

Página www

Página de Abertura

Contenido



Página 1 de 100

Regresar

Full Screen

Cerrar

Abandonar

Análisis de Datos Categóricos

Clase 1

Juan Carlos Correa
e-mail: jccorrea@unal.edu.co

7 de marzo de 2022

Página www

Página de Abertura

Contenido



Página 2 de 100

Regresar

Full Screen

Cerrar

Abandonar

Programa

- Introducción.
- Esquemas de muestreo.
- Resultados asintóticos.
- GSK.
- Aplicaciones.

Referencias

- Agresti, A. (1990) *Categorical Data Analysis*. New York: John Wiley
- Bishop, Y. M. M., Fienberg, S. E, y Holland, P. W. (1975) *Discrete Multivariate Analysis: Theory and Practice*. Cambridge: The MIT Press
- Christensen, R. (1990) *Log-Linear Models*. New York: Springer-Verlag
- Santner, T. J. y Duffy, D. E. (1989) *The Statistical Analysis of Discrete Data*. New York: Spriger-Verlag.
- Upton, G. J. G. (1978) *The Analysis of Cross-Tabulated Data*. Chichster: John Wiley & Sons.
- Correa, J. C. (2015) Notas de Clase.

Página www

Página de Abertura

Contenido

◀

▶

◀

▶

Página 3 de 100

Regresar

Full Screen

Cerrar

Abandonar

Página www

Página de Abertura

Contenido



Página 4 de 100

Regresar

Full Screen

Cerrar

Abandonar

Requisitos

- Un curso básico de Inferencia.
- Simulación.
- Manejar R (a nivel de programación).
- Manejar \LaTeX .
- Ganas...

Página www

Página de Abertura

Contenido



Página 5 de 100

Regresar

Full Screen

Cerrar

Abandonar

Horas de atención

Martes, Miércoles y Jueves de 8:00am a 9:00am. Con cita.

Página www

Página de Abertura

Contenido



Página 6 de 100

Regresar

Full Screen

Cerrar

Abandonar

Evaluación

- Dos parciales, cada uno del 20 %.
- Dos trabajos (que se entregan por escrito y se presentan en clase): el 30 % cada uno.

Introducción

- En su quehacer diario el analista de datos enfrenta bases de datos generalmente compuestas por variables de varios tipos.
- Dependiendo del tipo de variables y de su relación se pueden realizar diversos análisis estadísticos.
- No podemos en este momento dejar de establecer dos principios fundamentales del análisis de datos (Wickens, 1989):
 - Un análisis estadístico con sentido no puede realizarse sin información no estadística.
 - Entre más información estadística se posea, mejores serán las conclusiones que se puedan obtener.

Página www

Página de Abertura

Contenido



Página 8 de 100

Regresar

Full Screen

Cerrar

Abandonar

Tipos de Variables

Variable {
Continua
Discreta
Categorica {
Nominal
Ordinal

Definición 1 (Tabla de Contingencia) *Una tabla de contingencia es una tabla de clasificación cruzada de conteos.*

- Las unidades muestreadas de una población son clasificadas en la tabla de acuerdo a cada una de las variables categóricas o conjuntos de categorías tales como sexo (masculino, femenino), edad (joven, adulto, anciano).
- Cuando varias variables categóricas son consideradas a la vez ellas forman una tabla de contingencia multidimensional donde cada variable corresponde a una dimensión de la tabla.
- Tales tablas presentan problemas especiales para el análisis e interpretación.

Página www

Página de Abertura

Contenido



Página 10 de 100

Regresar

Full Screen

Cerrar

Abandonar

Definición 2 (Tabla Completa) *Decimos que una tabla de contingencia es una tabla completa si cada celda tiene probabilidad de ocurrencia diferente de cero, o sea que es posible, al menos en teoría, observar elementos en cada celda. Si lo anterior no es posible, entonces hablamos de tablas con ceros estructurales.*

Un ejemplo de la anterior situación puede ocurrir cuando cruzamos la edad categorizada de una mujer contra el número de hijos nacidos. Si tenemos mujeres menores de cinco años, es imposible que haya alguna con hijos a esta edad. Por lo tanto la celda que cruza estas dos situaciones tendrá obligatoriamente un cero estructural.

Las relaciones entre varias variables categóricas se estudian mediante el análisis de tablas de frecuencia o conteos. Algunos de los objetivos de los análisis son:

1. **Asociaciones entre variables:** Cuáles variables están asociadas o relacionadas entre sí y cuál es el tipo de asociación que se da?
2. **Efecto sobre una variable dependiente:** Cuando en una tabla tenemos una variable que es dependiente estamos interesados en determinar la influencia que sobre ella tienen el conjunto de covariables que conforman la tabla.
3. **Estimación de parámetros:**Cuál es el número esperado de observaciones para una combinación particular de variables?
4. **Importancia de los efectos:**
5. **Comparaciones específicas y análisis de tendencias:**

Página www

Página de Abertura

Contenido

◀

▶

◀

▶

Página 11 de 100

Regresar

Full Screen

Cerrar

Abandonar

Categorización de variables continuas

- La categorización de variables continuas o discretas es algo que se hace de forma permanente en la presentación de reportes, ya que para el lector puede ser más fácil la visualización de los resultados de esta forma.
- Categorizar variables continuas puede ser peligroso cuando se trata de realizar análisis de tipo estadístico y aún conducir a conclusiones erróneas, sin importar la técnica usada.
- Como una regla general, el investigador no debe discretizar ni combinar categorías y debe luchar por conservar tanta precisión como sea posible.

Tipos de Estudios

Podemos dividir en tres grandes clases los estudios realizados:

1. Retrospectivo o Caso-Control

En este tipo de trabajo se mira al pasado. Sería tratar de responder a la pregunta: Qué pasó?

2. Prospectivo

En este tipo de estudios podemos considerar los **Ensayos Clínicos**, donde los objetos se asignan aleatoriamente a los experimentos y los de **cohorte**, donde se hace un seguimiento a través del tiempo a un grupo interés. La pregunta a responder sería: Qué pasará?

