UNIDAD I. ESTADÍSTICA DESCRIPTIVA

La estadística es la ciencia que se encarga de recopilar, agrupar, presentar y analizar los datos.

La estadística puede dividirse en dos:

- Estadística descriptiva
- Estadística inferencial

La estadística descriptiva se dedica a recopilar, agrupar, presentar los datos. En ocasiones la serie de datos es muy difícil de analizar y de trabajar por el número relativamente grande de datos, por lo tanto conviene que cuando esta serie de datos exceda o sea igual a 25 se trabaje con clases o grupos de datos que se definirán por medio de un intervalo

Este proceso de agrupación se pude realizar mediante los siguientes pasos:

1. Encontrar el menor (m) y el mayor (M) de la serie de datos.
2. Encontrar el rango del conjunto de datos: R = M-m
3. Encontrar el número de clases que se utilizarán para la agrupación y estará determinado por: $\sqrt{n}$ donde n es el número de observaciones.
4. Definir la amplitud de cada uno de los intervalos:

5. Obtener los limites inferiores de cada intervalo
6. Realizar el conteo del número de datos que pertenecen a cada uno de los intervalos.

Ejemplo: Se tiene la siguiente serie de datos relacionada con la cantidad promedio de pesos que los alumnos gastan semanalmente por cuestión de gastos de transporte y alimentación:

10	235	256	178
180	153	145	299
155	130	123	245
399	139	145	158
201	200	210	209
111	121	156	123
201	222	123	189
98	197	099	100

Siguiendo el proceso o el procedimiento de agrupación señalado anteriormente tenemos

Paso 1

M = 399
m = 10

Paso 2

R = M- m = 399-10 = 389

Paso 3

Número de intervalos: $\sqrt{32}$ = 5.6569≈ 6 intervalos

Paso 4

Amplitud: $i = \frac{R}{No.deIntervalos} = \frac{389}{6} = 64.8\bar{3} \approx 65$

Nota: Este resultado se redondea al mismo máximo número de cifras significativas después del punto decimal de las observaciones en la serie.

Paso 5

Los límites inferiores son:

No. intervalo	Li
1	10
2	75
3	140
4	205
5	270
6	335

Y los límites inferiores se obtienen restando 1 valor significativo al límite inferior siguiente:

Los límites inferiores son:

No. intervalo	Li	Ls
1	10	74
2	75	139
3	140	204
4	205	269
5	270	334
6	335	399

Ahora se determinarán cuántas observaciones pertenecen a cada uno de los intervalos. Resultando:

No. intervalo	Li	Ls	fi
1	10	74	1
2	75	139	10
3	140	204	13
4	205	269	6
5	270	334	1
6	335	399	1

Una vez agrupada la serie de datos en intervalos o clases se podrán obtener de manera más sencilla las medidas descriptivas; estas se clasifican en: medidas de tendencia central y medidas de dispersión:

Como medidas de tendencia central podemos mencionar como las más usuales a la media aritmética, la moda y la mediana:

Y como medidas de variación o dispersión a el rango, la desviación media, la varianza y la desviación estándar.

La tabulación de información necesaria para encontrar dichas medidas es la siguiente:

No.	*Li*	*Ls*	*fi*	*xi*	*fixi*	*Li* Real	*Ls* Real	$\lvert xi - \bar{x} \rvert$	$fi \lvert xi - \bar{x} \rvert$	$(xi-\bar{x})^2$	$fi\,(xi-\bar{x})^2$	xi^2	$fixi^2$
1	10	74	1	42	42	9.5	74.5	127.96875	127.96875	16376.00098	16376.00098	1764	1764
2	75	139	10	107	1070	74.5	139.5	62.96875	629.6875	3965.063477	39650.63477	11449	114490
3	140	204	13	172	2236	139.5	204.5	2.03125	26.40625	4.125976563	53.63769531	29584	384592
4	205	269	6	237	1422	204.5	269.5	67.03125	402.1875	4493.188477	26959.13086	56169	337014
5	270	334	1	302	302	269.5	334.5	132.03125	132.03125	17432.25098	17432.25098	91204	91204
6	335	399	1	367	367	334.5	399.5	197.03125	197.03125	38821.31348	38821.31348	134689	134689
$\sum$			32		5439				1515.3125		139292.9688		1063753

Sabemos que:

$$\bar{x} = \frac{\sum fixi}{n} = \frac{5439}{32} = 169.96875$$

$$x = Li + \left(\frac{\frac{n}{2} - S}{f_{med}} \right) i = 139.5 + \left(\frac{\frac{32}{2} - 11}{13} \right) 65 = 139.5 + 25 = 164.5$$

$$x = Li + \left(\frac{d1}{d1 + d2}\right) i = 139.5 + \left(\frac{2}{2+7}\right) 65 = 139.5 + 14.4444 = 153.4444$$

Para las medidas de variación o dispersión tenemos que:

$$D.M. = \frac{\sum \text{fi} \left| \text{xi - } \bar{\text{x}} \right|}{\text{n}} = \frac{1515.3125}{32} = 47.35351563$$

$$S^2 = \frac{\sum \text{fi}\left(\text{xi} - \bar{\text{x}}\right)^2}{\text{n} - 1} = \frac{139292.9688}{31} = 4493.32157$$

Las 2 últimas columnas de la tabulación de la información únicamente son necesarias para utilizar la otra fórmula de la S; que a manera de ejemplo se mencionarán aquí:

$$S^2 = \frac{fixi^2 - \frac{\left(fixi\right)^2}{n}}{n-1} =$$

$$\frac{1063753 - \frac{5439^2}{32}}{31} = \frac{1063753 - 924460.0313}{31} = \frac{139292.9688}{31} = 4493.32157$$

$$S = \sqrt{S^2} = \sqrt{4493.321574} = 67.03222428$$

Ahora se dará otro ejemplo para que usted llene los espacios vacíos con el valor correspondiente

La siguiente serie de datos corresponde a los promedios de calificaciones de 50 estudiantes del CCH Naucalpan:

6.77	9.25	8.66	7.23	6.89
7.77	7.03	6.59	6.98	7.25
8.25	9.58	9.58	7.25	6.78
8.02	6.71	6.77	9.87	10.00
6.00	6.89	7.74	8.20	7.25
8.25	9.51	9.47	7.87	8.42
8.43	7.03	7.48	8.15	8.09
7.59	7.33	7.65	7.50	7.41
8.04	8.56	9.56	6.48	6.00
9.9	7.89	8.89	9.97	6.78

Realice la agrupación de datos para encontrar las medidas descriptivas.

Solución:

Paso 1

M = ________.

m =_________.

Paso 2

R = M-m = 4.00

Paso 3

$\sqrt{n} = \sqrt{\_\_} =$ _______ intervalos

Nota: Este resultado siempre se redondea a 0 decimales.

