

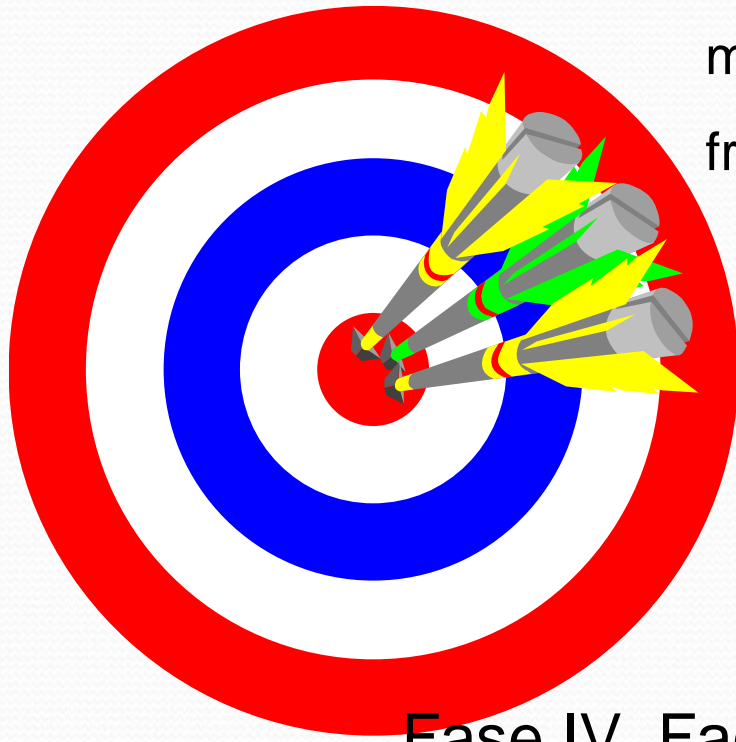
ANALISIS ESTADISTICO DE VARIABLES

Francisco Mendizábal Prem

Correo electrónico:

menprem@yahoo.com

franciscomendizabalprem@gmail.com



Fase IV. Facultad de Ciencias Médicas

Datos (Matriz de datos)

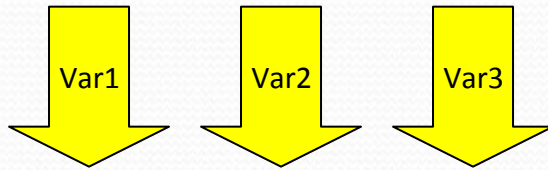
Es un arreglo o cuadrícula en la cual las filas representan registros de un individuo y las columnas variables

Datos (Matriz de datos)

Registro1 →

Registro2 →

Datos (Matriz de datos)



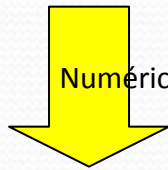
Datos (Matriz de datos)

La primera etapa del análisis de los datos es hacer análisis de A UNA VARIABLE también llamado estadística descriptiva

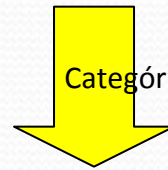
La segunda etapa del análisis de datos hacer un análisis de A DOS VARIABLES. En ese caso estamos buscando asociación entre dos variables

La tercera etapa del análisis de datos hacer un análisis de A DOS VARIABLES controlando por la tercera variable. En ese caso estamos buscando asociación entre dos variables pero controlamos la tercer variable; es decir, “restamos” su efecto.

Datos (Matriz de datos)



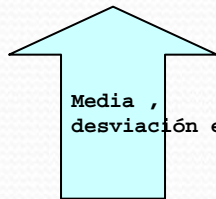
Numérica



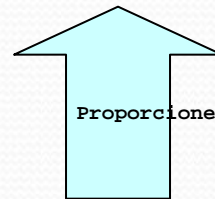
Categórica

Analisis de a
una variable

Peso			Genero			
152			F			
145			F			
132			M			



Media ,
desviación estandard



Proporciones

Datos (Matriz de datos)

Veamos de forma concreta como haríamos el análisis de un archivo.

En caso de que olvide los pasos exactos para llevar a cabo el siguiente análisis con Excel pruebe mi sitio de Internet:

<http://franciscomendizabalprem.googlepages.com/estadistica>

Y baje de ahí el material titulado:

“Como usar Excel para análisis de datos”

Datos (Matriz de datos)

The screenshot shows a Windows Internet Explorer browser window. The address bar displays the URL: <http://franciscohendizabalprem.googlepages.com/estadistica>. The page title is "franciscohendizabalprem - POSTGRADO DE LA FACULTAD DE CIENCIAS MEDICAS DE LA UNIVERSIDAD DE SAN CARLOS DE GUATEMALA".