3. Transversales

La pregunta sería: Qué está pasando? Los estudios de caso-control y los de cohorte son llamados *observacionales*. Los ensayos clínicos son experimentales. Tablas típicas con datos transversales son las presentadas en los anuarios estadísticos.

Página www

Página de Abertura

Contenido

◀

▶

◀

▶

Página 13 de 100

Regresar

Full Screen

Cerrar

Abandonar

Ejemplos

Es muy fácil encontrar ejemplos reales de tablas de contingencia ya que aparecen con frecuencia tanto en revistas como en periódicos. A continuación presentamos ejemplos de tablas de contingencia.

Lanzamiento de un par de dados

En un juego de parques se registraron los resultados del lanzamiento de un par de dados 130 veces. A partir de estos resultados quiere uno ver si los dados son conjuntamente buenos.

Resultado	2	3	4	5	6	7	8	9	10	11	12
Frecuencia	4	8	10	11	22	14	22	18	10	5	6

Nos podemos preguntar si con los datos anteriores podríamos jugar tranquilamente este juego de parques, o sea si los dados son buenos o están cargados.

Página *www*

Página de Abertura

Contenido

◀◀

▶▶

◀

▶

Página **14** de **100**

Regresar

Full Screen

Cerrar

Abandonar

Si el par de dados fueran perfectos, entonces el modelo teórico sería el que aparece en la siguiente tabla:

Resultado	2	3	4	5	6	7	8	9	10	11	12
Probabilidad	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

En los 130 lanzamientos de los dados esperaríamos hallar

Resultado	2	3	4	5	6	7	8	9	10	11	12
Esperada	3.61	7.22	10.83	14.44	18.06	21.67	18.06	14.44	10.83	7.22	3.61

Mes de Nacimiento de Estudiantes

En la siguiente tabla se presenta el mes de nacimiento de los estudiantes de la Universidad Nacional-Sede Medellín clasificados por sexo.

	01	02	03	04	05	06	07	08	09	10	11	12
H	541	454	470	435	482	485	511	490	501	508	504	497
M	315	262	270	286	321	266	279	340	329	293	285	315

Surgen preguntas tales como:

- Es la probabilidad de nacer en cualquier mes del año uniforme?
- Habrá diferencia por sexo para la distribución por mes?

Ejemplo de niños zurdos

La siguiente tabla presenta información sobre niños zurdos y el sexo:

	Zurdo	Diestro
Niño	79	202
Niña	57	138

La pregunta que nos surge es: Hay más niños zurdos que niñas? Será la proporción de niños zurdos igual a la proporción de niñas zurdas?

Página [www](#)

Página de Abertura

Contenido



Página **18** de **100**

Regresar

Full Screen

Cerrar

Abandonar

Padre	Madre	Edad	Zurdo	Diestro
Diestro	Diestra	0-10	15	4
		10-20	136	12
		20 ó más	133	6
Diestro	Zurda	0-10	0	0
		10-20	14	3
		20 ó más	6	3
Zurdo	Diestra	0-10	2	0
		10-20	6	7
		20 ó más	5	3
Zurdo	Zurdo	0-10	0	0
		10-20	1	0
		20 ó más	0	0

Aseo bucal de los primiparo del 2008 UNAL

Los estudiantes que entran a la universidad a pregrado deben responder a una encuesta de bienestar en la cual se les pregunta sobre muchos temas, entre ellos el de sus hábitos de aseo bucal. Cruzamos el uso de seda dental con el uso de enjuague bucal.

		Enjuague Bucal		
		Sin respuesta	Sí	No
Uso de seda dental	Sin respuesta	2	1	0
	Sí	11	270	301
	No	1	167	267

Cruzamos Género con el uso de seda dental y con el uso de enjuague bucal.

	Seda Dental	Enjuague Bucal		
Género		Sin respuesta	Sí	No
Masculino	Sin respuesta	1	1	0
	Sí	8	154	172
	No	1	125	192
Femenino	Sin respuesta	1	0	0
	Sí	3	116	129
	No	0	42	75

Razones y proporciones

Existen una gran variedad de herramientas numéricas que nos permiten entender mejor muchos datos. Para los categóricos tenemos las razones y proporciones como medidas descriptivas básicas, que usualmente son de fácil cálculo e interpretación.

Razón

Una razón es el resultado de dividir una cantidad por otra.

Razón

Una razón es utilizada para comparar las frecuencias de dos clases mutuamente excluyentes. Un ejemplo es el índice de masculinidad en una población o cohorte que se define como

$$\frac{\text{Número de hombres}}{\text{Número de mujeres}}$$

Página *www*

Página de Abertura

Contenido

◀◀

▶▶

◀

▶

Página *21* de *100*

Regresar

Full Screen

Cerrar

Abandonar

Como un ejemplo consideremos la población de Medellín en 2002 dividida por grupos de edades quinquenales

Grupo Etáreo	Hombres	Mujeres	Indice de masculinidad
0 - 4	95.883	92.075	1.04
5 - 9	94.691	91.337	1.04
10 - 14	88.804	87.667	1.01
15 - 19	81.013	86.598	0.94
20 - 24	76.451	92.877	0.82
25 - 29	76.216	96.462	0.79
30 - 34	78.885	95.059	0.83
35 - 39	79.251	94.668	0.84
40 - 44	66.950	84.117	0.80
45 - 49	53.694	71.391	0.75
50 - 54	42.010	58.332	0.72
55 - 59	31.031	42.885	0.72
60 - 64	22.067	31.797	0.69
65 - 69	15.845	24.744	0.64
70 y más	25.503	46.990	0.54

Fuente: Anuario Estadístico Metroplitano 2002

Un valor importante en el caso de tablas de contingencia es lo que se llama el *odds*, que es sencillamente

$$\frac{P(A)}{P(B)}$$

donde A y B son subconjuntos del mismo espacio muestral, y además $P(B) \neq 0$. Esta cantidad se puede interpretar como el número de apariciones del evento A por cada aparición del evento B . En el ejemplo de los dados tenemos que el odds del evento $A =$ La suma sea 3 con relación al evento $A =$ La suma sea 2 es $(2/36)/(1/36) = 2$, por lo tanto el evento de aparecer una suma de 3 en los dados es dos veces más frecuente que el de aparecer una suma de 2.

