

# TEMA I

## INTRODUCCIÓN

1. Datos e Incertidumbre
2. Estadística Descriptiva y Teoría de la Probabilidad
3. Presentación de las Herramientas Informáticas: Excel y SPSS
4. Utilidad de la Estadística
5. Variables y datos
6. Fuentes de Información estadística: Cómo se obtienen los datos

# TEMA I (Cont.)

## Datos e incertidumbre

2

### Estadística

Definición no rigurosa Ciencia que estudia:

- ▣ cómo debe emplearse la información
- ▣ cómo actuar en situaciones prácticas con incertidumbre

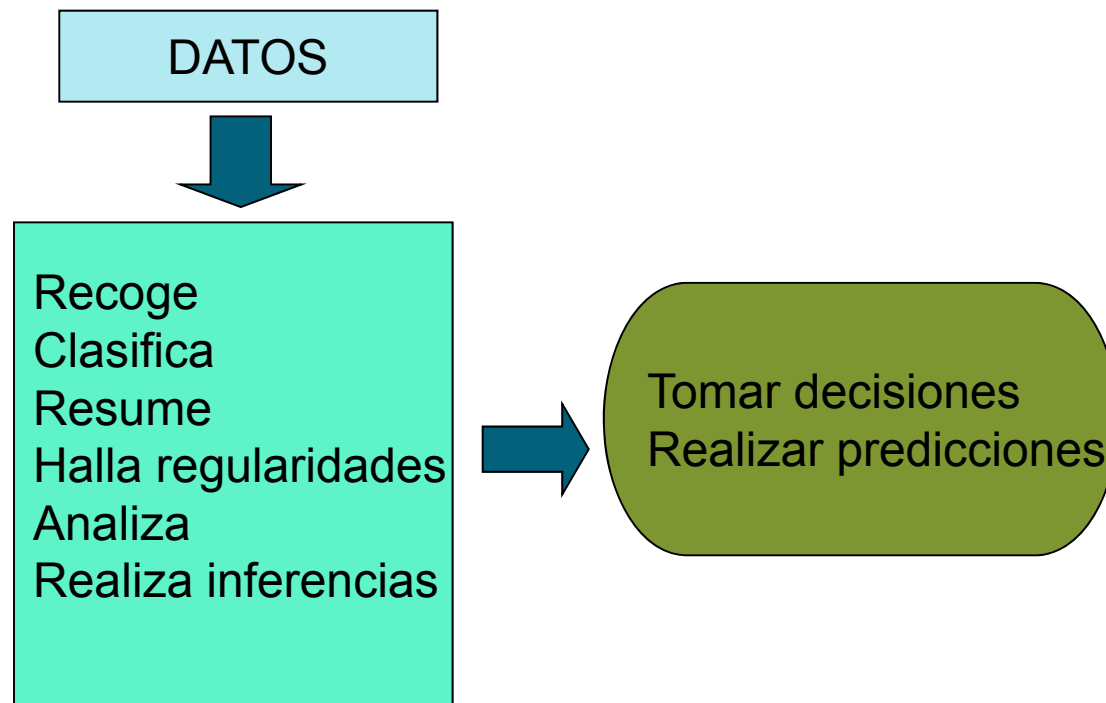
# TEMA I (Cont.)

## Datos e incertidumbre

3

### Definición rigurosa

Ciencia que estudia:



# TEMA I (Cont.)

## Estadística Descriptiva y Teoría de la Probabilidad

4

### □ Clasificación de la Estadística:

#### **Descriptiva:**

- ▣ Describe, analiza y representa un conjunto de datos
- ▣ Utiliza métodos numéricos y gráficos para resumir y presentar la información contenida en los datos

# TEMA I (Cont.)

## Estadística Descriptiva y Teoría de la Probabilidad

5

### Inferencial:

- ▣ Se apoya en el cálculo de probabilidades
- ▣ Se apoya en datos muestrales para efectuar estimaciones, decisiones, predicciones y otras generalizaciones sobre un conjunto mayor de datos (población)

