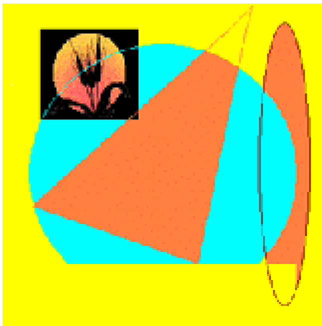
TODO MATEMÁTICAS

VOLUMEN 11

Estadística Descriptiva. Probabilidad. Distribuciones Discretas y Continuas



PROMOCIÓN
NO VENTA

Alejo González Criado

Profesor Numerario de Matemáticas

Destinado a

El Fígaro autodidacta:

Todo aquel que albergue algún interés por las Matemáticas y disfrute con su estudio.

Obra completa:

Formación básica,

Formación nivel medio

Formación nivel alto

Figuras y gráficos del autor

Edita: *El Autor*

Primera edición Mayo 2017

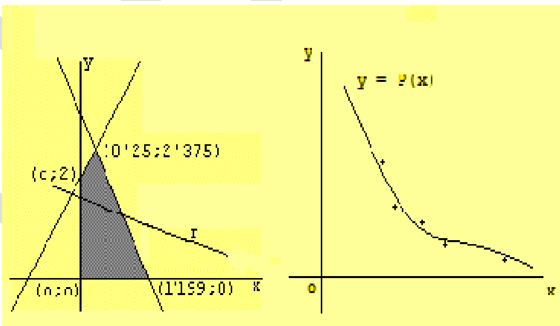
Editado en España

ISBN:
Depósito Legal:

VOLUMEN 11

Parte I

Parte II



Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

ÍNDICE

pág.:

Parte I

Tema 2 Estadística Descriptiva, dos variables
(Variable bidimensional (X,Y))

Tema 3 TEORÍA y Cálculo de Probabilidades

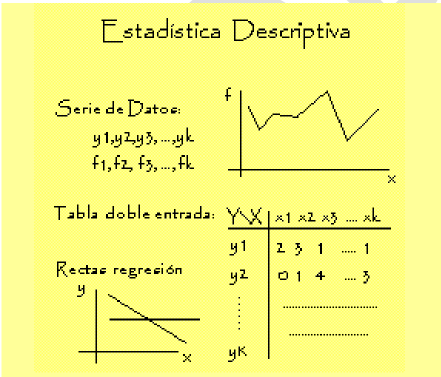
Tema 4 Variable aleatoria y Distribuciones

Tema 5 Distribuciones discretas

Tema 6 Distribución Continua

Tema 1

ESTADÍSTICA Descriptiva en una variable



Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

1.- ESTADISTICA Descriptiva de una variable

1.0.- INTRODUCCIÓN. Conceptos previos:

Los datos que se tratan en Estadística Descriptiva son extraídos de lo que llamamos ‘Población’, que no es necesariamente lo que normalmente entendemos por población, sino ‘el conjunto de individuos (en sentido amplio) sobre los que recae el estudio Estadístico a realizar.

Pero la población puede ser muy amplia y en la mayoría de los casos no será posible la consulta a cada uno de sus individuos. Por esta razón serán seleccionados un subconjunto de individuos de la población, que llamamos ‘Muestra’, y hacemos la consulta a cada uno de dichos individuos.

La consulta se referirá a un determinado ‘carácter’ o ‘circunstancia’ que afecta al individuo, como puede ser algún hecho o fenómeno natural o social.
De esta consulta obtenemos ‘valores’ en sentido amplio, los cuales puedan ser tratados matemáticamente.

Llamamos ‘variable’ estadística al ente X, que irá recibiendo los valores-resultados obtenidos en la consulta. Esto nos permite hacer referencia a la lista o serie de valores obtenidos, así:
x1, x2, x3, ..., xi, ..., xn , en caso

finito, ó bien
x1, x2, x3, ..., xi, ..., en caso

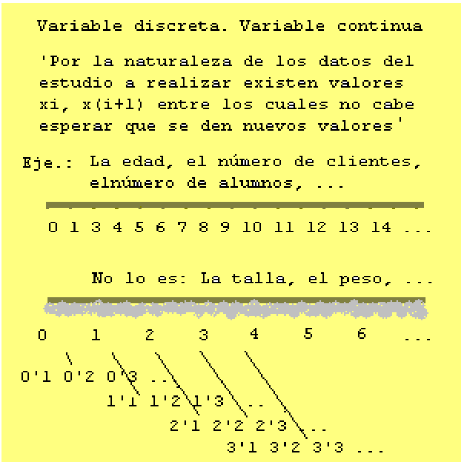
ilimitado numerable (Este caso no será tratado aquí).

El valor-resultado de la consulta, y por tanto el valor de X, puede ser de dos tipos:

a) Cualitativo: Por ejemplo una calificación del individuo u objeto así: bueno, malo; blanco, negro; sano, enfermo; ...
b) Cuantitativo: Cuando hacemos una valoración que permite traducirlo a valores numéricos.

En el caso de variable cuantitativa, esta puede ser de dos tipos:

b1) **Discreta:**



Los valores son enteros, y de forma que existen pares de valores, como 2 y 3, entre los cuales no es posible que resulte otro valor.

Ejemplo:

La edad computada en años, o computada en meses, o computada en días; El número de asignaturas aprobadas; …

b2) **Continua:**
Si para todo par de valores xi, xj, existe la posibilidad de presentarse otro valor.

Ejemplo:

Las tallas de un grupo de personas pueden darnos valores como 1'64 y 1'65, y si apreciásemos hasta las milésimas (de metro, como debe hacerse) sería 1'640, 1'650; puede presentarse un individuo que nos dé el valor 1'645.

Algo parecido ocurre cuando tomamos el peso de los recién nacidos durante una semana, por ejemplo.

La Muestra siempre será un conjunto finito, por tanto la lista de valores obtenidos también lo es.

1.1.- VARIABLE Cuantitativa discreta

1.1.1.- SERIES de Datos y sus características básicas.

Es una lista de N valores: $x_1, x_2, x_3, \ldots, x_n$ (1)

Estadística descriptiva: Una variable

xi	x1	x2	x3	...	xn
fi	f1	f2	f3	...	fn

Eje.:

xi	1	2	3	4	5	6	
fi	3	5	2	4	2	1	N= 17

xi	fi
x1	f1
x2	f2
x3	f3
.	.
.	.
	N= S(fi)

O bien de esta otra forma

FRECUENCIA absoluta de xi:
Es el número de veces que figura el valor xi en la serie anterior. La representamos por fi.

Los valores xi pueden estar repetidos; en este caso llamamos fi el número de veces que se presenta este valor xi. Decimos que los datos están 'agrupados', y designaremos por yi los valores distintos.

Si los datos están agrupados tenemos:

y1, y2, y3, ..., yk
f1, f2, f3, ..., fk

(2)

NOTA: En la figura 2, donde xi debemos ver yi
Si utilizamos la serie (1), con los valores sin agrupar, las frecuencias toman el valor 1:

f1 = 1, f2 = 1, f3 = 1, ..., fn = 1

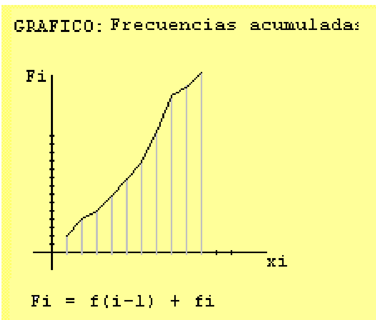
Observa que: N = f1 + f2 + f3 +... + fk, número total de valores de la serie.

FRECUENCIA relativa de yi:

Es el valor fri = $\frac{fi}{N}$

Observa que: fr1 + fr2 + fr3 + ... + frk = 1, fácil de comprobar.

FRECUENCIA absoluta acumulada de yi:



Es el valor Fi obtenido como sigue

$$Fi = f1 + f2 + f3 + ... + fi,$$

es decir, ‘suma de las frecuencias absolutas desde f1 hasta fi (Observa la figura)

1.1.2.- MEDIDAS DE CENTRALIZACIÓN

A los siguientes valores se les suele llamar ‘estadísticos’.

MEDIA aritmética:

Es el valor Mx obtenido a partir de la serie de datos (1), o de la serie (2), del siguiente modo:

$$Mx = \frac{x1 + x2 + x3 + \ldots + xn}{N} \qquad (3)$$

obien

$$Mx = \frac{f1.y1 + f2.y2 + f3.y3 + \ldots + fk.yk}{N} \qquad (3)'$$

MODA:

Es el valor yj cuya frecuencia fj es la más alta.
Puede tener más de un valor modal.

MADIA Cuadrática:

Es el valor Mcx obtenido como sigue:

$$Mcx = \sqrt{\frac{y_1^2.f_1 + y_2^2.f_2 + \cdots + y_k^2.f_k}{N}} \qquad (4)$$

Donde k es el número de valores distintos.

MEDIA Geométrica:

Es el valor Mgx obtenido como siguie:

$$\text{Mgx} = \sqrt[N]{y_1^{f1} \cdot y_2^{f2} \cdot \ldots \cdot y_k^{fk}} \qquad (5)$$

Recuerda que N= f1+f2+ ... +fk

MEDIA Armónica:

Es el valor Max obtenido como sigue:

Media de los inversos:

$$\text{Mix} = \frac{\frac{1}{y1} \cdot f1 + \frac{1}{y2} \cdot f2 + \cdots + \frac{1}{yk} \cdot fk}{N}$$

Ahora el valor Max es el ‘inverso’ de la media de los inversos:

$$\text{Max} = \frac{1}{Mix} \quad \text{(media armónica)} \qquad (6)$$

MEDIANA:

Es el valor Me de la variable X tal que el 50 % (es decir, la mitad) del total de los valores xi son menores o iguales que Me, y el otro 50 % son mayores que Me.

Para obtener su valor en circunstancias concretas debemos distinguir dos casos como sigue.

A) DATOS SIN AGRUPAR en frecuencias (fi = 1 para todo xi)

En variable discreta, Si los datos están ordenados crecientemente se admite aplicar lo siguiente:

a) Si N es impar, el valor de Me es el valor central de la serie:

$$Me = xh, \text{ donde } h = 1 + \frac{N-1}{2} \qquad (7)$$

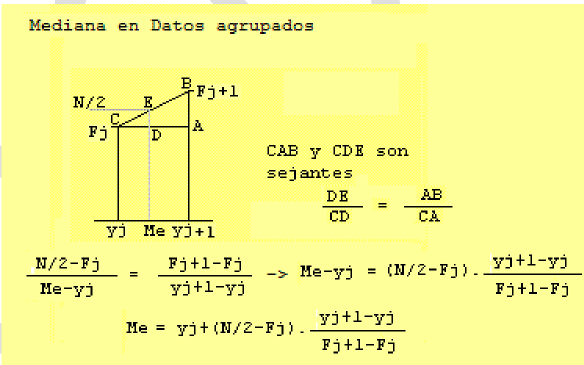
b) Si N es par, el valor de Me es la media de los dos valores centrales:

$$Me = \frac{x_h + x_{h+1}}{2}, \quad \text{donde } h = \frac{N}{2} \qquad (8)$$

B) DATOS AGRUPADOS tomando su frecuencia fi absoluta

En variable discreta, si los datos están ordenados crecientemente, se acepta la siguiente fórmula deducida como se muestra en la figura.
Calculamos la serie de frecuencias absolutas acumuladas Fi.

En primer lugar detectamos el valor Fj de las frecuencias acumuladas tal que Fj $\leq \frac{N}{2}$ y que F(j+1) $> \frac{N}{2}$.



Después tomamos el valor

$$Me = y_j + \left(\frac{N}{2} - F_j\right). \frac{(y_{j+1} - y_j)}{F_{j+1} - F_j} \qquad (9)$$

Observa que el segundo sumando es el resultado del ‘prorrateo’ mediante distribución lineal.

CUARTILES Q1, Q2, Q3:

Cuartil significa ‘cuarta parte’.

Q1 es el valor de X tal que 1/4 (25 %)de los valores son menores o iguales que Q1, y los otros 3/4 quedan por encima.

Q2 coincide con la mediana: Q2 = Me, ya ahora es el 50 % lo que se tiene en cuenta.

Q3 es el valor de X tal que 3/4 (75 %)de los valores son menores o iguales que Q3, y 1/4 quedan por encima.

La fórmula para obtener su valor es semejante a la de la media Me.

Cálculos:

$$Q1 = y_j + \left(\frac{N}{4} - F_j\right). \frac{(y_{j+1} - y_j)}{F_{j+1} - F_j}$$

donde F_j es tal que $F_j \leq N/4$ y $F_{j+1} > N/4$

$$Q3 = y_j + \left(\frac{3.N}{4} - F_j\right). \frac{(y_{j+1} - y_j)}{F_{j+1} - F_j}$$

donde F_j es tal que $F_j \leq 3.N/4$ y $F_{j+1} > 3.N/4$

DECILES: Decil significa ‘décima parte’.

‘Decil d es el valor xd de X tal que el $\frac{d}{10}$.100 por ciento de los valores son menores o iguales que xd’.

Su cálculo:

$$xd = y_j + \left(\frac{d.N}{10} - F_j\right). \frac{(y_{j+1} - y_j)}{F_{j+1} - F_j}$$

donde F_j es tal que $F_j \leq d.N/10$ y $F_{j+1} > d.N/10$

PERCENTILES: Percentil significa 'centésima parte'.

'Percentil p es el valor xp de X tal que el $\frac{d}{100}$.100 por ciento de los valores son menores o iguales que xp'

Su cálculo:

$$xp = y_j + \left(\frac{p.N}{100} - F_j\right). \frac{(y_{j+1} - y_j)}{F_{j+1} - F_j}$$

donde F_j es tal que $F_j \leq$ p.N/100 y $F_{j+1} >$ p.N/100

1.1.3.- MEDIDAS DE DISPERSIÓN

DESVIACIÓN media:

Es la media aritmética de los valores absolutos de las desviaciones respecto del valor Mx (media aritmética). Lo representamos por Dm (desviación media).

Su cálculo:

$$Dm = \frac{|y1-Mx|.f1 + |y2-Mx|.f2 + ... + |yk-mx|.fk}{N} \qquad (7)$$

VARIANZA: Es la media aritmética del cuadrado de las desviaciones respecto de Mx. Lo representamos por Vx.

Su valor:

$$Vx = \frac{(y1-Mx)^2.f1 + (y2-Mx)^2.f2 + ... + (yk-Mx)^2.fk}{N} \qquad (8)$$

DESVIACIÓN Típica: Es la raíz cuadrada de la varianza. Lo designamos por Sx.

Su cálculo:

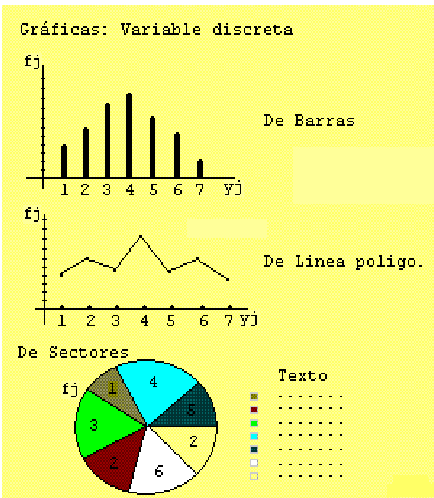
$$Sx = \sqrt{V_x} \quad \text{(tomando el valor positivo)} \qquad (9)$$

1.1.4.- GRAFICAS en variable discreta

En variable cuantitativa discreta los más usados son:

-Diagrama de barras
-Polígono de frecuencias
-Diagrama de Sectores

En variable cualitativa se utilizan los diagramas de sectores, de tarta, pictogramas, ...



Ejemplos/Ejercicios

1.- Serie estadística:

5, 3, 4, 1, 2, 8, 9, 8, 7, 6, 6, 7, 9, 8, 7, 7, 1, 0, 1, 5, 9, 9, 8, 0, 8, 8, 8, 9, 5, 7.

Determina:

a) Frecuencias absolutas
b) Frecuencias absolutas acumuladas
c) Frecuencias relativas
d) Frecuencias relativas acumuladas

Sol.:

a)

xi	0	1	2	3	4	5	6	7	8	9
fi	2	3	1	1	1	3	2	5	7	5

N = 30

b)

xi	0	1	2	3	4	5	6	7	8	9
Fi	2	5	6	7	8	11	13	18	25	30

c)

xi	0	1	2	3	4	5	6	7	8	9
fri	0’07	0’1	0’03	0’03	0’03	0’1	0’07	0’17	0’23	0’17
%	7	10	3	3	3	10	7	17	23	17

d)

xi	0	1	2	3	4	5	6	7	8	9
Fri	0’07	0’17	0’20	0’23	0’26	0’36	0’43	0’60	0’83	1
%	7	17	20	23	26	36	43	60	83	100

2.- Tomando la serie de datos tratados en el anterior (número 1) obtener:

a) La Media aritmética
b) La Mediana
c) La Moda

Sol.: a)

xi	0	1	2	3	4	5	6	7	8	9
fi	2	3	1	1	1	3	2	5	7	5
xi.fi	0	3	2	3	4	15	12	35	56	45

$$ma = \frac{\sum xi.fi}{N} = \frac{175}{30} = 5'833$$

b)

xi	0	1	2	3	4	5	6	7	8	9
Fi	2	5	6	7	8	11	13	18	25	30

$$N/2 = 15, \quad me = \frac{6+7}{2} = 6'50$$

c)

xi	0	1	2	3	4	5	6	7	8	9
fi	2	3	1	1	1	3	2	5	7	5

La frecuencia absoluta más alta es $f_8 = 7$, $mo = 8$

3.- Tomando la serie de datos tratados en el núm. 1, obtener:

a) La Desviación media
b) La Varianza
c) La Desviación típica

Sol.-

a) Recordamos que ma = 5’83

xi	fi	|xi-ma|	|xi-ma|.fi
0	2	5’83	11’66
1	3	4’83	14’49
2	1	3’83	3’83
3	1	2’83	2’83
4	1	1’83	1’83
5	3	0’83	2’49
6	2	0’17	0’34
7	5	1’17	5’85
8	7	2’17	15’19
9	5	3’17	15’85
	N=30		Suma = 74’36

$$Dm = \frac{\sum abs(xi - ma).fi}{N} = \frac{74'36}{30} = 2'48$$

b)

xi	fi	(xi-ma)	$(xi\text{-}ma)^2$	(xi-ma).fi
0	2	5’83	33’99	67’98
1	3	4’83	23’33	69’99
2	1	3’83	14’67	14’67
3	1	2’83	8’01	8’01
4	1	1’83	3’35	3’35
5	3	0’83	0’69	2’07
6	2	-0’17	0’03	0’06
7	5	-1’17	1’37	6’85
8	7	-2’17	4’71	32’97
9	5	-3’17	10’05	50’25
	N=30			Suma = 256’20

$$S^2 = \text{Var.} = \frac{\sum (xi - ma)^2 . fi}{30} = \frac{256'20}{30} = 8'54$$

c)

$$S = \text{Desviación típica} = \sqrt{Var.} = \sqrt{8'54} = 2'92$$

4.- Las puntuaciones obtenidas en un test aplicado a 20 alumnos son las siguientes:

16, 22, 21, 20, 23, 22, 17, 15, 13, 22, 17, 18, 20, 17, 22, 16, 23, 21, 22, 18

Obtener: La media, los cuartiles, la desviación típica

Sol.:

xi	fi	Fi	xi.fi	xi^2	$xi^2.fi$
13	1	1	13	169	169
15	1	2	15	225	225
16	2	4	32	256	512
17	3	7	51	289	867
18	2	9	36	324	648
20	2	11	40	400	800
21	2	13	42	441	882
22	5	18	110	484	2420
23	2	20	46	529	1058
N=20			Sum=385		Sum=7581

ma = 385/20 = 19'25

Cuartil Q1:

N/4 = 5, F4 = 7 > 5 y F3 < 5 y por tanto: Q1 = 17

Cuartil Q2:

2.N/4 = 10, F6 > 10 y F5 < 10 --> Q2 = 20

Cuartil Q3:

3.N/4 = 15, F8 > 15 y F7 < 15 --> Q3 = 22

Varianza:

$$S^2 = \frac{\sum xi^2.fi}{20} - ma^2 = \frac{7581}{20} - (19'25)^2 = 8'49$$

Desv. típica: S = 2'91

5.- En cierta Región, y durante cierto mes de 30 días, se registraron las siguientes temperaturas máximas:

32, 31, 28, 29, 29, 33, 32, 31, 30, 31, 31, 27, 28, 29, 30, 32, 31, 31, 30, 30, 29, 30 30, 30, 31, 30, 31, 34, 33, 33

Obtener:

a) Moda, la media ma, los percentiles 30 y 70

b) La varianza y la desv. típica

Sol.:

xi	fi	Fi	xi.fi	xi^2	xi^2.fi
27	1	1	27	729	729
28	2	3	56	784	1568
29	5	8	145	841	4205
30	7	15	210	900	6300
31	8	23	248	961	7688
32	3	26	96	1024	3072
33	3	29	99	1089	3267
34	1	30	34	1156	1156
N=30			Sum=915		Sum=27985

a) mo = 31, ma = 30'50,

p_{30} : 30.N/100 = 9'00, F3 = 9 --> p_{30} = 29

p_{70} : 70.N/100 = 21'00, F5 > 21 y F4 < 21 --> p_{70} = 31

a) $S^2 = \frac{7985}{30} - (30'50)^2 = 2'58$,

$S = \sqrt{2,58} = 1'61$

1.2.- VARIABLE Cuantitativa Continua

1.2.1.- Introducción: Definición

Es el caso en el que, por la naturaleza de la variable X a estudiar, ocurre que:

'Al tomar una muestra de Datos para X, entre dos valores a, b cualesquiera siempre cabe la posibilidad de que se presenten otros valores para X'.

Ejemplos:
Las tallas de un grupo de personas pueden darnos valores como 1'64 y 1'65, y si apreciásemos hasta las milésimas (de metro, como debe hacerse) sería 1'640, 1'650; Ocurre que, si seguimos tomando valores, pueden presentarse individuos que nos den los valores: 1'642, 1,644, 1'645, ...

Algo parecido ocurre cuando pretendemos construir una tabla de datos con los pesos de es colectivo.

Un caso habitual lo tenemos al tomar el peso, y la talla, de los recién nacidos durante una semana, o durante un mes, etc., por ejemplo.

Los mismos casos o análogos los tenemos si entramos en el colectivo más general de los animales, plantas, ...

1.2.2.- Series de Datos y conceptos básicos

En este caso de variable continua la serie de datos deben ser considerados como valores reales con expresión decimal finita (podrán ser valores enteros o fraccionarios, pero expresados como decimales).

NOTA:
Con el fin de evitar ambigüedades en el texto los expresaremos con la 'coma de decimal' arriba.

Ejemplo:

x1 , x2 , x3, x4, x5, …..
2'32, 3'42, 3'52, 4, 4'30, ...,

(lista finita)

En el caso de variable continua necesariamente hemos de agruparlos en intervalos:

(e0,e1], (e1,e2], (e2,e3], ..., (e(k-1),ek]

donde ei son los extremos de los intervalos. En la práctica, su amplitud di= e(i+1)-ei suele ser la misma para todos los intervalos, salvo en casos excepcionales que requieran un estudio 'por niveles'. Representaremos por L dicha amplitud:

$$L = e_k - e_{k-1}$$

MARCA de clase:

El valor central del intervalo: $c_i = e_{i-1} + \frac{L}{2}$

lo llamaremos 'marca de clase', y es el valor que 'representa' a los valores xi que quedan dentro de ese intervalo. Es el valor que tomaremos para el cálculo de los estadísticos.

Frecuencias:

La frecuencia absoluta fi, asociada al intervalo (e(i-1),ei], es el número de valores xi de X, que están dentro de este intervalo, esto es, que cumplen:

$$e_{i-1} < x_i \leq e_i$$

Fase previa al cálculo de los estadísticos:

En variable continua hemos de proceder como sigue.

-Observando la serie de valores xi fijamos la amplitud L que resulte más adecuada para los intervalos, y seguido obtenemos los extremos ei de estos intervalos.

Es muy conveniente comenzar con el valor

$$e_0 = x_0 – 0’05, \text{ ó } e_0 = x_0 – 0’005,$$

donde x_0 es el mínimo de los valores xi de la lista. Pero, claro está que cada caso real puede exigir modificaciones.

-Calculamos las marcas de clase ci.

-Calculamos las frecuencias absolutas fi y las frecuencias acumuladas Fi.