Paso 4

$$i = \frac{\qquad}{No.de\,\text{int}.} = \frac{}{7} = .57143 \approx .57$$

Nota: Se redondea a 2 porque máximo hay 2 cifras significativas después del punto decimal en cualquiera de las observaciones de la serie de datos.

Paso 5

Los límites inferiores son:

No.	Li
1	6
2	6.57
3	7.14
4	
5	8.28
6	
7	9.42

Y por lo tanto los límites superiores son:

No.	Li	Ls
1	6	6.56
2	6.57	7.13
3	7.14	
4		8.27
5	8.28	
6		
7	9.42	9.98

Este paso 6 nos proporciona un rango menor que el encontrado en el paso 2. Por lo tanto, se tendrá que utilizar i = 0.58 y se realizarán las modificaciones a los intervalos:

No. de intervalo	Li	Ls
1	6	6.57
2	6.58	
3		
4		8.31
5		
6		
7	9.48	10.05

Ahora tomamos en cuenta 5 posibles observaciones más, o sea, obtenemos un rango muy superior a lo estipulado en el paso 2, por lo tanto se realizará un equilibrio, este consiste en restar un cierto número de observaciones al primer límite inferior; en nuestro caso como son 5 las observaciones que se pasan, le restaremos 2 o 3; escogiendo 2 para restar.

No. de intervalo	Li	Ls
1	5.98	6.55
2	6.56	7.13
3	7.14	
4		
5		
6		
7		10.03

Con esto último se procederá a determinar cuantas observaciones pertenecen a cada uno de los intervalos y se realizarán todas las operaciones necesarias para encontrar las medidas descriptivas. Esta información también se muestra incompleta a manera de ejercicio:

No.	*Li*	*Ls*	*fi*	*xi*	*fixi*	*Li* **Real**	*Ls* **Real**	$\lvert xi - \bar{x} \rvert$	$fi \lvert xi - \bar{x} \rvert$	$(xi - \bar{x})^2$	$fi\,(xi - \bar{x})^2$	xi^2	$fixi^2$
1	5.98	6.55	3		18.795	5.975	6.555	1.6704	5.0112	2.7902	8.3707		
2	6.56		11	6.845	75.295			1.0904	11.9944				515.3943
3	7.14	7.71	10	7.425	74.25		7.715	0.5104		0.2605	2.6051		551.3063
4	7.72	8.29	11	8.005			8.295	0.0696	0.7656	0.0048	0.0533		
5	8.3	8.87	4	8.585	34.34	8.295	8.875					73.7022	
6	8.88	9.45	2		18.33	8.875		1.2296		1.5119	3.0238	83.9972	167.9945
7	9.46	10.03	9	9.745	87.705	9.455	10.035	1.8096	16.2864	3.2747	29.4719	94.9650	
$\sum$			50		396.77				44.2192		58.2914		3206.8201

Tome en cuenta el ejemplo anterior para encontrar las medidas descriptivas, cuyo valor se proporciona para su corroboración:

$\bar{x} = 7.9354$

$\tilde{x} = 7.7677$

$\hat{x} = 7.0706$

$\hat{x} = 7.7875$

$D.M. = 0.884384$

$S = 1.090697$

MEDIDAS DE TENDENCIA CENTRAL Y DISPERSIÓN PARA DATOS NO AGRUPADOS

1. Un profesor de aritmética quiere saber cuál es el promedio de las calificaciones de su grupo de 15 alumnos en el más reciente examen. Obtén el promedio si las calificaciones son las siguientes:

7.1 2.5 3.8 6.3 6.12
6.5 5.3 7.9 5.5 6.32
7.3 6.3 8.2 4.3 7.4

Solución: $\bar{x} = \frac{90.84}{} = 6.056$

2. Se pidió a 15 estudiantes del CCH Naucalpan seleccionados al azar que dijeran el número de horas que habían dormido la noche anterior. Los datos resultantes fueron: 5, 6, 6, 8, 7, 7, 9, 5, 4, 8, 10, 7, 5, 8, 7. Obtener lo siguiente:

a) la media $\bar{x}$; solución: $\bar{x} = \frac{}{15} = 6.8$.

b) la mediana $\tilde{x}$; solución: $\tilde{x} = 7$.

c) la moda. Solución: moda: 7.

3. Los siguientes datos son los aumentos de peso (en gramos) de pollos alimentados con una dieta rica en proteínas:

Aumento de peso	Frecuencia
12.5	2
12.7	6
13.0	22
13.1	29
13.2	12
13.8	4

Solución: $\bar{x} = \frac{}{} = 13.067$

$\tilde{x} = 13.1$

4. Obtener la media, la mediana y la moda para la siguiente distribución de frecuencias para datos no agrupados:

x	f
0	1
1	3
2	8
3	5
4	3

5. El número de respuestas incorrectas en una prueba de competencia de verdadero y falso para 15 estudiantes fueron los siguientes: 2, 1, 3, 0, 1, 3, 6, 0, 3, 4, 0, 5, 2, 6, 5. Encuentra la media, mediana y moda.

6. Dada el conjunto de datos 7, 6, 10, 7, 5, 9, 3, 7, 5, 13, obtener:

a) la varianza s^2;

Solución: $s^2 = \dfrac{73.6}{10-1} = \dfrac{73.6}{9} = 8.18$

b) la desviación estándar. Solución. $s = \sqrt{8.18} = 2.86$

7. Los contenidos de alquitrán de 8 marcas de cigarrillos seleccionadas al azar de la lista más reciente de la Procuraduría Federal del Consumidor (PFC) son los siguientes: 7.3, 8.6, 10.4, 16.1, 12.2, 15.1, 14.5 y 9.3 miligramos. Calcular:

a) la media;
b) la amplitud o rango;
c) la varianza y desviación estándar.

Solución:

a) $\bar{x} = \dfrac{93.5}{8} = 11.6875$

b) amplitud: $16.1 - 7.3 = 8.8$

c) $s^2 = \dfrac{75.42875}{8-1} = \dfrac{75.42875}{7} = 10.7755$; $s = \sqrt{10.7755} = 3.2826$

8. Obtener la media y la desviación estándar para cada uno de los siguientes conjuntos de datos:

a) 7.9, 9.2, 2, 1, 5, 4.5, 7.5, 6, 2.
b) 90, 87, 92, 81, 78, 85, 95, 80.

¿Qué efecto tendría sobre la media y desviación estándar el duplicar cada valor de un conjunto?

