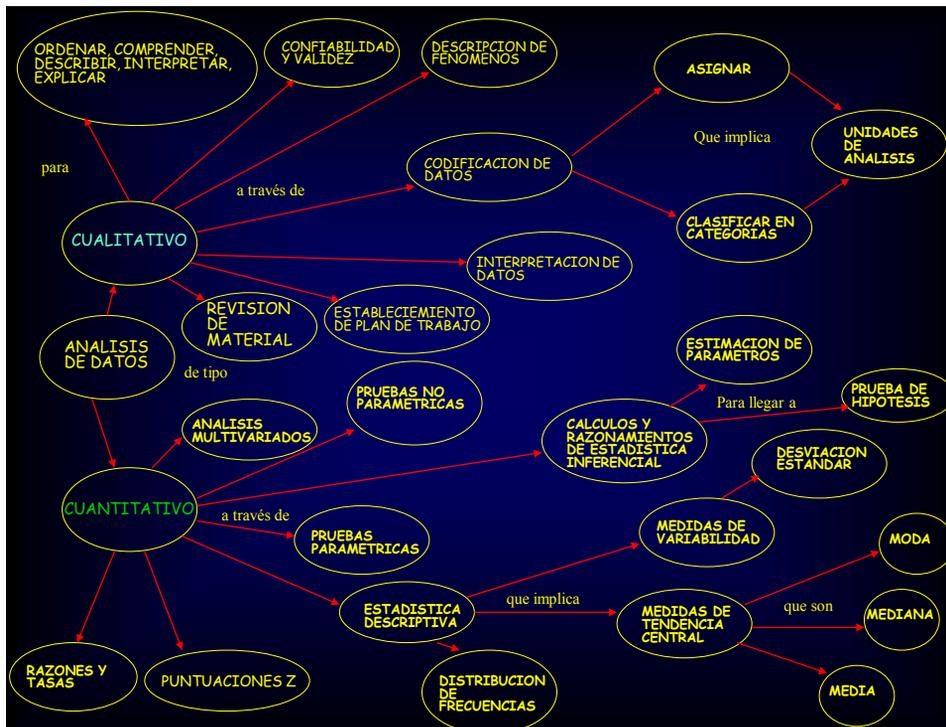


UNIVERSIDAD PERUANA DEL ORIENTE

SISTEMATIZACION DE BASE DE DATOS EN INVESTIGACION CIENTIFICA

• Jaime Alberto Mori Castro





¿Qué es Estadística?

Estadísticas: Es la ciencia de coleccionar, organizar, presentar, analizar, e interpretar datos numéricos con el propósito de una efectiva toma de decisiones.



Tipos de Estadísticas:

"coleccionar, organizar, presentar, e interpretar datos numéricos" es comúnmente conocida como **Estadística Descriptiva**.

Estadística Inferencial. Consiste en conocer algo acerca de una población basándose en una muestra tomada de esa población.

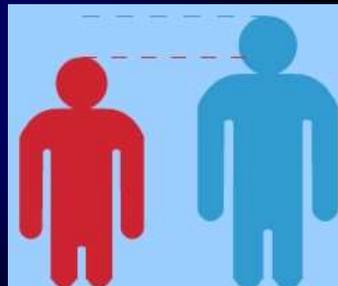
Una **población** puede consistir de **individuos**, tales como todos los estudiantes de la UNAS, todos los estudiantes que estudian Ingeniería.



Una población puede también consistir de **objetos**, tales como todos los peces en un estanque.



También puede consistir un grupo de **medidas**, tales como las estaturas del equipo de Fútbol del "Sport Loreto".



Población: Es una colección de todos los posibles individuos, objetos o medidas de interés.

Para inferir algo acerca de una población, usualmente tomamos una **muestra** de ella.



Tipos de Variables

Variable Cualitativa

Es cuando la variable es no numérica, o también llamada atributo (Ej. lugar de nacimiento, religión, color de ojos, etc.)

Variable Cuantitativa

Cuando la variable puede expresarse en forma numérica, (Ej. el saldo de una cuenta bancaria, la duración de una batería, la velocidad de los motocarros).



Variables Discretas

Es cuando cada una de las variables pueden ser números enteros (Ej. 1,2,3, etc.) pero no puede valer números fraccionarios (Ej. 1.5). Típicamente, las variables discretas resultan de un conteo.