The main content of the page is as follows:

POSTGRADO DE LA FACULTAD DE CIENCIAS MEDICAS DE LA UNIVERSIDAD DE SAN CARLOS DE GUATEMALA (Esta NO es una página oficial de la Facultad sino es una página de Francisco Mendizábal Prem, catedrático de Fase IV). MATERIALES SOBRE INVESTIGACION Y ESTADISTICA

Estimado Residente y estudiante de Medicina en general.
Aca podrá encontrar materiales relacionados tanto con diseños de investigación como con estadística.

ATENCION RESIDENTES:

NECESITA SABER ALGO SOBRE TAMAÑO DE MUESTRA ?

[Tamaño de Muestra Estudio Descriptivo.zip](#) (Este archivo comprimido contiene la explicación en un documento Word y una hoja Excel con las formulas para efectuar el cálculo en forma "automática")

[Tamaño de muestra para estudio comparativo.zip](#) (Este archivo comprimido contiene la explicación en un documento Word y una hoja Excel con las formulas para efectuar el cálculo en forma "automática")

Como manejar el turno nocturno y la falta de sueño:

Guia de preparación del turno nocturno (Preparada por el Colegio Real Británico Médico) [nightshiftbooklet1.pdf](#) (myr buen material sobre como hacer los turnos y como

The browser's taskbar at the bottom shows the Start button, several open applications including "Residentes III IGSS", "Windows Live Messen...", "Microsoft PowerPoint...", and two instances of "franciscohendizabalp...", along with system icons for network, volume, and battery, and the time "08:16 p.m.".

Datos (Matriz de datos)

Residencia Médica: [StrategiesForHandlingFatigue.pdf](#)

INVESTIGACION Y ESTADISTICA. ATENCION RESIDENTES II y III :

[Como usar Excel y Epi Info para analizar sus datos.doc](#) (Si ya están en la etapa de ingresar datos o de analizar datos esto les puede servir).

[COMO USAR EXCEL PARA ANALISIS DE DATOS.doc](#) (Este material contiene bastante sobre como usar funciones generales y estadísticas en Excel)

[Curso para Residentes.doc](#) (94 kb, buen comienzo para su trabajo de investigación como Residente I o Residente II)

[Como obtener Epi Info](#) (vinculo para "bajar" el programa desde el CDC)

"Cursito" de Estadística del British Medical Journal [StatisticsAtSquareOne.zip](#)

Metodos estadísticos para estudios de concordancia (concordancia de jueces, concordancia de diagnósticos, etc): <http://ourworld.compuserve.com/homepages/jsuebersax/agree.htm>

Si tiene dificultades de que prueba estadística usar y sobretodo tiene dificultades sobre como calcular dicha prueba use esta página la cual tiene una "calculadora" interactiva. Con lo anterior solamente le que a usted la tarea de interpretar:
<http://www.theresearchassistant.com/research/link.asp>

ATENCION RESIDENTES I y II:

Consejo de Investigación: Si va a hacer un trabajo piense en cual de las siguientes categorías podría caer:

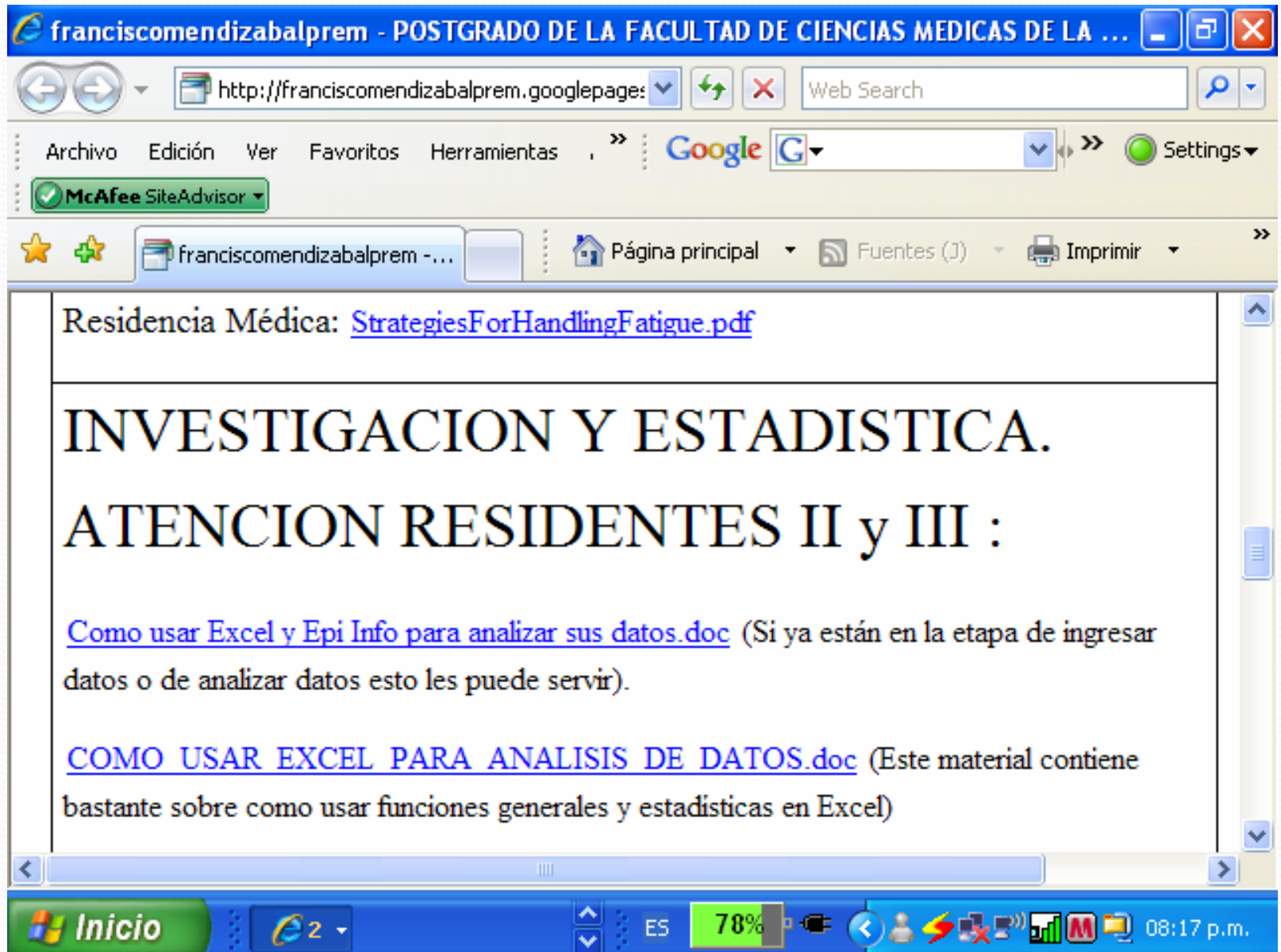
Tratamiento (debe ser un Ensayo Clínico Aleatorizado).
[\(Mas en este hipervinculo\)](#)

Diagnóstico (a ciegas la comparación de Dx con una amplia gama de pacientes y debe haber un "estándar de oro"). Vea este excelente artículo sobre la "arquitectura de la investigación diagnóstica":
[ArchitectureOfDiagnosticResearch.pdf](#)

Windows Internet Explorer pro
http://franciscomendizabalprem.googlepages.com/estadistica
McAfee SiteAdvisor
Franciscomendizabalprem - POSTGRADO DE LA FACUL...
Página principal Fuentes (3) Imprimir Página Herramientas Ayuda Agregar entrada Yahoo! Services Run DAP Messenger

Inicio Residentes III IGSS Windows Live Messen... Microsoft PowerPoint ... Franciscomendizabalp... Franciscomendizabalp... ES 77% 08:16 p.m.

Datos (Matriz de datos)



franciscomendizabalprem - POSTGRADO DE LA FACULTAD DE CIENCIAS MEDICAS DE LA ...

http://franciscomendizabalprem.googlepage: Web Search

Archivo Edición Ver Favoritos Herramientas Google G Settings

McAfee SiteAdvisor

franciscomendizabalprem -... Página principal Fuentes (J) Imprimir

Residencia Médica: [StrategiesForHandlingFatigue.pdf](#)

INVESTIGACION Y ESTADISTICA. ATENCION RESIDENTES II y III :

[Como usar Excel y Epi Info para analizar sus datos.doc](#) (Si ya están en la etapa de ingresar datos o de analizar datos esto les puede servir).

[COMO USAR EXCEL PARA ANALISIS DE DATOS.doc](#) (Este material contiene bastante sobre como usar funciones generales y estadísticas en Excel)

Inicio e 2 ES 70% 08:17 p.m.

Datos (Matriz de datos)

Veamos de forma concreta como haríamos el análisis de un archivo.

Calcularemos promedio y desviación estándar y percentiles en las variables de razón: edad, presión arterial sistólica al final del tratamiento y presión arterial diastólica al final del tratamiento.

Calcularemos proporciones de las variables categóricas: tratamiento recibido y género.



Datos (Matriz de datos)

En otros casos podemos usar estadística descriptiva para verificar la comparabilidad de un ensayo clínico aleatorizado.