Página *www*

Página de Abertura

Contenido



Página **24** de **100**

Regresar

Full Screen

Cerrar

Abandonar

Índice

Un índice es una medida que intenta reflejar la relación entre variables. Un ejemplo es el *índice de masa corporal* que se define como

$$IMC = \frac{\text{Peso -masa- en kilogramos}}{(\text{Estatura en metros})^2}$$

Página *www*

Página de Abertura

Contenido



Página *25* de *100*

Regresar

Full Screen

Cerrar

Abandonar

Proporción

La proporción es una clase especial de razón en la cual el numerador es el tamaño (en un sentido general) una parte del conjunto utilizado en el denominador, o sea,

$$p = \frac{a}{a + b}$$

donde a es el tamaño de un conjunto A y $a + b$ es el tamaño del conjunto AB .

Homicidios de niños en Medellín

Grupo Edad	Sexo	1990	1991	1992	1993	1994	1995	1996	1997
0	M	2	2	2	5	2	0	7	6
	F	1	2	0	3	1	2	2	4
1-4	M	1	5	7	6	2	2	3	1
	F	0	1	1	3	2	0	1	2
5-14	M	62	107	107	83	75	52	74	56
	F	9	24	19	25	12	13	12	17
15-17	M	911	1129	538	577	525	521	411	352
	F	42	73	38	41	41	29	36	26

Fuente: El Colombiano, Junio 7 de 1998.

Ejemplo sobre Tipo de Sangre

La siguiente tabla presenta la distribución aceptada como poblacional

Grupo	Tipo de Sangre			
	O	A	AB	B
Rh positivos	56.2 %	26.5	8.4	1.8
Rh negativos	5.1 %	2.7	0.7	0.3

El Tiempo, Abril 2 del 2000, pp. 2C

Los siguientes datos fueron tomados de la página WEB de la Cruz Roja norteamericana:

De 100 personas, alrededor de:

- 38 serán O positivo
- 7 serán O negativo
- 34 serán A positivo
- 6 serán A negativo
- 8 serán B positivo
- 2 serán B negativo
- 4 serán AB positivo
- 1 será AB negativo

Página *www*

Página de Abertura

Contenido

◀

▶

◀

▶

Página *27 de 100*

Regresar

Full Screen

Cerrar

Abandonar

Residencia	Sexo	Grupo Edad	Tipo de Sangre			
			O	A	AB	B
Medellín	Hombre	20-29 años	50	24	1	7
		30-39	36	17	0	3
		40-49	18	14	2	1
		50-59	7	3	0	3
Medellín	Mujer	20-29 años	55	30	1	9
		30-39	33	19	2	11
		40-49	21	11	2	3
		50-59	8	9	0	1
Oriente Antioqueño	Hombre	20-29 años	62	23	1	3
		30-39	26	27	0	0
		40-49	19	12	0	2
		50-59	9	3	0	2
Oriente Antioqueño	Mujer	20-29 años	65	28	1	8
		30-39	37	16	1	2
		40-49	22	8	0	2
		50-59	6	2	0	2

[Página www](#)

[Página de Abertura](#)

[Contenido](#)



[Página 28 de 100](#)

[Regresar](#)

[Full Screen](#)

[Cerrar](#)

[Abandonar](#)

Datos sobre accidentalidad vial

Accidentes por Hora: 1997		
Hora	Accidentes Solo Daños	Accidentes con Víctimas
0	225	473
1	212	418
2	199	334
3	144	264
4	131	217
5	233	274
6	709	446
7	1559	634
8	1600	659
9	1401	686
10	1698	876
11	1884	1021
12	2036	1084
13	1854	1051
14	2163	1105
15	2236	1127
16	2111	1182
17	2046	1330
18	1869	1297
19	1315	1178
20	801	1080
21	707	947
22	504	693
23	316	494

Fuente Fuente: Revista Accidentalidad Vial 1997
Secretaría de Transportes
y Tránsito de Medellín

Página *www*

Página de Abertura

Contenido



Página 29 de 100

Regresar

Full Screen

Cerrar

Abandonar

Ejemplo sobre Accidentalidad Laboral

La siguiente tabla presenta información sobre accidentalidad laboral en Antioquia. Se presenta el sexo, edad, tiempo en el oficio y parte del cuerpo afectada

Edad	Sexo	Tiempo en el Oficio	Parte Afectada		
			M. Inferior	M. Superior	Otra
Menor de 30 Años	Masculino	Menos de un año	537	913	860
		Uno a 4 años	186	449	338
		Cinco o más años	34	69	65
	Femenino	Menos de un año	43	140	37
		Uno a 4 año	22	109	41
		Cinco o más años	11	20	8
30 a 39 años	Masculino	Menos de un año	352	567	544
		Uno a 4 años	162	375	278
		Cinco o más años	115	226	183
	Femenino	Menos de un año	28	129	24
		Uno a 4 años	34	102	39
		Cinco o más años	38	83	32
40 o más años	Masculino	Menos de un año	179	300	262
		Uno a 4 años	76	181	157
		Cinco o más años	109	273	199
	Femenino	Menos de un año	19	51	27
		Uno a 4 años	16	64	20
		Cinco o más años	56	132	54

Ejemplo sobre salubridad

La próxima tabla presenta la información sobre partos en el departamento de Antioquia por regiones. Al ver una tabla de éstas nos podemos preguntar cosas como las siguientes:

- Es diferente la posibilidad de nacer vivo entre las diferentes regiones del departamento?
- Qué pasa con lo anterior si condicionamos por peso?
- Es la probabilidad de nacer con menos de 2500 gramos igual para cada zona?