# TEMA I (Cont.)

## Población y Muestra

6

- **Unidad:** El objeto más pequeño o individuo que puede ser investigado, la fuente de la información básica
- **Población o Universo:** Un grupo muy grande, infinito o no, de unidades con alguna característica en común
- **Muestra:** Subconjunto de unidades de la población o universo

# TEMA I (Cont.)

## Estadística Descriptiva y Teoría de la Probabilidad

7

- Etapas de un estudio estadístico:
  - 1- Planteamiento del problema
    - Objetivo
    - Población
  - 2- Recogida de la información muestral
  - 3- Análisis descriptivo

# TEMA I (Cont.)

## Estadística Descriptiva y Teoría de la Probabilidad

8

4- Inferencia

5- Diagnóstico

- Validez de los supuestos
- Interpretación
- Conclusiones



# TEMA I (Cont.)

## Estadística Descriptiva y Teoría de la Probabilidad

9

- Aplicaciones:
  - ▣ Recursos humanos
  - ▣ Marketing
  - ▣ Producción
  - ▣ Finanzas

# TEMA I (Cont.)

## Presentación Herramientas Informáticas

10

Desarrollo de las tecnologías de la información (informática) permite infinitas posibilidades de tratar la información con técnicas estadísticas

EXCEL

SPSS (Statistical Package for the Social Sciences)

# Utilidad de la estadística

11

Todo el que toma **decisiones** debe hacerlo bajo condiciones de **incertidumbre**,  
en mayor o menor grado.

Con el aumento de la competitividad, la administración de las  
instituciones requiere tomar decisiones cada vez con una mayor base de  
conocimiento para así **reducir la incertidumbre**.

## Datos no faltan

12

Aumento en la eficiencia de los computadores



Aumento en la capacidad de almacenar datos.

Pero los **datos** por si solos no sirven ...

... si la institución no es capaz de extraer **información** de ellos.

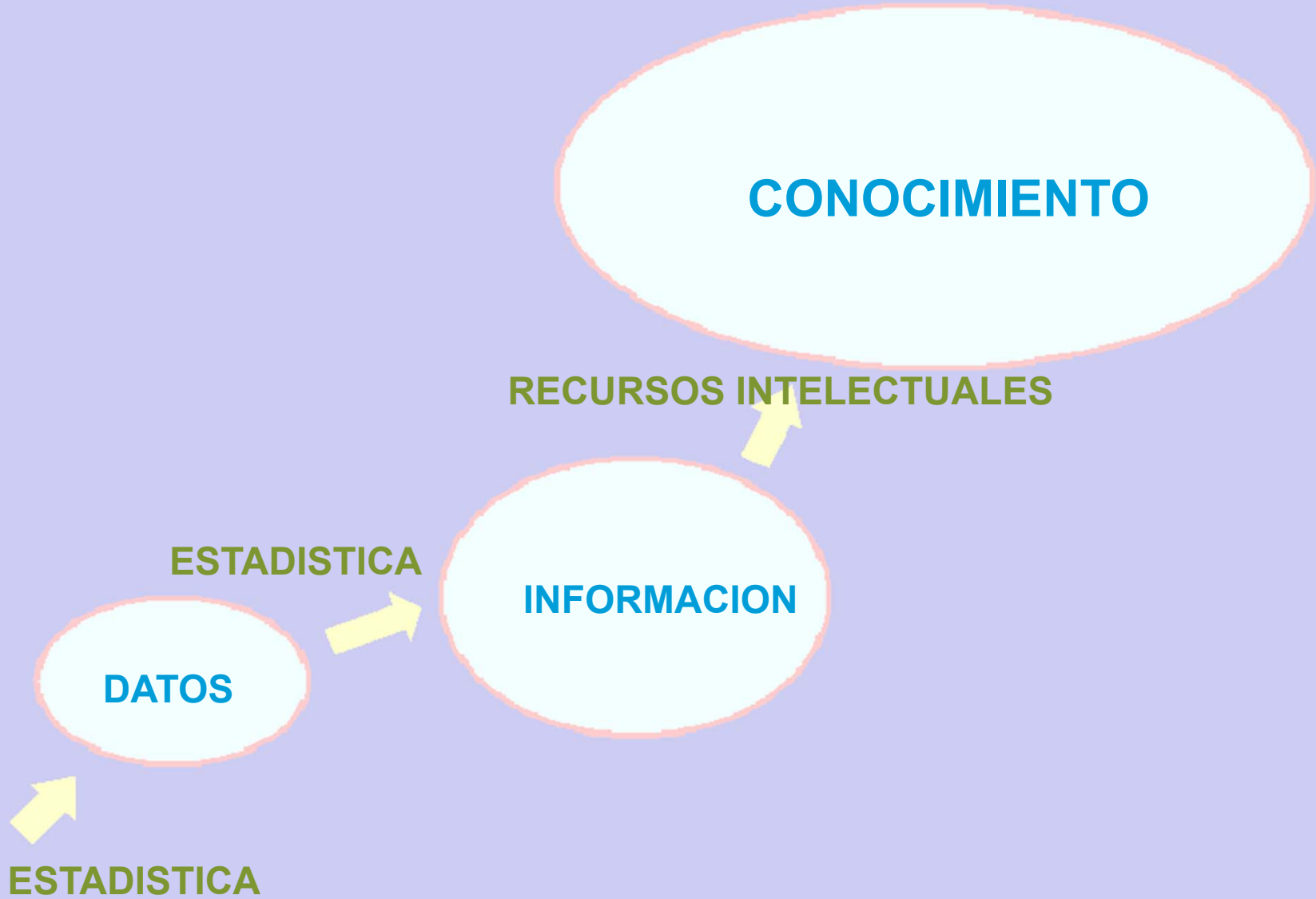
Y con los recursos necesarios, como competencias y experiencia, se puede convertir esta información en **conocimiento** ...