Solución:

a) $\bar{x} = \frac{45.1}{9} = 5.6375$; $s = \sqrt{\frac{71.080156}{9-1}} = \sqrt{\frac{71.080156}{8}} = 2.98077$

b) $\bar{x} = \frac{688}{8} = 86$; $s = \sqrt{\frac{260}{8-1}} = \sqrt{\frac{260}{7}} = 6.0945$

UNIDAD II. DATOS BIVARIADOS

La predicción estadística es una técnica que ayuda a predecir lo que ocurrirá en el futuro. El futuro, por lo general, no está determinado, es decir, está constituído por un conjunto de fenómenos aleatorios.

La predicción estadística se utiliza en diversos campos. Por ejemplo: la tendencia de precios de un producto, el desarrollo de productos competidores, predecir ventas, etc.

La predicción estadística se basa en una ecuación matemática obtenida pro medio de un análisis de regresión.

El análisis de regresión comprende el análisis de los datos muestrales para saber como se relacionan entre sí dos o más variables, trayendo consigo una ecuación matemática que describe dicha relación.

Existe también el análisis de correlación que generalmente resulta útil para tratar de determinar que variables tienen una mayor relación.

En el caso de un problema de dos variables, significaría que cada una de las observaciones proporcionará dos valores, uno por cada variable. Por ejemplo, dos variables de características físicas el peso y la estatura.

REGRESIÓN LINEAL

La regresión lineal simple comprende el intento de desarrollar una línea recta o una ecuación matemática lineal que describa la relación entre dos variables.

Una finalidad de una ecuación de regresión sería estimar los valores de una variable con base a los valores conocidos de la otra.

Otro uso de la ecuación de regresión es para predecir los valores futuros de una variable.

La forma general de una ecuación lineal es:

$$y = a + b\,x$$

donde a y b se obtendrán de los datos muestrales y de las siguientes fórmulas:

$$b = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{n\left(\sum x^2\right) - \left(\sum x\right)^2}$$

y

$$a = \frac{\sum y - b\sum x}{n}$$

Ejemplo 1. El presidente municipal de un pueblo del Estado de Tlaxcala tiene la siguiente información con respecto al número de habitantes de su población:

AÑO	NÚMERO DE HABITANTES
1998	12,000
1999	13,800
2000	15,700
2001	17,200
2002	18,900

Utilice esta información, para brindarle al presidente municipal una estimación o una predicción del número de habitantes que habrá para el año 2007 en esta comunidad.

Solución: Observe que la tabla anterior proporciona cantidades relativamente grandes para estar trabajándolas, por esta razón se propone utilizar la siguiente sustitución:

x (año)	Y (no. de habitantes en miles)
1	12
2	13.8
3	15.7
4	17.2
5	18.9

Esta sustitución pudo ser posible dado que el año tiene una diferencia constante entre un año y el próximo. Por lo tanto:

X	y	xy	x^2	y^2
1	12		1	144
2	13.8			
3	15.7	47.1		246.49
4	17.2		16	
5	18.9	94.5		357.21
$\sum x =$ _____	$\sum y =$ _____	$\sum xy = 250$	$\sum x^2 = 55$	$\sum y^2 = 1233.98$

$$b = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{n\left(\sum x^2\right) - \left(\sum x\right)^2}$$

$$b = \frac{5(\_\_\_\_\_\_\_) - (\_\_\_\_\_)(77.6)}{5(\_\_\_\_\_\_) - (\_\_\_\_\_\_)^2}$$

$$= \frac{1250 - \_\_\_\_\_}{\_\_\_\_\_ - 225}$$

$$= \frac{\_\_\_\_\_}{50} = 1.72$$

y

$$a = \frac{\sum y - b\sum x}{n}$$

$$a = \frac{\_\_\_\_\_\_ - 1.72(\_\_\_\_)}{5}$$

$$= \frac{77.6 - \_\_\_\_\_\_\_\_\_}{5} = \frac{\_\_\_\_\_\_}{5} = 10.36$$

Entonces :

$$y = 10.36 + 1.72x$$

Por lo tanto, si $\mathrm{x} = 10$ (año 2007)

$$\mathrm{y} = 10.36 + 1.72(10)$$
$$= 27.56$$

Se estima que para el año 2007 habrá 27,560 habitantes en esa población.

El coeficiente de correlación entre estas dos variables estará dado por:

$$\mathrm{r} = \frac{\mathrm{n}\left(\sum \mathrm{xy}\right) - \left(\sum \mathrm{x}\right)\left(\sum \mathrm{y}\right)}{\sqrt{\left[\mathrm{n}\left(\sum \mathrm{x}^2\right) - \left(\sum \mathrm{x}\right)^2\right]\left[\mathrm{n}\left(\sum \mathrm{y}^2\right) - \left(\sum \mathrm{y}\right)^2\right]}}$$

observe que muchas operaciones de la fórmula ya las realizó, por lo tanto no le será difícil obtener que

r = 0.99925718324

Ejemplo 2. CONASIDA tiene la siguiente información referente al número de infectados de sida en cierta comunidad del estado de Guanajuato.

x (año)	Y (número de infectados)
1997	280
1998	300
1999	358
2000	410
2001	500
2002	510

Utilice esta información para ajustar una recta y realizar una estimación del número de infectados que habrá en esa comunidad para el año 2010

Solución. Se realizará una modificación semejante a la anterior para agilizar las operaciones necesarias, usted llene los espacios vacíos con el valor correspondiente durante todo el proceso de solución.

X	y	xy	x^2	y^2
1	280	280		78400
2		600		
3	358			
4			16	
	500	2500		250000
6		3060	36	
$\sum x = 21$	$\sum y = 2358$	$\sum xy = 9154$	$\sum x^2 = 91$	$\sum y^2 = 974764$

$$b=\frac{n\left(\sum xy\right)-\left(\sum x\right)\left(\sum y\right)}{n\left(\sum x^2\right)-\left(\sum x\right)^2}$$

$$b=\frac{6(\_\_\_\_\_\_\_)-(\_\_\_\_\_)(2358)}{6(\_\_\_\_\_\_)-(\_\_\_\_\_\_)^2}$$

$$=\frac{54924-\_\_\_\_\_}{\_\_\_\_\_-441}$$

$$=\frac{5406}{105}=51.4857$$

y

$$a=\frac{\sum y-b\sum x}{n}$$

$$a=\frac{\_\_\_\_\_\_-51.4857(21)}{6}$$

$$=\frac{2358-\_\_\_\_\_\_\_\_\_}{6}=\frac{1276.8003}{6}=212.8$$

Entonces:

$$y=212.8+51.4857x$$

Por lo tanto, si $x=14$ (año 2010)

$$y=212.8+51.4857(\_\_\_)$$

$$=933.5998\approx 934$$

Entonces se estima que para el año 2010 habrá en esa comunidad 934 habitantes infectados de sida.

Y el coeficiente de correlación es r = 0.98235555555. Usted realice las operaciones correspondientes para corroborar.