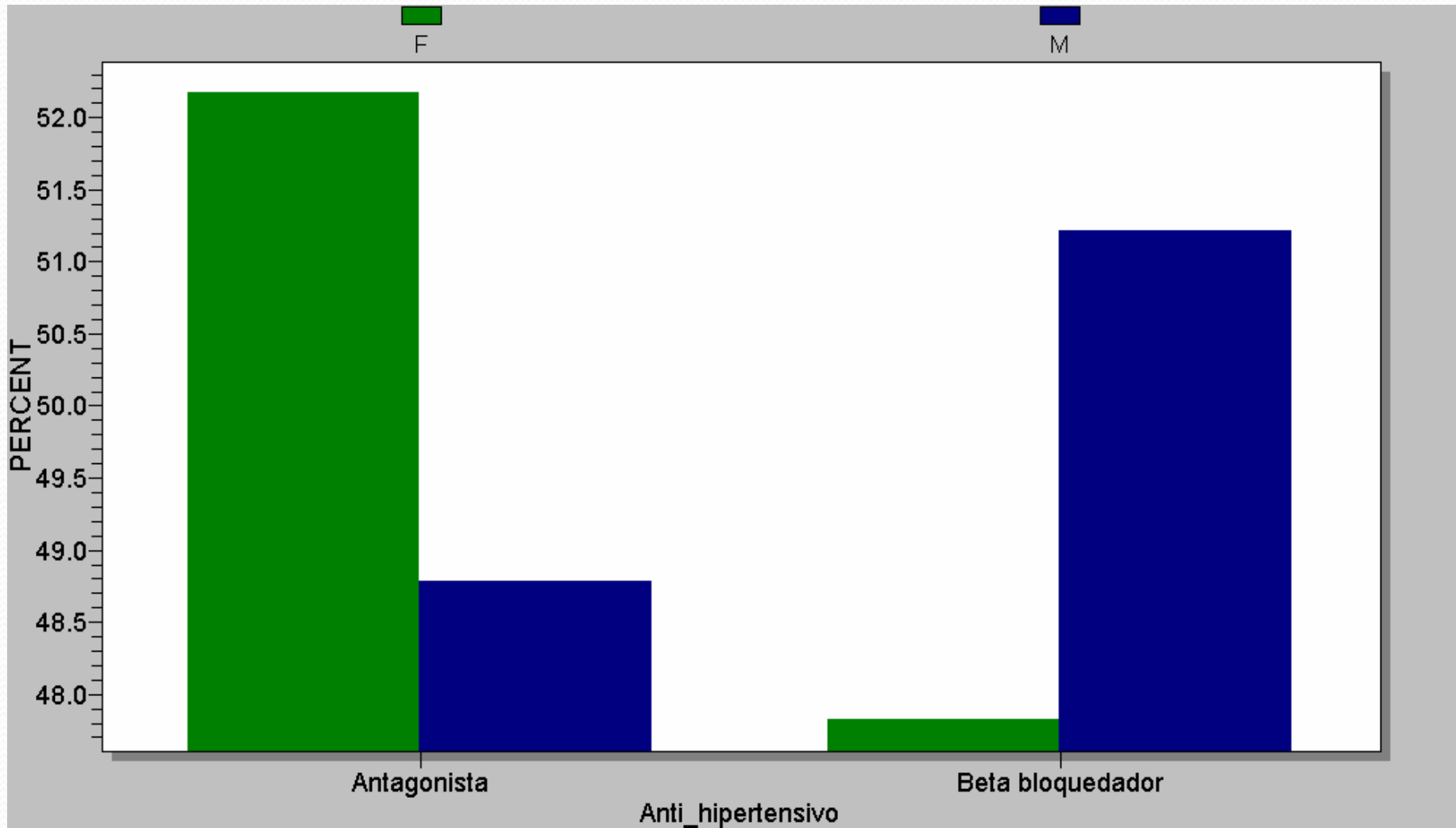
La pregunta es que pregunta clínica estaríamos respondiendo.

La primera: ¿ son los dos grupos balanceados ?