Hecho esto, el cálculo de los estadísticos se realiza del mismo modo que en el caso de variable discreta, salvo el cálculo de la Mediana, Cuartiles, Deciles y Percentiles, como veremos.

1.2.3.- Estadísticos

MEDIA aritmética:

Su definición y cálculo se realiza como en variable discreta sustituyendo yi por la marca de clase ci, así:

$$Mx = c1.f1 + c2.f2 + ... + ck.fk \qquad (10)$$

MEDIANA, Cuatiles, Deciles y Percentiles:

Su definición es la misma que la dada para variable discreta, pero al estar los datos agrupados en intervalos la técnica para su cálculo es algo diferente.

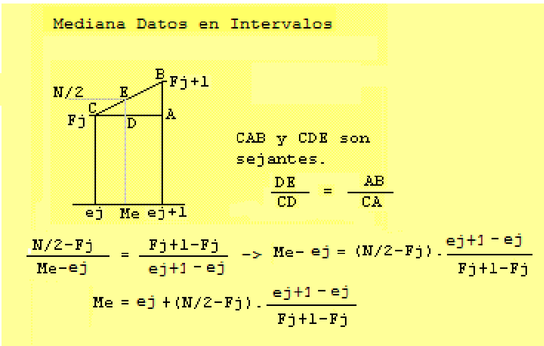
Mediana Me:

Su cálculo se realiza como sigue:

En primer lugar detectamos el ‘intervalo mediano’, que es el intervalo donde se encontrará el valor de la mediana. Para ello detectamos la frecuencia Fj tal que

$$\text{Fj} \leq \frac{N}{2} \quad \text{y} \quad \text{F(j+1)} > \frac{N}{2}$$

El intervalo mediano es (ej, e(j+1)], al cual corresponde la frecuencia Fj.



Tenemos

$$\text{Me} = e_j + \left(\frac{N}{2} - F_j\right) . \frac{(e_{j+1} - e_j)}{F_{j+1} - F_j} \qquad (11)$$

Observa que el segundo sumando es el resultado del prorrateo mediante distribución lineal.

Cuartiles:
Para su cálculo se aplica la fórmula (11) donde Fj se ha determinado como en el caso de la mediana pero tomando N/4 en lugar de N/2:

$$Q_1 = e_j + \left(\frac{N}{4} - F_j\right).\ \frac{(e_{j+1} - e_j)}{F_{j+1} - F_j} \qquad (12)$$

Q2 coincide con la mediana, ya que se tiene en cuenta el 50 % como en ésta.

‘Q3 es el valor de X tal que 3/4 (75 %) de los valores son menores o iguales que Q3’.

Para su cálculo se aplica la fórmula (11) donde Fj se ha determinado como en el caso de la mediana tomando $\frac{3.N}{4}$ en lugar de $\frac{N}{2}$:

$$Q_3 = e_j + \left(\frac{3.N}{4} - F_j\right).\ \frac{(e_{j+1} - e_j)}{F_{j+1} - F_j} \qquad (13)$$

Deciles:

‘Decil d es el valor xd de X tal que el (d/10).100 por ciento de los valores son menores o iguales que xd’

Su valor se obtiene análogamente a la mediana y los cuartiles. Detectamos F_j del mismo modo que se hace en el caso de la mediana, pero tomando $\frac{d.N}{10}$ en lugar de N/2, así:

$$xd = e_j + \left(\frac{d.N}{10} - F_j\right).\ \frac{(e_{j+1} - e_j)}{F_{j+1} - F_j} \qquad (14)$$

donde ahora tomamos $\frac{d.N}{10}$ en lugar de $\frac{N}{2}$

Percentil:

‘Percentil p es el valor xp de X tal que el $\frac{p.N}{100}$ por ciento de los valores son menores o iguales que xp’.

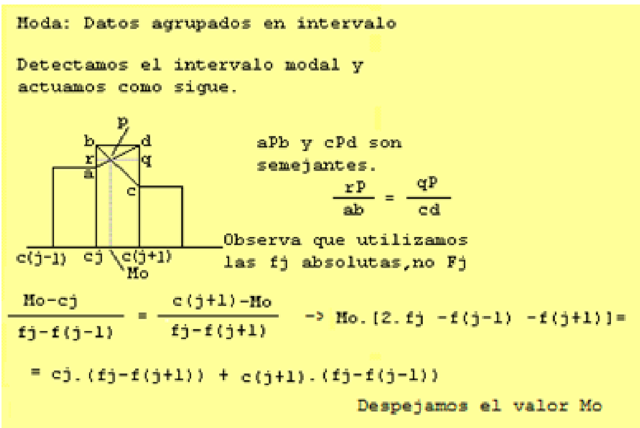
Su valor se obtiene análogamente a la mediana y los cuartiles, localizando el valor F_j de la misma forma pero tomando $\frac{p.N}{100}$ en lugar de $\frac{N}{2}$, así:

$$xp = e_j + \left(\frac{p.N}{100} - F_j\right). \frac{(e_{j+1} - e_j)}{F_{j+1} - F_j} \qquad (15)$$

donde ahora el valor $\frac{p.N}{100}$ sustituye a $\frac{N}{2}$

MODA:

Teniendo en cuenta que los datos están agrupados en intervalos, su cálculo se realiza mediante la siguiente fórmula, deducida (y aceptada) por las condiciones que se muestran en la figura.



En primer lugar detectamos el intervalo modal, que es aquel cuya frecuencia absoluta fj es la más alta. En la fórmula observa que se utilizan las marcas de clase y las frecuencias absolutas fj, y no las acumuladas como hacemos en el caso de la mediana y sus 'derivados'.

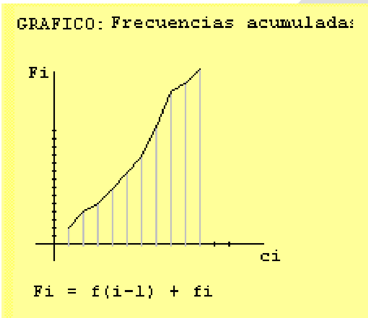
Su valor

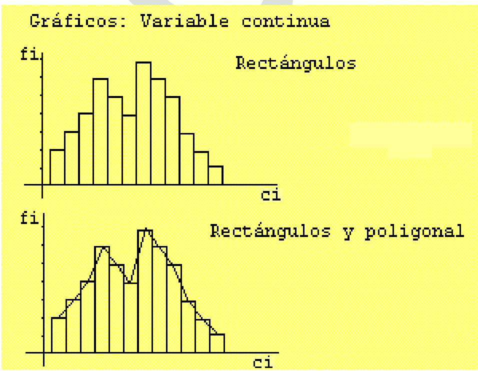
$$M_0 = \frac{c_j \,.\, (fj - f_{j+1}) \,+\, c_{j+1} \,.\, (fj - f_{j-1})}{2.fj - f_{j-1} - f_{j+1}} \qquad (16)$$

1.2.4.- GRAFICAS en variable continua

DATOS AGRUPADOS en intervalos

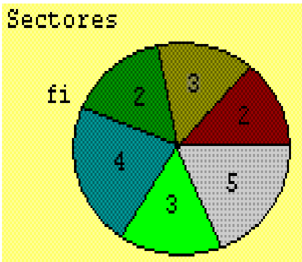
En el caso de variable continua los más significativos son el Diagrama poligonal, y en Diagrama de rectángulos.





En el diagrama de rectángulos, la base de cada rectángulo es el intervalo correspondiente: (e(i-1),ei], y su altura es fi.

En el de Sectores, lo más habitual es que el área del sector es proporcional a la frecuencia absoluta del correspondiente valor de la variable.



Ejemplos/Ejercicios

7.- Hemos registrado el peso de 50 recién nacidos, y agrupándolos en intervalos tenemos:

Peso kgs	núm. niños
[2’5-3’0)	6
[3’0-3’5)	23
[3’5-4’0)	12
[4’0-4’5)	9

Obtener: ma, los cuartiles, la desv. típica, el recorrido intercuartílico.

Sol.:

Clases	Marca xi	fi	Fi	xi.fi	xi^2.fi
[2’5-3’0)	2’75	6	6	16’50	45’38
[3’0-3’5)	3’25	23	29	74’75	242’94
[3’5-4’0)	3’75	12	41	45’00	168’75
[4’0-4’5)	4’25	9	50	38’25	162’56
		N=50		Sum=174’50	Sum=619’63

$$ma = \frac{174'50}{50} = 3'4$$

Q1: N/4 = 12’50 , F2 > 12’50 y F1 < 12’50

--> Intervalo: [3’0-3’5),

$$Q1 = 3+0'5.\frac{12'50-6}{29-6} = 3+0'5.0'28 = 3'14$$

NOTA: Fórmula para obtener Qi en el caso de variable continua

$$Qi = E_{i+1} + L.\frac{i.N/4-F_i}{F_{i+1}-F_i}$$

donde E_{i+1} es el extremo inferior del intervalo cuya frecuencia acumulada es la primera que supera (o iguala) el valor i.N/4, F_{i+1} es la referida frecuencia. L es la amplitud del intervalo.

Q3: 3.N/4 = 37’50, F3 > 37’50 y F2 < 37’50

--> Intervalo: [3’5-4’0)

$$Q3 = 3'5 +0'5.\frac{37'50-29}{41-29} = 3'5 +0'5.\frac{8'5}{12} = 3'85$$

Varianza: $S^2 = \dfrac{619'63}{50} - (3'49)^2 = 0'2125$

Desviación típica: $S = \sqrt{0'2125} = 0'461$

Recorrido intercuartílico:

Q3–Q1 = 3'85 – 3'14 = 0'71

8.- Damos el cuadro de datos y el alumno debe comprobar (haciendo los cálculos correspondientes) los resultados que damos:

Clases	fi
[0 -15)	10
[15-30)	15
[30-45)	25
[45-60)	20
[60-75)	20
[75-90)	10

Resultados:

N = 100, ma = 45'75,

$$mo = 30 + 15.\frac{25-15}{(25-15)+(25-20)} = 30 + 15.\frac{10}{15} = 40$$

NOTA:

La fórmula para la moda en el caso de variable continua es

$$mo = E_{i+1} + L.\frac{f_{i+1} - f_i}{(f_{i+1} - f_i) + (f_{i+1} - f_{i+2})}$$

donde E_{i+1} es el extremo izquierdo del intervalo modal (aquel con mayor frecuencia), f_{i+1} es la frecuencia más alta. L es la amplitud del intervalo.

Valor de me: (Mediana)

Fórmula: $me = E_{i+1} + L.\dfrac{N/2 - F_i}{F_{i+1} - F_i}$

Si N/2 coincide con una Fi, entonces

me = extremo superior del intervalo correspondiente.

En este caso: N/2 = 50 = F3 , por tanto me = 45

Rango: R = 90 – 0 = 90

Varianza: $S^2 = \dfrac{257625}{100} - (45'75)^2 = 483'19$

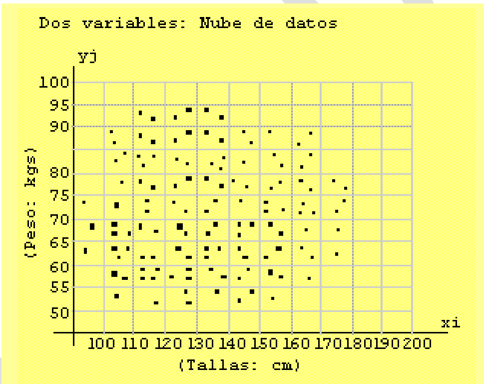
$S = 21'98$

$$$oOo$$$

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

Tema 2

Estadística Descriptiva en dos variables



Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

2.- ESTADÍSTICA Descriptiva, dos variables X,Y
(Variable bidimensional (X,Y))

2.0.- Introducción: Conceptos previos

Estamos interesados en estudiar, y obtener alguna consecuencia, la relación entre dos características de un colectivo de individuos (personas, animales, vegetales, etc. ...)

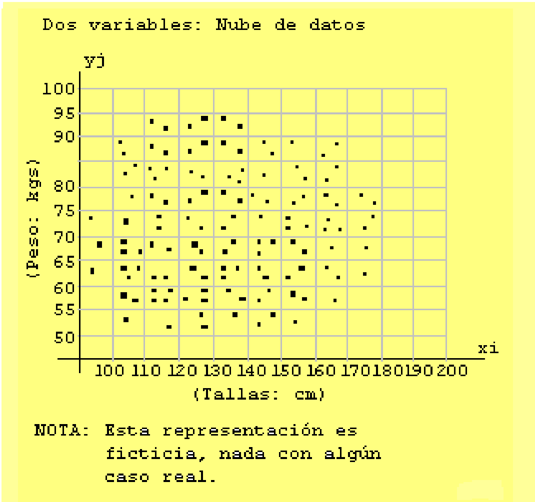
Fijadas las variables X, Y, correspondientes a cada una de las dos características objeto del estudio, tomaremos el valor de cada una de las variables X, Y, de modo que obtengo una lista de 'pares de valores' (xi,yi), correspondiendo cada par a un individuo.

Otra forma para la toma de datos es la tabla de doble entrada. Colocamos los valores de X en línea horizontal formando cabecera, y los valores de Y en línea-columna formando la izquierda de la tabla. Observa la fig. 10

Cada una de las variables podrá ser 'Discreta' o 'Continua', y según el caso procederemos como en el caso de una variable.

En la figura mostramos el caso de 'Tabla de doble entrada' para X e Y continuas y datos agrupados en intervalos.

La longitud del intervalo se fijará independientemente para cada variable, de modo que obtenemos así la 'malla' de casilleros en los que 'virtualmente' instalamos los pares de valores.



Lo que marcamos en cada casillero es la frecuencia absoluta fij del par (xi,yj).

2.1.- SERIES de Datos, sus formas, y Conceptos básicos

En el caso de dos variables la forma habitual de organizar los datos es una de las siguientes.

A) La forma directa, la más frecuente:

En la práctica sus valores suelen ser tomados de forma que resultan 'pareados' (xi,yi) por el hecho de que se toman consultando un mismo individuo y que además se presume que, de alguna manera, estos valores están 'ligados' entre sí.

Cada par de valores (xi,yi) puede presentarse fi veces, teniendo así su frecuencia absoluta.

En lo que sigue: N = f1 + f2 + ... + fk es el número total de pares de valores; k es el número de pares distintos (son diferentes desde el momento que uno de los componentes lo es).

Los pares de valores estarán ordenados en primer lugar por el valor de X.

Estadística Descrip. en dos variable.
Series de Datos

fi	xi	yi
f1	x1	y1
f2	x2	y2
f3	x3	y3
.	.	.

N = Sum(fi)

Datos en pareado
(o emparejados)

Eje.:

fi	xi	yi
4	0	1
2	1	2
3	2	3
1	2	1
3	3	0
4	3	1
2	3	2

N= 19

B) La forma Operativa, la más práctica:

Es la Tabla de doble entrada como mostramos en la figura.

De alguna forma podría resultar de la toma directa de datos, pero si no es así continuar como se indica a continuación.

Si tomamos la lista de pares mostrada antes y colocamos en cabecera los valores distintos de X, y en lateral izquierdo los valores distintos de Y, tenemos la tabla de doble entrada. En el interior, en cada casillero colocamos la frecuencia fij de cada par (xi,yj).

Hecho el 'conteo' o tabulación de las frecuencias absolutas podemos obtener con cierta facilidad lo siguiente:

-Frecuencias marginales:

$f_{xi.}$, frecuencia absoluta del valor x_i

$f_{.yj}$, frecuencia absoluta del valor y_j

-Media aritmética m_{xi} asociada al valor x_i

-Media aritmética m_{yj} asociada al valor y_j

Estadística Discreta con dos variables
Tabla de doble entrada

yj \ xi	x1 x2 x3 ...	f.j
y1		f.1
y2		f.2
y3	fij	f.3
.		.
.		.
fi.	f1. f2. f3. ..	N=Sum(fi.)= =Sum(f.j)

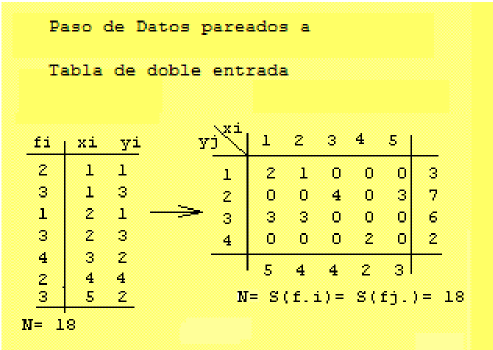
yj \ xi	0	1	2	3	4	5	6	7	f.j
0	2	3	1	0	2	1	2	0	11
1	0	1	2	3	0	2	1	1	10
2	3	2	0	1	3	0	2	3	14
3	1	0	3	0	1	3	0	2	10
4	2	1	2	3	2	1	2	1	14
5	0	3	1	0	0	2	1	0	7
6	3	0	2	1	2	0	3	2	13
fi.	11	10	11	8	10	9	11	9	79

N = 79

En todo caso N (o Nxy) cumple

$$N = \text{Sum}(fxi.) = \text{Suma}(f.yj)$$

Es interesante comprobar (siguiente figura) cómo pasar de 'datos pareados' a tabla de doble entrada.



Paso de Datos pareados a

Tabla de doble entrada

fi	xi	yi
2	1	1
3	1	3
1	2	1
3	2	3
4	3	2
2	4	4
3	5	2

N= 18

yj \ xi	1	2	3	4	5	
1	2	1	0	0	0	3
2	0	0	4	0	3	7
3	3	3	0	0	0	6
4	0	0	0	2	0	2
	5	4	4	2	3	

N= S(f.i)= S(fj.)= 18

2.2.- ESTADÍSTICOS en el caso de variable bidimensional (X,Y)

Para el análisis de estos datos, por ejemplo el estudio del grado de afinidad entre los valores del par (xi,yi), se utilizan los siguientes estadísticos.

Observación:

Por razones obvias, en lo que sigue escribiremos fxi en lugar de fxi. y fyj en lugar de f.yj

Si la tabla de doble entrada procede de una Serie de Datos pareados, entonces se cumple:

Nx = Sum(fxi) = Sum(fyj) = Ny, y coinciden con el valor Nxy (o N simplemente).

Pero en general, al tomar los datos mediante tabla de doble entrada puede resultar que

Nx $\neq$ Ny, y por consiguiente distintos de N_{xy}.

También, en general el número k de valores distintos de X será distinto que el número h de valores distintos de Y.

MEDIAS aritméticas:

Tomando las frecuencias marginales fxi , fyj , calculamos las llamadas 'Medias marginales' Mx, My, como veremos.

$$Mx = \frac{fx1.x1 + fx2.x2 + fx3.x3 + \ldots + fxk.xk}{Nx}$$

$$My = \frac{fy1.y1 + fy2.y2 + fy3.y3 + \ldots + fyh.yh}{Ny}$$

DESVIACIONES Típicas marginales:

Son las desviaciones respecto de las medias marginales:

$$S_x = \sqrt{\frac{(x1-Mx)^2.fx1 + (x2-Mx)^2\ .fx2 + \ldots + (xk-Mx)^2\ .fxk}{Nx}}$$

$$S_y = \sqrt{\frac{(y1-My)^2.fy1 + (y2-My)^2\ .fy2 + \ldots + (yk-My)^2\ .fyk}{Ny}} \qquad (22)$$

Observa que el radicando es siempre mayor que cero.
Por definición tomaremos el varo positivo de la raíz cuadrada.

COVARIANZA Sxy:
Diríamos también 'Covarianza típica' por estar referida y valorada respecto de las medias aritméticas.

Su valor es

$$S_{xy} = \frac{\sum_{i=1,2,..,k;j=1,2,\ldots,h} (xi-Mx).(yj-My).\ f_{ij}}{N_{xy}}$$

Este valor puede resultar negativo

COEFICIENTE de correlación lineal: $r = \dfrac{S_{xy}}{S_x \, . \, S_y}$

Este valor da una medida de la 'dependencia' o 'correlación' entre las variables X e Y.

Puede resultar negativo estando en el intervalo [-1;1] (Sin demostración).

REGRESIÓN: Rectas de regresión

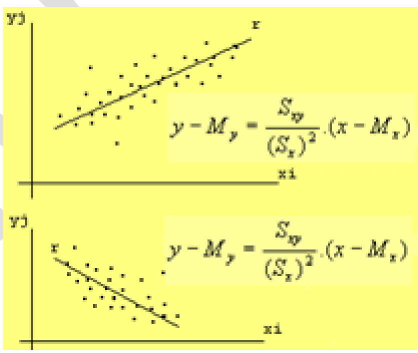
Son las rectas que mejor se aproximan a la nube de puntos que resulta sobre el plano al representar los valores (xi,yj) de la tabla de doble entrada.
Son dos, una expresando y = f(x), otra expresando x = g(y).

De y sobre x (y dependiendo de x):

$$y - M_y = \frac{S_{xy}}{(S_x)^2} . (x - M_x) \qquad (25)$$

De x sobre y (x dependiendo de y):

$$x - M_x = \frac{S_{xy}}{(S_y)^2} . (y - M_y)$$

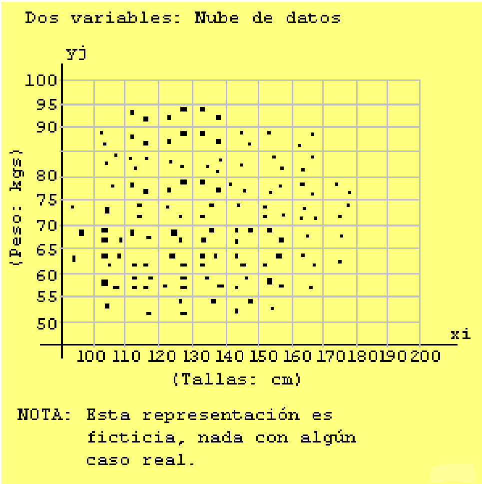


Observa su representación en la figura (nube de puntos, datos)

2.3.- GRÁFICAS en dos variables

En este caso de variable bidimensional tiene interés la 'nube de puntos' y la representación de las rectas de regresión.

Partiendo de la tabla de valores obtenemos en primer lugar la llamada 'nube de valores'. Estos son los valores de las frecuencias absolutas del par (xi, yj).



Ejemplos/Ejercicios:

1.- En una prueba sobre el número de bacterias en un cultivo, a medida que pasan las horas, obtenemos los siguientes datos:

Núm. de horas:	0	1	2	3	4	5
Núm. bacterias:	12	19	23	34	56	62

Se pide:

a) La media y desv. típica de las variables X = núm. de horas, Y = núm. de bacterias.

b) La covarianza de la variable bidimensional (X,Y)

Sol.: a)

xi	yi	fi	xi.fi	xi^2.fi	yi.fi	yi^2.fi	xi.yi.fi
0	12	1	0	0	12	144	0
1	19	1	1	1	19	361	19
2	23	1	2	4	23	529	46
3	34	1	3	9	34	1156	102
4	56	1	4	16	56	3136	224
5	62	1	5	25	62	3844	310
	N= 6		15	55	206	9170	701

$mx = \frac{15}{6} = 2'50$ $\qquad$ $my = \frac{206}{6} = 34'33$

$Vx = \frac{55}{6} - (2'50)^2 = 2'92,$ $\qquad$ $Sx = 1'71$

$Vy = \frac{9170}{6} - (34'33)^2 = 349'78,$ $\quad$ $Sy = 18'70$

b) S_{xy} = Covar(X,Y) = $\dfrac{\sum xi.yi.fi}{N} - mx.my$ =

= $\dfrac{701}{6} - 2'50.34'33$ = 31’01

2.- Al tomar datos de un estudio sobre la relación entre el número de horas semanales que ven televisión y los ingresos mensuales (en centenares de euros) de sus padres, en grupo de alumnos de la ESO, obtuvimos el siguiente cuadro de datos:

Horas sema.	Ingresos Mens. [4-10)	[10-16)	[16-22)	[22-28)
[10-20)	3	2	--	--
[20-30)	4	6	6	3
[30-40)	11	4	3	2
[40-50)	1	--	--	4

a) Calcula las medias y las desv. típicas

b) Calcula la covarianza

Sol.:

a) X = Ingresos, Y = Núm. de horas

x_i	y_i	$f_i=f_{x_iy_i}$	$x_i.f_i$	$x_i^2.f_i$	$y_i.f_i$	$y_i^2.f_i$	$x_i.y_i.f_i$
7	15	3	21	147	45	675	315
7	25	4	28	196	100	2500	700
7	35	11	77	539	385	13475	2695
7	45	1	7	49	45	2025	315
13	15	2	26	338	30	450	390
13	25	6	78	1014	150	3750	1950
13	35	4	52	676	140	4900	1820
13	45	0	0	0	0	0	0
19	15	0	0	0	0	0	0
19	25	6	114	2166	150	3750	2850
19	35	3	57	1083	105	3675	1995
19	45	0	0	0	0	0	0
25	15	0	0	0	0	0	0
25	25	3	75	1875	75	1875	1875
25	35	2	50	1250	70	2450	1750
25	45	4	100	2500	180	8100	4500
		N = 49	685	11833	1475	47625	21155

$mx = \frac{685}{49} = 13'98$, $\quad my = \frac{1475}{49} = 30'10$

$Sx^2 = \frac{11833}{49} - (13'98)^2 = 46'05$, $\quad Sx = 6'79$

$Sy^2 = \frac{21155}{49} - (30'10)^2 = 65'93$, $\quad Sy = 8'12$

b) $Sxy = \frac{21155}{49} - 13'98.30'10 = 10'94$

3.- El siguiente cuadro de datos registra los gastos (Cent. de e.) en publicidad de una empresa y las correspondientes ventas (cent. e.)