Ejemplo 3. En el CCH-N se seleccionaron al azar a 5 alumnas de sexto semestre para preguntarles su estatura y peso. La información que se recopiló fue la siguiente:

Estatura	Peso
1.57	52
1.55	51
1.62	59
1.55	50
1.69	65

Utilice esta información para realizar una estimación del peso que tendrá una estudiante de sexto semestre del CCH-N que tiene una estatura de 1.71 m.

Solución. Al igual que en los ejercicios anteriores se proporcionará el procedimiento y usted únicamente llenará los espacios vacíos con el valor correspondiente:

X	y	xy	x^2	y^2
1.57	52	81.64		
1.55	51		2.4025	
1.62	59			
1.55	50	77.5		2500
1.69	65		2.8561	
$\sum x =$ ____	$\sum y =$ ______	$\sum xy = 443.62$	$\sum x^2 = 12.7504$	$\sum y^2 = 15511$

$$b = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{n\left(\sum x^2\right) - \left(\sum x\right)^2}$$

$$b = \frac{5(\_\_\_\_\_\_) - (\_\_\_\_\_)(277)}{5(\_\_\_\_\_\_) - (\_\_\_\_\_\_)^2}$$

$$= \frac{2218.1 - \_\_\_\_\_}{\_\_\_\_\_ - 63.6804}$$

$$= \frac{\_\_\_\_\_\_}{0.0716} = 106.7039$$

y

$$a = \frac{\sum y - b\sum x}{n}$$

$$a = \frac{\_\_\_\_\_\_ - \_\_\_\_\_\_\_(7.98)}{5}$$

$$= \frac{277 - \_\_\_\_\_\_\_\_\_}{5} = \frac{-574.4971}{5} = -114.8994$$

Entonces:

$$y = -114.8994 + 106.7039x$$

Por lo tanto, si $x = 1.71$

$$y = -114.8994 + 106.7039(\_\_\_\_)$$
$$= 67.564269$$

Se estima que una alumna del CCH-N de sexto semestre que mida 1.71m de estatura tendrá un peso de 67.56 kgs.

Y el coeficiente de correlación entre estas dos variables es r = 0.99345185241. Usted realice las operaciones correspondientes para corroborar.

UNIDAD III. PROBABILIDAD

Probabilidad teórica

$$P(A) = \frac{\text{número de resultados favorables o asociados al evento A}}{\text{número total de resultados posibles}}$$

Probabilidad como frecuencia relativa

$$P(A) = \frac{\text{número de veces que ocurrió el evento A}}{\text{número total de ensayos realizados}}$$

cuando el número total de ensayos realizados es grande

Interpretación Subjetiva de probabilidad

La probabilidad subjetiva se interpreta como un grado de creencia o confianza que se tiene de la ocurrencia de un resultado al realizarse un fenómeno aleatorio. En este contexto, la probabilidad representa un juicio personal acerca de un fenómeno impredecible.

Espacio Muestral

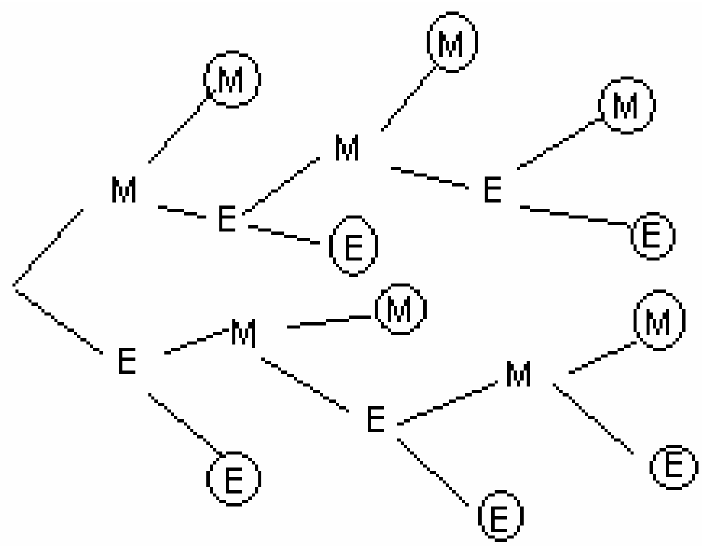
Es el conjunto total de posibles resultados de un fenómeno aleatorio

Técnicas de conteo

Diagramas de árbol

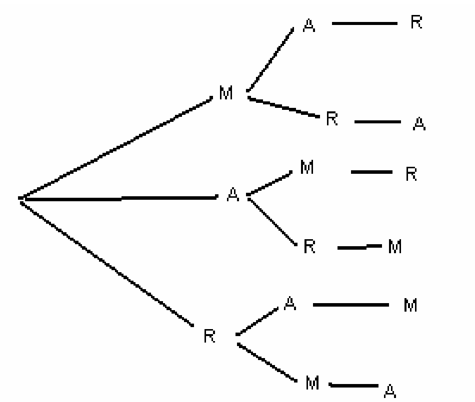
Un diagrama de árbol es el dibujo que se usa para enumerar todos los posibles resultados de una serie de experimentos en donde cada experimento puede suceder en un número finito de maneras.

Ejemplo:
Marcos y Enrique intervienen en un torneo de tenis. La primera persona que gane 2 juegos seguidos o que complete 3 gana el torneo.
¿cuáles son los posibles resultados?



Los posibles resultados se muestran en el diagrama de árbol y son 10.

Ejemplo: ¿Cuántas palabras se pueden formar con las letras de la palabra MAR sin que ninguna letra se repita?
(aunque carezcan de significado).
Solución:



Las palabras que se pueden formar son 6.

Notación Factorial

El factorial de un número n, entero y positivo es:

$$n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 2 \cdot 1$$

Se define a 0! como 1.

Ejemplos:

$$3! = 3 \cdot 2 \cdot 1 = 6$$

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

$$\frac{5!}{4!} = \frac{5 \cdot \cancel{4!}}{\cancel{4!}} = 5$$

$$\frac{n!}{(n+2)!} = \frac{n!}{(n+2)(n+2-1)(n+2-2)} = \frac{\cancel{n!}}{(n+2)(n+1)\cancel{n!}} = \frac{1}{(n+1)(n+2)}$$

Ejercicios para el alumno.
Realice la simplificación llenando los espacios vacíos:

$$\frac{(n-1)!}{(n+1)!} = \frac{(n-1)!}{(n+1)(n+1-\_\_)(n+1-2)!} = \frac{\cancel{(n-1)!}}{(n+\_)(n)\cancel{(n-1)!}} = \frac{1}{(n+1)n}$$

$$\frac{n!}{(n-1)!} = \frac{n(n-\_\_)!}{(n-\_\_)!} = n$$

$$\frac{23!}{20!\,3!} = \frac{23 \cdot \_\_ \cdot \_\_ \cdot \cancel{20!}}{\cancel{20!}\,3!} = \frac{23 \cdot 22 \cdot 21}{\_\_!} = \frac{10626}{\_\_} = 1771$$

Para calcular las probabilidades de varios eventos es necesario contar el número de resultados posibles de un experimento, o contar el número de resultados que son favorables a un evento dado. El proceso de conteo puede simplificarse mediante el empleo de 2 técnicas de conteo denominadas: permutaciones y combinaciones.