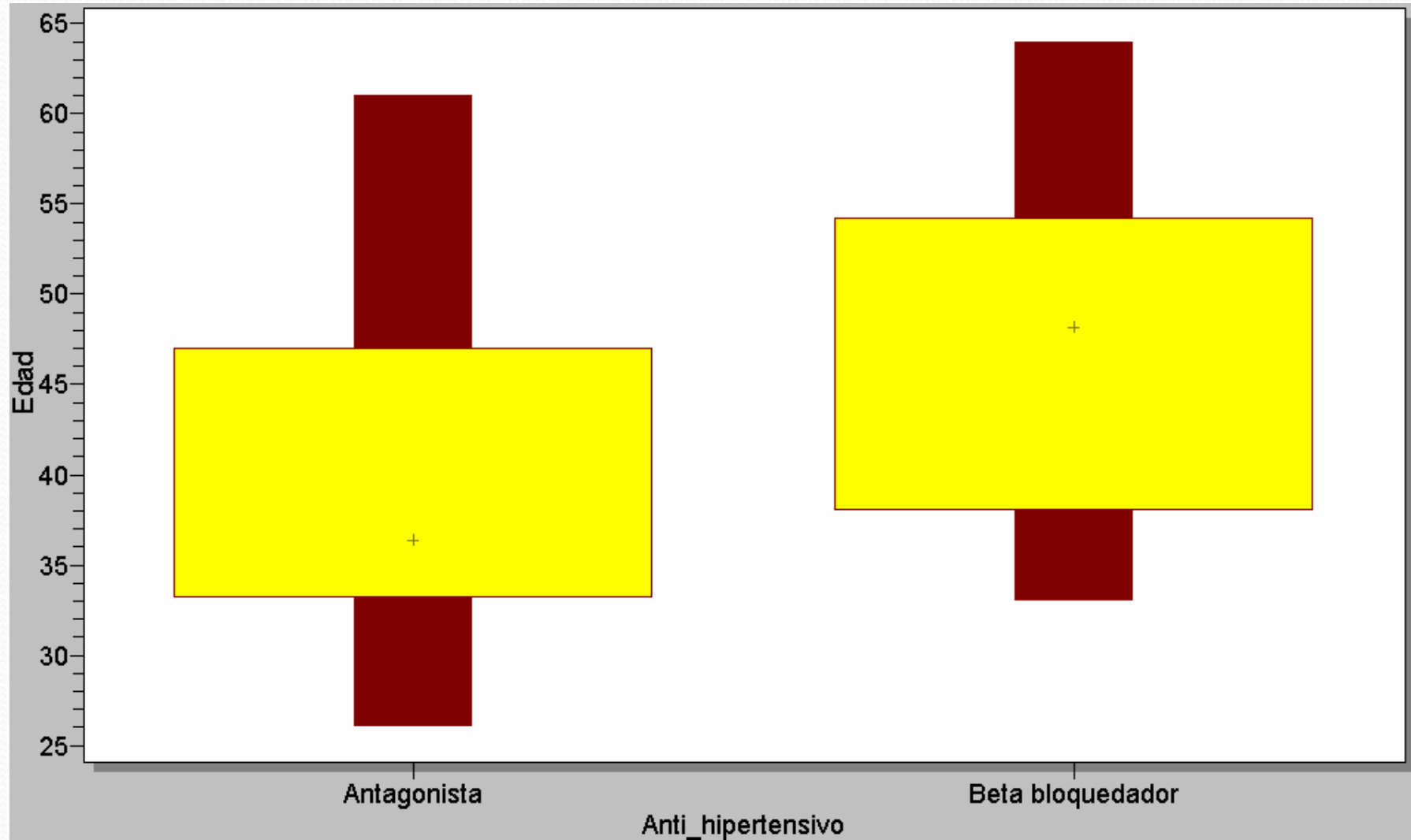
Veamos si están balanceados en género y en grupos de edad.



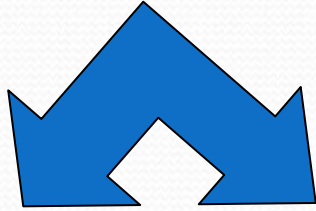
Datos (Matriz de datos)



Datos (Matriz de datos)



Datos (Matriz de datos)



Usamos medidas de asociación

Analisis de
A DOS
variable

Datos (Matriz de datos)

Veamos de forma concreta como haríamos el análisis de un archivo.

Vamos a usar el programa Epi Info para llevar a cabo el análisis.



Datos

The screenshot displays the SAS software interface. On the left is the 'Analysis' window with a tree view of 'Analysis Commands' including Data, Variables, Select/If, Statistics, and Advanced Statistics. The 'Analysis Output' window shows a table of data with columns: Genero, Edad, Anti_hipert, PAS, PAD, and reducción_. The 'Program Editor' window shows SAS code for reading an Excel file and displaying a grid table.

Analysis Output Table:

Genero	Edad	Anti_hipert	PAS	PAD	reducción_
M	34	Antagonista	220	140	SI
M	52	Beta bloquec	200	140	SI
M	25	Beta bloquec	140	90	SI
F	61	Beta bloquec	160	100	SI
F	35	Beta bloquec	200	150	SI
F	42	Antagonista	140	85	SI
F	48	Antagonista	160	100	SI
M	49	Antagonista	200	140	SI
M	54	Beta bloquec	140	90	SI
M	36	Beta bloquec	160	100	SI
F	28	Beta bloquec	200	150	NO
M	29	Beta bloquec	160	100	NO
F	35	Antagonista	140	100	SI
M	42	Beta bloquec	200	120	NO
F	48	Beta bloquec	180	120	NO
M	49	Beta bloquec	140	95	NO
F	34	Antagonista	200	100	SI
F	33	Antagonista	160	90	SI
F	38	Antagonista	185	100	NO
F	41	Antagonista	170	120	NO
M	51	Beta bloquec	200	120	NO
M	57	Beta bloquec	180	120	NO
M	47	Beta bloquec	140	95	NO
M	52	Beta bloquec	200	100	NO
M	64	Beta bloquec	160	90	NO
M	61	Beta bloquec	185	100	NO
M	34	Beta bloquec	170	120	NO
M	38	Antagonista	180	90	NO
M	26	Antagonista	140	100	NO