Nacimientos de más de 20 semanas, 1994				
	Menos de 2500 gms		Más de 2500 gms	
Zona	Vivos	Muertos	Vivos	Muertos
Valle de Aburrá	6388	555	55929	255
Bajo Cauca	417	100	3422	47
Magdalena Medio	83	11	1489	10
Nordeste	157	29	2462	22
Norte	246	31	2760	21
Occidente	143	16	2290	12
Oriente	781	98	8748	44
Suroeste	518	51	5660	31
Urabá	426	53	4169	58

Fuente: Anuario Estadístico de Antioquia, 1994.

Página *www*

Página de Abertura

Contenido

◀

▶

◀

▶

Página *31* de *100*

Regresar

Full Screen

Cerrar

Abandonar

Ejemplo sobre colegios en Antioquia

El ICFES clasifica los colegios de acuerdo al rendimiento obtenido en las pruebas nacionales como Alto, Medio o Bajo. La siguiente tabla presenta el resultado de la clasificación del IC-FES por Región (Area metropolitana, Otra),Jornada, Tipo de colegio (Femenino, Masculino o Mixto)

Area Metropolitana						
Jornada		Mañana			Unificada	
Tipo	Alto	Medio	Bajo	Alto	Medio	Bajo
Femenino	19	17	14	8	2	5
Masculino	5	4	6	5	3	0
Mixto	10	15	43	9	5	26
Jornada		Tarde			Noche	
Tipo	Alto	Medio	Bajo	Alto	Medio	Bajo
Femenino	1	4	9	0	0	0
Masculino	2	3	5	0	1	0
Mixto	4	7	51	0	1	45
Otra Area Diferente de la Metropolitana						
Jornada		Mañana			Unificada	
Tipo	Alto	Medio	Bajo	Alto	Medio	Bajo
Femenino	0	1	1	1	2	7
Masculino	1	0	2	1	3	6
Mixto	0	4	46	2	12	108
Jornada		Tarde		Noche		
Tipo	Alto	Medio	Bajo	Alto	Medio	Bajo
Femenino	0	0	0	0	0	0
Masculino	0	0	0	0	0	1
Mixto	1	1	23	0	1	34

Fuente: Antioquia 1995. Estudios Descriptivos
Doc. No. 5, S.N.P. 115, ICFES, 1996.

Página www

Página de Abertura

Contenido

◀

▶

◀

▶

Página 32 de 100

Regresar

Full Screen

Cerrar

Abandonar

Ejemplo sobre rendimiento académico

Las siguientes tablas provienen del estudio de Cabarcas y Sierra (1996) sobre factores que afectan el rendimiento académico a los estudiantes de la Universidad Nacional-Sede Medellín. Las variables son

1. **Rendimiento:** El rendimiento se mide como el porcentaje de materias aprobadas con relación al número de materias registradas.
2. **Sexo :** *H* : Hombre y *M* : Mujer.
3. **Edad :** Hace al grupo de edad (de menor a mayor) de pertenencia del individuo.

Edad x Sexo x Rendimiento									
		Edad							
		1		2		3		4	
	Sexo	H	M	H	M	H	M	H	M
Rendimiento	Alto	14	8	26	14	9	2	2	1
	Medio	9	5	32	29	19	15	11	2
	Bajo	3	1	8	8	13	5	7	3

Página www

Página de Abertura

Contenido



Página 34 de 100

Regresar

Full Screen

Cerrar

Abandonar

Análisis Exploratorio de Datos Categóricos

- Mucho del trabajo realizado con datos es de carácter exploratorio. En el caso de variables continuas el trabajo exploratorio ha sido tecnificado mediante los trabajos de Chatfield y de Tukey.
- Para las tablas de contingencia no es tan obvio qué hacer, pero dependiendo de los datos y del conocimiento del problema que originó los datos podemos utilizar funciones de los valores de la tabla que nos den un mejor entendimiento del problema.

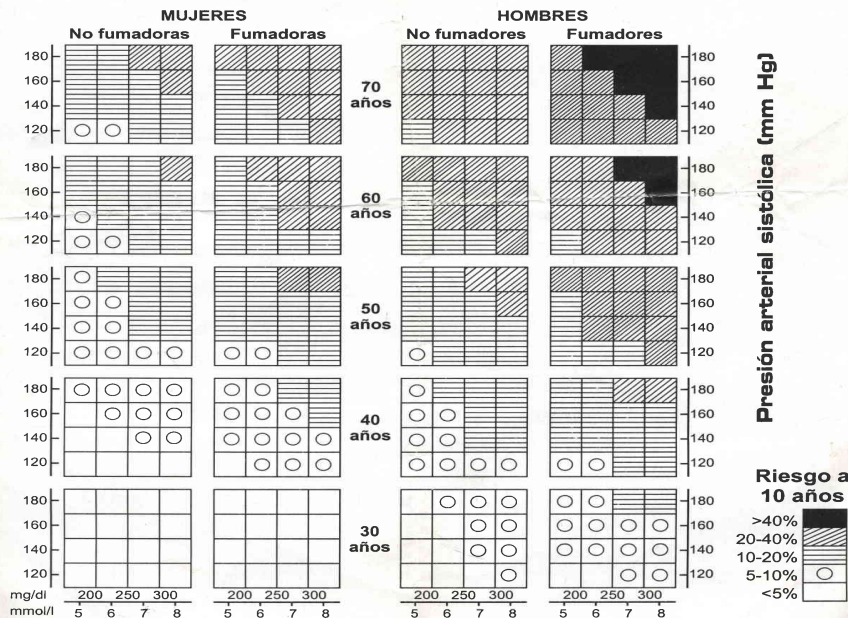
UNISALUD



EN UNISALUD LA PROMOCIÓN Y PREVENCIÓN
SON UNA OPCIÓN DE VIDA Y UN COMPROMISO DE TODOS



Gráfica de riesgo para enfermedad coronaria



Página www

Página de Abertura

Contenido



Página 36 de 100

Regresar

Full Screen

Cerrar

Abandonar

- El análisis exploratorio de tablas de contingencia consiste básicamente en el cálculo de porcentajes (proporciones)
 - marginales y
 - condicionales.
- Aunque hay propuestas gráficas no son muy populares, excepto por las más simples como son:
 - Gráficos de barras
 - Gráficos circulares (pastel)