Una permutación es un arreglo en un orden particular de los objetos que conforman un conjunto.

Una combinación de los objetos de un conjunto es una selección de estos sin importar el orden.

La diferencia entre una permutación y combinación es que en la primera el interés recae en contar todas las posibles selecciones y todos los arreglos de estas, mientras que en la segunda el interés solo recae en contar el número de selecciones diferentes.

Se denotará a una permutación de n elementos tomando r a la vez como:

$$P(n,r) = \frac{n!}{(n-r)!}$$

Se denotará a una combinación de n elementos tomando r a la vez como:

$$C(n,r) = \frac{n!}{(n-r)!r!}$$

Ejemplos:
1.- ¿Cuántos comités de 3 personas se pueden formar con 8 personas?
Solución. Sabemos que no nos importa el orden de las personas que seleccionemos para el comité y tampoco podemos seleccionar una persona más de una vez, por lo tanto, utilizaremos la fórmula de las combinaciones:

$$C(n,r)=\frac{n!}{(n-r)!r!}=\frac{8!}{(8-3)!3!}=\frac{8!}{5!3!}=\frac{8\cdot 7\cdot 6\cdot \cancel{5!}}{\cancel{5!}3!}=\frac{8\cdot 7\cdot 6}{3!}=\frac{8\cdot 7\cdot \cancel{6}}{\cancel{6}}=56$$

2.- ¿Cuántas permutaciones distintas se pueden formar con todas las letras de la palabra AMOR sin que ninguna letra se repita?
Solución. Sabemos que las palabras se forman dependiendo del número de letras que conforman la palabra y el orden en que estas estén colocadas, pro lo tanto, importa el orden y se utilizará la fórmula de las combinaciones:

$$P(n,r)=\frac{n!}{(n-r)!}=\frac{4!}{(4-4)!}=\frac{4!}{0!}=4!=24$$

Ejercicios para el alumno. Llene los espacios vacíos con el valor correspondiente:
3.- ¿De cuántas maneras puede escogerse un comité compuesto de 3 hombres y 2 mujeres de un grupo de 7 hombres y 5 mujeres?
Solución.

$$C(7,3)\cdot C(5,2)=\frac{7!}{(7-3)!3!}\cdot\frac{5!}{(5-2)!2!}=\frac{7!}{\_\_!3!}\cdot\frac{\_\_!}{3!\,\_\_!}=\frac{7\cdot 6\cdot 5\cdot \cancel{4!}}{\cancel{4!}\,\_\_!}\cdot\frac{5\cdot \_\_\cdot \cancel{3!}}{\cancel{3!}\,\_\_!}=\frac{7\cdot \cancel{6}\cdot 5}{\cancel{6}}\cdot\frac{5\cdot 4}{2}=35\cdot \_\_=350$$

Realice los siguientes ejercicios y corrobore su resultado con el que se proporciona:

4.- Un estudiante tiene que contestar 10 de 12 preguntas en un examen:
¿Cuántas maneras de escoger tiene?
Solución: No nos interesa el orden con el que contestamos las 10 preguntas, utilice la fórmula correcta para llegar al resultado de **66 maneras.**

¿Cuántas maneras de escoger tiene si las tres primeras son obligatorias?
Solución. Si las 3 primeras son obligatorias, le sobran 9 preguntas para seleccionar las otras 7 preguntas, realice las operaciones correspondientes para llegar al resultado de **36 maneras.**

¿Cuántas maneras de escoger tiene si tiene que contestar 4 de las 5 primeras?
Solución. Para contestar esta pregunta piense que él puede seleccionar 4 de las 5 primeras y, por lo tanto seleccionar 6 (que le faltan contestar) de las otras 7; complete el análisis y obtenga el resultado de **35 maneras.**

5.- Si no se permiten repeticiones:
   a) ¿Cuantos números de 3 dígitos se pueden formar con los dígitos 2, 3, 5, 6,7 y 9?
   Interesa el orden entonces utilizamos fórmula de _____________.**Respuesta: 120**

b) ¿cuántos de estos son menores a 400? **Respuesta: 40**
c) ¿cuántos son pares? **Respuesta:40**
d) ¿cuántos son impares? **Respuesta:80**
e) ¿cuántos son múltiplos de 5? **Respuesta: 20**

6.- Una persona olvidó los 3 últimos dígitos de un número telefónico de 8 dígitos, esta persona recuerda que ninguno de esos 8 dígitos se repite. Marcó todos los números que podían ser, atinándole hasta el último posible. ¿Cuántos números tuvo que marcar? Solución 60 números.

7.- Un profesor escoge a 100 alumnos de 103. ¿De cuantas formas posibles puede escogerlos? Solución. 176851.

Probabilidad Condicional

Sea E un evento arbitrario de un espacio muestral S con P (E)>0. La probabilidad de que un evento A suceda una vez que E haya sucedido o, en otras palabras, la probabilidad condicional de A dado E, se define como:

$$P(A \setminus B) = \frac{P(A \cap B)}{P(B)} \qquad \text{donde } P(B)>0$$

Ejemplo: Se lanzan 2 dados, si la suma es 6, hallar la probabilidad de que uno de los dados sea 5.

Solución: Los elementos del espacio muestral son:

+	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

La probabilidad de que las caras de los dados sumen 6 es de 5/36 ó P (suma sea 6)= P(A)=5/36, y la probabilidad de que la suma sea 6 y uno de los dados haya sido 5 es 2/36 ó P (A∩B)=2/36. Utilizando la fórmula tenemos que:

$$P(B \setminus A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{2}{36}}{\frac{5}{36}} = \frac{2 \cdot \cancel{36}}{5 \cdot \cancel{36}} = \frac{2}{5}$$

Ejemplo: Se lanza un dado, si el número es par, ¿cual es la probabilidad de que sea primo?

Solución.
Definimos al evento A = {número par} = {2, 4,6}
y a B = {número primo} = {2, 3,5}, por lo tanto A∩B= {2}
Entonces:

$$P(B \setminus A) = \frac{P(B \cap A)}{P(A)} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

Ejemplo: Se lanzan un par de dados, si los números que resultan son diferentes. Hallar la probabilidad de que su suma sea par.