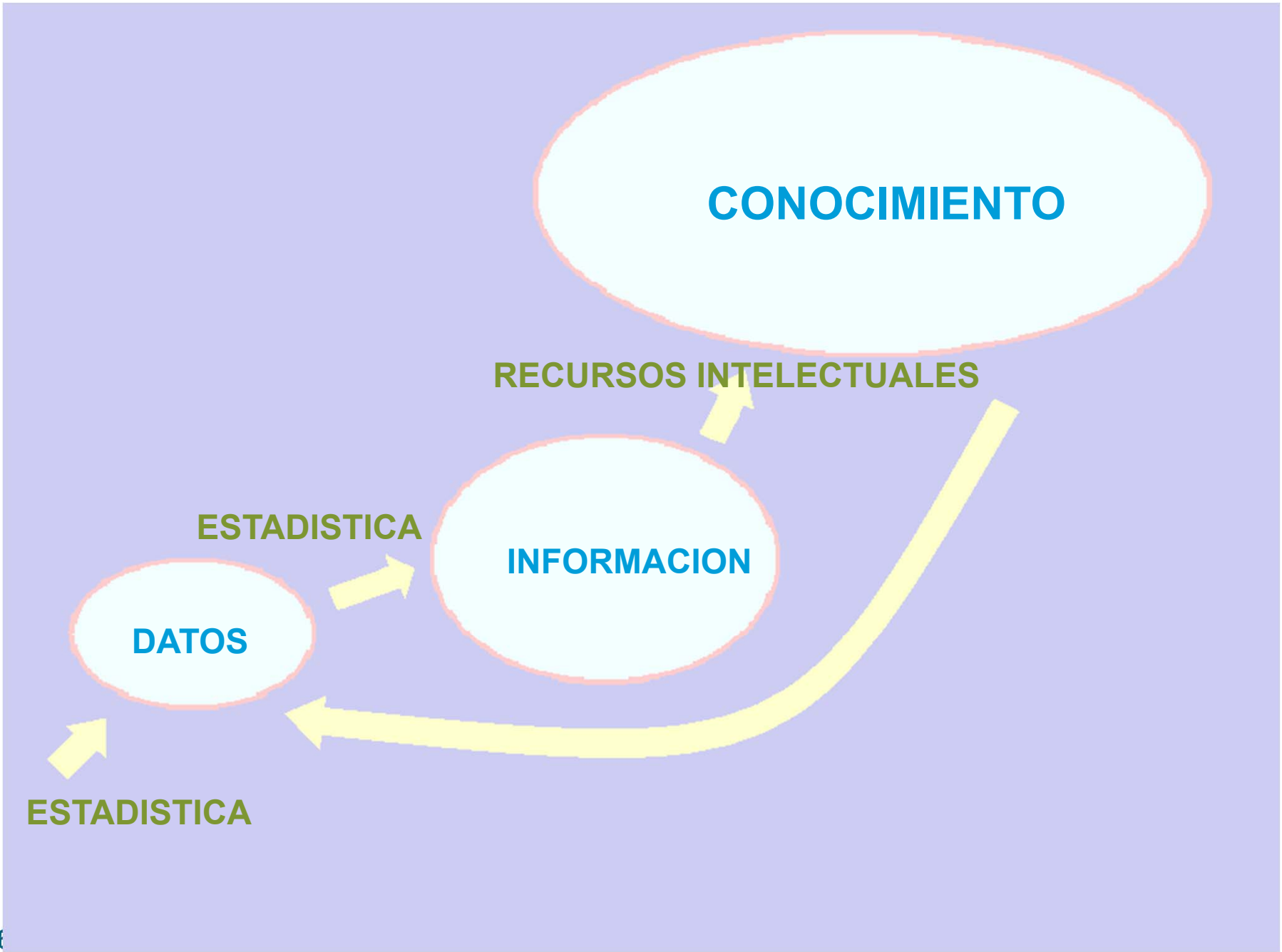
... para tomar las **decisiones** estratégicas, tácticas y operativas.