Mortalidad en Instituciones Públicas o Privadas

	Resultado	
	Vivo	Muerto
Oficial	4757	430
Privado	5148	464

```
> partos.dat<-array(c(4757,5148,430,464),c(2,2))
```

```
> partos.dat
      [,1] [,2]
[1,] 4757  430
[2,] 5148  464
```

```
> rownames(partos.dat)<-c('Oficial','Privado')
```

```
> colnames(partos.dat)<-c('Vivos','Muertos')
```

```
> partos.dat
      Vivos Muertos
Oficial  4757    430
Privado  5148    464
```

```
> par(mfrow=c(1,2))
```

```
> barplot(partos.dat)
```

```
> barplot(t(partos.dat))
```

```
> pie(partos.dat,labels=c('Vivos y Público','Vivos y Privado',
                          'Muertos y Público','Muertos y Privado'))
```

Página *www*

Página de Abertura

Contenido

◀

▶

◀

▶

Página *37* de *100*

Regresar

Full Screen

Cerrar

Abandonar

[Página www](#)

[Página de Abertura](#)

[Contenido](#)



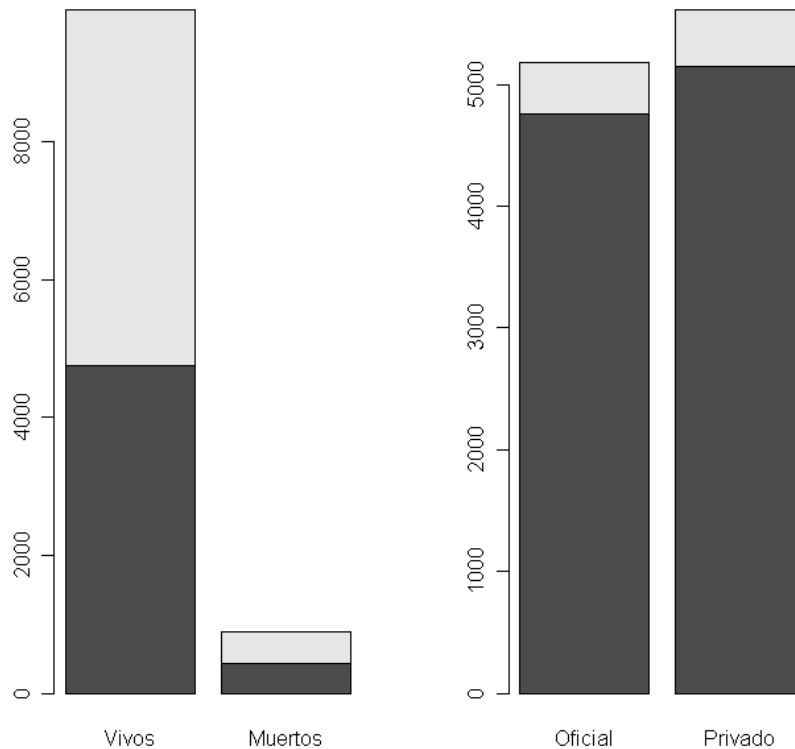
[Página 38 de 100](#)

[Regresar](#)

[Full Screen](#)

[Cerrar](#)

[Abandonar](#)



Página www

Página de Abertura

Contenido



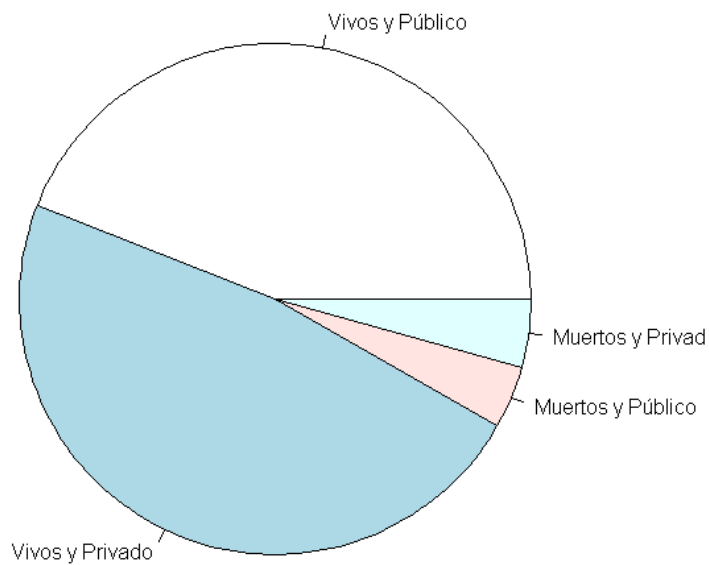
Página 39 de 100

Regresar

Full Screen

Cerrar

Abandonar



Página [www](#)

Página de Abertura

Contenido



Página **40** de **100**

Regresar

Full Screen

Cerrar

Abandonar

```
> prop.table(partos.dat)
      Vivos      Muertos
Oficial 0.4405038 0.03981850
Privado 0.4767108 0.04296694
> sum(prop.table(partos.dat))
[1] 1
>
```

```
> prop.table(partos.dat,1)
      Vivos      Muertos
Oficial 0.9171004 0.08289956
Privado 0.9173200 0.08267997
```

```
> prop.table(partos.dat,2)
      Vivos      Muertos
Oficial 0.4802625 0.4809843
Privado 0.5197375 0.5190157
>
```


Página www

Página de Abertura

Contenido



Página 41 de 100

Regresar

Full Screen

Cerrar

Abandonar

```
> matplot(t(prop.table(partos.dat,1)),type='b',
           ,ylab='Probabilidad',xlab='Vivo<-----> Muerto')
> legend(1.6,0.8,c('1: Oficial','2: Privado'))

> matplot(prop.table(partos.dat,1),type='b',
           ,ylab='Probabilidad',xlab='Oficial<-----> Privado')
> legend(1.6,0.8,c('1: Vivo','2: Muerto'))
```

