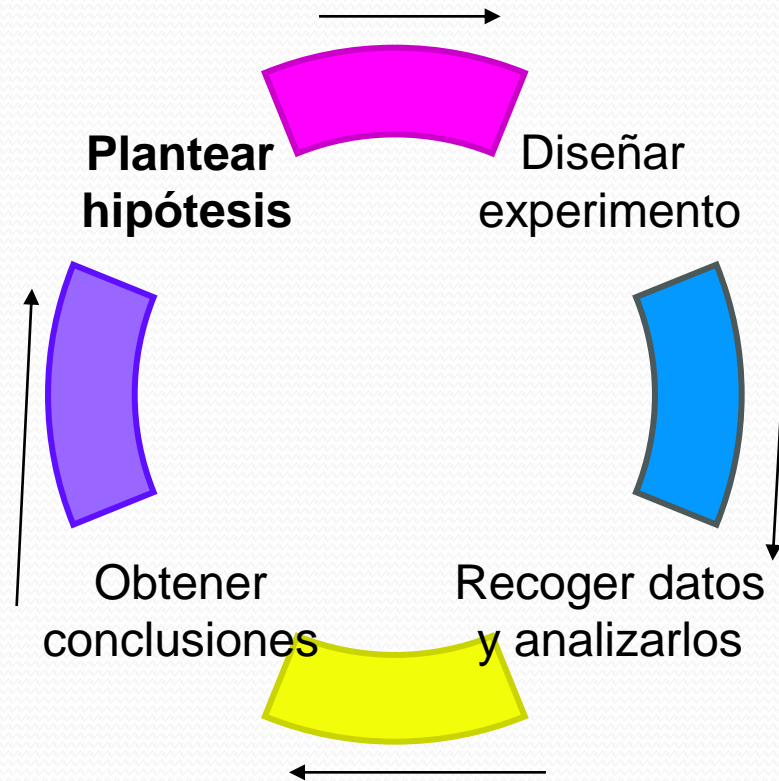
Pasos en un estudio estadístico

- Plantear **hipótesis** sobre una **población**
 - Los fumadores tienen “**más bajas**” laborales que los no fumadores
 - ¿En qué sentido? ¿Mayor número? ¿Tiempo medio?
- Decidir qué datos recoger (diseño de experimentos)
 - Qué individuos pertenecerán al estudio (**muestras**)
 - Fumadores y no fumadores en edad laboral.
 - Criterios de exclusión ¿Cómo se eligen? ¿Descartamos los que padecen enfermedades crónicas?
 - Qué datos recoger de los mismos (**variables**)
 - Número de bajas
 - Tiempo de duración de cada baja
 - ¿Sexo? ¿Sector laboral? ¿Otros factores?
- Recoger los datos (**muestreo**)
 - ¿Estratificado? ¿Sistemáticamente?
- **Describir** (resumir) los datos obtenidos
 - tiempo medio de baja en fumadores y no (**estadísticos**)
 - % de bajas por fumadores y sexo (**frecuencias**), gráficos,...
- Realizar una **inferencia** sobre la población
 - Los fumadores están de baja al menos 10 días/año más *de media* que los no fumadores.
- Cuantificar la confianza en la inferencia
 - **Nivel de confianza** del 95%
 - **Significación del contraste**: $p=2\%$



No tienes que entenderlo (aún)

Método científico y estadística



Población y muestra

- **Población** (*population*) es el conjunto sobre el que estamos interesados en obtener conclusiones (hacer inferencia).
 - Normalmente es demasiado grande para poder abarcarlo.
- **Muestra** (*sample*) es un subconjunto suyo al que tenemos acceso y sobre el que realmente hacemos las observaciones (mediciones)
 - Debería ser “representativo”
 - Esta formado por miembros “seleccionados” de la población (individuos, unidades experimentales).