14

Las etapa de **producción de datos** y de **extracción de información**, requiere de **métodos**, **técnicas** y **herramientas** de análisis.

El desarrollo de estos métodos, técnicas y herramientas se encuentra en una **ciencia** que se llama **ESTADISTICA**.

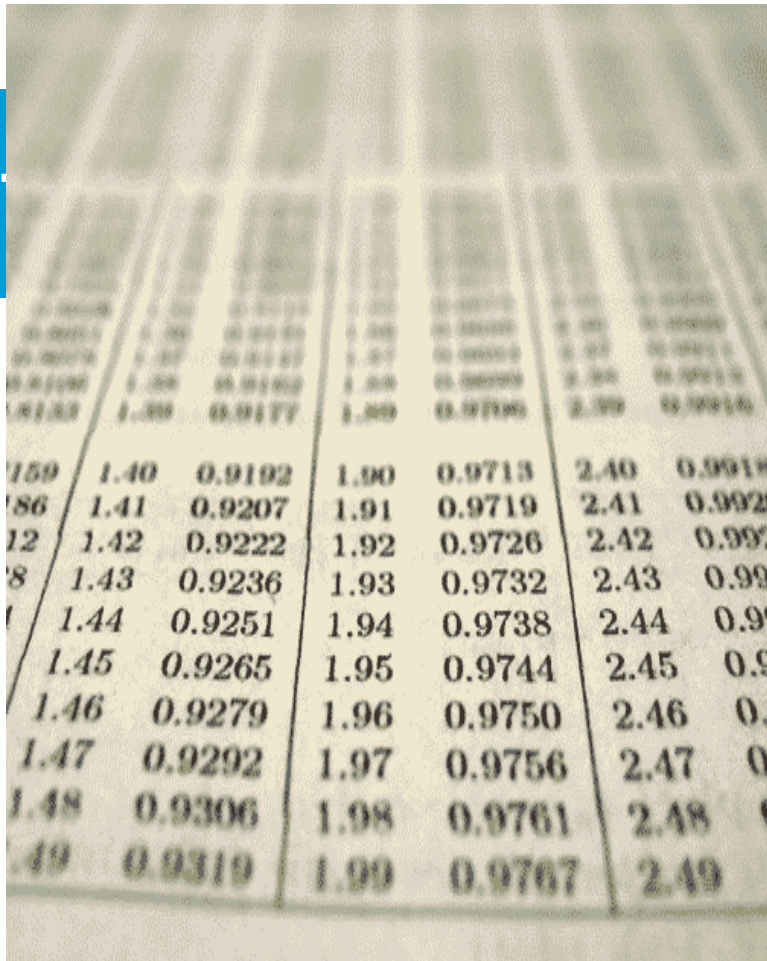






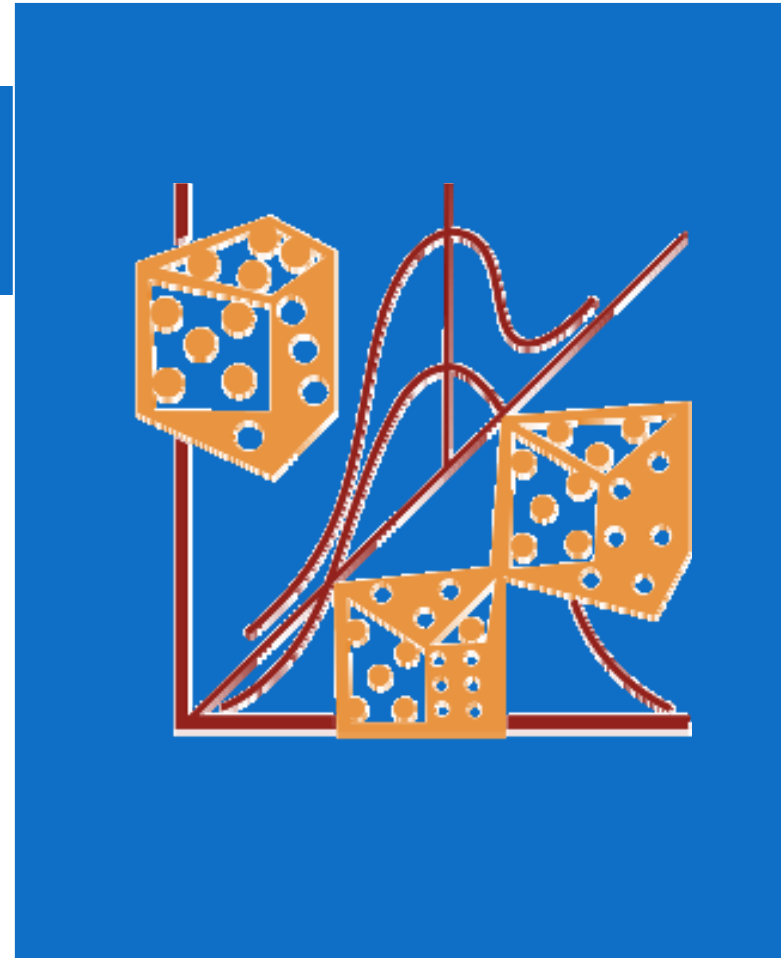
# ESTADISTICA

1



159	1.40	0.9192	1.90	0.9713	2.40	0.9918
186	1.41	0.9207	1.91	0.9719	2.41	0.9920
12	1.42	0.9222	1.92	0.9726	2.42	0.9920
8	1.43	0.9236	1.93	0.9732	2.43	0.9920
1	1.44	0.9251	1.94	0.9738	2.44	0.9920
	1.45	0.9265	1.95	0.9744	2.45	0.9920
	1.46	0.9279	1.96	0.9750	2.46	0.9920
	1.47	0.9292	1.97	0.9756	2.47	0.9920
	1.48	0.9306	1.98	0.9761	2.48	0.9920
	1.49	0.9319	1.99	0.9767	2.49	0.9920

Datos



Probabilidad

En años recientes ha habido un desarrollo de la Estadística de la mano del computador. Es así como se ha desarrollado la llamada

### **Computación Estadística.**

18

El computador permite realizar operaciones **repetitivas** a alta velocidad.

Cuando un problema es demasiado complejo como para encontrar soluciones analíticas, se pueden desarrollar métodos basados en la repetición.

Por ejemplo, métodos que buscan **aproximaciones a las soluciones óptimas** mediante la repetición.

Estos métodos dan origen a algoritmos computacionales que requieren mucho procesamiento.

Entre estos métodos, están los más conocidos, como el  
Bootstrap, de B. Efron.

20

El Algoritmo EM, de Dempster, Laird y Rubin.