Variables Continuas

Asumen valores dentro de un rango específico (Ej. El tiempo que se toma en viajar de Yarinacocha a Pucallpa, etc.). Típicamente, las variables continuas resultan de medir algo.

$$0.25 = \frac{25}{100}$$

NIVELES DE MEDICION

Nivel Nominal. Se establecen categorías que no implican un orden específico. Ejemplo, la categoría sexo con dos niveles, masculino (M) y femenino (F), los respondientes solo tienen que señalar su género, no se requiere de un orden real.

Nivel Ordinal. Son similares al nivel nominal con la diferencia de que en el nivel ordinal las categorías indican que unas son más que las otras.

Nivel Intervalar Puede ser nominal y ordinal. La escala se da a variables continuas pero carece de cero absoluto. Ejm. Una persona que en un examen que obtiene cero no significa que carezca de conocimientos, el cero es arbitrario por que sigue existiendo la característica medida.

Nivel Relativo o de razón Incluye los tres anteriores (nominal, ordinal e intervalo). Tiene un cero absoluto, es decir, en el punto cero no existe la característica que se mide. Las variables de ingreso, edad, número de hijos, etc. son ejemplos de este tipo de escala. El nivel de medición de razón se aplica tanto a variables continuas como discretas.



**ANALISIS CUANTITATIVO
O CUALITATIVO?**

La elección del tipo de análisis depende de los datos que hayamos recolectado

otra vez esa sensación



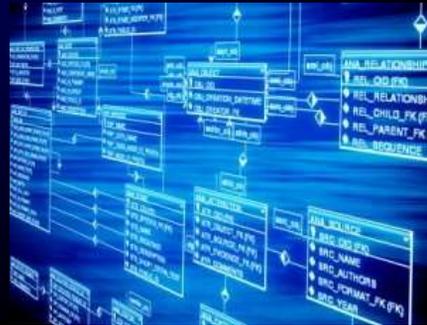
**ANALISIS CUALITATIVO
DE DATOS**

La elección del tipo de análisis depende de los datos que hayamos recolectado

ANALISIS CUANTITATIVO DE LOS DATOS (ANALISIS ESTADISTICO)

PROCEDIMIENTOS

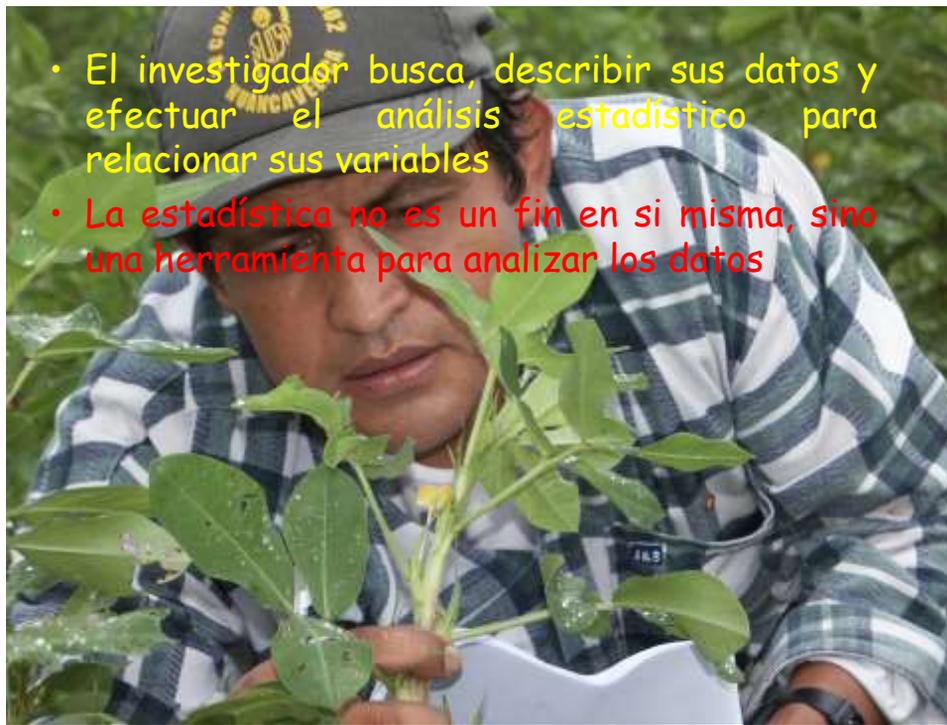
Codificado los datos, transferidas a una matriz y guardado en un archivo, se procede a analizarlos. Ya nadie lo hace de forma manual, en especial si hay un volumen considerado de datos.



