

# Seminario de Investigación (Tesis III- segunda Parte)

## Maestría en Ciencias: Administración (MBA)

DBA(c) Jhon Monroy Barrios

Universidad Nacional de San Agustín de Arequipa

2022



# UNSA

UNIVERSIDAD NACIONAL DE SAN AGUSTÍN DE AREQUIPA

# 1 Motivación



# Motivación



# ¿Qué es el método científico?





Consiste en manipular intencionalmente la variable independiente de un modelo para observar y medir sus efectos en la variable dependiente.



- Diseño pre-experimental
- Diseño cuasi experimental
- Diseño experimental verdadero



Es aquel que se realiza sin manipular deliberadamente las variables. Se basa fundamentalmente en la observación de fenómenos tal y como se dan en su contexto natural para después analizarlo.



## Diseño no experiemntal o Ex post facto

Se basa en categorías, conceptos, variables, sucesos, comunidades o contextos que ya ocurrieron o se dieron sin la intervención directa del investigador. Es por esto que también se le conoce como investigación **ex post facto** (hecho y variables que ya ocurrieron), al observar variables y relaciones entre estas en su contexto.



## Diseño no experiemntal o Ex post facto

En estos tipos de investigación no hay condiciones ni estímulos a los cuales se expongan los sujetos del estudio. los sujetos son observados en su ambiente natural y dependiendo en que se va a centrar la investigación, existen diferentes tipos de diseños en las que se puede basar el investigador.



# Tipos de diseño no experimentales



Transeccionales



Longitudinales



## Diseños transaccionales

Cuando la investigación se centra en analizar cuál es el nivel o estado de una o diversas variables en un momento dado o bien en cuál es la relación entre un conjunto de variables en **un punto en el tiempo**, se utiliza el diseño transaccional. en este tipo de diseño se recolectan datos en un solo momento, en un tiempo único.



# Diseños transaccionales

Pueden ser :

- Exploratorio
- Descriptivo
- Correlacional

## Diseño longitudinal

Se emplea cuando el interés del investigador es analizar **cambios a través del tiempo** en determinadas variables o en relaciones entre estas. Recolectan datos a través del tiempo en puntos o periodos específicos, para hacer inferencias respecto al cambio, sus determinantes y consecuencias.



# Diseño longitudinal

Pueden ser:

- De tendencia
- Evolutivo
- De Panel



## Confiabilidad

Grado en que un instrumento produce resultados consistentes y coherentes. Es decir en que su aplicación repetida al mismo sujeto u objeto produce resultados iguales. Kerlinger (2002).

### Ejemplo:

Si se midiera en este momento la temperatura ambiental usando un termómetro y este indicara que hay 22°C, un minuto más tarde 5°C, tres minutos después 40°C; dicho termómetro no sería confiable.

Intervalo al que pertenece el coeficiente alfa de Cronbach	Valoración de la fiabilidad de los ítems analizados
[0 ; 0,5[	Inaceptable
[0,5 ; 0,6[	Pobre
[0,6 ; 0,7[	Débil
[0,7 ; 0,8[	Aceptable
[0,8 ; 0,9[	Bueno
[0,9 ; 1]	Excelente

# Validez

Grado en el que un instrumento en verdad mide la variable que se busca medir.

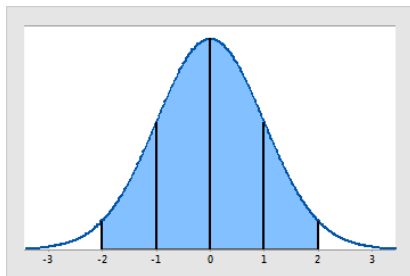
## Ejemplo:

Un instrumento válido para medir la inteligencia debe medir la inteligencia y no la memoria.