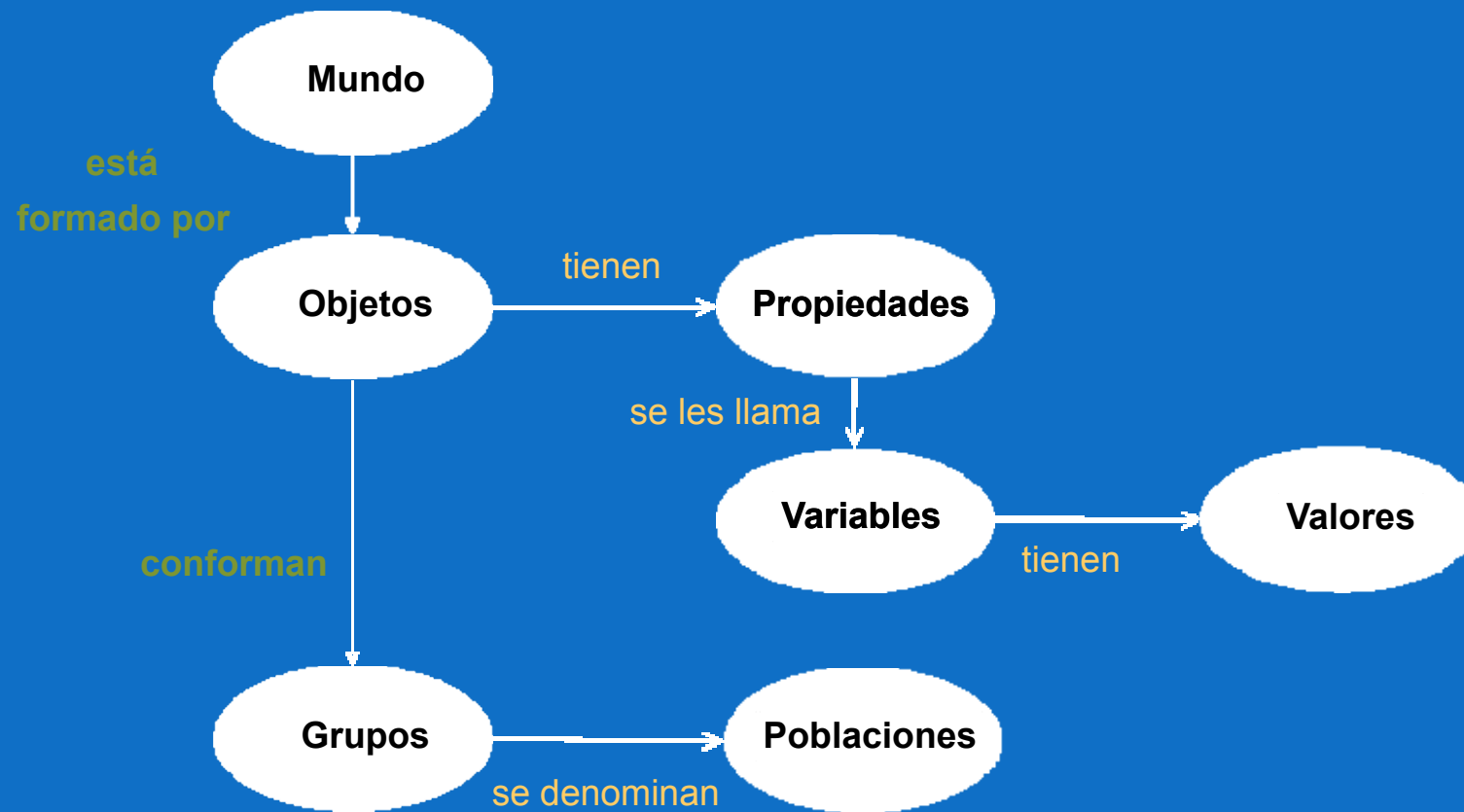
El jackknife, la validación cruzada (cross validation), el Gibbs  
sampling.

Y muchos más, que aún están por descubrirse.....