Solución. El espacio muestral se puede representar de la siguiente manera:

+	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

P (números diferentes) = P (D) = 30/36
P (suma sea par) = P (T) = 18/36

Ejemplo: Se escogen al azar dos dígitos diferentes del 1 al 9.

a) Si la suma de los 2 números escogidos es impar, ¿cuál es la probabilidad de que el dos sea uno de los 2 números escogidos? Sol. 1/3
b) Si el dos es uno de los números seleccionados, ¿cuál es la probabilidad de que la suma sea impar? Sol. 6/11

Eventos independientes

Se dice que un evento A es independiente de un evento E si la probabilidad de que A suceda no está influenciada porque E haya o no sucedido. En otras palabras:

$$P(A \setminus B) = P(A)$$
$$entonces:$$
$$\frac{P(A \cap B)}{P(B)} = P(A)$$
$$\therefore P(A \cap B) = P(A) \cdot P(B)$$

Ejemplo: La probabilidad de que A dé en el blanco es de ¼ y la de B es de 2/5. Si A y B disparan,

a) ¿cuál es la probabilidad de que se pegue en el banco?

Solución. Se busca la probabilidad de que de menos 1 de en el blanco, esto es:

$$\begin{aligned}
P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
&= P(A) + P(B) - P(A)P(B) \\
&= \frac{1}{4} + \frac{2}{5} - \frac{1}{4} \cdot \frac{2}{5} \\
&= \frac{1}{4} + \frac{2}{5} - \frac{2}{20} \\
&= \frac{5+8-2}{20} \\
&= \frac{11}{20}
\end{aligned}$$

b) ¿Cuál es la probabilidad de que no se dé en el blanco?

Solución. La probabilidad deseada es $P(A^C \cap B^C)$, y se lee la probabilidad de que A **no** de en el blanco **y** B **no** de en blanco, esto es:

$$\begin{aligned}
P(A^C \cap B^C) &= P\left[(A \cup B)^C\right] && \text{Ley de D' Morgan} \\
&= 1 - P(A \cup B) && \text{Por axioma de probabilidad} \\
&= 1 - \left[P(A) + P(B) - P(A \cap B)\right] \\
&= 1 - \left[P(A) + P(B) - P(A)P(B)\right] \\
&= 1 - \left[\frac{11}{20}\right] \\
&= \frac{9}{20}
\end{aligned}$$

Este resultado también se pudo obtener sabiendo que:

$$P(A^C) = \frac{3}{4} \quad \text{y} \quad P(B^C) = \frac{3}{5}$$

entonces como A y B son eventos independientes sus complementos también lo son, entonces:

$$P(A^C \cap B^C) = P(A^C) \cdot P(B^C)$$
$$= \frac{3}{4} \cdot \frac{3}{5}$$
$$= \frac{9}{20}$$

c) ¿Cuál es la probabilidad de que los dos acierten en el blanco?

Solución. La probabilidad de que los dos acierten esta representado por: $P(A \cap B)$ y se lee la probabilidad de que A acierte **y** B acierte.

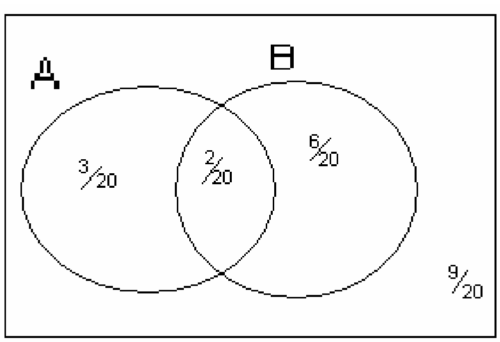
$$P(A \cap B) = P(A) \cdot P(B)$$
$$= \frac{1}{4} \cdot \frac{2}{5} = \frac{2}{20}$$
$$= \frac{1}{10}$$

Las probabilidades muchas veces se pueden obtener y representar por medio de un diagrama de Venn, revise y compare los resultados con su representación:

Solución. Primero manejaremos el mínimo común denominador de las probabilidades proporcionadas en el enunciado:

$$\frac{1}{4} = \frac{5}{20}$$
$$\frac{2}{5} = \frac{8}{20}$$

Entonces:



Cheque que al sumar las probabilidades dentro de los círculos la suma es: $\frac{11}{20}$

Ejercicio. La probabilidad de que Hugo pase su examen de aritmética es de $\frac{2}{7}$ y de que Ciro pase ese mismo examen de $\frac{3}{5}$. Encuentre la probabilidad de que...

a) ninguno pase el examen. Sol. 10/35
b) los dos pasen el examen. Sol. 6/35
c) de menos uno pase el examen. Sol. 25/35
d) solamente uno pase el examen. Sol. 19/35

Ejercicio. La probabilidad de que Ricardo Mora termine su carrera es $\frac{4}{5}$ y la de Enrique Cuenca es de $\frac{3}{4}$. La vida de estas dos personas son totalmente independientes entre sí, ¿cuál es la probabilidad de que...

a) las dos personas la concluyan? Solución 12/20
b) Ninguno de los dos la concluya? Solución 1/20
c) De menos uno la concluya? Solución 19/20
d) Solamente uno de los dos la concluya? Solución 7/20

PROBLEMAS DE PROBABILIDAD

1. Sea P(A) = 0.2, P(B) = 0.3 y A y B eventos independientes. Obtenga:

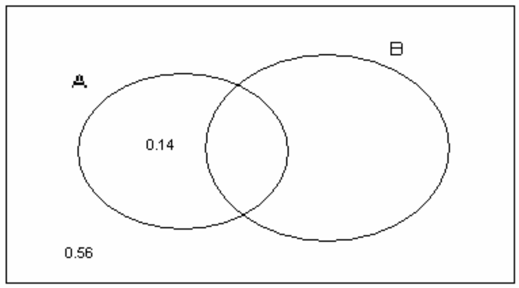
a)$P(A \cap B)$

b)$P(A \cup B)$

c)$P(A^C \cap B^C)$

d) Representa por medio de un diagrama de Venn.