Publicidad | 1 2 3 4 5 6 7 8

Ventas	15 16 14 17 20 18 18 19

a) El alumno debe comprobar los siguientes resultados.

b) Obtener el coeficiente de correlación.

c) Expresa las rectas de regresión:

y = f(x), x = g(y)

Sol.: a) X = Gasto en publicidad, Y = Producto ventas

mx = 4’50, my = 17’13

Sx^2 = 5’25, Sy^2 = 3’61

Sx = 2’29, Sy = 1’90

Covarianza de (X,Y): Sxy = 3’31

b) Coef. de correlación: $r = \dfrac{Sxy}{Sx.Sy}$

$$r = \frac{3'31}{2'29.1'90} = 0'76$$

Correlación débil positiva

c) Rectas de regresión:

$$y\text{-}my = \frac{Sxy}{Sx^2}.(x - mx),\quad x\text{-}mx = \frac{Sxy}{Sy^2}.(y - my)$$

y = 17’13 + 0’63.(x-4’50)

y = 0’63.x + 14’30

x = 4’50 + 0’92.(y-17’13)

x = 0’92.y – 11’26

4.- El siguiente cuadro de datos se refiere al conteo (tabulación) del número de bacterias dependiendo del tiempo transcurrido:

Núm. horas	0	1	2	3	4	5
Núm. bacterias	12	19	23	34	56	62

a) Comprueba los siguientes resultados
b) Obtener el coef. de correlación
c) Expresa las rectas de regresión

Sol.: a) X = número de horas, Y = núm. de bacterias

mx = 2’50, my = 34’33

Sx^2 = 2’92, Sx = 1’71

Sy^2 = 349’69, Sy = 18’70

Sxy = 31

b) $r = \dfrac{Sxy}{Sx.Sy} = \dfrac{31}{1'71.18'70}$ = 0’97

c) $y\text{-}my = \dfrac{Sxy}{Sx^2}.(x-mx)$, $x\text{-}mx = \dfrac{Sxy}{Sy^2}.(y-my)$

y = 10’62.x + 7’78, x = 0’09.y – 0’59

5.- Damos el siguiente cuadro de datos (doble entrada), y el alumno debe comprobar los resultados que adjuntamos:

Habilidad verbal

[10-20) [20-30) [30-40) [40-50)

Razonamiento abstracto

[15-25)	5	3	0	0
[25-35)	2	6	1	0
[35-45)	0	1	4	2
[45-55)	0	0	3	3
[55-65)	0	0	1	2

Sol.: Resultados que el alumno debe comprobar

X = Habilidad verbal, Y = Razonamiento

mx = 29’85, my = 36’06
Sx^2 = 109’83, Sx = 10’48
Sy^2 = 163’27, Sy = 12’78
Sxy = 110’00

r = 0’82

y = x + 6’21
x = 0’67.y + 5’69

¿Qué puntuación obtiene en razonamiento abstracto una persona que obtuvo 45 en habilidad verbal?

x = 45 -- > y = 45 + 6’21 = 51’21

Observa: La recta y = x + 6’21 tiene pendiente 1, lo que significa que y crece igual que x.

El valor r = 0’82 desdice apreciablemente el hecho anterior. Este valor de r debía estar más próximo al valor 1.

$$$oOo$$$

Tema 3

Teoría y Cálculo de Probabilidades

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

3.- TEORÍA de la Probabilidad

3.1.- Introducción: Conceptos básicos

EXPERIMENTO o fenómeno aleatorio:

Es aquel experimento que puede producir varios resultados, ninguno de ellos predecible.

Ejemplo:

Si lanzamos una moneda u otro objeto pesado al 'aire', podemos estar seguros que caerá al suelo.

Si lanzamos la moneda, y nos interesamos por el resultado, presentará hacia arriba la 'cara' o la 'cruz'. (En adelante diremos 'resultar cara' o 'resultar cruz').

En el primer caso es un 'experimento determinista', puesto que con seguridad caerá al suelo. En el segundo es un 'experimento aleatorio', puede resultar cara ó cruz.

Otro caso muy frecuente consiste en 'lanzar un dado' (de seis caras, por ejemplo). No podemos asegurar cuál será el resultado: Es un experimento aleatorio.

Si consideramos el suceso A= 'Obtener número par', los casos favorables a A los tenemos cuando resulte alguna de las caras 2, 4, 6. Expresaremos:

$$A = \{2, 4, 6\}$$

Del estudio de estos experimentos aleatorios se ocupa la 'Teoría de Probabilidades'.

Si nos preguntamos por 'múltiplos de 3', el suceso es $A = \{3, 6\}$

Ejemplo:

-En el caso del dado, de entre 6 posibles valores, el valor 3 tiene una 'probabilidad' de 1/6, ya que tenemos 1 de entre 6.

-La probabilidad del suceso A = 'obtener valor par', es: P(A) = 3/6, como parece dictar la lógica común, ya que de 6 posibilidades tiene 3 a su favor.

-La probabilidad de A = 'Múltiplo de 3' es P(A) = 2/6, ya que A = {3, 6}

NOTA: En ocasiones simbolizaremos por P la palabra 'probabilidad' (por simplificar la expresión)

ESPACIO Muestral:

Todo experimento aleatorio lleva asociado el 'conjunto M de todos los casos posibles'.

Lo llamaremos 'Espacio muestral'. Sus elementos serán calificados como 'Sucesos elementales'.

Ejemplos:

-En el caso de la moneda es M = {cara, cruz}, o {C, X}. Sucesos elementales: cara, cruz

-En el caso del dado es M= {1, 2, 3, 4, 5, 6}. Sucesos elementales: 1, 2, 3, 4, 5, 6

SUCESO:
Cuando lanzamos un dado ya sabemos cuál es su Espacio muestral. Ahora podemos hacernos varias preguntas:

-¿Obtendremos valor par?
-¿Obtendremos valor impar?
-¿Obtendremos múltiplo de 3?
....

A Cada uno de estos 'hechos' lo llamamos 'suceso'.

Los anteriores sucesos podemos expresarlos como subconjuntos de M, así:

$$A = \{2,4,6\},\ B = \{1,3,5\},\ C = \{3,6\}$$

Este razonamiento nos lleva al conjunto P(M) cuyos elementos son los subconjuntos de M. Decimos que es la 'familia' de subconjuntos de M.

En P(M) incluimos el vacío $\emptyset$ (no contiene elementos)

3.2.- OPERACIONES con sucesos

Son las mismas operaciones que se realizan con los conjuntos, en este caso considerados como subconjuntos de M (este es el Conjunto universo).

Son las siguientes operaciones:

a) $C = A \cup B$, unión de A y B
Un elemento está en C si está en A ó en B, ó en los dos

b) $C = A \cap B$, intersección de A y B
Un elemento está en C si está en A y en B simultáneamente

c) $C = A - B$, suceso diferencia A menos B
Un elemento está en C si está en A y no está en B

d) A’ = M – A, complementario de A
Un elemento de M está en A’ siempre que no esté en A. Además M = A∪A’, A∩A’ = ∅

Observa que el conjunto ‘vacío’, interpretado como suceso, es aquel que ‘no ocurre nunca’. Lo llamaremos ‘suceso imposible’ (representado por ∅).

Observa que el conjunto total (universo M), interpretado como suceso, ocurre siempre. Lo llamaremos ‘suceso seguro’.

Diremos que A y B son sucesos ‘incompatibles’ si A∩B = ∅.

SUCESO Elemental:

Definición:
Llamamos ‘suceso elemental’ a cada ‘subconjunto unitario’ {ei} del espacio muestral M.

Este subconjunto unitario no permite expresión como unión de otros subconjuntos.

3.3.- CONCEPTO clásico de PROBABILIDAD

Sea el espacio muestral M, finito:

$$M = \{e1,e2,e3,...,en\},$$

y sea A un ‘suceso’= ‘subconjunto de M’.

CASOS posibles del experimento:

En número de ‘casos posibles del experimento’ es el número de elementos ei que constituyen el espacio muestral M. Coincide con el número de sucesos elementales.

CASOS favorables del suceso A:

Llamamos número de 'casos favorables del suceso A' al número de sucesos elementos ei que están dentro del subconjunto A.

PROBABILIDAD de que ocurra el suceso A:

Si los sucesos elementales ei son equiprobables (que son equiposibles), entonces la lógica común nos lleva a tomar como válida la siguiente definición para la 'Probabilidad del suceso A':

$$p(A) = \frac{Casos\ favorables\ de\ A}{Casos\ posibles\ del\ experiemnto}$$

Los Clásicos tomaron esta igualdad como definición de 'probabilidad del suceso A', y ha sido bautizada como 'Regla de Laplace' (en honor del que fue uno de los primeros estudiosos del tema).

REGLA de Laplace:

Definición:

Si los sucesos {ei} son equiprobables, definimos P(A) mediante la siguiente igualdad:

$$P(A) = \frac{Casos\ favorables\ de\ A}{Casos\ posibles\ del\ experiemnto}$$

Propiedades:

a) $0 <= p(A) <= 1$
b) p(vacío) = 0, p(M) = 1
c) $p(A \cup B) = p(A) + p(B) - p(A \cap B)$, para todo par de sucesos A y B.

Ejemplo: Caso de lanzamiento de un dado.

Espacio muestral M = {1,2,3,4,5,6}. Casos posibles = 6

Considero el suceso 'obtener múltiplo de 3':
A = {3,6}. Casos favorables = 2.
Entonces: p(A) = 2/6 = 1/3

A = 'obtener número impar', P(A) = 3/6 = 1/2

NOTA:
Teniendo en cuenta los destinatarios del presente trabajo, renuncio a exponer el concepto de probabilidad desde la teoría de la frecuencia relativa f(A) y su tendencia ... .

3.4.- Probabilidad Condicionada

Motivación que justifica la introducción de este concepto:

Observemos la siguiente situación (que puede ser real)

Un colectivo social compuesto por 200 personas se distribuye por su afinidad política como muestra la siguiente Tabla:

	Del PP	Del PSE	Totales
Hombres:	80	60	140
Mujeres	25	35	60
Totales	105	95	200

H = Hombre M = Mujer = H'
A = 'Ser del PP', A' = 'Ser del PSE

P(H) = $\frac{140}{200}$,   P(M) = $\frac{60}{200}$

P(PP)= $\frac{105}{200}$,   P(PSE)= $\frac{95}{200}$

P(H ∩PP) = $\frac{80}{200}$,   P(H ∩PSE) = $\frac{60}{200}$

P(M ∩PP) = $\frac{25}{200}$,   P(M ∩PSE) = $\frac{35}{200}$

Hago notar su significado:

H∩PP = 'Probabilidad de se hombre y del PP', y de forma análoga en los otros casos.

Pero ahora deseo obtener:
'Probabilidad de ser del PP de entre los hombres', es decir:

'De los hombres, probabilidad de ser del PP'

NOTACIÓN:
Este suceso lo representamos por PP/H , que significa

'ser del PP condicionado a que es hombre'

Cuidado!: Nunca representará una fracción salvo que el contexto lo requiera.

Aplicando la fórmula de Laplace obtengo

P(PP/H) = $\frac{80}{140}$

Por otro lado tenemos:

$$\frac{\mathrm{P(H \cap PP)}}{P(H)} = \frac{\frac{80}{200}}{\frac{140}{200}} = \frac{80}{140}$$

es decir $P(PP/H) = \frac{P(H \cap PP)}{P(H)}$

Llamamos ‘probabilidad condicionada de PP condicionada a H’ al valor

$$P(PP/H) = \frac{P(H \cap PP)}{P(H)}$$

Es muy importante la relación

$$P(H \cap PP) = P(H).P(PP/H)$$

Generalización:

Def.:
Llamamos ‘Probabilidad del suceso B condicionada al suceso A’ al valor P(B/A), que resulta de la igualdad

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

3.5.- INDEPENDENCIA de Sucesos

De la definición de probabilidad condicionada

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

obtenemos

$$p(A \cap B) = p(A).p(B/A)$$

y también

$$p(A \cap B) = p(B).p(A/B)$$

Def.:
Diremos que A y B son ‘sucesos independientes’ si se cumple

$$p(A \cap B) = p(A).p(B)$$

de donde

$$p(B/A) = p(B),$$

lo cual indica que:

‘No existe dependencia de B respecto de A’ (que ocurra o no el suceso A No influye en la ocurrencia de B).

Consecuencia:

Si los sucesos A y B son independientes se cumple

$$p(A\cap B) = p(A).p(B)$$

Observa que:
En el caso de que fuesen ‘físicamente independientes’ gozarían de esta propiedad-condición de sucesos independientes.

Podemos extenderla al caso de k sucesos:

Si los sucesos A1, A2, A3, ..., Ak, son independientes entre sí, se cumple

$$p(A1\cap A2\cap ...\cap Ak) = p(A1).p(A2)....p(Ak)$$

(Para lo cual es suficiente que sean ‘físicamente independientes’)

3.6.- PROBABILIDAD Compuesta

Consideremos el caso de tres sucesos A, B, C.

Supongamos que existe condicionamiento entre ellos, es decir, No podemos afirmar que sean independientes.

Tenemos

$$p(A\cap B\cap C) = p((A\cap B)\cap C) = p(A\cap B).p(C/(A\cap B)) =$$

$$= p(A).p(B/A).p(C/(A\cap B))$$

En el caso de cuatro sucesos tenemos

$p(A \cap B \cap C \cap D) = p((A \cap B \cap C) \cap D) = p(A \cap B \cap C).p(D/(A \cap B \cap C))=$

$= p(A).p(B/A).p(C/(A \cap B)).p(D/(A \cap B \cap C))$

En caso general, si tenemos k sucesos

$A1, A2, ..., Ak,$

Fórmula de la Probabilidad compuesta:

$p(A1 \cap A2 \cap A3 \cap ... \cap A_{k-1} \cap Ak) =$

$= p(A1).p(A2/A1).p(A3/(A1 \cap A2)) ... p(Ak/(A1 \cap A2 \cap \ldots \cap A_{k-1}))$

A esta última igualdad la llamamos:

‘Fórmula de la Probabilidad compuesta’

para otros: **‘Teorema de la p. compuesta’**

3.7.- Probabilidad Total

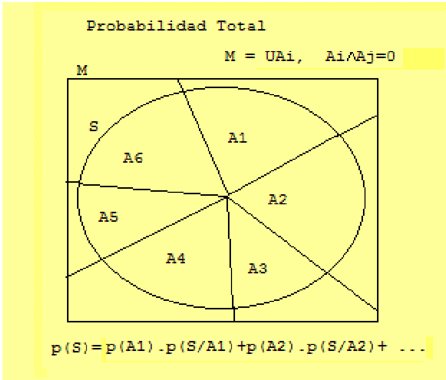
Sea M el espacio muestral de un experimento que admite una familia de n subconjuntos Ai que se comporta como una partición de M:

a) Cumplen que $Ai \cap Aj = \varnothing$, siempre que $i \neq j$

(Son independientes dos a dos)

b) $\cup Ai = M$, (la unión, ó reunión, de todos los Ai nos da M)

Observa las figuras siguientes.



Bajo estas condiciones, para cualquier suceso S del experimento se cumple:

$$p(S)= \sum_{i=1,\ldots,n} p(S \cap Ai) = \sum_{i=1,\ldots,n} p(Ai).p(\frac{S}{Ai})$$

Esta nos da la probabilidad total de S.

La llamamos:

'Fórmula de la probabilidad total'

para otros: **'Teorema de la p. total'**

Por la Probabilidad condicionada tenemos también, para los sucesos S y Ai

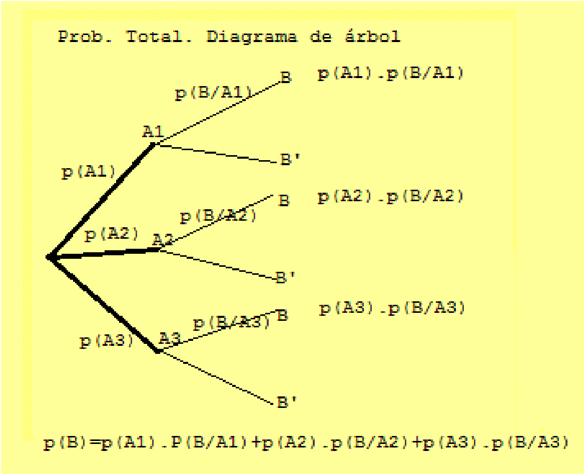
$$p(S\cap Ai)= p(S).p(Ai/S),$$

de donde

$$p(S) = \frac{p(S\cap Ai)}{p(\frac{Ai}{S})}, \quad p(Ai/S)= \frac{p(S\cap Ai)}{p(S)}$$

y teniendo en cuenta la probabilidiad total, tengo, para cada suceso Ak

$$\frac{p(S\cap Ak)}{p(\frac{Ak}{S})} = \sum_{i=1,\ldots,n} p(Ai).p(\frac{S}{Ai})$$



$$p(Ak/S) = \frac{p(S\cap Ak)}{\sum_{i=1,\dots,n} p(Ai).p(\frac{S}{Ai})}$$

la cual justifica el siguiente Teorema de Bayes.

3.8.- TEOREMA de Bayes

Teniendo en cuenta que

$$p(S\cap Ak) = p(S).p(ak/S)$$

y llevándolo a la anterior obtenemos el llamado

Teorema de Bayes:

$$p(Ak/S) = \frac{p(Ak).p(S/Ak)}{\sum_{i=1,\dots,n} p(Ai).p(\frac{S}{Ai})}$$

Decimos que p(Ak/S) es la ‘Probabilidad a posteriori’.
p(Ak) es la llamada Probabilidad a priori.

3.9.- Experimentos Compuestos

Experimento compuesto es el Experimento resultante de la ejecución consecutiva de dos o más experimentos simples.

Ejemplos:
a)Si 'Lanzamos una moneda' y a continuación 'Tomo una bola de la urna U', puedo preguntar por el Suceso 'Obtener cara y bola blanca'. Puesto que no hemos impuesto condición alguna y son ´físicamente' independientes, tengo

$$P = \frac{1}{2} \cdot \frac{4}{10}$$

(Supuesto: La urna contiene 4 blancas de un total de 10)

b)Lanzamos una moneda y, si resulta cara tomo una bola de la urna U1, pero si resulta cruz la tomo de la urna U2, siendo la composición de las urnas

U1 = {4 blancas, 3 rojas, 3 negras}
U2 = {3 blancas, 2 rojas, 3 negras}

Deseo saber la probabilidad de obtener bola roja

Resultado: p(R) = p(C).p(R/C) + p(X).p(R/X) =

$$= \frac{1}{2} \cdot \frac{3}{10} + \frac{1}{2} \cdot \frac{2}{8} = \frac{3}{20} + \frac{2}{16} = \frac{3}{20} + \frac{1}{8} = \frac{44}{160}$$

En este ejemplo evidentemente No son independientes.

En general, y a priori, podemos establecer la igualdad

p(S)= p(C).p(B/U1) + p(C').p(B/U2)

y después, al realizar los cálculos se verá si hay o no dependencia.

3.10.- FUNCIÓN de Probabilidad

Recordamos que P(M) representa la familia de todos los subconjuntos de M, incluido el vacío (suceso imposible) y el suceso seguro M.

La siguiente definición permite introducir (definir) el concepto de 'probabilidad de S' de un modo diferente a la definición de Laplace.

Def.:
Llamamos 'función de probabilidad' a toda función

$$p: P(M) \dashrightarrow [0,1]$$

con Dominio P(M) e Imagen en [0,1] que cumpla las siguientes condiciones:

a) $p(A) \geq 0$, para todo suceso A de P(M)
b) $p(A \cup B) = p(A) + p(B) - p(A \cap B)$
c) $p(M) = 1$

CONSECUENCIA:

Se deducen de forma inmediata los siguientes resultados:

a) Si $A \cap B = \emptyset$, entonces $p(A \cup B) = p(A) + p(B)$

b) $p(A') = 1 - p(A)$, ya que
$A \cup A' = M$, y $A \cap A' = \emptyset$

c) $p(\emptyset) = 0$

d) Si B es subconjunto de A, $p(B) \leq p(A)$

Ejemplos/Ejercicios: De Probabilidades

1.- Expresa el espacio muestral de los siguientes experimentos:

a) Lanzar una moneda

b) Lanzar el dado de quinielas (6 caras: 2 con el 1, dos con el 2, dos con x)

c) Lanzar el dado de 6 caras (del 1 al 6)

d) Tomar una carta de la baraja española, y considerar el número obtenido del 1 al 12

e) Lanzar dos monedas simultáneamente

f) Lanzar dos dados sucesivamente, y anotando en cada lanzamiento

g) Lanzar dos dados normales simultáneamente

h) Lanzar dos dados normales sucesivamente, y anotando en cada lanzamiento

Sol.: a) {C,X}

b) {1,2,X}

c) {1,2,3,4,5,6}

d) {1,2,3,3,4,5,6,7,8,9,10,11,12}

e) {CC,CX,XX}

f) {CC,CX,XC,XX}

g) {11,12,13,14,15,16,22,23,24,25,26,33,34,...,55,56,66}

h) {11,12,13,14,15,16,21,22,23,...,31,32,33,34,..., ...,61,62,...,66}

2.- A) Lanzamos un dado de 6 caras (del juego de dados). Expresa los siguientes sucasos:

a) “Obtener par”

b) “Obtener impar”

c) “Obtener múltiplo de 3”

B) Calcular la probabilidad de cada uno de estos sucesos

Sol.: A)

a) S1 = {2,4,6}, S2 = b) {1,3,5},

c) S3 = {3,6}

B) Aplicamos la Regla de Laplace

P(S1) = 3/6 = 1/2, P(S2) = 3/6 = 1/2

P(S3) = 2/6 = 1/3

3.- Tomamos una carta de la baraja española de 40 cartas. Calcula la probabilidad de los siguientes sucesos:

a) S1= “Obtener rey”

b) S2= “Obtener rey o un oros”

c) S3= “Obtener figura de copas o una espada”

Sol.: a) Casos posibles: 40, Casos favorables: 4

P(S1) = 4/40 = 1/10 = 0’1

b) c.p.: 40, c.f.: 4 + 9 = 13, P(S2) = 13/40

c) c.p.: 40, c.f.: 3 + 10 = 13,P(S3) = 13/40

4.- Lanzamos 3 monedas simultáneamente. Calcula la probabilidad de los siguientes sucesos:

a) S1 = “Obtener menos de 3 caras”

b) S2 = “Obtener al menos 1 caras”

c) S3 = “Obtener al menos 2 caras”

d) S4 = “Obtener ninguna cara”

Sol.: Espacio muestral M = {CCC,CCX,CXX,XXX}

Casos posibles: 4

a) P(S1) = 3/4, b) P(S2) = 3/4

c) P(S3) = 2/4, d) P(S4) = 1/4

Observa que S2 y S4 son sucesos contrarios

5.- Tengo una bolsa con 10 bolas: 4 rojas, 3 verdes, 2 blancas, 1 negra.

A) Tomo una bola al azar. Calcula:

a) Probabilidad de bola roja

b) Probabilidad de bola verde o blanca

B) Tomo dos bolas simultáneamente al azar. Calcula:

a) Probabilidad de dos rojas

b) Probabilidad de roja y verde

Sol.: A) Casos posibles: 10

a) C. favor.: 4 --> P = 4/10 = 2/5

b) C. favor.: 5 --> P = 5/10 = 1/2

B) Si las tomamos simultáneamente ‘no procede’ considerar el orden de salida, por tanto son combinaciones.