Program Editor - New Program

```
File Edit View Fonts Run Help
New Open Save Print Run Run This Command
READ "Excel 8.0" 'H:\Curso Investigación Residentes I 2008\Datos1.xls':'Hoja1$'
FILESPEC HDR="YES" END
LIST * GRIDTABLE
MEANS Edad
TABLES Anti hipertensivo reducción_de hipertensión PERCENTS=(-)
LIST * GRIDTABLE
```

Datos

C:\Epi_Info\OUT2.htm

Previous Next Last History Open Bookmark Print Restore

[Forward](#)

REDUCCIÓN DE HIPERTENSIÓN

Anti-hipertensivo	NO	SI	TOTAL
Antagonista	8	10	18
Beta bloqueador	30	6	36
TOTAL	38	16	54

Single Table Analysis

	Point Estimate	95% Confidence Interval	
		Lower	Upper
PARAMETERS: Odds-based			
Odds Ratio (cross product)	0.1600	0.0446	0.5742 (T)
Odds Ratio (MLE)	0.1667	0.0431	0.5945 (M)
		0.0366	0.6796 (F)
PARAMETERS: Risk-based			
Risk Ratio (RR)	0.5333	0.3118	0.9123 (T)
Risk Difference (RD%)	-38.8889	-64.8731	-12.9047 (T)

(T=Taylor series; C=Cornfield; M=Mid-P; F=Fisher Exact)

STATISTICAL TESTS	Chi-square	1-tailed p	2-tailed p
Chi-square - uncorrected	8.7039		0.0031763827
Chi-square - Mantel-Haenszel	8.5428		0.0034701515
Chi-square - corrected (Yates)	6.9387		0.0084360407
Mid-p exact		0.0026437536	
Fisher exact		0.0046639430	

Microsoft PowerPoint - [Variables [Modo de compatibilidad]]

Windows Live Mess... Curso Investigación ... Epi Info Epi Info Analysis Microsoft PowerPoi... ES 11:53 a.m.

Datos (Matriz de datos)

Veamos de forma concreta como haríamos el análisis de un archivo.

Vamos a usar el programa SPSS para llevar a cabo el análisis.

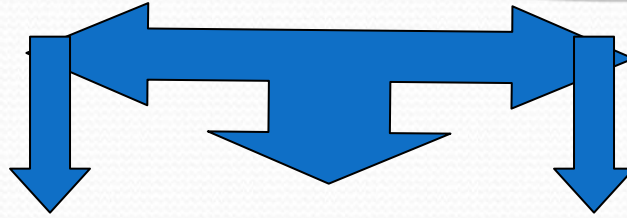


Datos (Matriz de datos)

Podemos usar nuevamente Excel para hacer un analisis de regresión



Datos (Matriz de datos)



Analisis de
A DOS
variable

Controland
o una
tercer
variable

Datos

The screenshot shows the Statcalc software window with the following content:

Window Title: C:\Epi_Info\STATCALC.EXE

Header: EpiInfo Version 6 Statcalc November 1993

Table:

	+ Disease	-	
+	254	125	379
-	451	258	709
E x p o s u r e	705	383	1088

***** Stratified Analysis *****
Summary of 2 Tables

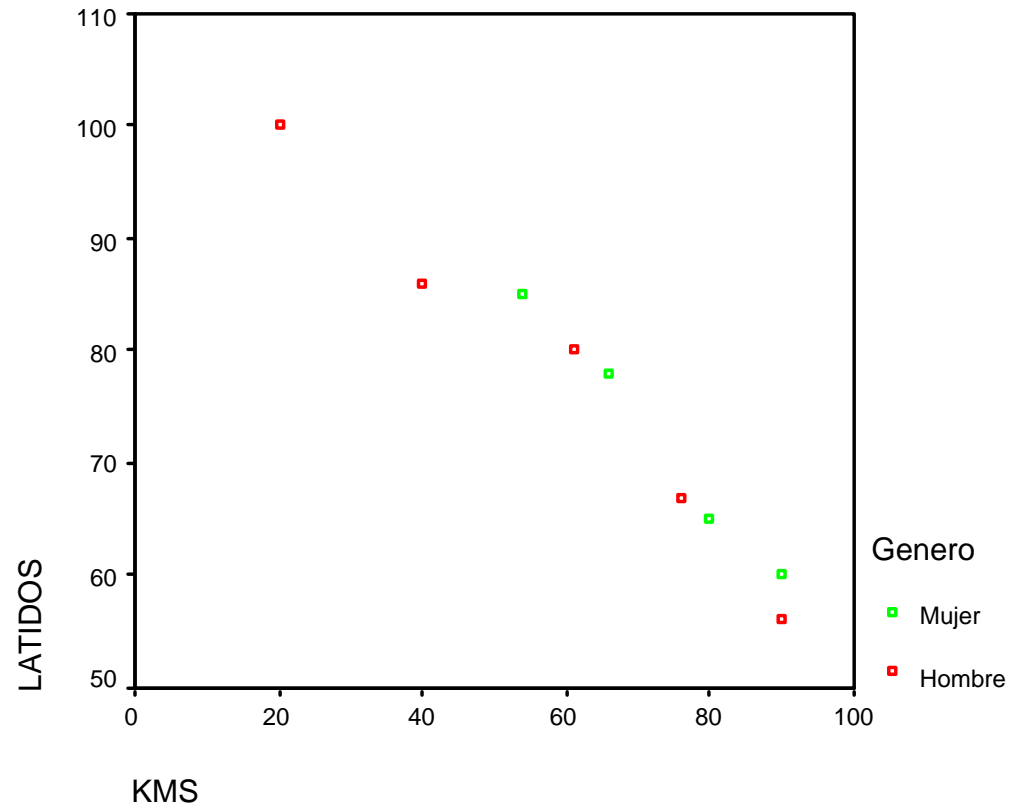
Crude odds ratio for all strata = 1.44
Mantel-Haenszel Weighted Odds Ratio = 1.47
Cornfield 95% Confidence Limits
1.16 < 1.47 < 1.85
Mantel-Haenszel Summary Chi Square = 10.45
P value = 0.00122721 ←