<b>0,53 a menos</b>	<b>Validez nula</b>
<b>0,54 a 0,59</b>	<b>Validez baja</b>
<b>0,60 a 0,65</b>	<b>Válida</b>
<b>0,66 a 0,71</b>	<b>Muy válida</b>
<b>0,72 a 0,99</b>	<b>Excelente validez</b>
<b>1.0</b>	<b>Validez perfecta</b>

## Distribucion normal

La distribución normal es una distribución con forma de campana donde las desviaciones estándar sucesivas con respecto a la media establecen valores de referencia para estimar el porcentaje de observaciones de los datos. Estos valores de referencia son la base de muchas pruebas de hipótesis, como las pruebas Z y t.



# Distribución normal

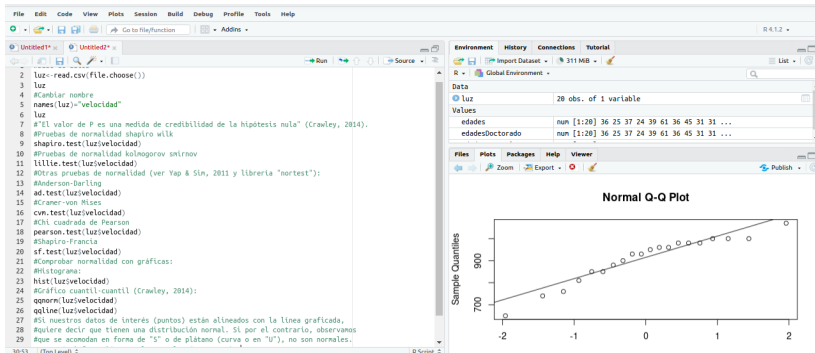
Los datos se pueden "distribuir"(esparcir) de diferentes maneras.



[www.mathsisfun.com/data/standard-normal-distribution.html](http://www.mathsisfun.com/data/standard-normal-distribution.html)