Aplicando la combinatoria: C. posibles $C_{10,2}$ = (10.9)/(2.1) = 45

a) c. fav.: $C_{4,2}$ = (4.2)/(1.2) = 4 --> P = 4/45

b) C. fav.: 4.3 = 12 --> 12/45 = 4/15

6.- Tengo una bolsa con 10 bolas: 4 rojas, 3 verdes, 2 blancas, 1 negra.

Tomo dos bolas sucesivamente (una tras otra) al azar. Calcula:

a) Probabilidad de dos rojas

b) Probabilidad de roja y verde

Sol.: Lo tratamos de dos formas:

A) Aplicando la combinatoria (variaciones porque sí influye el orden):

C. posibles $V_{10,2}$ = 10.9 = 90

a) c. fav.: $V_{4,2}$ = 4.3 = 12 --> P = 12/90 = 2/15

b) C. fav.: 4.3 = 12 --> 12/90 = 2/15

B) Tratándolo como suceso compuesto (extraemos una a una y tenemos en cuenta que la segunda está condicionada al resultado de la primera)

a) $P = \frac{4}{10} \cdot \frac{3}{9} = \ldots = \frac{2}{15}$

b) $P = \frac{4}{10} \cdot \frac{3}{9} = \ldots = \frac{2}{15}$

7.- Lanzo dos dados simultáneamente y anoto el resultado.

Se pide:

a) Probabilidad de obtener 2 y 5. (observa que no hay orden)

b) Probabilidad de que sumen 6

Sol.: C. posibles: $CR_{6,2} = C_{6+2-1,2} = C_{7,2}$ = (7.6)/(1.2) =21

(Otra forma: 6+5+4+3+2+1 = 21)

a) c. fav.: (2,5), (5,2) --> 2, --> P = 2/21

b) c. fav.: 1+5, 2+4, 3+3 --> 3, --> P = 3/21

8.- Lanzo dos dados uno tras otro y anoto los resultados.

Se pide:

a) Probabilidad de obtener el 2 y el 5 en este orden.

b) Probabilidad de que sumen 6

Sol.: C. posibles: $VR_{6,2} = 6^2 = 36$

a) c. fav.: (2,5) --> 1, --> P = 1/36

b) c. fav.: 1+5, 2+4, 3+3, 4+2, 5+1 -- > 5

P = 5/36

Probabilidad condicionada

9.- Tengo dos bolsas con bolas: U1 con 4 rojas, 3 verdes, 2 blancas, 1 negra. U2 con 3 rojas, 2 verdes, 4 blancas, 3 negras.

Lanzo una moneda, y, si resulta cara tomo una bola de U1, y si sale cruz la tomo de U2. Se pide:

a) S1 = "Obtener cara y bola verde"

b) S2 = "Obtener bola negra"

c) S3 = "Obtener bola roja"

NOTA: Es importante señalar que aquí el símbolo ^ significa la copulativa 'y'. Es preferible esta notación a utilizar el símbolo ∩ de 'intersección' de sucesos.

Sol.: a) P(S1) = $\frac{1}{2} \cdot \frac{3}{10} = \frac{3}{20}$

b) S2 = [C^(N/U1)]U[X^(N/U2)]

P(S2) = $[\frac{1}{2} \cdot \frac{1}{10}] + [\frac{1}{2} \cdot \frac{3}{12}] = \ldots$

c) S3 = [C^(R/U1)]U[X^(R/U2)]

P(S3) = $[\frac{1}{2} \cdot \frac{4}{10}] + [\frac{1}{2} \cdot \frac{3}{12}] = \ldots. = \frac{1}{5} + \frac{1}{8} = \frac{13}{40}$

Aplicación Teorema de Bayes

10.- Tengo dos bolsas con bolas: U1 con 4 rojas, 3 verdes, 2 blancas, 1 negra. U2 con 3 rojas, 2 verdes, 4 blancas, 3 negras.

Lanzo una moneda, y, si resulta cara tomo una bola de U1, y si sale cruz la tomo de U2.

Realizado el experimento resultó bola blanca. Calcula la probabilidad de que la moneda resultase cruz.

Sol.: P(X/R) = $\dfrac{P(X).P(R/X)}{P(C).P(R/C)+P(X).P(R/X)}$

P(X).P(R/X) = $\frac{1}{2}\cdot\frac{3}{12} = \ldots = \frac{1}{8}$

P(C).P(R/C) = $\frac{1}{2}\cdot\frac{4}{10} = \ldots = \frac{1}{5}$

Por tanto P(X/R) = $\dfrac{\frac{1}{8}}{\frac{1}{5}+\frac{1}{8}} = \dfrac{\frac{1}{8}}{\frac{13}{40}} = \dfrac{40}{13.8} = \dfrac{5}{13}$

11.- Blanca y María escriben cada una en su cuaderno una vocal, sin acuerdo previo. Se pide:

a) Describe el Espacio muestral de este experimento. ¿Cuántos elementos tiene?
b) Halla la probabilidad de que escriban la misma vocal.

c) Halla la probabilidad de que escriban vocales diferentes.

Sol.: a)E= $\{aa, ae, ai, \ldots, ea, ee, ei, \ldots, ia, ie, ii, \ldots, \ldots.\}$

Número = 5^2 = 25 (Son $VR_{5,2}$)

b) S= $\{aa, ee, ii, oo, uu\}$, cf = 5, P(S) = $\frac{5}{25} = \frac{1}{5}$

c) P(S’) = 1-P(S) = $\frac{4}{5}$

12.- El Programa de un examen consta de 30 temas. Un alumno ha aprendido de memoria 20 temas, y el examen consiste en tomar dos temas al azar y aprueba si contesta correctamente a los dos. Se pide:

a) Probabilidad que tiene de aprobar.

b) Probabilidad de que de los dos temas sólo uno haya aprendido de memoria.

Sol.: a) S= "Aprueba"

$$P(S)= \frac{C_{20,2}}{C_{30,2}} = ... = \frac{38}{87} = 0'43678...$$

b) S= "Sabe solo uno de memoria"

$$Cp = C_{30,2} = \frac{30.29}{2} = 435$$

$$Cf = C_{20,1} \,.\, C_{10,1} = 20 \,.\, 10 = 200$$

$$P(S) = \frac{200}{435} = \dots = 0'459770...$$

Otra forma (experimento compuesto y probabilidad condicionada):

$$a)\ P = \frac{20}{30} \,.\, \frac{19}{29} = \dots$$

$$b)\ P = \frac{20}{30} \,.\, \frac{10}{19} + \frac{10}{30} \,.\, \frac{20}{29} = \dots$$

13.- La Ciudad A tiene doble de habitantes que la Ciudad B. Se sabe que el 30% de los habitantes de B leen literatura, mientras que en A lo hacen solo el 10%. De una persona tomada al azar se sabe que es habitante de A o de B. Se pide:

a) La probabilidad de que lea literatura.

b) Después de comprobar que lee literatura, halla la probabilidad de que sea habitante de B.

Sol.: En primer lugar hacemos unas observaciones:

A= “Es habitante de A”, B= “Es habitante de B”

P(A) = 2.P(B), ya que A tiene el doble habitantes.

Por otro lado: P(A) + P(B) = 1, ya que sabemos que es de A ó de B y que $A \cap B = \emptyset$.

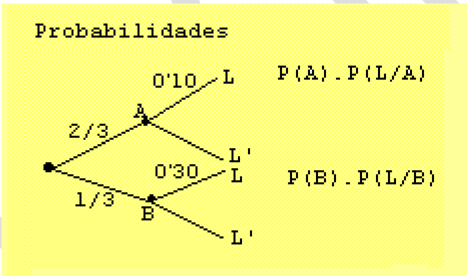
Entonces   1 = 3.P(B), P(B) = 1/3,   P(A) = 2/3

a) L = “Lee literatura”

L = $(A \cap L) \cup (B \cap L)$

P(L) = P(A).P(L/A) + P(B).P(L/B)

$= \frac{2}{3}.0{,}10 \quad + \frac{1}{3}.0{,}30 = 0'1666...$



b)$P(B/L) = \frac{P(B)\,.\,P\left(\frac{L}{B}\right)}{P(L)} = \frac{1}{3}\,.\,0'30 : 0'1667 \;=\; 0'5998\ldots$

$$$oOo$$$

Ejercicios/Ejemplos:

Tema 1

1.- Serie estadística:

5, 3, 4, 1, 2, 8, 9, 8, 7, 6, 6, 7, 9, 8, 7, 7, 1, 0, 1, 5, 9, 9, 8, 0, 8, 8, 8, 9, 5, 7.

Determina:

a) Frecuencias absolutas

b) Frecuencias absolutas acumuladas

c) Frecuencias relativas

d) Frecuencias relativas acumuladas

Sol.: a)

xi	0	1	2	3	4	5	6	7	8	9
fi	2	3	1	1	1	3	2	5	7	5

N = 30

b)

xi	0	1	2	3	4	5	6	7	8	9
Fi	2	5	6	7	8	11	13	18	25	30

c)

xi	0	1	2	3	4	5	6	7	8	9
fri	0’07	0’1	0’03	0’03	0’03	0’1	0’07	0’17	0’23	0’17
%	7	10	3	3	3	10	7	17	23	17

d)

xi	0	1	2	3	4	5	6	7	8	9
Fri	0’07	0’17	0’20	0’23	0’26	0’36	0’43	0’60	0’83	1
%	7	17	20	23	26	36	43	60	83	100

2.- Tomando la serie de datos tratados en el anterior (número 1) obtener:

a) La Media aritmética

b) La Mediana

c) La Moda

Sol.: a)

xi	0	1	2	3	4	5	6	7	8	9
fi	2	3	1	1	1	3	2	5	7	5
xi.fi	0	3	2	3	4	15	12	35	56	45

$$\text{ma} = \frac{\sum xi.fi}{N} = \frac{175}{30} = 5'833$$

b)

xi	0	1	2	3	4	5	6	7	8	9
Fi	2	5	6	7	8	11	13	18	25	30

$$N/2 = 15, \quad me = \frac{6+7}{2} = 6'50$$

c)

xi	0	1	2	3	4	5	6	7	8	9
fi	2	3	1	1	1	3	2	5	7	5

La frecuencia absoluta más alta es $f_8 = 7$,
mo = 8

3.- Tomando la serie de datos tratados en el núm. 1, obtener:

a) La Desviación media

b) La Varianza

c) La Desviación típica

Sol.-

a) Recordamos que ma = 5'83

xi	fi	|xi-ma|	|xi-ma|.fi
0	2	5'83	11'66
1	3	4'83	14'49
2	1	3'83	3'83
3	1	2'83	2'83
4	1	1'83	1'83
5	3	0'83	2'49
6	2	0'17	0'34
7	5	1'17	5'85
8	7	2'17	15'19
9	5	3'17	15'85
	N=30		Suma=74'36

$$Dm = \frac{\sum abs(xi - ma).fi}{N} = \frac{74'36}{30} = 2'48$$

b)

xi	fi	(xi-ma)	$(xi\text{-}ma)^2$	(xi-ma).fi
0	2	5’83	33’99	67’98
1	3	4’83	23’33	69’99
2	1	3’83	14’67	14’67
3	1	2’83	8’01	8’01
4	1	1’83	3’35	3’35
5	3	0’83	0’69	2’07
6	2	-0’17	0’03	0’06
7	5	-1’17	1’37	6’85
8	7	-2’17	4’71	32’97
9	5	-3’17	10’05	50’25
	N=30			Suma=256’20

$$S^2 = \text{Var.} = \frac{\sum (xi - ma)^2.fi}{30} = \frac{256'20}{30} = 8'54$$

c)

S = Desviación típica = $\sqrt{Var.} = \sqrt{8'54} = 2'92$

4.- Las puntuaciones obtenidas en un test aplicado a 20 alumnos son las siguientes:

16, 22, 21, 20, 23, 22, 17, 15, 13, 22, 17, 18, 20, 17, 22, 16, 23, 21, 22, 18

Obtener: La media, los cuartiles, la desviación típica

Sol.:

xi	fi	Fi	xi.fi	xi^2	xi^2.fi
13	1	1	13	169	169
15	1	2	15	225	225
16	2	4	32	256	512
17	3	7	51	289	867
18	2	9	36	324	648
20	2	11	40	400	800
21	2	13	42	441	882
22	5	18	110	484	2420
23	2	20	46	529	1058
N=20			Sum=385		Sum=7581

$ma = \frac{385}{20} = 19'25$

Cuartil Q1:

N/4 = 5, F4 = 7 > 5 y F3 < 5 y por tanto
Q1 = 17

Cuartil Q2:

2.N/4 = 10, F6 > 10 y F5 < 10 --> Q2 = 20

Cuartil Q3:

3.N/4 = 15, F8 > 15 y F7 < 15 --> Q3 = 22

Varianza:

$$S^2 = \frac{\sum xi^2 . fi}{20} - ma^2 = \frac{7581}{20} - (19'25)^2 = 8'49$$

Desviación típica: S = 2'91

5.- En cierta Región, y durante cierto mes de 30 dias, se registraron las siguientes temperaturas máximas:

32, 31, 28, 29, 29, 33, 32, 31, 30, 31, 31, 27, 28, 29, 30, 32, 31, 31, 30, 30, 29, 30 30, 30, 31, 30, 31, 34, 33, 33

Obtener:

a) Moda, la media ma, los percentiles 30 y 70

b) La varianza y la desv. típica

Sol.:

xi	fi	Fi	xi.fi	xi^2	xi^2.fi
27	1	1	27	729	729
28	2	3	56	784	1568
29	5	8	145	841	4205
30	7	15	210	900	6300
31	8	23	248	961	7688
32	3	26	96	1024	3072
33	3	29	99	1089	3267
34	1	30	34	1156	1156
N=30			Sum=915		Sum=27985

a) mo = 31, ma = 30’50,

p_{30} : $\frac{30\,.\,N}{100}$ = 9’00, F3 = 9 --> p_{30} = 29

p_{70} : $\frac{70\,.N}{100}$ = 21’00, F5 > 21 y F4 < 21--> p_{70} = 31

b)$S^2 = \frac{27985}{30}$ - $(30’50)^2$ = 2’58, S = $\sqrt{2,58}$ = 1’61

De Variable continua

7.- Hemos registrado el peso de 50 recién nacidos, y agrupándolos en intervalos tenemos:

Peso kgs	núm. niños
[2’5-3’0)	6
[3’0-3’5)	23
[3’5-4’0)	12
[4’0-4’5)	9

Obtener: ma, los cuartiles, la desv. típica, el recorrido intercuartílico.

Sol.:

Clases	Marca xi	fi	Fi	xi.fi	xi^2.fi
[2’5-3’0)	2’75	6	6	16’50	45’38
[3’0-3’5)	3’25	23	29	74’75	242’94
[3’5-4’0)	3’75	12	41	45’00	168’75
[4’0-4’5)	4’25	9	50	38’25	162’56
		N=50		Sum=174’50	Sum=619’63

$$ma = \frac{174'50}{50} = 3'4$$

Q1: $\frac{N}{4} = 12'50$, F2 > 12’50 y F1 < 12’50 -->

Intervalo: [3’0-3’5),

$$Q1 = 3+0'5.\frac{12'50-6}{29-6} = 3 + 0'5.0'28 = 3'14$$

NOTA:

Fórmula para obtener Qi en el caso de variable continua

$$Qi = E_{i+1} + L.\frac{\frac{i.N}{4} - F_i}{F_{i+1} - F_i}$$

donde E_{i+1} es el extremo inferior del intervalo cuya frecuencia acumulada es la primera que supera (o iguala) el valor i.N/4, F_{i+1} es la referida frecuencia. L es la amplitud del intervalo.

Q3: 3.N/4 = 37’50, F3 > 37’50 y F2 < 37’50

--> Intervalo: [3’5-4’0)

$$Q3 = 3'5 + 0'5.\frac{37'50 - 29}{41 - 29} = 3'5 + 0'5.\frac{8'5}{12} = 3'85$$

Varianza: $S^2 = \dfrac{619'63}{50} - (3'49)^2 = 0'2125$

Desviación típica: $S = \sqrt{0'2125} = 0'461$

Recorrido intercuartílico:

$$Q3–Q1 = 3'85 – 3'14 = 0'71$$

8.- Damos el cuadro de datos y el alumno debe comprobar (haciendo los cálculos correspondientes) los resultados que damos:

Clases	fi
[0 -15)	10
[15-30)	15
[30-45)	25
[45-60)	20
[60-75)	20
[75-90)	10

Resultados:

N = 100, ma = 45’75,

mo = $30 + 15.\dfrac{25-15}{(25-15)+(25-20)} = 30 + 15.\dfrac{10}{15} = 40$

NOTA:
La fórmula para la moda en el caso de variable continua es

$$mo = E_{i+1} + L.\frac{f_{i+1} - f_i}{(f_{i+1} - f_i) + (f_{i+1} - f_{i+2})}$$

donde E_{i+1} es el extremo izquierdo del intervalo modal (aquel con mayor frecuencia), f_{i+1} es la frecuencia más alta. L es la amplitud del intervalo.

Mediana me:

NOTA:
Fórmula para la mediana

$$me = E_{i+1} + L.\frac{N/2 - F_i}{F_{i+1} - F_i}$$

Si N/2 coincide con una Fi, entonces

me = extremo superior del intervalo correspondiente.

En este caso: N/2 = 50 = F3 , por tanto me = 45

Rango: R = 90 – 0 = 90

Varianza: $S^2 = \dfrac{257625}{100} - (45'75)^2 = 483'19$

S = 21’98

Tema 2

1.- En una prueba sobre el número de bacterias en un cultivo, a medida que pasan las horas, obtenemos los siguientes datos:

Núm. de horas:	0	1	2	3	4	5
Núm. bacterias:	12	19	23	34	56	62

Se pide:

a) La media y desv. típica de las variables X = núm. de horas, Y = núm. de bacterias.

b) La covarianza de la variable bidimensional (X,Y)

Sol.: a)

xi	yi	fi	xi.fi	xi^2.fi	yi.fi	yi^2.fi	xi.yi.fi
0	12	1	0	0	12	144	0
1	19	1	1	1	19	361	19
2	23	1	2	4	23	529	46
3	34	1	3	9	34	1156	102
4	56	1	4	16	56	3136	224
5	62	1	5	25	62	3844	310
		N=6	15	55	206	9170	701

mx = 15/6 = 2’50 my = 206/6 = 34’33

Vx = 55/6-(2’50)^2 = 2’92, Sx = 1’71

Vy = 9170/6-(34’33)^2 = 349’78, Sy = 18’70

b) S_{xy} = Covar(X,Y) = $\dfrac{\sum xi.yi.fi}{N} - mx.my$ =

= $\dfrac{701}{6} - 2'50.34'33$ = 31'01

2.- Al tomar datos de un estudio sobre la relación entre el número de horas semanales que ven televisión y los ingresos mensuales (en centenares de euros) de sus padres, en grupo de alumnos de la ESO, obtuvimos el siguiente cuadro de datos:

Horas sema.	Ingresos Mens. [4-10)	[10-16)	[16-22)	[22-28)
[10-20)	3	2	--	--
[20-30)	4	6	6	3
[30-40)	11	4	3	2
[40-50)	1	--	--	4

a) Calcula las medias y las desv. típicas

b) Calcula la covarianza

Sol.:

a) X = Ingresos, Y = Núm. de horas

xi	yi	fi	xi.fi	xi^2.fi	yi.fi	yi^2.fi	xi.yi.fi
7	15	3	21	147	45	675	315
7	25	4	28	196	100	2500	700
7	35	11	77	539	385	13475	2695
7	45	1	7	49	45	2025	315
13	15	2	26	338	30	450	390
13	25	6	78	1014	150	3750	1950

13 35	4	52	676	140	4900	1820
13 45	0	0	0	0	0	0
19 15	0	0	0	0	0	0
19 25	6	114	2166	150	3750	2850
19 35	3	57	1083	105	3675	1995
19 45	0	0	0	0	0	0
25 15	0	0	0	0	0	0
25 25	3	75	1875	75	1875	1875
25 35	2	50	1250	70	2450	1750
25 45	4	100	2500	180	8100	4500
	N = 49	685	11833	1475	47625	21155

$mx = \frac{685}{49} = 13'98$, $\quad$ $my = 1475/49 = 30'10$

$Sx^2 = \frac{11833}{49} - (13'98)^2 = 46'05$, $\quad Sx = 6'79$

$Sy^2 = \frac{47625}{49} - (30'10)^2 = 65'93$, $\quad Sy = 8'12$

c) $Sxy = \frac{21155}{49} - 13'98.30'10 = 10'94$

3.- El siguiente cuadro de datos registra los gastos (Cent. de e.) en publicidad de una empresa y las correspondientes ventas (cent. e.)

Publicidad	1	2	3	4	5	6	7	8
Ventas	15	16	14	17	20	18	18	19

a) El alumno debe comprobar los siguientes resultados.
b) Obtener el coeficiente de correlación.
c) Expresa las rectas de regresión:

$y = f(x)$, $x = g(y)$

Sol.: a) X = Gasto en publicidad, Y = Producto ventas mx = 4’50, my = 17’13

Sx^2 = 5’25, Sy^2 = 3’61

Sx = 2’29, Sy = 1’90

Covarianza de (X,Y): Sxy = 3’31

b) Coef. de correlación: r = $\frac{Sxy}{Sx.Sy}$

r = $\frac{3'31}{2'29.1'90}$ = 0’76

Correlación débil positiva

c) Rectas de regresión:

y-my = $\frac{Sxy}{Sx^2}.(x-mx)$, x-mx = $\frac{Sxy}{Sy^2}.(y-my)$

y = 17’13 + 0’63.(x-4’50)

y = 0’63.x + 14’30

x = 4’50 + 0’92.(y-17’13)

x = 0’92.y – 11’26

4.- El siguiente cuadro de datos se refiere al conteo (tabulación) del número de bacterias dependiendo del tiempo transcurrido:

Núm. horas	0	1	2	3	4	5
Núm. bacterias	12	19	23	34	56	62

a) Comprueba los siguientes resultados

b) Obtener el coef. de correlación

c) Expresa las rectas de regresión

Sol.: a) X = número de horas, Y = núm. de bacterias

mx = 2’50, my = 34’33

Sx^2 = 2’92, Sx = 1’71

Sy^2 = 349’69, Sy = 18’70

Sxy = 31

b) $r = \dfrac{Sxy}{Sx.Sy} = \dfrac{31}{1'71.18'70}$ = 0’97

c) $y\text{-}my = \dfrac{Sxy}{Sx^2}.(x-mx)$, $x\text{-}mx = \dfrac{Sxy}{Sy^2}.(y-my)$

y = 10’62.x + 7’78

x = 0’09.y – 0’59

5.- Damos el siguiente cuadro de datos (doble entrada), y el alumno debe comprobar los resultados que adjuntamos:

Habilidad verbal / Razonamiento abstracto	[10-20)	[20-30)	[30-40)	[40-50)
[15-25)	5	3	0	0
[25-35)	2	6	1	0
[35-45)	0	1	4	2
[45-55)	0	0	3	3
[55-65)	0	0	1	2

Sol.: Resultados que el alumno debe comprobar

X = Habilidad verbal, Y = Razonamiento

a) mx = 29’85, my = 36’06

Sx^2 = 109’83, Sx = 10’48

Sy^2 = 163’27, Sy = 12’78

Sxy = 110’00

b) r = 0’82

c) y = x + 6’21

x = 0’67.y + 5’69

d) ¿Qué puntuación obtiene en razonamiento abstracto una persona que obtuvo 45 en habilidad verbal?

x = 45 --> y = 45 + 6’21 = 51’21

Observa: La recta y = x + 6’21 tiene pendiente 1, lo que significa que y crece igual que x.

El valor r = 0’82 desdice apreciablemente el hecho anterior. Este valor de r debía estar más próximo al valor 1.