Página www

Página de Abertura

Contenido



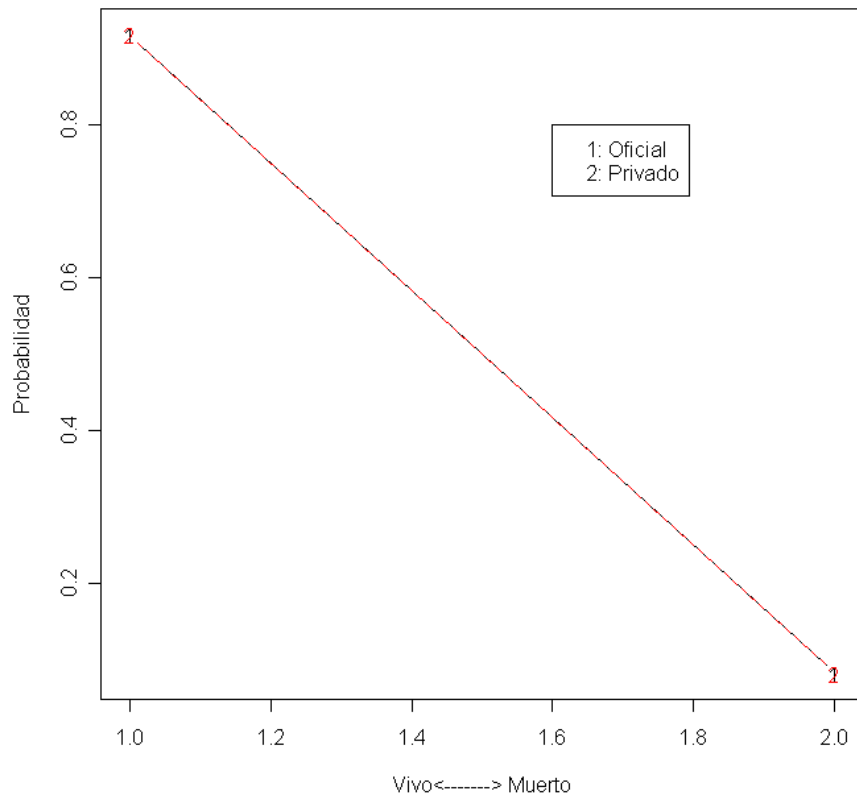
Página 42 de 100

Regresar

Full Screen

Cerrar

Abandonar



Página www

Página de Abertura

Contenido



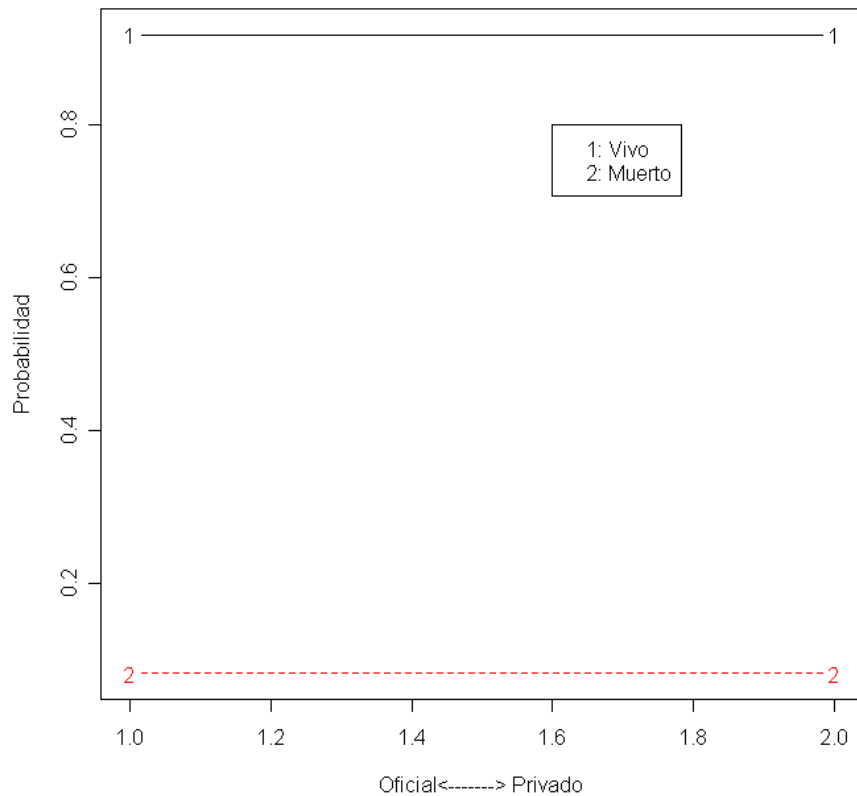
Página 43 de 100

Regresar

Full Screen

Cerrar

Abandonar



Página www

Página de Abertura

Contenido



Página 44 de 100

Regresar

Full Screen

Cerrar

Abandonar

Tipos de Sangre

```
datos<-array(c(
50,24,1,7,
36,17,0,3,
25,17,2,4,
55,30,1,9,
33,19,2,11,
29,20,2,4,
62,23,1,3,
26,27,0,0,
28,15,0,4,
65,28,1,8,
37,16,1,2,
28,10,0,4),c(4,3,2,2),
list(c('0','A','AB','B'),c('20-29','30-39','40-59'),
c('Hombre','Mujer'),c('Valle de Aburrá','Oriente')))
```