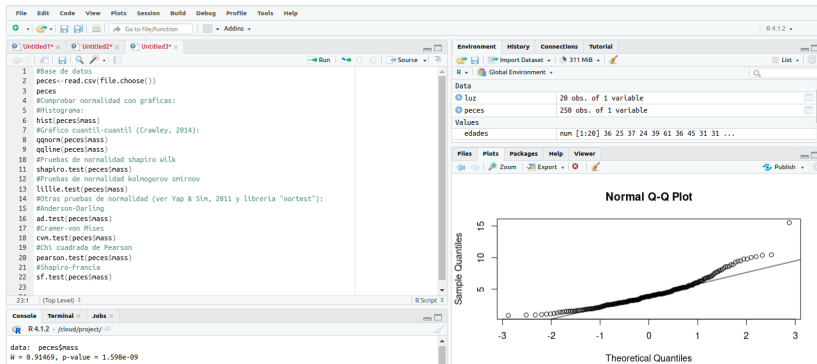
# Ejemplo 1

## Base de datos de luz



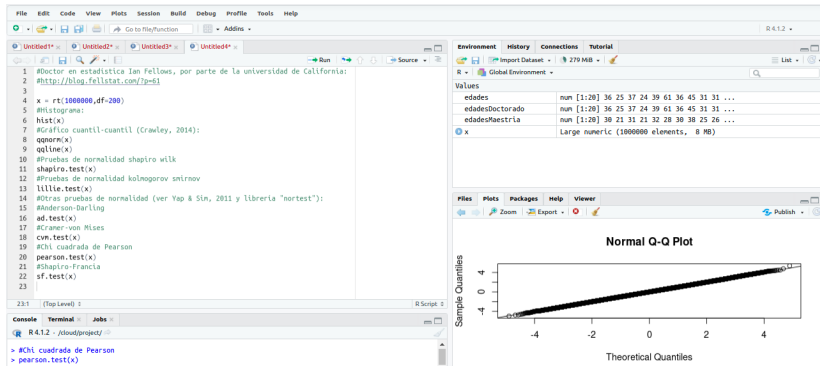
## Ejemplo 2

### Base de datos de pescados



## Ejemplo 3

### Base de datos generada aleatoriamente



# Homocedasticidad varianzas

```

File Edit Code View Plots Session Build Debug Profile Tools Help
+ + + + + Go to file/function + + Addins +
Untitled1* x Untitled2* x ozono x Untitled3* x rio x Untitled4* x
Run + + + Source
3 #Base de datos
4 ozono = read.table("./f.test.data.txt", header=T)
5 jardines = read.table("./gardens.txt", header=T)
6 refugio = read.table("./refuge.txt", header=T)
7 #Varianza
8 ozono
9 var(ozono$jardinA)
10 var(ozono$jardinB)
11 #Grados de libertad (degrees of freedom, df)
12 length(ozono$jardinA)-1
13 length(ozono$jardinB)-1
14 #Prueba de F de Fisher (2 muestras)
15 var.test(ozono$jardinA, ozono$jardinB)
16 var(ozono$jardinA)/var(ozono$jardinB)
17 #Las varianzas del ozono entre los jardines A y B son iguales (F=1; P=1)
18 #Pruebas para más de 2 muestras.
19 #Observamos la base de datos
20 jardines
21 #Creamos los vectores de la variable de respuesta (ozono.jar) y de la variable explicativa (tipo.jar)
22 ozono.jar = c(jardines$gardenA, jardines$gardenB, jardines$gardenC)
23 tipo.jar = factor(rep(c("A","B","C"), c(10,10,10)))
24 #Queremos saber si la variación en la concentración de ozono difiere entre los 3 jardines.
25 #Prueba de Bartlett
26 bartlett.test(ozono.jar~tipo.jar)
27 #Las varianzas en la concentración de ozono entre los jardines son diferentes (K=16.758; P=0.0002)
28 #Prueba de Fligner-Killeen
29 fligner.test(ozono.jar~tipo.jar)
30 #Las varianzas son iguales (X^2=1.8; P=0.4)
31 #¿Por qué tenemos resultados diferentes?
32 #La prueba de Bartlett es sensible a los datos atípicos, la de Fligner-Killeen no (Crawley, 2013).

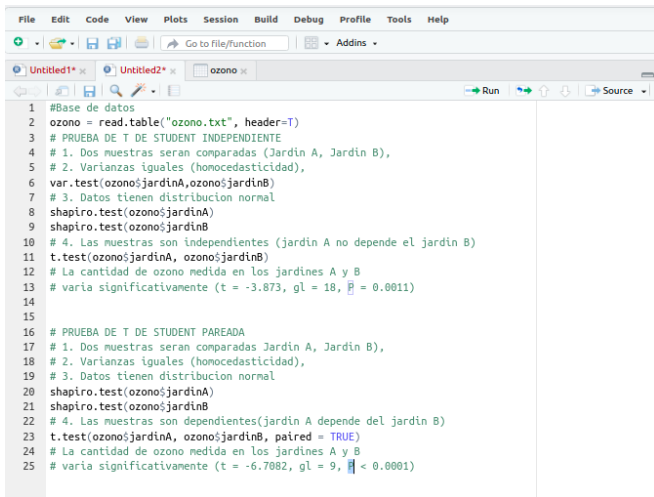
```