Tema 3

1.- Expresa el espacio muestral de los siguientes experimentos:

a) Lanzar una moneda

b) Lanzar el dado de quinielas (6 caras: 2 con el 1, dos con el 2, dos con x)

c) Lanzar el dado de 6 caras (del 1 al 6)

d) Tomar una carta de la baraja española, y considerar el número obtenido del 1 al 12

e) Lanzar dos monedas simultáneamente

f) Lanzar dos dados sucesivamente, y anotando en cada lanzamiento

g) Lanzar dos dados normales simultáneamente

h) Lanzar dos dados normales sucesivamente, y anotando en cada lanzamiento

Sol.: a) {C,X}

b) {1,2,X}

c) {1,2,3,4,5,6}

d) {1,2,3,3,4,5,6,7,8,9,10,11,12}

e) {CC,CX,XX}

f) {CC,CX,XC,XX}

g) {11,12,13,14,15,16,22,23,24,25,26,33,34,...,55,56,66}

h) {11,12,13,14,15,16,21,22,23,...,31,32,33,34,..., ...,61,62,...,66}

2.- A) Lanzamos un dado de 6 caras (del juego de dados). Expresa los siguientes sucasos:

a) "Obtener par"

b) "Obtener impar"

c) "Obtener múltiplo de 3"

B) Calcular la probabilidad de cada uno de estos sucesos

Sol.: A)

a) S1 = {2,4,6}, S2 = b) {1,3,5},

c) S3 = {3,6}

B) Aplicamos la Regla de Laplace

P(S1) = 3/6 = 1/2, P(S2) = 3/6 = 1/2

P(S3) = 2/6 = 1/3

3.- Tomamos una carta de la baraja española de 40 cartas. Calcula la probabilidad de los siguientes sucesos:

a) S1= "Obtener rey"

b) S2= "Obtener rey o un oros"

c) S3= "Obtener figura de copas o una espada"

Sol.: a) Casos posibles: 40, Casos favorables: 4

P(S1) = 4/40 = 1/10 = 0'1

b) c.p.: 40, c.f.: 4 + 9 = 13, P(S2) = 13/40

c) c.p.: 40, c.f.: 3 + 10 = 13,P(S3) = 13/40

4.- Lanzamos 3 monedas simultáneamente. Calcula la probabilidad de los siguientes sucesos:

a) S1 = “Obtener menos de 3 caras”

b) S2 = “Obtener al menos 1 caras”

c) S3 = “Obtener al menos 2 caras”

d) S4 = “Obtener ninguna cara”

Sol.: Espacio muestral M = {CCC,CCX,CXX,XXX}

Casos posibles: 4

a) P(S1) = 3/4, b) P(S2) = 3/4

c) P(S3) = 2/4, d) P(S4) = 1/4

Observa que S2 y S4 son sucesos contrarios

5.- Tengo una bolsa con 10 bolas: 4 rojas, 3 verdes, 2 blancas, 1 negra.

A) Tomo una bola al azar. Calcula:

a) Probabilidad de bola roja

b) Probabilidad de bola verde o blanca

B) Tomo dos bolas simultáneamente al azar. Calcula:

a) Probabilidad de dos rojas

b) Probabilidad de roja y verde

Sol.: A) Casos posibles: 10

a) C. favor.: 4 --> P = 4/10 = 2/5

b) C. favor.: 5 --> P = 5/10 = 1/2

B) Si las tomamos simultáneamente ‘no procede’ considerar el orden de salida, por tanto son combinaciones.

Aplicando la combinatoria: C. posibles $C_{10,2}$ = (10.9)/(2.1) = 45

a) c. fav.: $C_{4,2}$ = (4.2)/(1.2) = 4 --> P = 4/45

b) C. fav.: 4.3 = 12 --> 12/45 = 4/15

6.- Tengo una bolsa con 10 bolas: 4 rojas, 3 verdes, 2 blancas, 1 negra.

Tomo dos bolas sucesivamente (una tras otra) al azar. Calcula:

a) Probabilidad de dos rojas

b) Probabilidad de roja y verde

Sol.: Lo tratamos de dos formas:

A) Aplicando la combinatoria (variaciones porque sí influye el orden):

C. posibles $V_{10,2}$ = 10.9 = 90

a) c. fav.: $V_{4,2}$ = 4.3 = 12 --> P = 12/90 = 2/15

b) C. fav.: 4.3 = 12 --> 12/90 = 2/15

B) Tratándolo como suceso compuesto (extraemos una a una y tenemos en cuenta que la segunda está condicionada al resultado de la primera)

a) P = $\frac{4}{10} \cdot \frac{3}{9}$ = 12/90 = 2/15

b) P = $\frac{4}{10} \cdot \frac{3}{9}$ = 12/90 = 2/15

7.- Lanzo dos dados simultáneamente y anoto el resultado.

Se pide:

a) Probabilidad de obtener 2 y 5. (observa que no hay orden)

b) Probabilidad de que sumen 6

Sol.: C. posibles: $CR_{6,2} = C_{6+2-1,2} = C_{7,2}$ = (7.6)/(1.2) =21

(Otra forma: 6+5+4+3+2+1 = 21)

a) c. fav.: (2,5), (5,2) --> 2, --> P = 2/21

b) c. fav.: 1+5, 2+4, 3+3 --> 3, --> P = 3/21

8.- Lanzo dos dados uno tras otro y anoto los resultados.

Se pide:

a) Probabilidad de obtener el 2 y el 5 en este orden.

b) Probabilidad de que sumen 6

Sol.: C. posibles: $VR_{6,2} = 6^2 = 36$

a) c. fav.: (2,5) --> 1, --> P = 1/36

b) c. fav.: 1+5, 2+4, 3+3, 4+2, 5+1 --> 5

P = 5/36

De Probabilidad condicionada

9.- Tengo dos bolsas con bolas: U1 con 4 rojas, 3 verdes, 2 blancas, 1 negra. U2 con 3 rojas, 2 verdes, 4 blancas, 3 negras.

Lanzo una moneda, y, si resulta cara tomo una bola de U1, y si sale cruz la tomo de U2. Se pide:

a) S1 = "Obtener cara y bola verde"

b) S2 = "Obtener bola negra"

c) S3 = "Obtener bola roja"

Sol.: a) P(S1) = 1/2.3/10 = 3/20

b) S2 = [C^(N/U1)]U[X^(N/U2)]

P(S2) = [1/2.1/10] + [1/2.3/12] = 1/20 + 3/24

c) S3 = [C^(R/U1)]U[X^(R/U2)]

P(S3) = [1/2.4/10] + [1/2.3/12] = 4/20 + 3/24 = 1/5 + 1/8 = 13/40

De Teorema de Bayes

10.- Tengo dos bolsas con bolas: U1 con 4 rojas, 3 verdes, 2 blancas, 1 negra. U2 con 3 rojas, 2 verdes, 4 blancas, 3 negras.

Lanzo una moneda, y, si resulta cara tomo una bola de U1, y si sale cruz la tomo de U2.

Realizado el experimento resultó bola blanca. Calcula la probabilidad de que la moneda resultase cruz.

Sol.: P(X/R) = $\dfrac{P(X).P(R/X)}{P(C).P(R/C)+P(X).P(R/X)}$

$P(X).P(R/X) = \frac{1}{2} \cdot \frac{3}{12} = 3/24 = 1/8$

$P(C).P(R/C) = \frac{1}{2} \cdot \frac{4}{10} = 4/20 = 1/5$

Por tanto $P(X/R) = \dfrac{1/8}{1/5+1/8} = \dfrac{1/8}{13/40} = \dfrac{40}{13.8} = 5/13$

11.- Blanca y María escriben cada una en su cuaderno una vocal, sin acuerdo previo. Se pide:

a) Describe el Espacio muestral de este experimento. ¿Cuántos elementos tiene?
b) Halla la probabilidad de que escriban la misma vocal.

c) Halla la probabilidad de que escriban vocales diferentes.

Sol.: a)E= $\{aa, ae, ai, \ldots, ea, ee, ei, \ldots, ia, ie, ii, \ldots, \ldots.\}$

Número = $5^2 = 25$ (Son $VR_{5,2}$)

b) S= $\{aa, ee, ii, oo, uu\}$, cf = 5, P(S) = 5/25 = 1/5

c) P(S') = 1-P(S) = 4/5

12.- El Programa de un examen consta de 30 temas. Un alumno ha aprendido de memoria 20 temas, y el examen consiste en tomar dos temas al azar y aprueba si contesta correctamente a los dos. Se pide:

a) Probabilidad que tiene de aprobar.

b) Probabilidad de que de los dos temas sólo uno haya aprendido de memoria.

Sol.: a) S= “Aprueba”

$$P(S)=\frac{C_{20,2}}{C_{30,2}} = ... = 38/87 = 0'43678...$$

b) S= “Sabe solo uno de memoria”

$$Cp = C_{30,2} = \frac{30\,.29}{2} = 435$$

$$Cf = C_{20,1}\,.\,C_{10,1} = 20\,.\,10 = 200$$

$$P(S) = \frac{200}{435} = \frac{40}{87} = 0'459770...$$

Otra forma (experimento compuesto y probabilidad condicionada):

$$a)\ P = \frac{20}{30}\,.\frac{19}{29} = ...$$

$$b)\ P = \frac{20}{30}\,.\frac{10}{19} + \frac{10}{30}\,.\frac{20}{29} = ...$$

13.- La Ciudad A tiene doble de habitantes que la Ciudad B. Se sabe que el 30% de los habitantes de B leen literatura, mientras que en A lo hacen solo el 10%. De una persona tomada al azar se sabe que es habitante de A o de B. Se pide:

a) La probabilidad de que lea literatura.

b) Después de comprobar que lee literatura, halla la probabilidad de que sea habitante de B.

Sol.: En primer lugar hacemos unas observaciones:

A= “Es habitante de A”, B= “Es habitante de B”

P(A) = 2.P(B), ya que A tiene el doble habitantes.

Por otro lado: P(A) + P(B) = 1, ya que sabemos que es de A ó de B y que $A \cap B = \varnothing$.

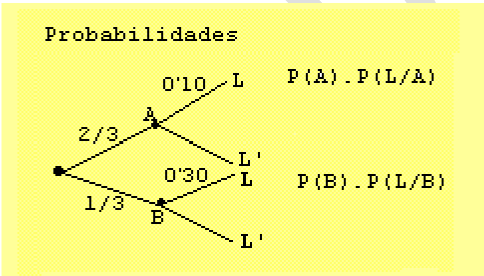
Entonces 1 = 3.P(B), P(B) = 1/3, P(A) = 2/3

a) L = “Lee literatura”

L = $(A \cap L) \cup (B \cap L)$

P(L) = P(A).P(L/A) + P(B).P(L/B)

= 2/3.0,10 + 1/3.0,30 = 0’1666...



b) P(B/L)= $\dfrac{P(B).P(L/B)}{P(L)} = \dfrac{1/3.0'30}{0'1667}$ = 0’5998...

Problemas:

1.- Lanzo dos dados al aire. Calcula la probabilidad de:

a)Suma par
b)Suma mayor que siete
c)Suma múltiplo de tres

(Res.: 1/2, 5/12, 1/3)

2.- Tomo un dado y pinto de blanco las caras 1,2,4,5, y de verde las caras 3,6. Lo lanzo al aires dos veces. Calcula la probabilidad de

a)Las dos veces verde
b)La primera verde y la segunda blanco
c)La primera blanco y la segunda verde

(Res.: 1/9, 2/9, 2/9)

3.- Calcula las siguientes probabilidades

a)Tomo 5 cartas simultáneamente, deseo tres ases y las otras dos iguales entre sí.

b)En una bolsa tengo 6 bolas blancas, 8 verdes. Tomo 4 simultáneamente, y deseo que no sean las cuatro blancas.

(Res.: a) C.posibles = $C_{40,5}$, C.f.= 9.$C_{4,3}$,
P = $\frac{1}{18278}$)

b)P = -$\frac{C_{6,4}}{C_{14,4}} = \frac{986}{1001}$)

4.- Calcula las siguientes probabilidades

a)Tengo en una bolsa siete bolas numeradas del 1 a 7. Tomo dos simultáneamente y deseo que las dos sean impares, o que sean de la misma paridad.

b)El Profesor ha escrito cinco números coincidiendo con los temas del texto. El alumno, con los ojos cerrados, ha de hacer coincidir cada número con el Tema correspondiente.

Res.: a) P1 = 2/7, P2 = 3/7
b)c.f.= 1, c.p. = $VR_{5,2}$, P = 1/25)

Tema 4

Experimento aleatorio y Variable aleatoria

Función de probabilidad asociada

Función de Distribución asociada

Distribuciones Binomial

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

4.1.- Variable aleatoria

Necesariamente hemos de 'activar' nuestra intuición. Lo conseguiremos mediante dos ejemplos.

Ejemplo 1:

Lanzamos tres monedas y anotamos el número de caras. (Da igual como las lancemos ya que las monedas son exactamente iguales y perfectamente equilibradas).

Los posibles resultados los tenemos en el espacio muestral

M = {CCC, CCX, CXC, XCC, XXC, XCX, CXX, XXX}

El número de caso posibles es VR(2;3) = 2^3 = 8

El número de caras posibles lo tenemos en
{0, 1, 2, 3}

Pues bien, en este ejemplo, tenemos una 'Ley' que a cada elemento del espacio muestral M le asocia un valor:

CCC --> 3, CCX --> 2, XCX --> 1, XXX --> 0

Variable aleatoria:

Llamamos 'Variable aleatoria' a esta ley así definida, y la representamos por X,

X: M ----> R
m ----> "número de caras en m"

Tenemos Im(X) = {0,1,2,3}

NOTA: En otros contestos, o por conveniencia de notación, la representaremos por fa y haremos referencia a ella como 'función aleatoria'.

Ejemplo 2:

Lanzamos dos dados y anotamos la suma de los resultados. (Por tratarse de dos dados exactamente iguales, no cambia nada si los lanzamos simultáneamente o si lo hacemos uno tras otro).

El espacio muestral es

M = {11,12,21,13,31, ..., 61,...,22,23,32,...,55,56,65,66}

El número de casos posibles es VR(6;2) = 6^2 =36

Variable aleatoria:

Ley que a cada m de M le asocia la suma:

X: M ----> R
m ----> "Suma de los dígitos de m"

Im(X) = {2,3,4,5,6,7,8,9,10,11,12}

En general:

Tenemos un Experimento aleatorio y su Espacio muestral M. Toda 'Ley' que a cada elemento de M le asocie un valor real la llamamos 'Variable aleatoria', que se representa por X (también función aleatoria fa):

X: M ----> R
m ----> r

4.2.- Función de Probabilidad asociadas a una variable aleatoria

Volvemos a los ejemplos.

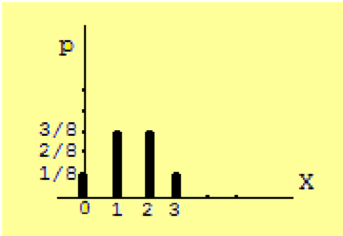
Ejemplo 1: Im(X) = {0,1,2,3}

Teniendo en cuenta

M = {CCC, CCX, CXC, XCC, XXC, XCX, CXX, XXX}

la probabilidad de la ‘ocurrencia’ de cada uno de estos valores es:

$p(0) = 1/8$, $p(1) = 3/8$, $p(2)=3/8$, $p(3)=1/8$



Observa que su suma es 1.

Tenemos así una nueva función

p: Im(X) ----> R

r ----> p(X = r) = “Probabilidad de que X tome el valor r”

Es la **‘Función de probabilidad asociada a X’**.

La designaremos por P_x, o simplemente p si no existe ambigüedad.

4.3.- Función de distribución asociada a la variable aleatoria X

Nos preguntamos ahora por la “Probabilidad de que X tome valor $\leq$ r ”, donde r es un valor real.

Ejemplo 2:

$p(X \leq 0) = 1/8$,

$p(X \leq 1) = 1/8 + 3/8 = 4/8$

$p(X \leq 2) = 1/8 + 3/8 + 3/8 = 7/8$

$p(X \leq 3) = 1/8 + 3/8 + 3/8 + 1/8 = 1$

Este concepto es análogo al de 'frecuencia acumulada' en Estadística.

Definimos una nueva función así:

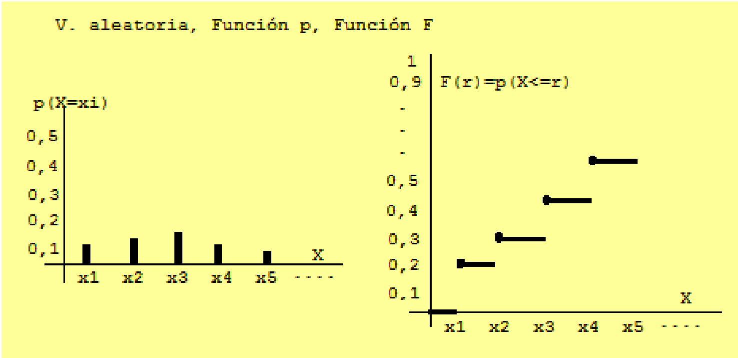
F: R ----> R
r ----> $p(X \leq r)$ = "Probabilidad de que X tome un valor $\leq$ r"

x ---> $F(x) = p(X \leq x)$

Son evidentes las siguientes

Propiedades:

a) $0 \leq F(x) \leq 1$
b) $F(x) = 0$ para todo x < 'mínimo valor de X'
c) $F(x) = 1$ para todo $x \geq$ 'máximo valor de X'
d) Si X es discreta, su gráfica es creciente escalonada, produciendo 'salto' cuando x = 'un valor de X'.



4.4.- Parámetros asociados a la v.a. X: Media, Varianza, Esperanza matemática

Media y Varianza de X:

Si Im(X) = {x1,x2,...,xn}, y sus probabilidades son pi = p(X = xi), tenemos la Media m (media aritmética) y la Varianza de forma análoga a como hacíamos en Estadística. El papel de pi es el mismo que el de la frecuencia relativa.

Media:

m = x1.p1 + x2.p2 +...+ xn.pn

Varianza:

$S^2 = (x1-m)^2.p1 + (x2-m)^2.p2 + ...+ (xn-m)^2.pn$

Ejemplo 1:

m = 0.1/8 + 1.3/8 + 2.3/8 + 3.1/8 = (3+6+3)/8 = 12/8 = 1,5

$S^2 = (0-1{,}5)^2.1/8 +(1-1{,}5)^2.3/8 +(2-1{,}5)^2.3/8 +(3-1{,}5)^2.1/8$ = (2,25 +0,25.3 +0,25.3 +2,25)/8 = (2,25 + 0,75 + 0,75 + 2,25)/8 = 6/8 = 0,75

$S^2 =$ **0,75**

Esperanza matemática:

Def.:

Es la Suma de los productos que resultan al tomar cada valor xi de la v.a. X por la probabilidad pi de que X tome el valor xi.

Su valor viene dado por la siguiente igualdad:

$$E[X] = \sum_{i=1,2,\ldots,n} xi.p(X = xi) \text{ , o bien}$$

$$E[X] = \sum_{i=1,2,\ldots,n} xi.pi$$

(Fórmula general de la Esperanza matemática)

NOTA:
Cuando tenemos una variable aleatoria X y hemos llegado a 'determinar' su función de distribución asociada, decimos que tenemos una 'Distribución …'

4.5.- Introducción a un caso muy frecuente: Distribución Binomial

NOTA: Este concepto será tratado otra vez en el Tema 5.

Sea un experimento aleatorio en el cual sólo es posible un suceso A y su contrario A' .

Supongamos además que si repetimos el experimento, el resultado no está influenciado por el resultado de las pruebas precedentes (Son físicamente independientes).

Tenemos:

Espacio muestral M = {A, A'} y el espacio de sucesos P(M) = {$\emptyset$, A, A', M}

Sean sus probabilidades los valores:

p = p(A), q = p(A') = 1-p(A) = 1-p

Lo que sigue resulta más explícito e inteligible si desarrollo un ejemplo.

Ejemplo:

Lanzo 3 monedas sucesivamente. El resultado de un lanzamiento no influye en el resultado del siguiente: Son 'físicamente independientes'.

NOTA: Es indiferente lanzarlas sucesivamente o lanzarlas simultáneamente.

El espacio muestral es

M = {CCC,CCX,CXC,XCC,CXX,XCX,XXC,XXX}

Su número coincide con el resultado que nos da la fórmula de las variaciones con repetición:

$$VR(2;3) = 2^3 = 8$$

(dos elementos tomados de 3 en 3, con repetición)

Tenemos aquí una Variable aleatoria:

X: M ----> R,

X(m) = 'número de caras en m'

Im(X) = {0,1,2,3}

Nos preguntamos por la 'Probabilidad de que resulten 2 caras':

p(X=2)

(con lo cual tenemos ya la función de probabilidad asociada)

Casos favorables:

De los n = 3 lugares, las dos caras pueden ocupar dos lugares de todas las formas posibles. Este número viene dado ‘combinando’ de dos en dos esos tres lugares.

Por tanto

Número casos f. = C(3;2) = (3.2.1)/(2.1) = 3

Como el número de casos posibles es 8, Tengo

p(X=2) = C(3;2)/VR(2;3) = 3/8

Otra forma (siguiendo con el ejemplo):

Considerándolo como experimento compuesto por una sucesión de experimentos.

Al lanzar una moneda tengo p = p(C) = 1/2, y q = p(X)= 1/2. Recuerda la notación:

p = p(A), q = p(A’)

Uno de los casos con 2 caras es CCX. Teniendo en cuenta la independencia de las pruebas, la probabilidad de que ocurra este caso es

$p(CCX) = p^2.q$, y la misma para un caso como CXC.

Ahora bien, estos casos con 2 caras pueden ocurrir tantas veces como C(3;2), ya que tenemos ‘tres lugares’ que pueden ser ocupados por C ‘de dos en dos’ de todas las formas posibles.

Por tanto, llegamos a que la probabilidad de dos caras es

$$p(X=2) = C(3;2).p^2.q$$

En este ejemplo concreto, C(3;2) = 3, y

$$p(X=2) = 3.(1/2)^2.(1/2) = 3/8$$

Generalización:

Si en lugar de 3 monedas tenemos n, y si preguntamos por la probabilidad de obtener k caras, entonces llego a:

$$p(X= k) = C(n;k).p^k.q^{n-k}$$

Concretamente, siendo p = 1/2, q = 1/2, si hago n = 8 y k = 5, entonces tengo

$$p(X=5) = C(8;5).(1/2)^5.(1/2)^3 = (*)$$

C(8;5) = V(8;5)/P(5) = (8.7.6.5.4)/(5.4.3.2.1) = (8.7.6)/(3.2) = 8.7 = 56

(*) = $56.(1/2)^8$ = 56/256 = 0,2188

Fórmula caso general:

Si de una Población de individuos sabemos que p(A) = p, y por tanto p(A') = q = 1-p, al tomar una muestra de n individuos, simultáneamente o sucesivamente con recuperación, y nos preguntamos por el número de éstos que son del tipo A, tenemos la siguiente fórmula:

$$p(X=k) = C(n;k).p^k.q^{n-k} \qquad (1)$$

A continuación tenemos la función de distribución:

$$F(k) = P(X \leq k)$$

Definición:

Decimos que X es una Variable aleatoria binomial, y que tenemos una ‘Distribución binomial’. La representamos por B(n;p), donde n y p son los parámetros que la determinan.

$$$$oOo$$$$

Tema 5

Distribuciones discretas

Frecuencias relativas y Concepto de Probabilidad

Variable aleatoria y Función de densidad

Función de distribución

La Distribución Binomial

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

5.1.- Experimento aleatorio: Frecuencias relativas y Concepto de Probabilidad

Experimento:

Espacio Muestral y Álgebra de sucesos

En un experimento aleatorio llamamos ‘Espacio muestral’ al conjunto M = {Todos los sucesos elementales posibles}.

La familia ‘partes de M’, P(M), de todos los subconjuntos de M, incluido el vacío $\emptyset$, dotado de las operaciones entre conjuntos:

-Unión e Intersección
-Complementario

cumple las condiciones para que $P(M;\cup,\cap)$ tenga estructura de Álgebra de Boole, y la llamaremos ‘Álgebra de sucesos’ del experimento.

Frecuencia relativa $F_r(A)$ de la ocurrencia de un suceso y **Concepto de probabilidad, p(A):**

Sea un experimento aleatorio y A un suceso que puede ocurrir al realizar el experimento. Ejecutamos el experimento n veces y anotamos el número de veces que ocurre el suceso A. Sea n_A este valor y lo llamamos frecuencia absoluta de A. Su frecuencia relativa es $f_r(A) = \frac{n_A}{n}$.

Cuando hacemos n arbitrariamente grande el valor $f_r(A)$ tiende hacia un valor real, que designamos por p(A), y al que llamamos ‘Probabilidad de A’.

Significado y uso práctico del valor p(A):

Al ejecutar de nuevo el experimento, aceptamos que la esperanza (probabilidad) de que ocurra A queda expresada mediante el valor p(A).

5.2.- Experimento aleatorio: Función de probabilidad, Función aleatoria, y Función de distribución asociadas (a la f.p. y a la f.a.)