21



Los datos son una  
representación del mundo



# ELEMENTOS, VARIABLES Y DATOS



Variable

Dato

Elemento



# Variables y Datos

25

- **Variable:** Es una característica que puede tomar diferentes valores (no necesariamente numéricos) en las distintas unidades de observación. Por ejemplo, edad, sexo, tiempo de evolución de la diabetes, tipo de tratamiento administrado, etc. Las propiedades o *características susceptibles de tomar distintos valores o intensidades es lo que se conoce con el nombre de variables.*
- **Dato:** Son los valores observados o medidos de una variable y constituyen la información disponible.

# TIPOS DE VARIABLES (¿QUÉ EXPRESAN?)

- Categóricas (cualitativas)
  - Expresan clasificación
  - Ej. Estado civil
- Numéricas (cuantitativas)
  - Expresan valor
  - Discretos: Número de ítems
  - Continuos: Tiempo de espera

# NIVELES DE MEDICIÓN



**Escala de Razón**

Ej. Ingresos, tiempo

**Escala de Intervalos**

Ej. Escalas de temperatura, coeficiente de inteligencia

**Escala Ordinal**

Ej. Nivel Educativo, rankings

**Escala Nominal**

Ej. Sexo, Tipo de cliente

# UN ESQUEMA ÚTIL

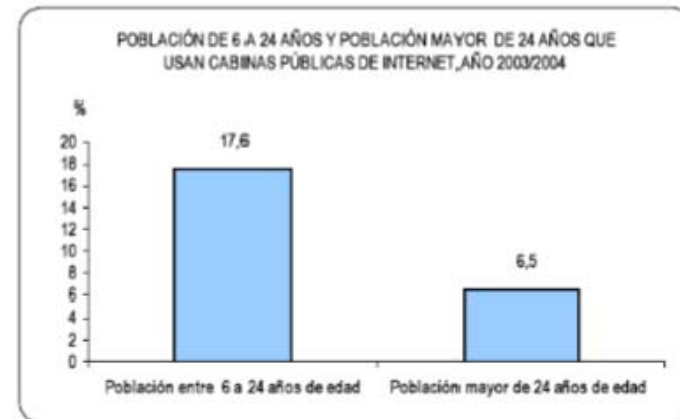
28



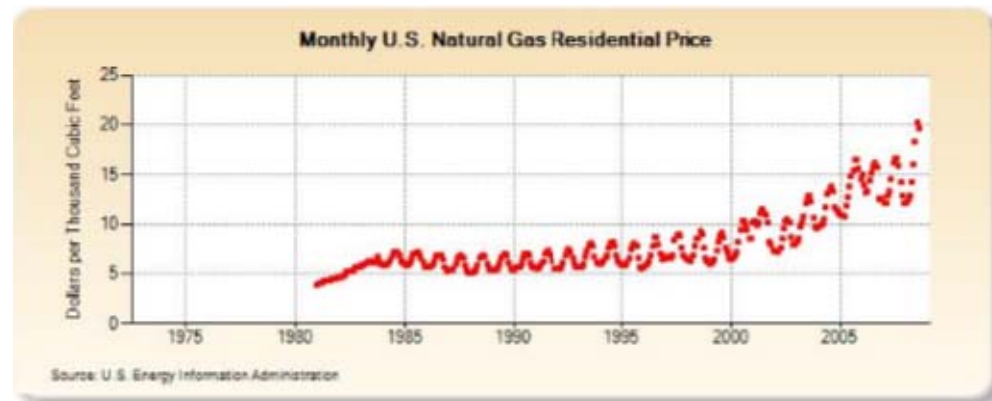
# Otra Clasificación de los Datos

29

- Datos de Corte Transversal
- Series de Tiempo
- Panel de datos



Fuente: INEI-Encuesta nacional de hogares, año 2003/2004.



# FUENTES DE INFORMACIÓN ESTADÍSTICA: Cómo se obtienen los datos

30

## Fuentes de datos



Datos de fuentes secundarias



Datos de fuentes primarias

# DATOS DE FUENTES PRIMARIAS



- Se recogen específicamente para el análisis realizado
- Su principal ventaja es la concordancia con nuestros objetivos
- Su principal desventaja es el coste

# DATOS DE FUENTES SECUNDARIAS



- Fueron compilados con anterioridad
- Su principal ventaja es el coste
- Su debilidad es que pueden no corresponder con nuestros objetivos