> datos
, , Hombre, Valle de Aburrá

	20-29	30-39	40-59
O	50	36	25
A	24	17	17
AB	1	0	2
B	7	3	4

, , Mujer, Valle de Aburrá

	20-29	30-39	40-59
O	55	33	29
A	30	19	20
AB	1	2	2
B	9	11	4

, , Hombre, Oriente

	20-29	30-39	40-59
O	62	26	28
A	23	27	15
AB	1	0	0
B	3	0	4

, , Mujer, Oriente

	20-29	30-39	40-59
O	65	37	28
A	28	16	10
AB	1	1	0
B	8	2	4

Página *www*

Página de Abertura

Contenido



Página *45* de *100*

Regresar

Full Screen

Cerrar

Abandonar

```

> margin.table(datos,1)
  0   A  AB   B
474 246  11  59
> prop.table( margin.table(datos,1))
           0           A           AB           B
0.60000000 0.31139241 0.01392405 0.07468354
> margin.table(datos,2)
20-29 30-39 40-59
  368   230   192
> prop.table( margin.table(datos,2))
      20-29      30-39      40-59
0.4658228 0.2911392 0.2430380
> margin.table(datos,3)
Hombre  Mujer
   375    415
> prop.table( margin.table(datos,3))
      Hombre      Mujer
0.4746835 0.5253165
> margin.table(datos,4)
Valle de Aburrá      Oriente
              401              389
> prop.table( margin.table(datos,4))
Valle de Aburrá      Oriente
      0.5075949      0.4924051
>

```

Página www

Página de Abertura

Contenido

◀

▶

◀

▶

Página 46 de 100

Regresar

Full Screen

Cerrar

Abandonar

Página www

Página de Abertura

Contenido



Página 47 de 100

Regresar

Full Screen

Cerrar

Abandonar

```
par(mfrow=c(2,2))
  pie( margin.table(datos,1))
  pie( margin.table(datos,2))
  pie( margin.table(datos,3))
  pie( margin.table(datos,4))
```

Página [www](#)

Página de Abertura

Contenido



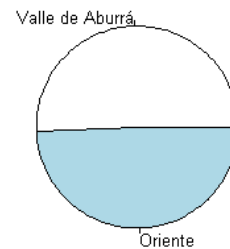
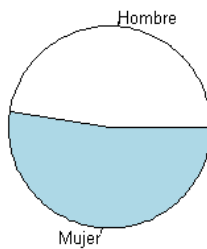
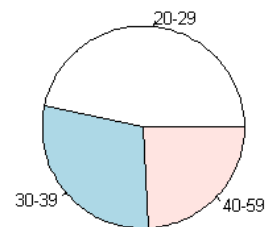
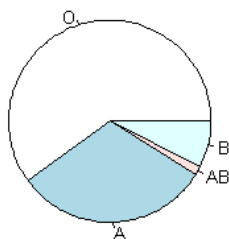
Página **48** de **100**

Regresar

Full Screen

Cerrar

Abandonar




```

> temp<-margin.table(datos,c(1,2))
> temp
      20-29 30-39 40-59
O      232   132   110
A      105    79    62
AB       4     3     4
B       27    16    16
> prop.table(temp)
      20-29      30-39      40-59
O  0.293670886 0.167088608 0.139240506
A  0.132911392 0.100000000 0.078481013
AB 0.005063291 0.003797468 0.005063291
B  0.034177215 0.020253165 0.020253165
> prop.table(temp,1)
      20-29      30-39      40-59
O  0.4894515 0.2784810 0.2320675
A  0.4268293 0.3211382 0.2520325
AB 0.3636364 0.2727273 0.3636364
B  0.4576271 0.2711864 0.2711864
> prop.table(temp,2)
      20-29      30-39      40-59
O  0.63043478 0.57391304 0.57291667
A  0.28532609 0.34347826 0.32291667
AB 0.01086957 0.01304348 0.02083333
B  0.07336957 0.06956522 0.08333333
>

```

Página www

Página de Abertura

Contenido

◀ ▶

◀ ▶

Página 49 de 100

Regresar

Full Screen

Cerrar

Abandonar

Página www

Página de Abertura

Contenido



Página 50 de 100

Regresar

Full Screen

Cerrar

Abandonar

```
> matplot(t(prop.table(temp,2)),type='b',  
          ylab='Probabilidad',xlab='Grupo de Etáreo')  
> legend(2,0.5,c('1: Tipo 0','2: Tipo A',  
                 '3: Tipo AB','4: Tipo B'))
```

[Página www](#)

[Página de Abertura](#)

[Contenido](#)



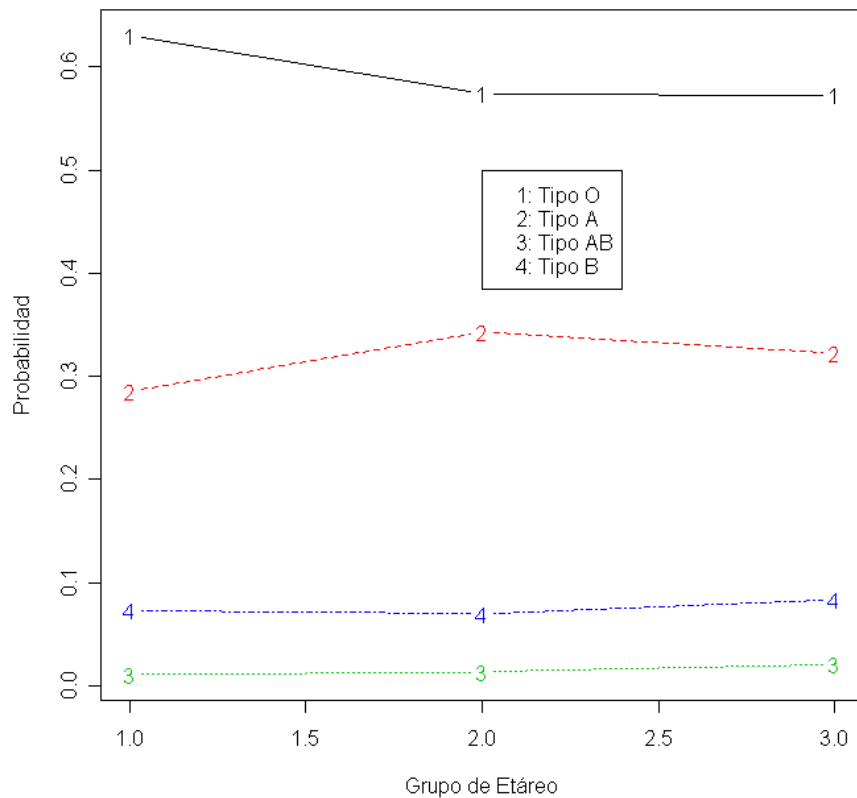
[Página 51 de 100](#)