Definiciones:

Sea un experimento aleatorio en cual tenemos M y P(M). Definimos una función con valores reales

$$p: P(M) \dashrightarrow R$$

$$A \dashrightarrow p(A)$$

que cumpla las siguientes condiciones:

a) $0 \leq p(A) \leq 1$, para todo A

b) $p(M) = 1$

c) $p(\emptyset) = 0$

d) Para todo par A, B, cumple

$$p(A\cup B) = p(A)+ p(B)- p(A\cap B)$$

En estas condiciones la llamamos ‘Función de probabilidad’ (f.p.)

Consecuencias:

-Si $A\cap B = \emptyset$, entonces $p(A\cup B) = p(A) + p(B)$

-Si A’ es el complementario da A, $p(A') = 1-p(A)$, la cual se deduce de que $A\cup A' = M$,

Función aleatoria asociada (f.a.):

NOTA:
Este concepto coincide con el de 'variable aleatoria' dado en otro lugar.

Sea un experimento aleatorio con espacio muestral M y su álgebra de sucesos P(M).

Definimos en M una función real:

fa: M ----> R
m ----> fa(m)

que cumpla la condición:

-Para cada valor real r, el subconjunto Ar = {m;fa(m) $\leq$ r}

es un elemento de P(M).

La llamamos 'Función aleatoria' (f.a.)

NOTA: La expresión {m;fa(m) $\leq$ r} debe ser interpretada así:

'Conjunto de los elementos m tales que fa(m) $\leq$ r'

NOTA:
Por coherencia en la notación la designamos fa (función aleatoria), en lugar de X (variable aleatoria) como se ve en algunos textos. Del mismo modo seguiremos expresando Ar en lugar de [X $\leq$ r] como se ve en aquellos.

Función de Distribución asociada a la f.p. y la f.a. anteriores:

Es equivalente a la Frecuencia acumulada en Estadística.

Después de construir las funciones p y fa, definimos además la función real

$$F: R \dashrightarrow R$$
$$r \dashrightarrow F(r) = p(fa \leq r) = p(Ar)$$

Propiedades:

a) Por definición es $0 \leq F(r) <= 1$

b) Cuando r ---> $+\infty$, F(r) ---> 1

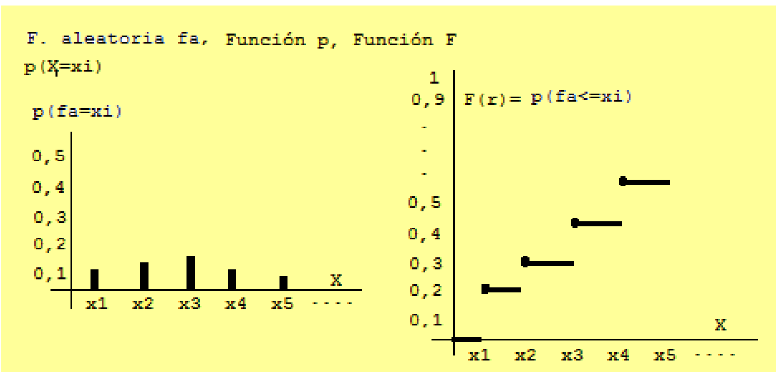
c) Cuando r ---> $-\infty$, F(r) ---> 0

d) Para todo par de valores r1, r2, si $r1 \leq r2$ entonces $F(r1) \leq F(r2)$

e) Si r ---> a, entonces F(r) ---> F(a)

f) Si $r1 \leq r2$, entonces $F(r2) – F(r1) = p(fa \leq r2) - p(fa \leq r1)$,

que expresamos también así: $p(r1 < fa \leq r2)$



Demostración:

Primero veamos que este último es un suceso de P(M). Tenemos

$\{m;r1 < fa(m) \leq r2\} = \{m;fa(m) \leq r1\}' \cap \{m;fa(m) \leq r2\}$

donde $\{m;fa(m) \leq r1\}'$ es el complementario de $\{m;fa(m) \leq r1\}$.

El resultado es un suceso ya que los dos componentes están en P(M).

Esto nos dice que $\{m;r1 < fa(m) \leq r2\}$ no tiene puntos comunes con $\{m;fa(m) \leq r1\}$

Por tanto tenemos

$\{m;fa(m) \leq r2\} = \{m;fa(m) \leq r1\}$ U $\{m;r1 < fa(m) \leq r2\}$

y por la afirmación anterior podemos concluir que

$p(fa \leq r2) = p(fa \leq r1) + p(r1 < fa \leq r2)$,

de donde $p(r1 < fa \leq r2) = p(fa \leq r2) - p(fa \leq r1)$

que es la que deseábamos demostrar.

F(r) es continua por la derecha, basta observar su gráfica. Por la izquierda experimenta 'saltos'.

5.3.- Estudio de Variable (o función) aleatoria discreta. Funciones de densidad y de Distribución asociadas

Seguimos en un experimento con espacio muestral M = $\{casos\ elementales\}$

Sea una función aleatoria (variable aleatoria) como se definió en apartados anteriores

$$fa: M \dashrightarrow R$$
$$m \dashrightarrow fa(m)$$

Decimos que fa es 'discreta' si el conjunto de valores Im(fa) es finito o infinito numerable.

Si Im(fa) = {r1, r2, ... , r_n, ...}, entonces, evidentemente, para r ≠ ri será

$$p(fa = r) = 0$$

Designamos pi = p(fa = ri)

Función de densidad asociada (f.d):

En realidad desempeña el mismo papel que la Función de probabilidad.

Después de definir la función aleatoria fa definimos una nueva función, asociada a fa, así:

fd: R ----> R
r ----> fd(r) = p(fa = r)

NOTA:
Observa que ahora la probabilidad de que tal ocurra viene dada por vía 'naturaleza del experimento'.

Por definición se cumple:

a) $0 \leq fd(r) \leq 1$

b) $fd(r) = \begin{cases} pi, & \text{si } r = ri \text{ para algún i} \\ 0, & \text{si } r \neq ri \text{ para todo i} \end{cases}$

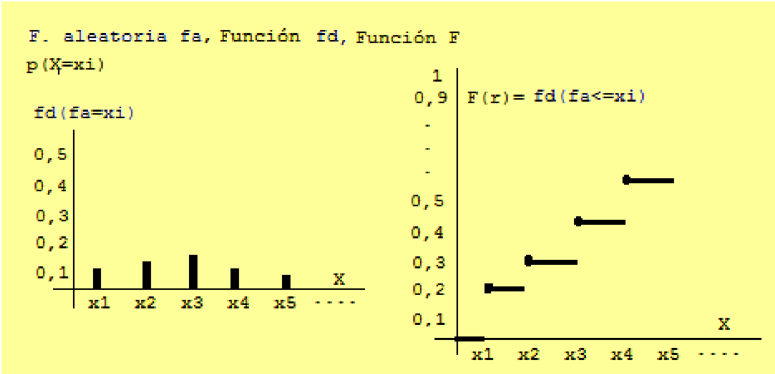
Función de Distribución asociada a fa:

F: R ----> R
r ---> F(r) = p(fa ≤ r)

Propiedades de fd y F:

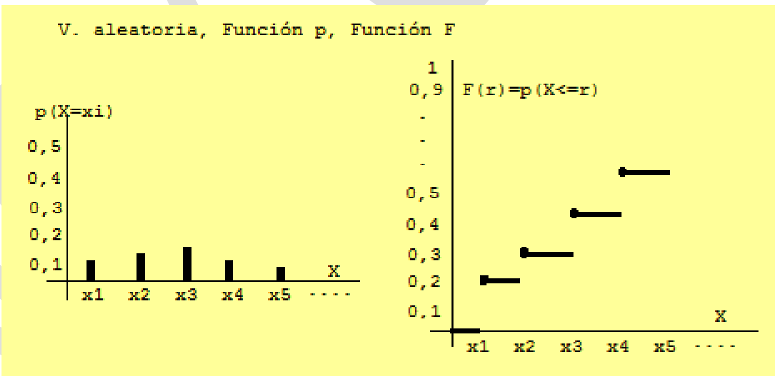
a) $\sum_{i=1,2,\ldots,n} f_d(ri) = 1$

b) $F(r) = \sum_{ri \leq r} f_d(ri)$



Ejemplo:

El alumno puede reconstruir el caso del lanzamiento de dos dados (o de tres dados), y tomar como v. aleatoria la suma de puntos.



Ejemplo:

La Dirección de una Empresa tiene constancia de que todos los años, al menos una vez, se reúne el Consejo de Dirección, pero que algunos años se ha reunido en más ocasiones, incluso hasta cinco veces.

Los siguientes datos muestran cuántas veces se ha reunido cada año

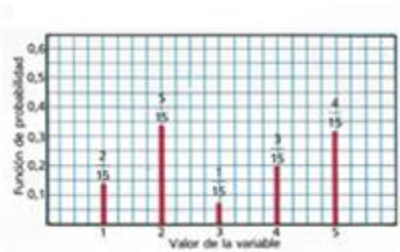
número veces	número de años
1	2
2	5
3	1
4	3
5	4

Gráficas:

Observa que en este ejemplo Im(X) = {1,2,3,4,5}

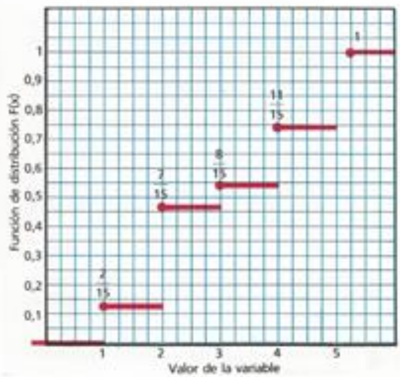
a) Función de probabilidad:

X = xi	$p_i = p(X = x_i)$
1	$\frac{2}{15}$
2	$\frac{5}{15}$
3	$\frac{1}{15}$
4	$\frac{3}{15}$
5	$\frac{4}{15}$



b) Función de distribución:

Valor de x	$F(x) = p(X \leq x)$
$x < 1$	0
$1 \leq x < 2$	$\frac{2}{15}$
$2 \leq x < 3$	$\frac{7}{15}$
$3 \leq x < 4$	$\frac{8}{15}$
$4 \leq x < 5$	$\frac{11}{15}$
$5 \leq x$	1



Para los próximos cálculos formamos la siguiente tabla:

$X = x_i$	p_i	$x_i \cdot p_i$	$x_i^2 \cdot p_i$
1	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{2}{15}$
2	$\frac{5}{15}$	$\frac{10}{15}$	$\frac{20}{15}$
3	$\frac{1}{15}$	$\frac{3}{15}$	$\frac{9}{15}$
4	$\frac{3}{15}$	$\frac{12}{15}$	$\frac{48}{15}$
5	$\frac{4}{15}$	$\frac{20}{15}$	$\frac{100}{15}$
	1	$\frac{47}{15}$	$\frac{179}{15}$

Se piden:

a) Función de Probabilidad y su representación
b) Función de distribución y representación
c) Media, Varianza y Desviación típica
d) Elegido un año al azar, Probabilidad de que se celebren más de dos sesiones

Sol.:

a) y b) las tenemos en el gráfico adjunto m = $\frac{47}{15}$ = 3,1333

$S^2 = \frac{179}{15} - 3{,}1333 = 2{,}1158$

S = 1,4546

p(X>2) = 1 – p(X ≤ 2) =

= 1 – [p(X=1) + p(X=2)] = 1 – [$\frac{2}{15} + \frac{5}{15}$]

= 1 – 0,4667 = 0,5333

Media:

$$m(X) = \sum_{i=1,2,\ldots,n} ri.p(X = ri)$$

Varianza (asociada a X):

Está definida por el valor

$$Var(X) = \sum_{i=1,2,\ldots,n} ri^2.p(X = ri)$$

Desviación Típica (asociada a fa):

$$S(X) = \sqrt{Var(X)}$$

5.4.- El caso de Distribución Binomial

Sea un experimento aleatorio en el cual sólo ocurren dos sucesos: A y A’. Supongamos conocidas las probabilidades

p = p(A), q = p(A’)

q = 1-p como es bien sabido.

Repetimos el experimento N veces, y consideramos el espacio muestral M de los casos posibles (Sucesos elementales):

M = {Todas las formas posibles como se puede presentar una n-tupla compuesta por A y A'}

Designo por m los elementos de M. Este elemento m es de la forma

m = AA...A'A'....
(n componentes)

Función aleatoria fa asociada al experimento:

Podemos definir, y definimos, la función aleatoria
fa: M ----> R
m ----> fa(m)
donde
fa(m) = número de veces que se presenta A
en el suceso elemental m

Im(fa) = {0,1,2,...,N}, evidentemente

Función de densidad:

Mediante aplicación de la combinatoria (Como vimos en otro lugar) y por la definición de Laplace, llegamos a que

$$p(fa = r) = C(n;r).p^r.q^{N-r}$$

Decimos que fa sigue una 'Ley binomial', y la llamamos 'Distribución binomial'. La representamos por

B(N,p), N, p son sus parámetros

NOTA:
En este caso se aprecia perfectamente la interpretación que debemos hacer de la expresión p(fa = r). Significa: 'Probabilidad de que la función fa tome el valor r'

Función de Distribución asociada a fa:

Su Función de distribución asociada queda así:

$$F(x) = \sum_{ri \leq x} p(fa = ri)$$

Esperanza Matemática (asociada a fa):

El concepto de 'esperanza matemática' es equivalente al de 'media aritmética' definido en Estadística.

Def.:
Es la Suma de los productos que resultan al tomar cada valor ri de la función aleatoria fa por la probabilidad $p(fa = ri)$ de que fa tome este valor.

La fórmula general es

$$E(fa) = \sum_{i=1,2,\ldots,N} ri.\, p(fa = ri)$$

En el caso particular de la Distribución binomial queda

$$E(fa) = \sum_{i=1,2,\ldots,N} ri.\, C(N; ri).\, p^{ri}.\, q^{N-ri}$$

Si continuamos desarrollando el sumatorio de la expresión anterior tengo

$\sum_{i=1,..,N} ri.\frac{N!}{ri!.(N-ri)!}.\,p^{ri}.\,q^{N-ri}$ = (simplifico el factor ri, separo factores N y p)

$= \sum_{i=1,\ldots,N} \frac{(N-1)!}{(r_i-1)!.(N-ri)!}.\,N.\,p.\,p^{ri-1}.\,q^{N-ri} =$

(expreso: N-ri = (N-1)-(ri-1)

= N.p.$\sum_{i=1,\ldots,N} \frac{(N-1)!}{(ri-1)!.((N-1)-(ri-1))!}.p^{ri-1}.q^{(N-1)-(ri-1)}$ =

(la fracción es $C_{(N-1;ri-1)}$)

= N.p.$\sum_{i=1,2,\ldots,N} C(N-1;ri-1).p^{ri-1}.q^{(N-1)-(ri-1)}$ =

(el sumatorio es el desarrollo de $(p+q)^{N-1}$)

= N.p.$(p+q)^{N-1}$

(puesto que p+q = 1, queda)

E(fa) = N.p

(Habitualmente n.p, n en lugar de N)

Varianza (asociada a fa):

Es el valor definido por

Var(fa) = $\sum_{i=1,2,\ldots,n} ri^2.p(fa=ri)$

Por cálculos y transformaciones análogas a las realizadas para la Esperanza matemática llegamos a que

Var(fa) = n.p.q

Desviación Típica (asociada a fa):

S(fa) = $\sqrt{n.p.q}$

5.5.- Distribución de Poisson

En el caso de una distribución binomial pero tratándose de ‘sucesos raros’, lo cual ocurre cuando n es grande y n.p es muy pequeño (se estima así cuando n.p < 15), la Distribución binomial queda muy sesgada, y además su aproximación mediante una distribución normal no es apropiada.

El matemático Siméon-Denis Poisson estudió con detenimiento estas situaciones determinando a dónde nos lleva la situación límite cuando n -> ∞ y p -> 0 simultáneamente, y cumpliéndose además que n.p -> constante.

Llegó a la conclusión de que la función de densidad

$$p(fa = r) = C(n;r).p^r.q^{n-r}$$

de la binomial se 'reconvierte' en la siguiente, llamada 'distribución de Poisson':

$$P(fa = r) = e^{-\lambda}.\frac{\lambda^r}{r!} \; ,$$

donde $\lambda = \lim_{n\to\infty, p\to 0}(n.p)$

Los valores 'media' y 'varianza' de esta distribución de Poisson coinciden con el valor λ.

Aclaraiones:

a) En realidad, la fórmula de Poisson es una aproximación de la binomial en los casos 'raros', tomando para la media el valor λ.

b) La experiencia ha demostrado que la mejor aproximación se consigue cuando:

   $n \geq 50$, $p \leq 0{,}1$ (o bien que $n.p \leq 5$)

c) Conociendo λ = media, mediante la fórmula de Poisson podemos calcular el valor (una buena aproximación) de la probabilidad como si aplicásemos la binomial. Esta es la razón por la que λ es el 'parámetro' que determina la distribución de Poisson.

d) Esta distribución de Poisson es de fundamental importancia en el estudio de fenómenos: Físicos, Biológicos, Astronómicos, Químicos, …

e) Ejemplos de Variables aleatorias para las que se adaptan a una distribución de Poisson son:

Número de llamadas que entran en una central telefónica en ‘cierto’ período de tiempo; Cantidad de plancton por ‘unidad’ de volumen de agua; Recuento en una placa de laboratorio del número de … ; Número de partículas radiactivas emitidas en unidad de tiempo; etc, etc, …

Existen tablas con valores tabulados para estas distribuciones: Binomial y De Poisson.

Para la binomial, la que conocemos, para valores de n del 2 al 10, y, fijado el valor n, para valores de r del 0 al n, y valores de p del 0,01 al 0,50, escalando 0,05 centésimas.

Para la de Poisson, la que conocemos, para valores de λ del 0,10 al 9,00, y valores de r entre 0 y 24.

5.6.- Ajuste de una serie de datos a una Distribución binomial

Sea una serie (lista) de N datos ordenados de menor a mayor, donde puede haberlos repetidos,

$$x_1, x_2, x_3, \ldots , x_N$$

Extraemos la lista de los k valores distintos yj y su frecuencia absolutas f(yj)

$$y_1, \quad y_2, \quad \ldots, \quad y_k$$

$$f(y1), f(y2), \ldots , f(yk), \qquad \sum_j f(yj) = N$$

Escribiremos f_j para las frecuencias.

Obtengo la media aritmética $m_x = \frac{\sum_j yj.fj}{N}$ de la serie de datos, y tomo el valor de p

$$p = \frac{mx}{k} , \qquad q = 1\text{-}p$$

Tomando n = N, aplico la fórmula de la Distribución binomial B(n,p), variable aleatoria Y tal que Im(Y) = {y1,y2,…, y_k}, y obtengo una nueva lista de valores probabilísticos, dados por

p(i) = p(Y = yi), i = 0, 1, 2, 3, …,k

(k = nº valores distintos)

Recuerda que en una variable X discreta y binomial ocurren sólo dos sucesos: A, A' , y que

p(X= h) = "Probabilidad de que el suceso A ocurra h veces"

$$p(X = h) = C(n;h).p^h.q^{n-h}$$

Tomando como dijimos $p = \frac{mx}{k}$, y aplicando esta igualdad obtengo una lista de valores p(i)

p(i)	p0	p1	p2	p3	….	pk

Recuerda que la probabilidad pi representa el mismo papel que la frecuencia relativa. En este caso

$f_{rj} = \frac{fj}{k}$, y por tanto $f_j = k.f_{rj}$

Esto nos exige obtener la siguiente lista y compararla con la lista de frecuencia absolutas:

k.p(i)	k.p0	k.p1	k.p2	k.p3	….	k.pk
	0	f1	f2	f3	….	fk

Contrastando la lista de los valores k.p(i) con la lista de los valores f_j podemos decidir si aquella serie de datos se 'adapta' o no,(admitiendo alguna desviación) a la distribución binomial B(n,p).

En caso afirmativo podremos aplicar la fórmula

$$p(X= h) = C(n;h).p^{h}.q^{n-h}$$

para obtener las frecuencias fj.

La tesis será aceptada si el máximo de las diferencias abs(fi-k.pi) es menor que una cota prefijada, que suele ser un valor 'pequeño', como por ejemplo

$$máx\{|f_i - k.p_i|,\ i = 0,1,\dots,n\} < 0,5$$

5.7.- Distribución Hipergeométrica

Sea una población finita de tamaño N. Supongamos que sus individuos pertenecen a dos clase (respecto del carácter objeto de estudio): Tipo A, tipo A'.

Supongamos que se conoce que al tomar al azar un elemento de las probabilidades son:

$$p = p(A),\ q = p(A') = 1\text{-}p$$

Si Na = número de individuos del tipo A, entonces

$$p = \frac{N_a}{N}, \quad q = \frac{N-N_a}{N}$$

y también Na = p.N, Na' = q.N

Tomamos al azar una muestra con n individuos, tomándolos sucesivamente sin reposición (equivalente a tomarlos simultáneamente), y nos preguntamos por la probabilidad de que el Tipo A figure k veces.

Dentro de los Na elementos del tipo A, podemos seleccionar k elementos de tantas formas distintas como indica el valor

$$C(Na;k) = C(p.N;k)$$

mientras que los del tipo A’ pueden presentarse n-k veces de tantas formas distintas como

$$C(Na';n-k) = C(q.N;n-k)$$

Por consiguiente, al tomar n individuos, una muestra con k del tipo A y (n-k) del tipo A’:

AA..(k ..AA’A’..(n-k ..A’

puede presentarse de tantas formas distintas como indica el valor

$$C(p.N;k).C(q.N;n-k)$$

El número de muestras distintas posibles es

$$C(N:n)$$

Una vez más tenemos aquí una ‘Variable aleatoria’

X: M ----> R
m ---> k = número de veces que se presenta el suceso A

Lógicamente, el suceso A’ se presentará n-k veces

Por tanto

$$p(X= k) = \frac{C(p.N; k).C(q.N; n-k)}{C(N;n)}$$

Ejemplo: Vamos a tratar el ejemplo anterior de la Binomial como si se tratase de un Grupo social ‘limitado’ con N = 500 individuos.

Calculamos las mismas probabilidades que allí.

En primer lugar p.N = 0,4.500 = 200, q.N = 300,

$$C(500;10) = \frac{500.499.498.497.496.495.494.493.492.491}{10.9.8.7.6.5.4.3.2.1} =$$

$$= \frac{10.499....491}{9.8.7.6.4.3.2} = \frac{5.499.249.497.248.165.247.493.246.491}{9.4.7.6} =$$

$$= 1{,}2290529.10^{20}$$

$$C(200;4)= \frac{200.199.198.197}{4.3.2} = \frac{50.199.99.197}{3} = 64684950$$

$$C(300;6)= \frac{300.299.298.297.296.295}{6.5.4.3.2} = \frac{150.299.149.297.148.295}{6.5.3} =$$

$$= 9{,}6282285.10^{11}$$

Por tanto
$p(X=4) = 64684950.9{,}6282285 : 1{,}2290529).10^{-9} =$

$= 5{,}0673285.10^{8}.10^{-9} = 0{,}5067$

NOTA:
Como podemos ver hay una gran diferencia según que lo tratemos por uno u otro procedimiento.

Si el colectivo es ‘reducido’, y consideramos que 500 individuos lo es, será más fiable el resultado obtenido por la Hipergeométrica, mientras que si es un colectivo ‘ilimitado’ o difícil de determinar su número, será aplicable la Binomial.

$$$$oOo$$$$

Ejemplos/Ejercicios:

1.-
En un grupo social (muy grande) se sabe que el 40 % son votantes del PP. Tomamos al azar una muestra de 10 personas.

Analiza si se trata de una distribución binomial o no. En caso afirmativo calcula las siguientes probabilidades.

Al seleccionar 'personas', si los hacemos una a una (experimento aleatorio simple) no tiene sentido la reposición, como no tiene sentido que figure una persona dos veces. Por tanto No es d. binomial.

Si hiciésemos los cálculos tratándola como distribución binomial (d.b.) obtendríamos los siguientes resultados, muy diferentes a los obtenidos tratándola como distribución Hipergeométrica (d.h.) (Ver Apéndice)

Tratándola como d.b.:

a) Exactamente 4 votan al PP
b) Menos de 4 votan al PP
c) Al menos 4 votan al PP

Sol.: Al tomar al azar una persona del grupo social, una de dos, vota o no vota al PP, y p = p(A)=40/100 = 0,4, es constante. Podemos suponer que el grupo social es suficientemente amplio como para que p no varíe al realizar la siguiente extracción.