## Instrucciones para elegir entre pruebas estadísticas para dos muestras:

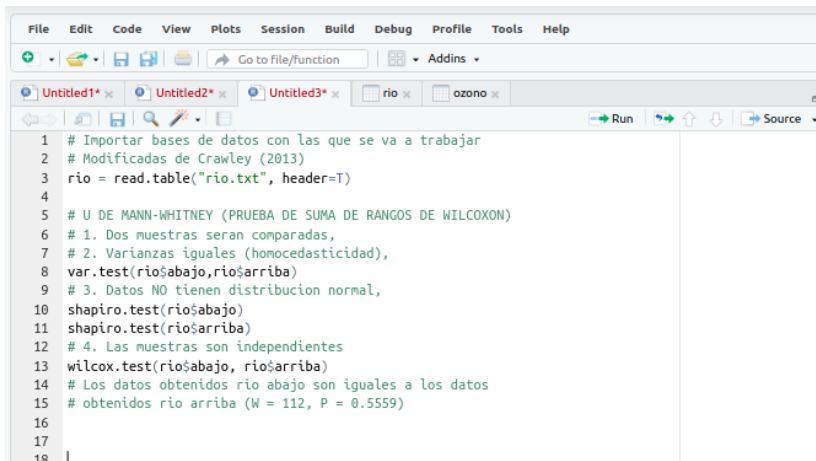
- ① Las muestras son normales. Las muestras no son normales.  
`Shapiro.test()`, o visualmente `qqnorm()`; `qqplot()`.
- ② Las varianzas son homogéneas ( $P$  mayor a 0.05). Pruebas de homocedasticidad.
- ③ Las muestras son independientes. T de Student independiente  
`t.test(paired=F)`. Las muestras son dependientes T de Student pareada `t.test(paired=T)`.
- ④ Las muestras son independientes. U de Mann Whitney (Prueba de suma de rangos de Wilcoxon)  
`wilcox.test(paired=F)`. Las muestras no son dependientes ...  
Prueba de rangos con signo de Wilcoxon  
`wilcox.test(paired=T)`.

# Ejemplo 1: T de Student



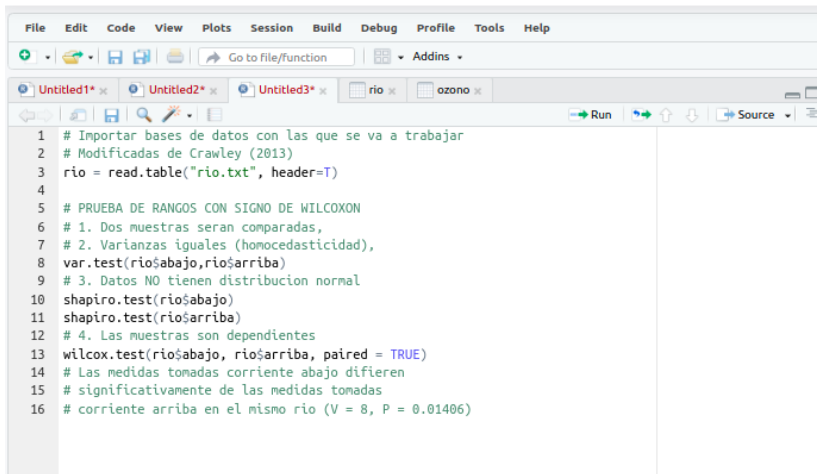
```
1 #Base de datos
2 ozono = read.table("ozono.txt", header=T)
3 # PRUEBA DE T DE STUDENT INDEPENDIENTE
4 # 1. Dos muestras seran comparadas (Jardin A, Jardin B),
5 # 2. Varianzas iguales (homocedasticidad),
6 var.test(ozono$jardinA,ozono$jardinB)
7 # 3. Datos tienen distribucion normal
8 shapiro.test(ozono$jardinA)
9 shapiro.test(ozono$jardinB)
10 # 4. Las muestras son independientes (jardin A no depende el jardin B)
11 t.test(ozono$jardinA, ozono$jardinB)
12 # La cantidad de ozono medida en los jardines A y B
13 # varia significativamente (t = -3.873, gl = 18, p = 0.0011)
14
15
16 # PRUEBA DE T DE STUDENT PAREADA
17 # 1. Dos muestras seran comparadas Jardin A, Jardin B),
18 # 2. Varianzas iguales (homocedasticidad),
19 # 3. Datos tienen distribucion normal
20 shapiro.test(ozono$jardinA)
21 shapiro.test(ozono$jardinB)
22 # 4. Las muestras son dependientes(jardin A depende del jardin B)
23 t.test(ozono$jardinA, ozono$jardinB, paired = TRUE)
24 # La cantidad de ozono medida en los jardines A y B
25 # varia significativamente (t = -6.7082, gl = 9, p < 0.0001)
```