Solución:
a) 0.06
b) 0.44
c) 0.56
d) Se proporciona el diagrama de Venn de manera parcial. Usted complete la representación:



2. Sea P(A) = 0.4 y P(B) = 0.2, y A y B eventos mutuamente excluyentes. Obtenga:

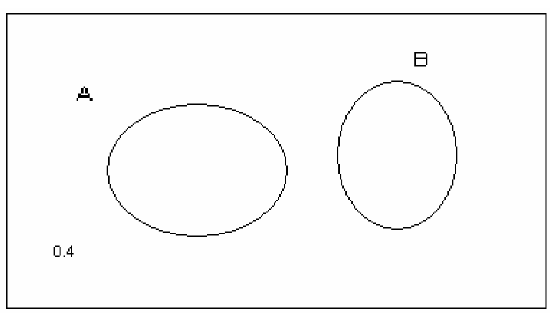
a) $P(A \cup B)$

b) $P(A \cap B)$

c) $P(B^C)$

d) Represente por medio de un diagrama de Venn

Solución:
a) 0.6
b) 0
c) 0.8
d) Se proporcionará de manera parcial:



3. Sea P(A) = 0.4, P(B) = 0.2 y $P(A \cap B) = 0.1$. Obtenga:

a) $P(A \cup B)$

b) $P(A^C \cap B^C)$

c) Representa por medio de un diagrama de Venn

Solución:
a) 0.5

b) 0.5

4. Sea P(A) = 0.5, P(B) = 0.3 y $P(A \cup B) = 0.6$. Obtenga:

a) $P(A \cap B)$

b) $P(A^C \cap B^C)$

c) $P((A \cap B)^C)$

d) Represente por medio de un diagrama de Venn

Solución:
a) 0.2
b) 0.4
c) 0.8

5. Un furgón contiene siete sistemas electrónicos complejos. Sin que el comprador lo sepa, hay tres defectuosos. Se seleccionan dos de siete para someterlos a pruebas exhaustivas y se clasifican como defectuosas o no defectuosas.

a) Haga una lista de los puntos muestrales para este experimento. Sol. Habrá 21 puntos muestrales, enlístelos.
b) Sea A el suceso de que la selección no incluye defectuosos. Haga una lista de los puntos muestrales en A. Sol. Habrá 6 elementos en el evento A, enlístelos.
c) Asigne probabilidades a los puntos muestrales y encuentre P(A). Solución. 0.2857

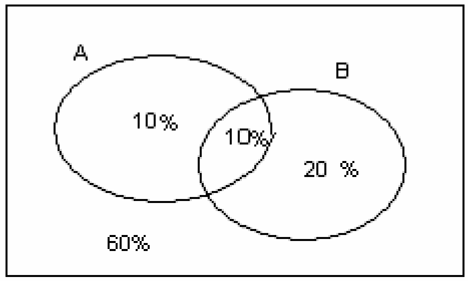
6. Un dado se lanza dos veces. ¿Cuál es la probabilidad de que la suma de los números observados sea mayor que 9? Solución. 6/36.

7. Dos ambulancias se mantienen listas para emergencias. Debido a la demanda y a la posibilidad de fallas mecánicas, la probabilidad de que una ambulancia específica esté disponible cuando se necesite es 9/10. La disponibilidad de una ambulancia es independiente de la otra.

a) En caso de catástrofe, ¿cuál es la probabilidad de que ambas ambulancias estén disponibles? Sol. 0.81
b) ¿Cuál es la probabilidad de que ninguna esté disponible? Sol. 0.01
c) Si se necesita una ambulancia en una emergencia, ¿cuál es la probabilidad de que al menos haya una disponible? Solución. 0.99
d) Si se necesita una ambulancia en una emergencia, ¿cuál es la probabilidad de que haya solamente una disponible? Sol. 0.18

8. Se debe examinar un grupo de personas respecto a dos sistemas comunes de cierta enfermedad. Se considera que 20% de las personas presentan el síntoma A, 30% tiene el síntoma B, 10% tiene ambos síntomas y el resto no tiene síntoma alguno. Para una persona escogida al azar de este grupo, encuentre las probabilidades de los eventos siguientes:

a) Que la persona no presente síntoma alguno. Sol. 0.60
b) Que la persona presente al menos un síntoma. Sol. 0.40
c) Que la persona presente ambos síntomas, dado que presenta el síntoma B. Sol. 0.3333

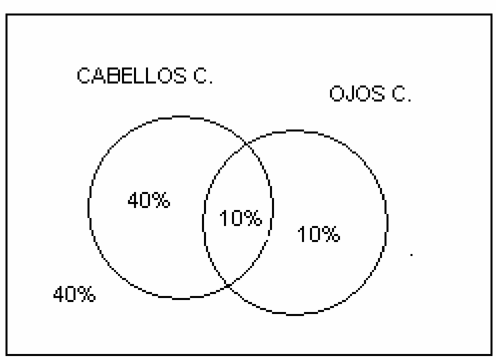


9. De 10 niñas de clase, 3 tienen ojos azules. Si se escogen dos niñas al azar, ¿cuál es la probabilidad de que...
(i) las dos tengan ojos azules? Sol. 3/45
(ii) ninguna tenga ojos azules? Sol. 21/45
(iii) una por lo menos tenga ojos azules? Sol. 24/45

10. En cierta ciudad, 50% de la población tienen cabellos castaños, 20% tiene ojos castaños y 10% tiene cabellos y ojos castaños. Se escoge una persona al azar

i) Si tiene cabellos castaños, ¿cuál es la probabilidad de que también tenga ojos castaños? Sol. 0.25
ii) Si tiene ojos castaños, ¿cuál es la probabilidad de que no tenga cabellos castaños? Sol. 0.5
iii) ¿Cuál es la probabilidad de que no tenga cabellos castaños ni ojos castaños? Sol. 0.40

Diagrama de Venn que representa la situación.



11.En una escuela particular de bachillerato se tiene la siguiente información con respecto a los idiomas que estudian sus 125 alumnos:

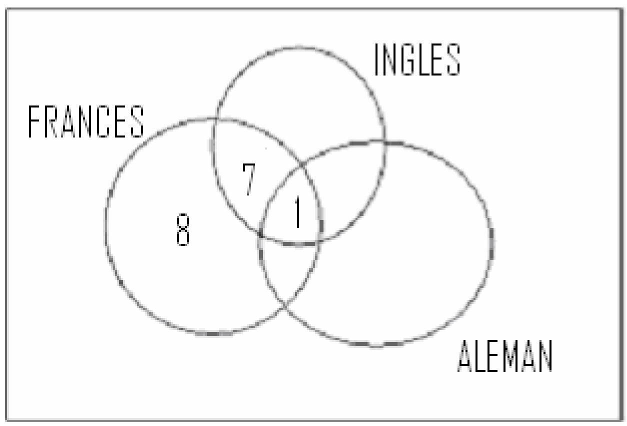
1 solamente estudia los tres idiomas que se imparten en la escuela.
5 estudian francés y alemán.
10 estudian inglés y alemán.
8 estudian francés e inglés.
9 estudian solamente inglés.
22 estudian francés.

Y 75 de esos 125 que son el total de los alumnos en la escuela no estudian algún idioma.