Es una B(10; 0,4)

$p(X= 4) = C(10;4).(0,4)^4.(0,6)^6 = 210.0,0256.0,0467 = 0,2511$

$p(X < 4) = p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3)$

$p(X= 0) = C(10;0).(0,4)^0.(0,6)^{10} = 1.1.0,0060 = 0,0060$

$p(X= 1)= C(10;1).(0,4)^1.(0,6)^9 = 10.0,4.0,0101 = 0,0404$

$p(X= 2) = C(10;2).(0,4)^2.(0,6)^8 = 45.0,16.0,0168 = 0,1210$

$p(X= 3) = C(10;3).(0,4)^3.(0,6)^7 = 120.0,0640.0,0280 = 0,2150$

Por tanto $p(X< 4) = 0,3824$

$p(X \geq 4) = 1- p(X<4) = 0,6176$

El alumno terminará los cálculos y hará las comprobaciones pertinentes.

2.-
En un colectivo que padece de sida, se sabe que la probabilidad de mejorar por un tratamiento contra el sida es p = 0,2. Tomamos 6 personas al azar y se les aplica el tratamiento. Se supone que para cualquiera de los pacientes si no mejora es que empeora. Calcula:

a) Probabilidad de que los 6 mejoren.
b) Probabilidad de que al menos dos empeoren

Sol.: La v. aleatoria es X = "número de pacientes que mejoran"

Por el mismo razonamiento que hicimos en el anterior diríamos que no satisface d. Binomial.

Pero aplicamos un razonamiento análogo basado en la sucesión de pruebas independientes.

Elegido un paciente al azar, la probabilidad de mejora es p

Tomamos el siguiente, y la probabilidad de que los dos mejoren es $p.p = p^2$. La p. de que sólo uno mejore es: $p.q + q.p = C(2;1).p.q$

Tomo el siguiente; la p. de que los tres mejoren es p^3. La p. de que sólo dos mejoren es $p^2.q +p.q.p +q.p^2 = C(3;2).p^2.q$.

La p. de que sólo uno mejore es: $p.q^2 + q.p.q + q^2.p = C(3;1).p.q^2$

Por el mismo razonamiento, si tomamos hasta 6 pacientes tenemos:

$$p(X = 3) = C(6;3).p^3.q^3$$

$$p(X = 6) = C(6;6).p^6.q^0$$

Que al menos dos emperoren equivale a que a lo más 4 mejoran:

$$p(X \leq 4) = 1 - p(X>4) = 1 - [p(X=5) + p(X=6)] =$$

$$= 1 - [C(6;5).p^5.q + C(6;6).p^6] = 1 - [6.0{,}2^5.0{,}8 + 1.0{,}2^6] =$$

$$= 1 - [0{,}001536 + 0{,}000064] = 1 - 0{,}0016 = 0{,}9984$$

(p. de que al menos dos no mejoran).

Tema 6

Distribuciones continuas

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

6.1.- Función aleatoria continua.

Tenemos un experimento y su espacio muestral M como se ha explicado en el Tema 5. Allí definimos el concepto de 'función aleatoria fa', que es equivalente al concepto de variable aleatoria también explicado.

Def.:
Dada la función aleatoria

$$\begin{aligned} fa:\ & M \dashrightarrow R \\ & m \dashrightarrow fa(m) \end{aligned}$$

si su imagen Im(fa) es no numerable, por ejemplo un subconjunto denso de los reales R, decimos que es 'función aleatoria continua'. Este es el caso cuando Im(fa) = [a, b]

Supondremos que Im(fa) = [a, b], un intervalo de R, (y no la unión de intervalos disjuntos).

Ejemplo:
Es el caso cuando tallamos un grupo de jóvenes, o cuando anotamos el peso de los recién nacidos, o el peso de una persona cualquiera. Decimos que es continua porque entre dos valores cualesquiera, por ejemplo r1 = 3'3420, r2 = 3'3410, siempre cabe la posibilidad de presentarse otro valor intermedio: 3'3415.

6.2.- Funciones de densidad y de Distribución asociadas a fa

La función de densidad va a jugar el mismo papel que el concepto de probabilidad. Lo que sucede es que en el caso de f.a. continua no viene dado de forma intrínseca (como sí ocurre en el caso de f.a. discreta) el valor de probabilidad. Hemos de utilizar otros instrumentos que nos proporcionen un valor equivalente. Como sigue.

En el caso de función aleatoria continua no será fácil obtener la ‘Ley’ que determine la ‘función de probabilidad’.

Hemos de suponer también que para todo r de Im(fa) se cumple que $fa^{-1}(r)$ es un subconjunto de M, y por tanto un elemento de P(M).

Llamamos $Ar = fa^{-1}(r)$ (Antiimagen de r)

Supongamos que se ha podido determinar una función fd

fd: R ----> R
r ---> fd(r) = p(Ar), donde $Ar = fa^{-1}(r)$

Decimos que fd es una ‘Función de densidad’ asociada a fa. Si además cumple las siguientes condiciones:

a) $fd(t) \geq 0$, y $fd(t) \leq 1$, para todo t de R
b) $\int_{-\infty}^{+\infty} fd(t).dt = 1$
c) fd(t) es continua en R
d) $p(a < fa \leq b) = \int_{a}^{b} fd(t).dt$, $a < b$

Función de Distribución:

Si hemos podido determinar una f.d., utilizando la integración que sustituirá al sumatorio, podemos definir la siguiente función real, cuyos valores representan el mismo papel que el de la frecuencia acumulada.

Def.:

F: R ----> R
x ----> $F(x) = \int_{-\infty}^{x} fd(t).dt$

Si esto es posible, decimos que la función aleatoria fa es ‘Absolutamente continua’.

NOTA: Por el Teorema fundamental del cálculo sabemos que, siendo fd(t) continua, se cumplirá

F'(x) = fd(x), ó bien F'(r) = fd(r)

También, por definición de la Integral Definida sabemos que

$$\int_{-\infty}^{b} fd(t).dt - \int_{-\infty}^{a} fd(t).dt = \int_{a}^{b} fd(t).dt$$

y por tanto, por la propiedad d) de fd tenemos

p(a < fa ≤ b) = F(b) – F(a)

OBSERVA:
Puesto que $\int_{a}^{a} fd(t).dt = 0$, tenemos p(fa = r) = 0, para todo valor real r. Es decir, la probabilidad de que fa tome un valor 'puntual' concreto es cero. Tendríamos que calcular

p(r-e < fa < r+e) = $\int_{r-e}^{r+e} fd(t).dt$

donde e > 0 es un valor 'pequeño'.

Esperanza Matemática (asociada a fa):

Este concepto es equivalente a la 'Media aritmética' en Estadística.

El valor de la frecuencia relativa de xi es sustituido por el valor fd(xi)

En el caso de una distribución continua, la suma $\sum_{i=1,\dots,n} xi.p(fa = xi)$ que utilizamos en el caso de variable discreta, es sustituida por la integral, de modo que ahora tenemos

E(fa) = $\int_{-\infty}^{+\infty} x.fd(x).dx$

Varianza (asociada a fa):

Por las mismas razones expuestas antes definimos

$$S^2 = \int_{-\infty}^{+\infty} \left(x - E(fa)\right)^2 . fd(x).\, dx$$

$S = \sqrt{S^2}$ es la Desviación típica

6.3.- La Distribución Normal

La llamada Distribución normal tiene sus raíces en un tipo de variable aleatoria continua X (volvemos a esta notación: X por fa) que se aplica constantemente en las Ciencias Aplicadas.

Sus funciones de densidad y de distribución asociadas han sido estudiadas con detalle en las citadas ciencias, por métodos experimentales.

INTRODUCCIÓN

Por estudios realizados durante siglos con resultados reconocidos por estudiosos con renombre (De Moivre, Gauss, Laplace, ...), se ha aceptado que en la mayoría de los Fenómenos Naturales, los sucesos ocurren aleatoriamente según una ‘Ley’ a la que bautizaron con el nombre de ‘Distribución Normal’ (con más precisión: Distribución continua Normal).

Dependiendo del tipo de fenómeno natural, y mediante multitud de observaciones, se ha podido determinar para muchos de esto fenómenos el valor de la Media m y de la Desviación típica s, que determinarán, como veremos, la expresión de su función de densidad fd, siendo estos valores los parámetros para ese caso de fenómeno.

Eminentes estudiosos como los mencionados más arriba consiguieron establecer, para este tipo de fenómenos y conocidas m y s, la forma que toma su Función de densidad, y, fijada la expresión de ésta, su Función de distribución.

El siguiente paso consistió en 'tabular' construyendo la tabla correspondiente, el valor de la Función de distribución para una lista de valores de su variable x, dando como resultado las llamadas Tablas de la Normal.

En este apartado nos interesa principalmente el cálculo de probabilidades mediante el uso de la Tabla de la Normal estándar (también llamada tipificada).

Continuamos exponiendo la expresión de cada una de las referidas funciones.

Suponiendo conocidos los valores m, s, se ha establecido que la Función de densidad fd(t) toma la forma:

$$fd(t) = \frac{1}{s\,.\,\sqrt{2.pi}}.\,e^{\frac{-1}{2}.\frac{(t-m)^2}{s^2}}$$

Si hacemos $c = \frac{1}{\sqrt{2.pi.s^2}}$, d = -1/2, entonces queda

$$fd(t) = c.e^{d.(\frac{t-m}{s})^2} \qquad (e = 2{,}71828...)$$

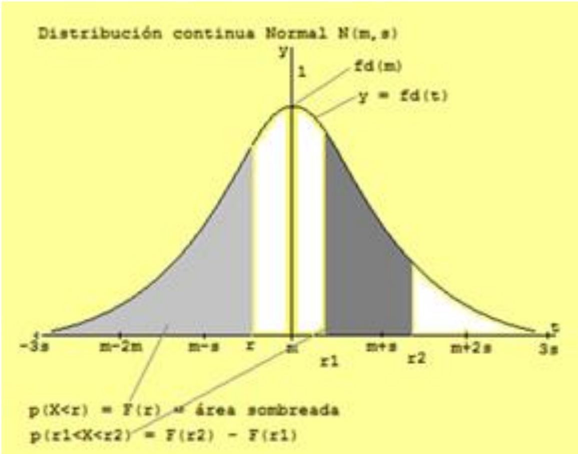
Cumple las condiciones de función de densidad:

a) fd(t) ≥ 0, y fd(t) ≤ 1, para todo t de R

b) $\int_{-\infty}^{+\infty} fd(t).dt = 1$

c) fd(t) es continua en R

d) $p(a < X \leq b) = \int_a^b fd(t).dt$, a < b donde p(a < X ≤ b) significa: Probabilidad de que X: M ---> R tome valores a< X(m) ≤ b



Su gráfica tiene forma de campana (Campana de Gauss), simétrica respecto a la recta x = m, por encima del eje ox y tal que se aproxima a éste eje ox cuando t --> +∞ y cuando t --> - ∞

Función de Distribución:

Su función de distribución asociada la definimos como en cualquier función aleatoria continua

$$F(r) = \int_{-\infty}^{r} fd(t).dt \;=c.\int_{-\infty}^{r} e^{d.(\frac{t-m}{s})^2}.dt$$

F(r) nos da el valor del área bajo la curva de fd(t) desde - ∞ hasta t = r

Importa hacer notar que

p(X ≤ r) = F(r) = Area bajo la gráfica de fd

Obtenidas fd y F como acabamos de ver, decimos que tenemos "Una Distribución Normal" (continua), y la designamos por

N(m;s),

donde m y s son los valores concretos de la Media y la Desviación típica, determinados previamente por la 'experiencia' u obtenidos mediante pruebas empíricas actuales.

6.4.- Distribución Normal estándar

En el apartado anterior, 6.3, hemos tratado el caso general de Distribución normal.

Cuando m = 0 y s = 1, tenemos un caso concreto, N(0;1). Es la llamada Distribución Normal estándar, y utilizada como referencia constante en el uso de la Tabla de la normal. Hasta el punto que, como veremos, mediante un cambio de variable llamado 'tipificación', cualquier otro caso en el que m y s no tomen los valores anteriores (0 y 1, respectivamente) se consigue transformarla en otra equivalente cuyos parámetros toman el valor 0, 1, y por tanto adecuada para utilizar las citada tabla de la normal estándar.

Su variable aleatoria subyacente es representada por Z.

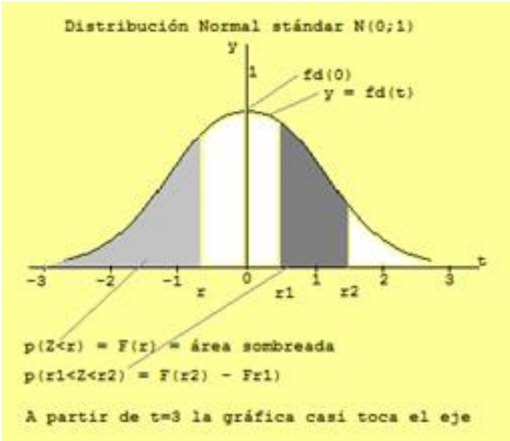
En este caso la función de densidad fd(t) queda así:

$$\text{fd(t)} = \frac{1}{\sqrt{2.pi}} . e^{-\frac{1}{2}.t^2} \quad , \quad \text{t recorriendo R}$$

y la de Distribución: $\text{F(x)} = \frac{1}{\sqrt{2.pi}} . \int_{-\infty}^{x} e^{-\frac{1}{2}.t^2} . dt$

Ahora

p(Z ≤ x) = F(x)



NOTA:
Es muy importante el hecho que cualquier variable aleatoria X continua, con su distribución normal N(m;s) asociada, puede ser transformada mediante un cambio de variable:

X --- > Z (proceso de tipificación)

cuya función de distribución asociada es estándar:
N(0;1).

Esta posibilidad permite utilizar, para cualquier caso N(m; s), la Tabla de valores obtenida mediante la tabulación de N(0; 1).

6.5.- Tipificación de una Distribución normal N(m;s)

Sea una distribución N(m; s); m y s son sus parámetros.

Deseamos transformarla en otra distribución normal estándar N(0,1) equivalente.

Gráficamente significa trasladar la gráfica de N(m; s) de modo que su eje de simetría coincida con el eje oy. Esto se consigue mediante el cambio de variable z = x- m, que significa hacer la citada traslación.

Además, conseguir que el valor s se convierta en el valor 1 equivale a ‘contraerla’ o ‘dilatarla’. Esto se consigue haciendo un cambio de escala dividiendo por el valor s.

Unificando los dos cambios nos lleva a realizar el cambio

$$z = \frac{X-m}{s}$$

Es el momento de pasar a practicar el uso de la Tabla obtenida para la distribución normal standar N(0;1), también llamada ‘Tipificada’.

Representaremos por Z la variable aleatoria cuando se trate de la normal tipificada N(0;1), y por X en otro caso N(m;s) cualquiera.

6.6.- Uso de la Tabla de la Normal tipificada

Lo veremos con ejemplos que traten todas las posibilidades situaciones, o al menos lo pretendemos.

Ejemplos:

Sea Z una variable aleatoria que sigue una distribución N(0,1). Calculamos las siguientes probabilidades. Hacemos notar que la Tabla está construida sólo para valores positivos.

a) $p(Z \leq 1{,}36)$

En la primer columna localizo 1,3, y en la fila-cabecera 0,06, en la intersección tengo el valor

$$p(Z \leq 1{,}36) = 0{,}9131.$$

b) $p(Z \leq -2{,}54)$

El valor de $p(Z<a) = F(a)$ coincide con el área limita por la gráfica y la recta $x=a$ desde $-\infty$ hasta $x = a$. Teniendo en cuenta que el área total vale 1, y la simetría de la gráfica, se cumple que $F(-2{,}54) = 1 - F(2{,}54)$. Por tanto

$p(z \leq -2{,}54) = 1- p(z<2{,}54) = 1- 0{,}9945 = 0{,}0055$

c) $p(Z > 2{,}54)$

La tabla registra valores de $F(x)$, los cuales vienen dados por el área desde $-\infty$ hasta x. Este valor $F(x)$ nos da el valor de $p(Z<x)$ y de $p(Z \leq x)$. Recuerda que $p(Z=x) = 0$.

Cuando tenemos $p(Z > x)$ ó $p(Z \geq x)$ tenemos que recurrir a la simetría de la gráfica.

$p(Z >2{,}54) = 1-p(Z<= 2{,}54) =1-0{,}9945 = 0{,}0055$

d) $p(0{,}25 < Z <= 1{,}35)$

Atendiendo a la gráfica y lo explicado en los casos anteriores tenemos

$p(0{,}25< Z \leq 1{,}35) = p(Z \leq 1{,}35) - p(Z< 0{,}25) = 0{,}9115 - 0{,}5987 = 0{,}3128$

e) $p(-3{,}25 \leq Z < -1{,}50)$

Atendiendo a las áreas tenemos

$p(-3{,}25 \leq Z < -1{,}50) = p(Z <-1{,}50) - p(Z <-3{,}25) =$

$= (1 - p(Z< 1{,}50) - (1 - p(Z< 3{,}25) = (1 - 0{,}9332) - (1 - 0{,}9994) = 0{,}0668 - 0{,}0006 = 0{,}0662$

f) $p(-3{,}25 < Z \leq 1{,}50)$

Atendiendo a las áreas tenemos

$p(-3,25 < Z \leq 1,50) = p(Z \leq 1,50) - (1 - p(Z<3,25))=$

$0,9332 - (1 - 0,9994) = 0,9332 - 0,0006 = 0,9326$

Casos de N(m,s) no tipificada:

Sea el caso N(2,3), por ejemplo.

Cambio: $Z = \frac{X-2}{3}$

$p(X< 2,25) = p(Z < \frac{2,25-2}{3}) = p(Z< 0,0833) = 0,5319$

De forma análoga en cualquier otro caso.

6.7.- Aproximación de una Binomial mediante una Distribución Normal

Tenemos una distribución B(n,p). Sabemos que

Media $m = n.p$

Varianza = $n.p.q$

Desviación típica $s = \sqrt{n.p.q}$

Tomando estos valores m y s, aproximamos la binomial B(n,p) mediante una Normal N(m,s), lo cual es ‘aceptable’ si se cumple alguna de las siguientes condiciones:

Si [np > 5 y nq > 5] --> La aproximación es muy buena.

Si [np > 3 y nq > 3] --> Aproximación buena.

6.8.- Ajuste de una serie de datos a una Distribución normal

Sea una serie (lista) de N datos ordenados de menor a mayor, donde pueden presentarse repetidos,

x1, x2, x3, … , xN

Extraemos la lista de valores distintos yj y su frecuencia f(yj):

y1, y2, …., yk

f(y1), f(y2), … , f(yk), $\sum_j f(yj)$ = N

Escribiremos fj para las frecuencias. Tengo K valores distintos.

Obtengo la media aritmética m = $\frac{\sum_j yj.fj}{K}$, la varianza Var = $\frac{\sum_j (yj.fj-m)^2}{K}$, y la desviación típica s = $\sqrt{Var}$

Tomo la Distribución normal N(m,s).

Construyo los intervalos $[y_j, y_{j+1}]$, j = 0, 1, 2, …, K-1, (tomando $y_0 = 0$), y aplicando N(m; s) obtengo la lista de valores p(j), j = 0, 2, 3, …, K-1. A partir de esta construyo la lista qi = $\frac{f_j}{K}$ y la lista de las diferencias d_j :

p_j	p0	p1	p2	p3	…	p_{k-1}
$\frac{f_j}{K}$	q0	q1	q2	q3	…	q_{k-1}
$\frac{f_j}{K} - p_j$	d0	d1	d2	d3	…	d_{k-1}

Contrastamos las dos listas de valores y_j y p_j, mediante el valor absoluto de las diferencias d_j.

Calculo el valor $d = \text{máx.}\{ d_j ; j = 0, 1, \ldots, K-1\}$.

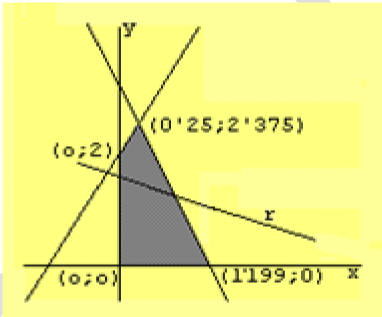
Si este valor cumple la condición impuesta como que $/d/ < \varepsilon$, siendo $\varepsilon > 0$ un valor fijado 'pequeño', podemos aceptar el ajuste como una aproximación válida.

$$$oOo$$$

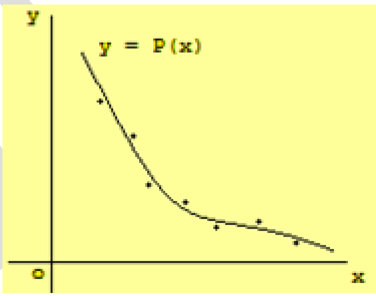
Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

VOLUMEN 11 Parte II

PROGRAMACIÓN Lineal



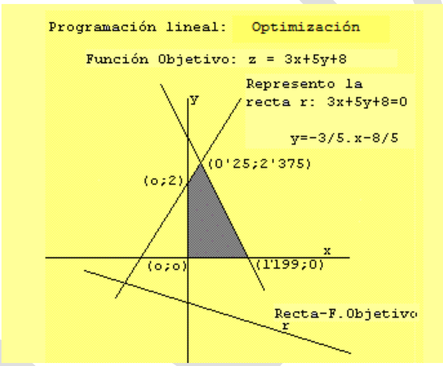
INTERPOLACIÓN



Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

Tema 1

PROGRAMACIÓN LINEAL



Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

1.1.- **Introducción:** El Problema

Si tenemos, por ejemplo, la función z = F(x1,x2,x3) que representa la productividad de una empresa, donde x1, x2, x3 son los parámetros o variables básicos de los que depende su productividad, un buen Directivo buscará la manera de que esta productividad sea máxima.

En otros casos puede resultar más beneficioso para el buen funcionamiento de la Empresa o "negocio" conseguir que los Costes sean mínimos, en cuyo caso aquella función representaría los costes.

La clave: Productividad máxima con costes mínimos.

Las variables x1, x2, x3 representarán, por ejemplo, "los gastos en personal", "el coste de materias primas", "el gasto en administración", etc., etc.

Evidentemente, dependiendo del tipo de Empresa el problema a resolver puede exigir mayor número de variables.

Normalmente, casi con toda seguridad, el rango de valores que puedan tomar cada uno de los parámetros xi será finito. Más aún, los valores de estos parámetros estarán "condicionados" entre sí. Por ejemplo, a más personal de producción tendremos mayor coste en Administración; o posiblemente, a más gasto en maquinaria menos coste en personal.

Por tanto, parece muy claro que los parámetros quedan "ligados" entre sí. Lo que llamamos "condiciones" y/o "restricciones".
Aquí hablaremos de restricciones.

Matemáticamente estas restricciones vienen expresadas mediante "ecuaciones" o "inecuaciones" (desigualdades) donde intervienen las variables, tales como:

$$\begin{cases} a11.x1 + a12.x2 + a13.x3 \ < b1 \\ a21.x1 + a22.x2 + a23.x3 > b2 \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. \end{cases}$$

El “Problema planteado” en una Empresa concreta puede ser en extremo difícil de resolver.

En resumen se trata de un Problema de “Optimización”, bien de maximizar los beneficios o de minimizar las pérdidas.

También hemos de decir que este tipo de Porblemas no es exclusivo del mundo de la Empresa.

En Matemáticas se han desarrollado también Métodos de Optimización “No Lineales”. No viene al caso tratarlos en el presente trabajo.

El objetivo del presente trabajo es presentar al Alumno material suficiente que le ayude a entender este tipo de problemas y adquirir la destreza y práctica suficientes para resolver “casos básicos”, planteados matemáticamente.

Esto quiere decir que, quizás la tarea más ardua como es traducir el problema real a “términos matemáticos” se nos dará ya resuelta. Las herramientas y datos necesarios para conseguir la referida traslación al modelo Matemático sólo puede proporcionarlos la propia Empresa a través de su Departamento Económico-Comercial.

1.2.- Programación lineal

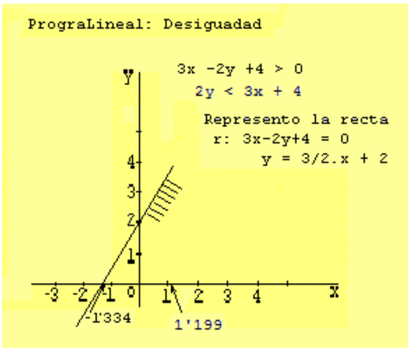
Representamos y resolveremos modelos matemáticos en dos variables, ya que estos modelos en dos variables permiten representación en el plano 0xy, y por tanto visualizar fácilmente su significado.