[Regresar](#)

[Full Screen](#)

[Cerrar](#)

[Abandonar](#)



```
> datos[,,,2]
, , Hombre
```

	20-29	30-39	40-59
O	62	26	28
A	23	27	15
AB	1	0	0
B	3	0	4

```
, , Mujer
```

	20-29	30-39	40-59
O	65	37	28
A	28	16	10
AB	1	1	0
B	8	2	4

```
> datos[,,, 'Oriente']
, , Hombre
```

	20-29	30-39	40-59
O	62	26	28
A	23	27	15
AB	1	0	0
B	3	0	4

```
, , Mujer
```

	20-29	30-39	40-59
O	65	37	28
A	28	16	10
AB	1	1	0
B	8	2	4

```
>
```

[Página www](#)

[Página de Abertura](#)

[Contenido](#)

[◀◀](#)

[▶▶](#)

[◀](#)

[▶](#)

[Página 52 de 100](#)

[Regresar](#)

[Full Screen](#)

[Cerrar](#)

[Abandonar](#)

```
> temph<-temp[,1]
```

```
> temph
```

	20-29	30-39	40-59
O	62	26	28
A	23	27	15
AB	1	0	0
B	3	0	4

```
> prop.table(temph,2)
```

	20-29	30-39	40-59
O	0.69662921	0.490566	0.59574468
A	0.25842697	0.509434	0.31914894
AB	0.01123596	0.000000	0.00000000
B	0.03370787	0.000000	0.08510638

```
> tempm<-temp[,2]
```

```
> prop.table(tempm,2)
```

	20-29	30-39	40-59
O	0.637254902	0.66071429	0.6666667
A	0.274509804	0.28571429	0.2380952
AB	0.009803922	0.01785714	0.0000000
B	0.078431373	0.03571429	0.0952381

```
>
```

Página www

Página de Abertura

Contenido



Página 54 de 100

Regresar

Full Screen

Cerrar

Abandonar

```
matplot(cbind(temph.prop,tempm.prop),type='b',  
        xlab='O...A...AB...B ',ylab='Probabilidades')  
title(main='Distribución tipos de sangre: Oriente Antioqueño')  
legend(3,0.65,c('1: Hombre 20-29','2: Hombre 30-39',  
                '3:Hombre 40-59','4: Mujer 20-29','5: Mujer 30-39',  
                '6: Mujer 30-39'))
```

[Página www](#)

[Página de Abertura](#)

[Contenido](#)



[Página 55 de 100](#)

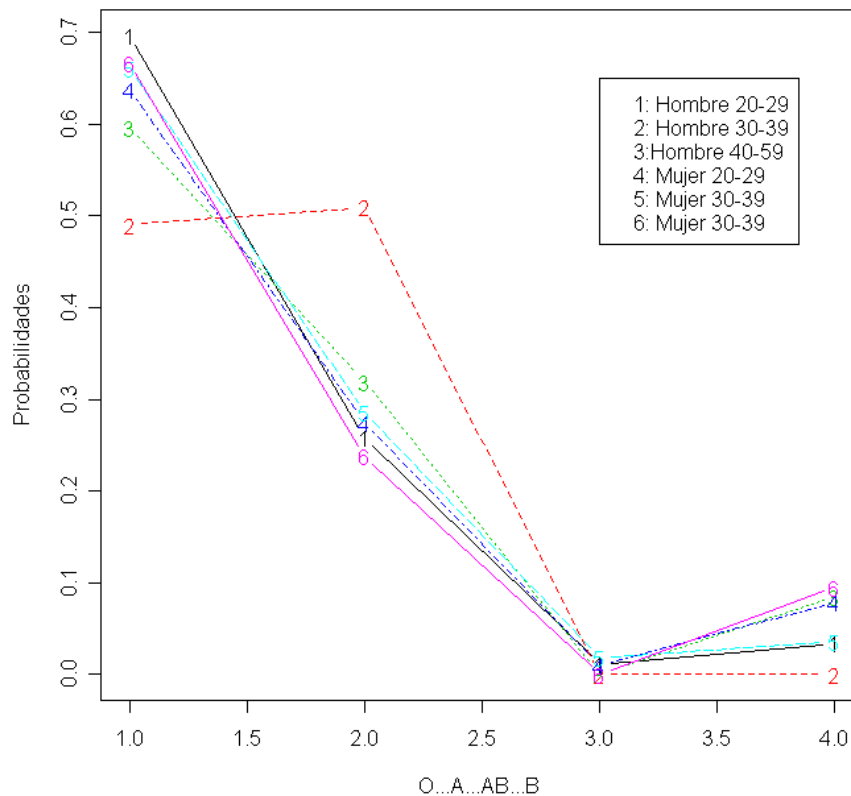
[Regresar](#)

[Full Screen](#)

[Cerrar](#)

[Abandonar](#)

Distribución tipos de sangre: Oriente Antioqueño



Página www

Página de Abertura

Contenido



Página 56 de 100

Regresar

Full Screen

Cerrar

Abandonar

En R existe la librería `vcd`, desarrollada por Michael Friendly para representar datos categóricos visualmente.


```
> tile(table(Sexo,P40_2,P40_3))}
```

vcd1.png

Página www

Página de Abertura

Contenido



Página 57 de 100

Regresar

Full Screen

Cerrar

Abandonar

```
> mosaic(table(Sexo,P40_2,P40_3))}
```

vcd2.png

Página www

Página de Abertura

Contenido



Página 58 de 100

Regresar

Full Screen

Cerrar

Abandonar