Si se selecciona a un alumno al azar de esta escuela, ¿Cuál es la probabilidad de que...

a) estudie francés? Solución: 0.176
b) No estudie algún idioma? Solución: 0.6
c) Estudie alemán? Solución: 0.2
d) De que estudie de menos algún idioma? Solución: 0.4
e) Estudie solamente alemán? Solución: 0.072

Se proporcionará parcialmente el diagrama de Venn que representa la situación, usted tendrá la labor de completarlo:



12. Un vendedor de autos caros tiene 23 clientes de los cuales 17 son millonarios, 8 jubilados, incluidos 4 que también son millonarios. Si seleccionamos de manera aleatoria a uno de sus clientes, ¿Cuál es la probabilidad de que el cliente...

a) sea millonario o jubilado? Sol. 21/23
b) no sea millonario ni jubilado? Sol. 2/23

13. De un grupo de 48 personas que asisten a una fiesta, 30 fuman, 25 consumen bebidas alcohólicas y 10 ni fuman ni consumen bebidas alcohólicas. Al seleccionar a una persona de manera aleatoria de esta fiesta, ¿Cuál es la probabilidad de que...

a) fume e ingiera bebidas alcohólicas? Sol. 17/48
b) únicamente fume? Sol. 13/48
c) únicamente ingiera bebidas alcohólicas? Sol. 8/48

14. Al SILADIN han asistido 100 alumnos a presentar los exámenes de Matemáticas y Física. Sabiendo que las matemáticas la han aprobado 54 alumnos en total, Física 75 alumnos y 40 han aprobado ambas asignaturas, ¿cuál es la probabilidad de que al seleccionar a un alumno al azar...

a) no haya aprobado ninguna asignatura? Sol. 0.11
b) haya aprobado una de las dos asignaturas? Sol. 0.49
c) haya aprobado las dos asignaturas? Sol. 0.40

15. Un grupo de “cecehacheros” realizó una encuesta a los estudiantes varones del CCH Naucalpan para obtener la siguiente información referente al tipo de persona que dicen ser:

20% dice ser tranquilo únicamente
15% dice ser tranquilo y estudioso pero no “chambeador”
5% dice ser tranquilo, estudioso y “chambeador”
7% dice ser tranquilo y “chambeador” pero no estudioso
20% dice ser estudioso y “chambeador” pero no tranquilo
10% dice no tener ninguna de esas características
53% dice tener la característica de ser estudioso

Retomando esta información y al seleccionar a un estudiante varón del CCH Naucalpan, ¿Cuál es la probabilidad de que...

a) Sea estudioso? 0.53
b) Sea estudioso pero no tranquilo ni “chambeador”? Sol. 0.13
c) Sea “chambeador” y tranquilo, pero no estudioso? Sol. 0.07
d) Sea únicamente “chambeador”? Sol. 0.10
e) Sea “chambeador”? 0.42

16. En una encuesta de 100 estudiantes se observó la siguiente información: 39 alumnos estudian Inglés, 39 Francés, 36 Portugués, 15 estudian Inglés y Francés, 18 Francés y Portugués, 19 Inglés y Portugués y 5 estudian los tres idiomas.

a) Construya el diagrama de Venn correspondiente.
b) ¿Cuál es la probabilidad de que al seleccionar de manera aleatoria a un estudiante de esos 100 éste estudie 2 idiomas exactamente? Sol. 0.37
c) ¿ Cuál es la probabilidad de que al seleccionar de manera aleatoria a un estudiante de esos 100 éste estudie 1 idioma exactamente? Sol. 0.27
d) Cuál es la probabilidad de que al seleccionar de manera aleatoria a un estudiante de esos 100 éste estudie Francés y Portugués? Sol. 0.18

17. En un grupo de 60 personas, se sabe que 40 trabajan, 24 estudian y 27 tiene coche. Si 55 trabajan o tienen coche, 49 trabajan o estudian, 40 estudian o tiene coche y 2 de esas personas trabajan, estudian y tienen coche. Si seleccionamos a una persona de este grupo, ¿Cuál es la probabilidad de que...

a) trabaje, estudie y tenga coche? Sol. 2/60
b) tenga coche pero no estudie ni trabaje? Sol. 6/60
c) no trabaje, ni estudie y tampoco tenga coche? Sol. 5/60

d) no trabaje, pero si estudie y tenga coche? Sol. 9/60

18. Un investigador del área de Psicología aplica un Test a 100 individuos, 30 adolescentes y el resto adultos. La prueba permite determinar si un individuo es neurótico o no lo es. De las 100 personas, 20 adultos y 10 adolescentes presentaron rasgos de neurosis. Si se definen los eventos:

N = {persona neurótica}
A = {persona adulta}
D = {persona adolescentes}

Obtener:

a) $P(N \cup D)$, Sol. 0.50
b) $P(N \cap D^C)$, Sol. 0.20
c) $P(N \cup D^C)$, Sol. 0.80
d) $P(A \cap D^C)$, Sol. 0.70
e) $P(N^C \cup D^C)$, Sol. 0.90

19. Los empleados de una compañía se encuentran separados en tres divisiones: Administración (A), Ventas (V) y Bodega (B). La siguiente tabla indica el número de empleados en cada división clasificados por sexo:

	Mujer (M)	Hombre (H)	Total
Administración (A)	120		
Ventas (V)	90		180
Bodega (B)		100	
Total		220	500

Si se elige aleatoriamente un empleado:

a) ¿Cuál es la probabilidad de que sea mujer? Sol. 280/500
b) ¿Cuál es la probabilidad de que sea mujer dado que no es un empleado de Bodega? Sol. 210/330
c) ¿Cuál es la probabilidad de que sea mujer y trabaje en Ventas? Sol. 90/500
d) ¿Cuál es la probabilidad de que sea hombre o trabaje en Administración? Sol. 340/500

e) ¿Cuál es la probabilidad de que no sea mujer o trabaje en Ventas o trabaje en Bodega? Sol. 380/500

Obtenga:

f) $P(AUM^C)$, Sol. 340/500
g) $P(A^C \cap M^C)$, Sol. 190/500
h) $P((AUM)^C)$, Sol. 390/500
i) $P(H \setminus B)$, Sol. 100/170
j) $P(A^C \setminus H^C)$, Sol. 160/280
k) $P(A \setminus V)$, Sol. 0
l) $P(V^C \setminus H)$, Sol. 130/220

20. En una investigación que se realizó en los antros se obtuvo la siguiente información:

	Fuma (F)	No Fuma	Total
Mayor de edad (M)			42%
Menor de edad (N)		5%	
Total	82%		100%

Obtenga:

a) $P(FUM^C)$, Sol. 0.87
b) $P(F^C \cap M^C)$, Sol. 0.05
c) $P((FUM)^C)$, Sol. 0.05
d) $P(N \setminus F)$, Sol. 0.6463
e) $P(N^C \setminus F^C)$, Sol. 0.72
f) $P(F \setminus N)$, Sol. 0.9138