Crude RR for all strata = 1.12
Mantel-Haenszel Weighted Relative Risk
of Disease, given Exposure = 1.13
Greenland/Robins Confidence Limits =
1.05 < MHRR < 1.21

<Enter> for more; F10 to quit.

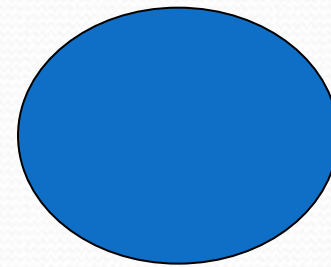
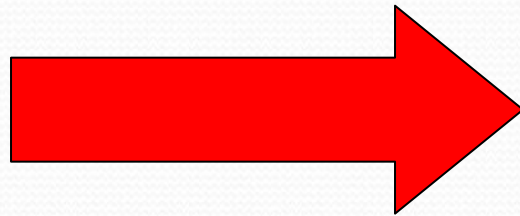
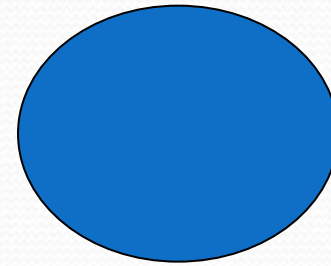
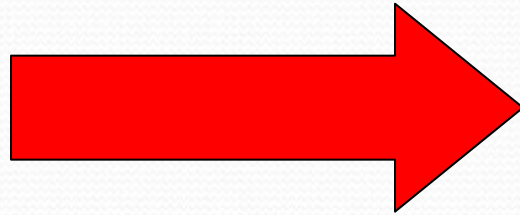
Footer: F1-Help F5-Print F6-Open File F10-Done

Datos



Datos

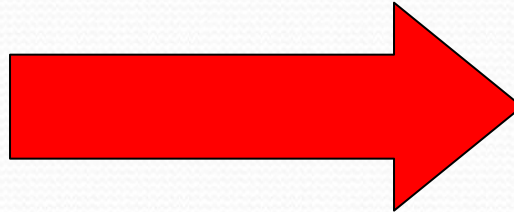
Variable dependiente de razón o intervalo:



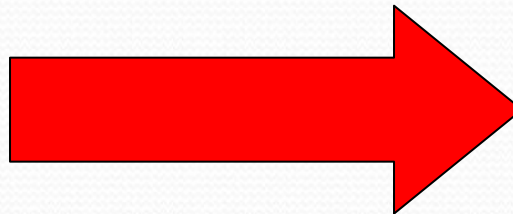
Datos

Variable dependiente de razón o intervalo:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$



$$\bar{x}_1, S_1$$



$$\bar{x}_2, S_2$$

Datos

Variable dependiente de razón o intervalo:

C:\Epi_Info\OUT34.htm

Previous Next Last History Open Bookmark Print Restore

B 18 1495.0000 83.0556 462.1732 21.4982

	Minimum	25%	Median	75%	Maximum	Mode
A	85.0000	90.0000	100.0000	110.0000	160.0000	100.0000
B	55.0000	65.0000	85.0000	90.0000	140.0000	85.0000

ANOVA, a Parametric Test for Inequality of Population Means

(For normally distributed data only)

Variation	SS	df	MS	F statistic
Between	5170.6124	1	5170.6124	10.9836
Within	17888.7626	38	470.7569	
Total	23059.3750	39		

T Statistic = 3.3142

P-value = 0.0020

Bartlett's Test for Inequality of Population Variances

Bartlett's chi square= 0.0050 df=1 P value=0.9437

A small p-value (e.g., less than 0.05 suggests that the variances are not homogeneous and that the ANOVA may not be appropriate.