¿QUE ANALISIS PUEDEN EFECTUARSE EN LOS DATOS?

Los análisis dependen de tres factores:

- El *nivel de medición* de las variables
- La manera como se hayan formulado la *hipótesis*
- El *interés del investigador*



- El investigador busca, describir sus datos y efectuar el análisis estadístico para relacionar sus variables
- La estadística no es un fin en si misma, sino una herramienta para analizar los datos

ANÁLISIS DE LOS DATOS

Procedimientos para efectuar análisis estadísticos por computador.

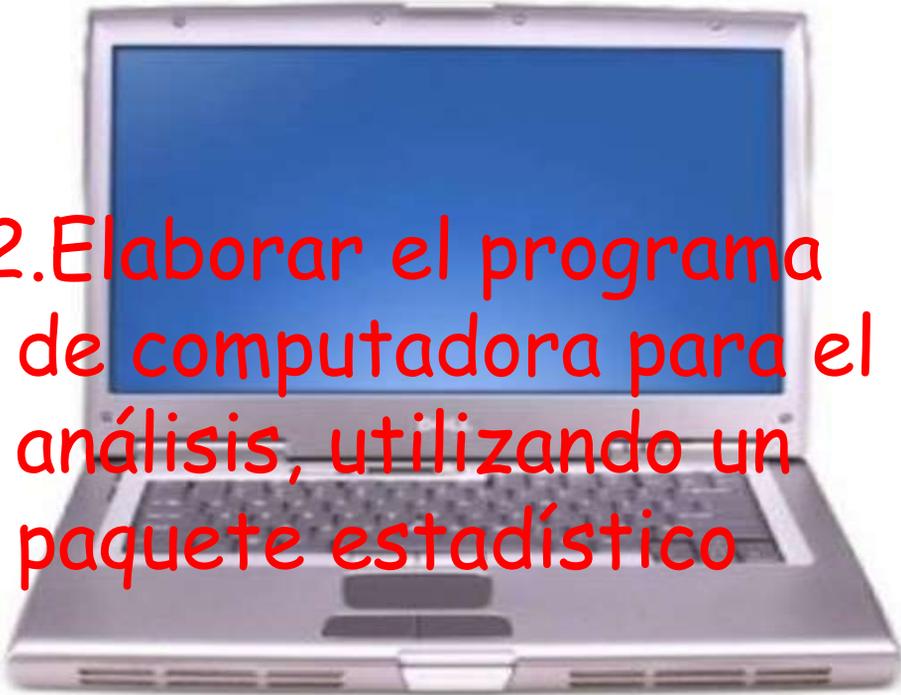
Se analizaran y se ejemplificaran las pruebas utilizados en la ciencia, incluyendo estadísticas descriptivas, análisis paramétricos, no paramétricos y multivariados.



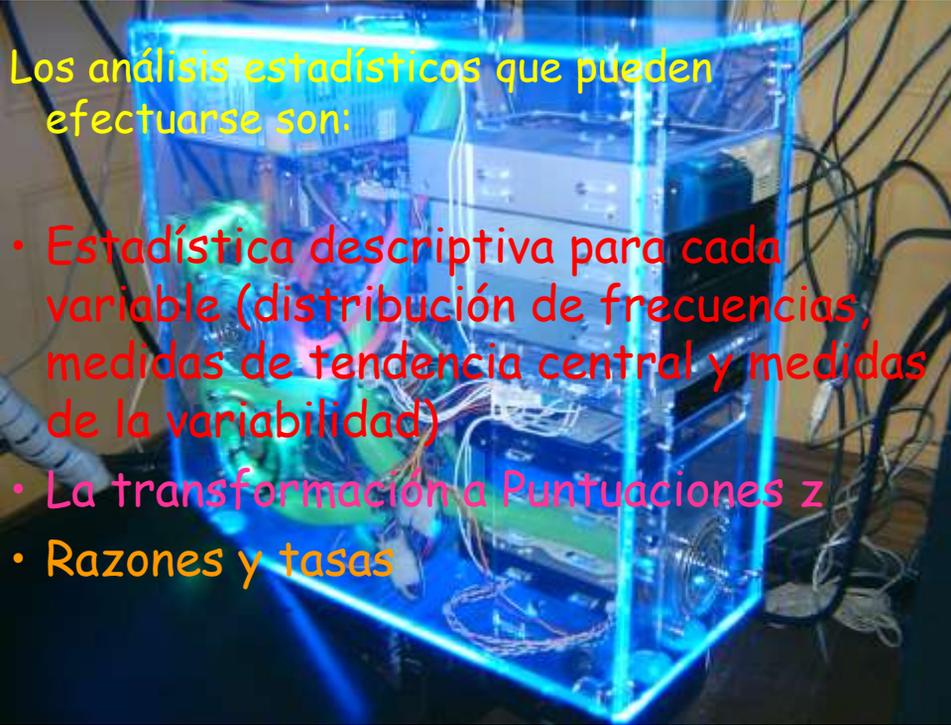
La mayor parte de estos análisis se centrara en el uso y la interpretación de la prueba mas que en el procedimiento de calcular estadísticas, ya que actualmente los análisis se realizan con la ayuda del computador no manualmente.