## Ejemplo 2: U de Mann-Whitney



```
1 # Importar bases de datos con las que se va a trabajar
2 # Modificadas de Crawley (2013)
3 rio = read.table("rio.txt", header=T)
4
5 # U DE MANN-WHITNEY (PRUEBA DE SUMA DE RANGOS DE WILCOXON)
6 # 1. Dos muestras seran comparadas,
7 # 2. Varianzas iguales (homocedasticidad),
8 var.test(rio$abajo, rio$arriba)
9 # 3. Datos NO tienen distribucion normal,
10 shapiro.test(rio$abajo)
11 shapiro.test(rio$arriba)
12 # 4. Las muestras son independientes
13 wilcox.test(rio$abajo, rio$arriba)
14 # Los datos obtenidos rio abajo son iguales a los datos
15 # obtenidos rio arriba (W = 112, P = 0.5559)
16
17
18 |
```

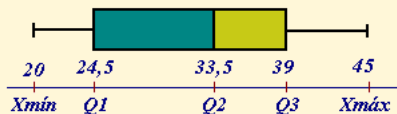
## Ejemplo 3: PRUEBA DE RANGOS CON SIGNO DE WILCOXON



```
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Untitled1* x Untitled2* x Untitled3* x rio x ozono x
Run Source
1 # Importar bases de datos con las que se va a trabajar
2 # Modificadas de Crawley (2013)
3 rio = read.table("rio.txt", header=T)
4
5 # PRUEBA DE RANGOS CON SIGNO DE WILCOXON
6 # 1. Dos muestras seran comparadas,
7 # 2. Varianzas iguales (homocedasticidad),
8 var.test(rio$abajo, rio$arriba)
9 # 3. Datos NO tienen distribucion normal
10 shapiro.test(rio$abajo)
11 shapiro.test(rio$arriba)
12 # 4. Las muestras son dependientes
13 wilcox.test(rio$abajo, rio$arriba, paired = TRUE)
14 # Las medidas tomadas corriente abajo difieren
15 # significativamente de las medidas tomadas
16 # corriente arriba en el mismo rio (V = 8, P = 0.01406)
```

## Diagrama de caja y bigotes

Es un resumen gráfico que permite visualizar, para un conjunto de datos, la tendencia central, la dispersión y la presencia posible de datos atípicos. Para realizarlo se necesita calcular la mediana, el primer cuartil, y el tercer cuartil de los datos.



El bigote de la izquierda representa al colectivo de edades ( $X_{\min}$ ,  $Q_1$ )

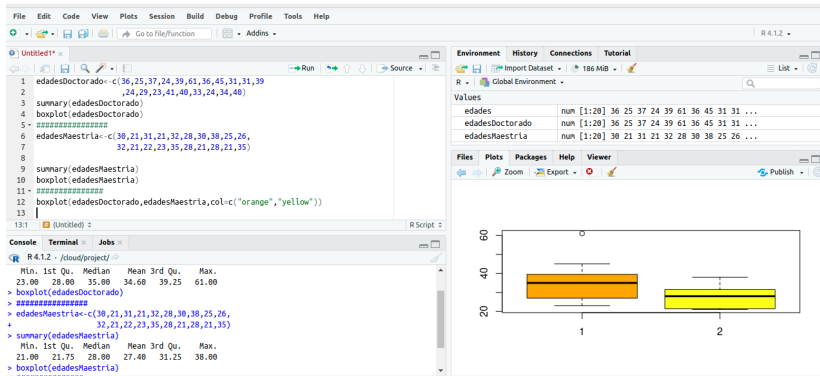
La primera parte de la caja a ( $Q_1$ ,  $Q_2$ ),

La segunda parte de la caja a ( $Q_2$ ,  $Q_3$ )

El bigote de la derecha viene dado por ( $Q_3$ ,  $X_{\max}$ ).

[www.estadisticaparatodos.es/taller/graficas/cajas.html](http://www.estadisticaparatodos.es/taller/graficas/cajas.html)

# ejemplo





*Gracias!*