Mann-Whitney/Wilcoxon Two-Sample Test (Kruskal-Wallis test for two groups)

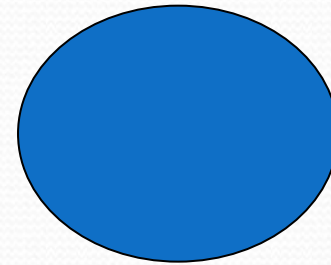
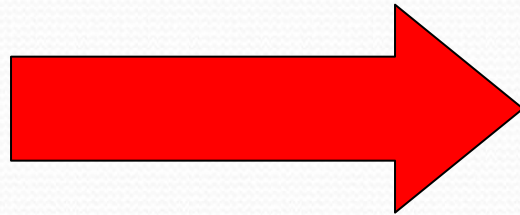
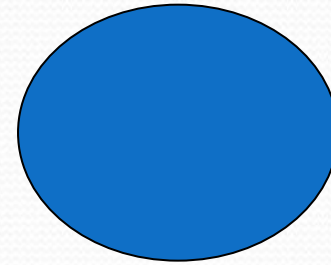
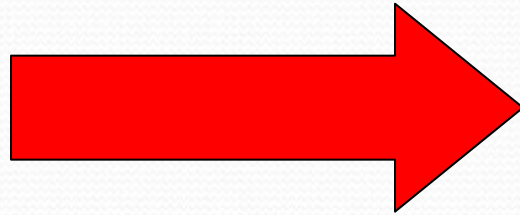
Kruskal-Wallis H (equivalent to Chi square) = 13.7001

Degrees of freedom = 1

P value = 0.0002

Datos

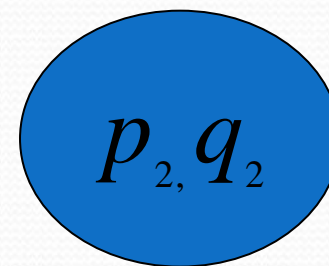
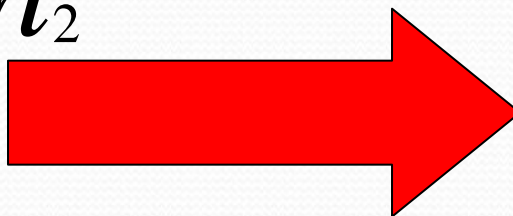
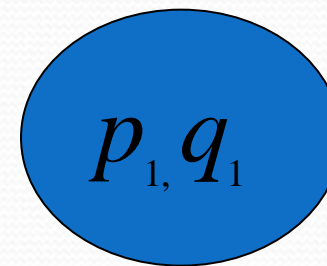
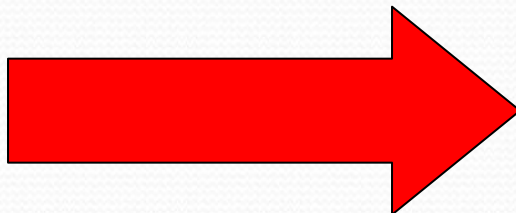
Variable dependiente de categórica o nominal:



Datos

Variable dependiente de razón o intervalo:

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$



Datos

Variable dependiente de razón o intervalo:

C:\Epi_Info\OUT34.htm

Previous Next Last History Open Bookmark Print Restore

ALVIODOLOR

Tratamiento	NO	SI	TOTAL
A	7	15	22
Row %	31.8	68.2	100.0
Col %	29.2	93.8	55.0
B	17	1	18
Row %	94.4	5.6	100.0
Col %	70.8	6.3	45.0
TOTAL	24	16	40
Row %	60.0	40.0	100.0
Col %	100.0	100.0	100.0

Single Table Analysis

	Point Estimate	95% Confidence Interval	
		Lower	Upper
PARAMETERS: Odds-based			
Odds Ratio (cross product)	0.0275	0.0030	0.2496 (T)
Odds Ratio (MLE)	0.0305	0.0012	0.2176 (M)
		0.0006	0.2650 (F)
PARAMETERS: Risk-based			
Risk Ratio (RR)	0.3369	0.1809	0.6274 (T)
Risk Difference (RD%)	-62.6263	-84.7803	-40.4722 (T)
(T=Taylor series; C=Cornfield; M=Mid-P; F=Fisher Exact)			
STATISTICAL TESTS			
	Chi-square	1-tailed p	2-tailed p
Chi square - uncorrected	16.1785		0.0000588153
Chi square - Mantel-Haenszel	15.7740		0.0000725462
Chi square - corrected (Yates)	13.6742		0.0002185852