En la actualidad las formulas ayudan a entender los conceptos estadísticos, pero no a calcular estadísticas

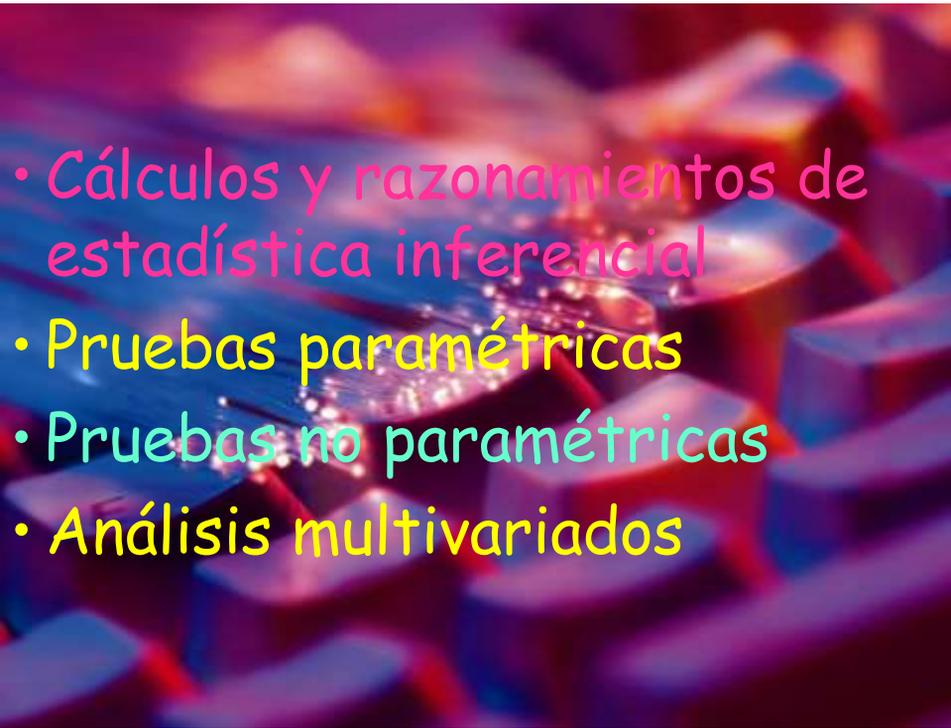


2. Elaborar el programa de computadora para el análisis, utilizando un paquete estadístico



Los análisis estadísticos que pueden efectuarse son:

- Estadística descriptiva para cada variable (distribución de frecuencias, medidas de tendencia central y medidas de la variabilidad)
- La transformación a Puntuaciones z
- Razones y tasas

- 
- Cálculos y razonamientos de estadística inferencial
 - Pruebas paramétricas
 - Pruebas no paramétricas
 - Análisis multivariados

Los analisis o pruebas estadísticas paramétricas más utilizadas son:

PRUEBA	TIPOS DE HIPOTESIS
Coefficiente de correlación de Pearson	Correlacional
Regresión Lineal	Correlacional / Causal
Prueba "t"	Diferencia de grupos
Contraste de la diferencia de proporciones	Diferencia de grupos
Análisis de varianza (ANOVA):	Diferencia de grupos / Causal
Análisis de covarianza (ANCOVA)	Correlacional / Causal

En todas las pruebas estadísticas paramétricas las variables están medidas en un nivel por intervalos o razón.

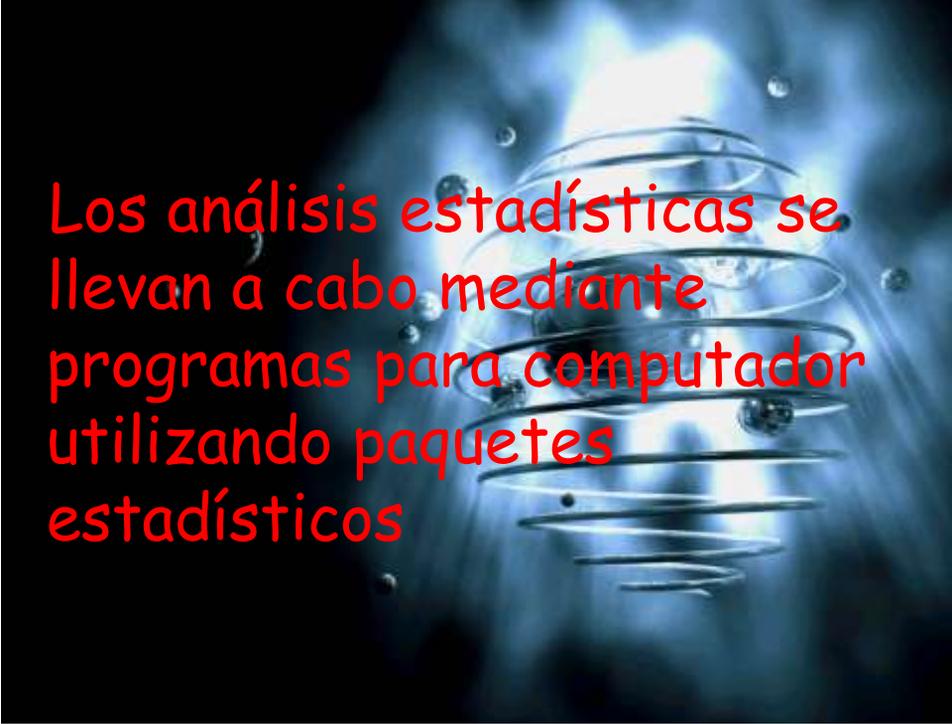
Los analisis o pruebas estadísticas no paramétricas más utilizadas son:

PRUEBA	TIPOS DE HIPOTESIS
Ji cuadrada (χ^2).	Diferencias de grupos para establecer correlación.
Coefficientes de correlación e independencia para tabulaciones cruzadas: Phi, C de Pearson, V de Cramer, Lambda, Gamma, Taub, D de Sommers y Eta.	Correlacional
Coefficientes Spearman y Kendall.	Correlacional

Las pruebas no paramétricas se utilizan con variables nominales u ordinales.

Los analisis multivariados más utilizados son:

PRUEBA	VARIABLES INVOLUCRADAS Y NIVELES DE MEDICION	TIPOS DE HIPOTESIS
Regresión múltiple	Una dependiente (intervalos o razón) y dos o más independientes (cualquier nivel de medición).	Correlacional / Causal $X_1 \quad X_2 \quad X_k \quad Y$
Análisis Lineal path	Varias, secuencia causal (cualquier nivel de medición).	Correlacional / Causal $X \rightarrow W \rightarrow L$ $Y \quad Z \rightarrow N \rightarrow H$
Análisis de factores	Varias (intervalos o razón).	Correlacional / Causal $X_1 \rightarrow X_2 \rightarrow X_k$
Análisis multivariado de varianza (MANOVA) y correlación canónica.	Varias independientes y varias dependientes (intervalos o razón).	$X_1 \rightarrow Y_1$ $X_2 \rightarrow Y_2$ $X_k \rightarrow Y_k$
Análisis discriminante	Varias independientes (intervalos o razón) y una dependiente (nominal u ordinal).	$X_1 \quad X_2 \quad X_3 \quad Y$



Los análisis estadísticas se llevan a cabo mediante programas para computador utilizando paquetes estadísticos



GRACIAS