Las relaciones entre las variables y las restricciones vendrán dadas por inecuaciones (desigualdades). Por ejemplo, una desigualdad como

$$a.x + b.y + c < 0$$

relaciona las variables x, y, y expresa una restricción de esta relación. Esta desigualdad representa una región del Plano (un Semiplano).

Ejemplo: $3x - 2y + 4 > 0$



Despejando y tengo $y < 3/2.x + 2$

Represento la igualdad (una recta)

$$r: \; y = 3/2.x + 2$$

Pues bien, aquella desigualdad (1) representa o determina el semiplano de los puntos (x’,y’) en los cuales se cumple

$$y' < 3/2.x' + 2,$$

Esto significa que el punto con coordenadas (x’,y’) está “por debajo” de la recta r, en la vertical que pasa por x = x’. De hecho el semiplano que determina es el de todos los puntos del plano que están por debajo de r.

Un semiplano es una “región” del plano que es abierta cuando tenemos desigualdad estricta (<), o es “cerrada” si tenemos <=. En aquel caso quedan excluidos los puntos de la recta. En este caso quedan incluidos.

Consideremos ahora el siguiente sistema de inecuaciones:

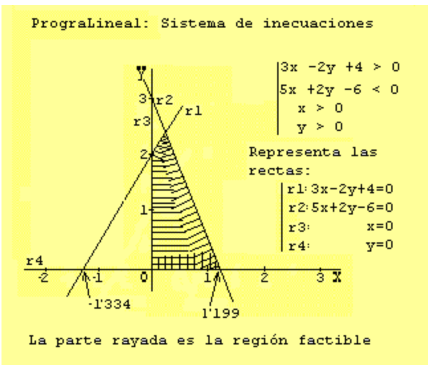
$$\begin{cases} 3x-2y+4>0 \\ 5x+2y-6<0 \\ x>0 \\ y>0 \end{cases} \qquad (1)$$

Despejo y de cada una:

y < 3/2.x + 2 (puntos por debajo de r1)
y < -5/2.x + 3 (puntos por debajo de r2)

Las convierto en igualdades y represento las rectas:

r1: y = 3/2.x + 2
r2: y = -5/2.x + 3
r3: x = 0
r4: y = 0



Teniendo en cuenta que los puntos P(x’,y’) que cumplan el sistema (3) deben cumplir las condiciones del sistema (3), queda así delimitada la región de los “puntos factibles”, representada por la zona rayada en fig. 2.

Puntos factibles:

Aquellos puntos P(x’,y’) cuyas coordenadas cumplen el sistema (1).

Región factible:

Conjunto de todos los puntos factibles (Región rayada en fig. 2).

Normalmente esta región será el interior de un polígono, que como tal tiene sus lados y vértices.

Solución del Problema:

La solución para el problema de optimización, bien se trate de la producción o bien de minimizar los gastos, es uno o varios puntos de la región factible.

Se puede demostrar (aunque en este trabajo el objetivo es otro) que este punto, cuando sea único, es uno de los vértices del citado polígono. Y que en el caso de más de un punto-solución estos puntos coinciden con los puntos de alguno de los lados del polígono.

Estos casos los veremos reflejados en algunos ejemplos.

Ejemplo 2:
Deseamos optimizar el valor de la función

$$z = 3x+5y+8$$

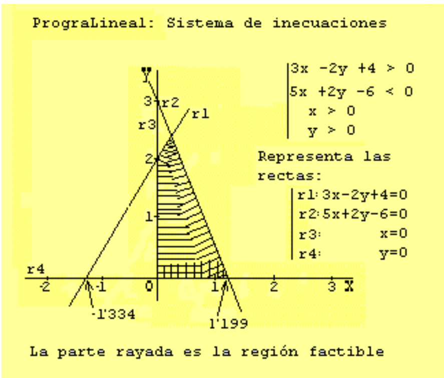
sometida a las siguientes restricciones:

$$\begin{cases} 3x-2y+4>0 \\ 5x+2y-6<0 \\ x>0 \\ y>0 \end{cases} \text{ de donde: } \begin{cases} y<3/2.x+2 \\ y<-5/2.x+3 \\ x>0 \\ y>0 \end{cases}$$

Resolución:

Representamos el conjunto de rectas

$$\begin{cases} r:3x-2y+4=0 \\ r2:5x+2y-6=0 \\ r3:x=0 \\ r4:y=0 \end{cases}$$



y represento la región factible teniendo en cuenta cada una de las siguientes condiciones:

x > 0 rayado hacia la derecha
y > 0 rayado hacia arriba
y > .... rayado hacia arriba
y < .... rayado hacia abajo

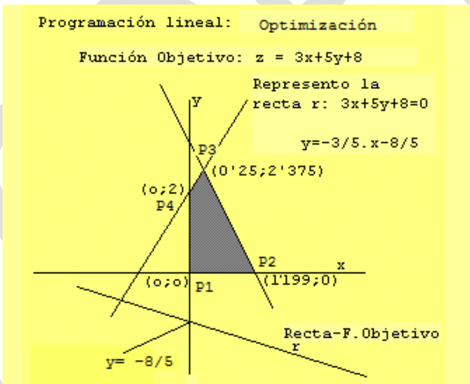
A continuación represento la función objetivo tomando la expresión F(x,y) e igualándola a 0, obteniendo la expresión de una recta r. De esta recta lo que realmente interesa es su pendiente.

En la figura 4, si trasladamos r hacia arriba, el primer vértice encontrado corresponde a ‘mínimo, y el último a ´máximo.

Si la función objetivo fuese

$$z = 3x + 5y - 15$$

al representarla cortaría a oy por encima del vértice más alto. Trasladaríamos hacia abajo, y el primer vértice corresponde a un máximo y el último a mínimo.

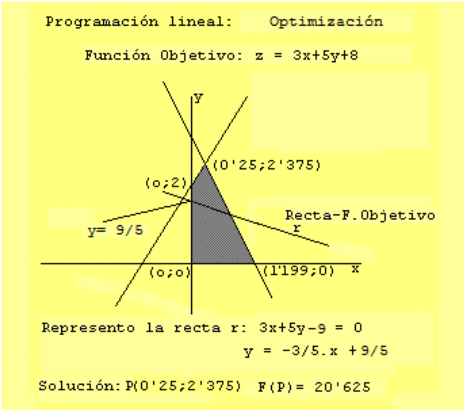


Lo más adecuado será lo siguiente (siguiente fig.)

Prescindimos del término constante y a la expresión 3x+5y le sumo una constante c de tal forma que la recta

$$r: 3x + 5y + c = 0$$

corte la región factible.



Después tener en cuenta que:

- Si deseo maximizar traslado ascendiendo y el último vértice encontrado es la solución.

- Si deseo minimizar traslado descendiendo y el último vértice encontrado es la solución.

Ejemplos resueltos:

1.- Tenemos el sistema de restricciones

$$\begin{cases} 2x + 3y - 4 < 0 \\ 3x - 2y + 2 < 0 \\ x > 0 \\ y > 0 \end{cases}$$

a)Obtener los puntos de corte de las restricciones y los vértices de la región factible.

b)Determina los valores de x, y que hacen máximo el valor de la función z = x – 2y

Sol.: Tenemos:

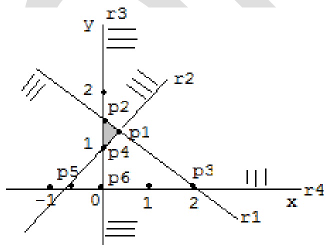
y < -2/3.x + 2
y > 3/2.x + 1

Sistema asociado $\begin{cases} 2x + 3y - 4 = 0 \\ 3x - 2y + 2 = 0 \\ x = 0 \\ y = 0 \end{cases}$

Puntos de corte dos a dos:

p1(0,153; 1,23), p2(0; 1,333), p3(2; 0), p4(0; 1),

p5(-0,667; 0), p6(0, 0)



Vértices: p1, p2, p4

2.- Tenemos el sistema de restricciones

$$\begin{cases} 2x - 3y - 5 < 0 \\ x + 2y - 6 > 0 \\ x > 0 \\ y > 0 \end{cases}$$

a)Obtener los puntos de corte de las restricciones y los vértices de la región factible.

b)Determina los valores de x, y que hacen máximo el valor de la función f(x,y) = x – 2y

Sol.: Observa $\begin{cases} y > \frac{1}{3}.(2x-5) \\ y < 1/2.(-x+6) \end{cases}$, $\begin{cases} r1: 2x-3y-5=0 \\ r2: x+2y-6=0 \\ r3: x=0 \\ r4: y=0 \end{cases}$

Puntos de corte:

r1^r2 : $\begin{cases} 2x-3y-5=0 \\ x+2y-6=0 \end{cases}$, x = 6 - 2y -- > 2(6-2y) -3y -5 = 0,

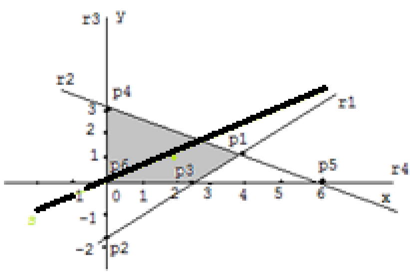
-7y = -7 -> y = 1 -> x = 4, p1(4; 1)

r1^r3: $\begin{cases} 2x-3y-5=0 \\ x=0 \end{cases}$, x = 0 -> y = -5/3, p2(0; -5/3)

r1^r4: $\begin{cases} x+2y-6=0 \\ y=0 \end{cases}$, y = 0 ->x = 6, p3(6; 0)

r2^r3: $\begin{cases} x+2y-6=0 \\ x=0 \end{cases}$, x = 0 -> y = 3, p4(0; 3)

r2^r4: $\begin{cases} x+2y-6=0 \\ y=0 \end{cases}$, y = 0 -> x = 6, p2(6; 0)



Vértices: p1(4; 1), p3(5/2; 0), p4(0; 3), p6(0; 0)

b) z = x -2y,

Valor en cada vértice: z(p1) = 2, z(p3) = 2,5, z(p4) = -6, z(p6) = 0

Valor máximo cuando x = 5/2, y = 0

3.- Tenemos el sistema de restricciones

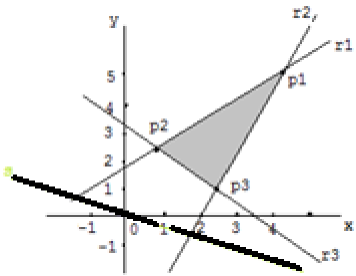
$$\begin{cases} 2x - 3y + 6 > 0 \\ 2x - y - 4 < 0 \\ 2x + 2y - 7 > 0 \end{cases}$$

a)Obtener los puntos de corte de las restricciones y los vértices de la región factible.

b)Determina los valores de x, y que hacen máximo el valor de la función f(x,y) = x + 2y

Sol.: Tenemos: $\begin{cases} y < \frac{1}{3}.(2x+6) \\ y > 2x - 4 \\ y > 1/2.(-2x+7) \end{cases}$

$\begin{cases} r1: 2x - 3y + 6 = 0 \\ r2: \ 2x - y - 4 = 0 \\ r3: \ 2x + 2y - 7 = 0 \end{cases}$, Puntos de corte dos a dos:



r1^r2: $\begin{cases} 2x - 3y + 6 = 0 \\ 2x - y - 4 = 0 \end{cases}$, y = 2x-4 -> 2x-3(2x-4) + 6 = 0

-4x+18 = 0, x = 9/2 -> y = 5, p1(9/2;5)

r1^r3: $\begin{cases} 2x - 3y + 6 = 0 \\ 2x + 2y - 7 = 0 \end{cases}$, Restándolas: -5y +13 = 0,

y = 13/5 -> 2x + 26/5 -7 = 0, 2x = 9/5, x = 9/10

p2(9/10; 26/10)

r2^r3: $\begin{cases} 2x - y - 4 = 0 \\ 2x + 2y - 7 = 0 \end{cases}$, Restándolas: -3y +3 = 0,

y = 1 -> 2x -1-4 = 0, 2x = 5, x = 5/2, p3(5/2; 1)

Vértices: p1(9/2; 5), p2(9/10; 26/10), p3(5/2; 1)

Función a optimizar: z = x + 2y,

Valor en cada vértice:
Vértices: p1(9/2; 5), p2(9/10; 26/10), p3(5/2; 1)

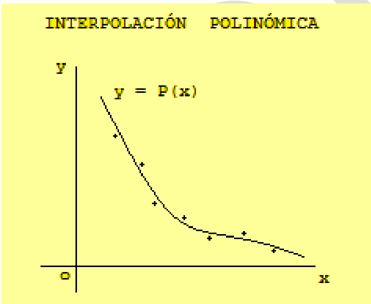
z(p1) = 9/2 +10 = 29/2 = 14,5
z(p2) = 9/10 +52/10 = 61/10 = 6,1
z(p3) = 5/2 + 2 = 9/2 = 4,5

El valor máximo lo toma cuando x = 9/2, y = 5

$$$$oOo$$$$

Tema 2

INTERPOLACIÓN Polinómica



Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

2.1.- Introducción

En las Ciencias experimentales es necesario obtener la ley de comportamiento que da respuesta, si es posible la mejor respuesta, al resultado de una serie (finita) de experimentos. Para conseguirlo es necesario realizar de forma ordenada una serie de pruebas empíricas concretas suficientes para llegar a inducir la citada 'Ley de comportamiento'.

Tratamos aquí los casos en los que esa serie de pruebas generan una lista de pares (xi,yi) de valores.   La ley de correspondencia que deseamos inducir a partir de esto pares de valores se concretará en una expresión explícita del tipo: $y = f(x)$, que ha de cumplir: $y_i = f(x_i)$, para todos los pares de la lista.

Ejemplos:

1.- Complamos manzanas, por ejemplo 5 personas la misma mañana y en el mismo mercado de modo que el precio es el mismo para los cinco. Precio: 1,20 e./kg. La siguiente tabla muestra las compras realizadas

kgs:	1	2	4	6	8
Coste	1,20	2,40	4,80	7,20	9,60

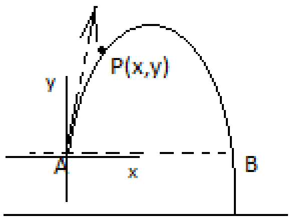
Si representamos los pares (x;y), donde x = cantidad, y = coste, evidentemente debe resultar una recta, ya que en todos los casos se cumple

$$y = x.p, \quad (p = \text{precio/kg})$$

Esta expresión muestra una 'proporcionalidad directa. Dejo para el alumno la representación gráfica de

$$y = 1,20.x$$

2.- Supongamos ahora el caso ‘lanzamiento de una bola pesada’ hacia arriba, con cierta inclinación y hacia adelante. Sabemos que debido a la acción de la gravedad alcanzará el punto más alto posible y después comentará a descender hasta chocar contra el suelo. Existe un instante en el pasará a la misma altura que el punto de lanzamiento. Describe lo que llamamos ‘parábola’.



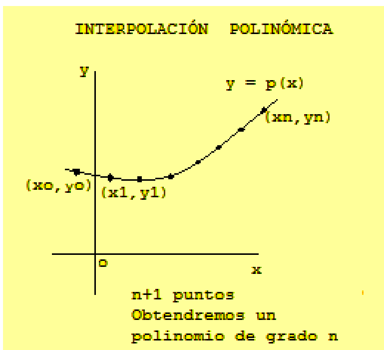
x	0	1	2	3	4	5	6
y	0	1,2	2	6	2	1,2	0

Si tuviésemos la posibilidad de medir el avance x y la altura y, por ejemplo en cinco instante, obtendríamos una tabla de valores, como la siguiente, que nos permitiría obtener una expresión

$$y = f(x)$$

que satisface los valores obtenidos.

2.2.- Interpolación Polinómica



VISION gráfica del Problema:

Tenemos en el plano los pares de valores:

(xi, yi), i = 1, 2, 3, ... , n

Deseamos obtener la expresión y = f(x) más sencilla posible que cumpla yi = f(xi) en todos los pares de valores.

Se ha podido demostrar que existe una función polinómica: y = p(x), que da respuesta al problema. No es precisa otra demostración que expresar las igualdades

$$\begin{cases} y1 = a0 + a1.x1 + a2.x1^2 + \cdots.+an.x1^n \\ y2 = a0 + a1.x2 + a2.x2^2 + \cdots.+an.x2^n \\ \qquad\qquad \ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

que dan lugar a un sistema lineal donde las incógnitas son los coeficientes a0, a1, …, an del polinomio p(x). Si los puntos (xi, yi) son distintos el sistema anterior tiene solución única.

Si el grado p(x) fuese m, y por tanto tiene m+1 coeficientes, para obtener el valor de estos coeficientes necesitamos resolver m+1 igualdades (ecuaciones con m+1 incógnitas), y para obtener dicho sistema necesitamos m+1 puntos del plano, es decir pares de valores.

En consecuencia, si tenemos n+1 puntos

(xi, yi), i = 0, 1, 2, ..., n,

el grado del polinomio es n, y tendrá n+1 coeficientes:

ao, a1, a2, ..., an.

Conviene hacer notar que el polinomio obtenido, cualquiera que sea el camino o método utilizado, es único.

2.3.- METODO: Interpolación Parabólica progresiva
(Método aproximado)

Tomando dos puntos (x0, y0), (x1, y1), de la lista dada, obtenemos un polinomio de grado 1 como sigue:

p1(x) = a0 + a1.(x-x0),

y0 = a0 + a1.(x0-x0) -> a0 = y0

y1 = a0 + a1.(x1-x0) = y0 + a1.(x1-x0)

de donde: a1 = $\frac{y1-y0}{x1-x0}$
El polinomio es:

$$P1(x) = a0 + \frac{y1-a0}{x1-x0} \,.\, (x - x0) \qquad (3)$$

de donde a1 = $\frac{y1-a0}{x1-x0}$

Añadimos otro punto, el punto (x2, y2), para obtener el polinomio de grado 2, conservando los coeficientes a0,a1 ya obtenidos, como sigue:

p2(x) = a0 + a1.(x-x0) + a2.(x-x0).(x-x1),

o bien:

p2(x) = p1(x) + a2.(x-x0).(x-x1).

Entonces ha de cumplirse:

y2 = a0 + a1.(x2-x0) + a2.(x2-x0).(x2-x1),

o mejor

y2 = p1(x2) + a2.(x2-x0).(x2-x1),

de donde

$$a2 = \frac{y2 - p1(x2)}{(x2-x0).(x2-x1)} \qquad (4)$$

El polinomio

$$p2(x) = a0 + a1.(x\text{-}x0) + a2.(x\text{-}x0).(x\text{-}x1),$$

pasa por los tres primeros puntos.

En efecto, $p2(x0) = a0 = y0$

$$p2(x1) = a0 + a1.(x1\text{-}x0) = y0 + \frac{y1-a0}{x1-x0}.(x1-x0) =$$
$$= y0 +(y1\text{-}y0) = y1$$

Para obtener a3:

$$p3(x) = p2(x) + a3.(x\text{-}x0).(x\text{-}x1).(x\text{-}x2)$$

de donde obtenemos:

$$a3 = \frac{y3-p2(x3)}{(x3-x0).(x3-x1).(x3-x2)} \qquad (5)$$

De forma análoga a como hicimos para p2(x), podemos comprobar que este polinomio

$$p3(x) = a0 + a1.(x\text{-}x0) + a2.(x\text{-}x0).(x\text{-}x1) +$$
$$+ a3.(x\text{-}x0).(x\text{-}x1).(x\text{-}x2)$$

pasa por los cuatro primeros puntos.

En general, para el polinomio de grado n

$$p_n(x) = a0+a1.(x\text{-}x0)+a2.(x\text{-}x0)(x\text{-}x1)+ \ldots +$$

$$+ an.(x\text{-}x0)(x\text{-}x1)\ldots(x\text{-}x(n\text{-}1))$$

Es inmediato que la fórmula para obtener el valor an es:

$$an = \frac{yn - p_{n-1}(xn)}{(xn-x0).(xn-x1).(xn-x2)....(xn-x_{n-1})} \qquad (6)$$

2.4.- METODO DE Lagrange

Tengo como antes la lista de pares de valores así

(xi, yi), i = 0,1,2,...,n,

donde tenemos n+1 puntos.

El polinomio será de grado n, con coeficientes:

a0, a1, a2, ..., an (n+1 valores).

El Método consiste en plantear, y después resolver, la siguiente igualdad:

Pn(x) = mo.(x-x1).(x-x2)...(x-xn) +
se ha excluido (x-x0)

+ m1.(x-xo).(x-x2)...(x-xn) +
se ha excluido (x-x1)
+ ... +
mk.(x-xo).(x-x1) ... (x-xn) +
se ha excluido (x-xk)
+ ... +
+ mn(x-xo).(x-x1)...($x-x_{n-1}$),
se ha excluido (x-xn)

donde m_i, i = 0, 1, 2, ..., n son valores indeterminados y que hemos de determinar.

Las condiciones impuestas son:

yk = mo.(xk-x1).(xk-x1)...(xk-xk)...(xk-xn) +

+ m1.(xk-xo).(xk-x2)...(xk-xk)...(xk-xn) +

+ ... +

+ mk.(xk-xo).(xk-x1)...(xk-x_{n-1}).(xk-xn)+
(excluido (xk-xk))
+ ... +

+ mn.(xk-xo).(xk-x1)...(xk-xk)...(xk-x_{n-1})

donde k = 0,1,2,...,n.

Observa que se hacen cero todos los términos salvo el término del coeficiente mk, resultando

yk = + mk.(xk-xo).(xk-x1)...(xk-x_{n-1}).(xk-xn)
(excluido (xk-xk))

De la anterior obtengo

$$mk = \frac{yk}{(xk-x0).(xk-x1)....(xk-xn)},$$
(excluido (xk-xk))

Observa que los puntos dados han de tener valores xi (abscisas) no coincidentes.

Ejemplos:

1.- Obtener el polinomio P(x) cuya gráfica pase por los siguientes puntos

q1(-5; 0), q2(-3; 1), q3(0; 3), q4(2; 0), q5(3; -2)

Sol.:

p(x) = 0 + 0,5.(x+5) + 0,03.(x+5)(x+3) - 0,07.(x+5)(x+3)(x-0) +

+ 0,01.(x+5)(x+3)(x-0)(x-2)

p(x) = 2,95 - 0,61.x - 0,54.x^2 -0,01.x^3 + 0,01.x^4

2.- Obtener el polinomio P(x) cuya gráfica pase por los siguientes puntos

q1(-3; 1), q2(1; 0), q3(3; 2)

Sol.: Por el método ‘aproximación parabólica’ obtenemos

1 + - 0,25.(x+3) + 0,2.(x+3)(x-1), operando obtengo

P(x) = - 0,35 + 0,15.x + 0,2.x^2 (comprobado)

Por método de Lagrange

m0 = $\frac{y0}{(-3-1).(-3-3)} = \ldots = \frac{1}{24} = 0{,}042$

m1 = $\frac{y1}{(1+3).(1-3)} = \frac{0}{-8} = 0$

m2 = $\frac{y2}{(3+3).(3-1)} = \frac{2}{12} = 0{,}167$

(x-1).(x-3) = x^2 -4x +3 -> 0,042.(x^2-4x+3)

(x+3).(x-3) = x^2 -9 = -> 0.(x^2-9) = 0

(x+3).(x-1) = x^2 +2x -3 -> 0,167.(x^2+2x-3)

Resultado: P(x) = 0,209.x^2 +0,166.x -0,375

3.- Obtener el polinomio P(x) cuya gráfica pase por los siguientes puntos

q1(1; 0,20), q2(2,3; 1,06), q3(3,5; 2,45)

Sol.: Por Método de aproximaciones sucesivas

Tengo tres puntos, obtendré un polinomio de grado dos. Obtengo

P_0 = y0 = 0,20 -> a0 = 0,20

$P_1(x)$ = a0 + a1.(x-1)

1,06 = P1(2,3) = 0,20 +a1.(2,3-1) ->

$a1 = \frac{1,06-0,20}{1,3} = 0,662$

$P_2(x)$ = 0,20 + 0,662.(x-1) + a2.(x-1).(x-2,3)

2,45 = P2(3,5) = 0,20 + 0,662.(3,5-1) +

+ a2.(3,5-1).(3,5-2,3) ->

$a2 = \frac{2,45 - 0,20 - 1,655}{3} = \frac{0,595}{3} = 0,198$

Por tanto

P(x) = 0,20 + 0,662.(x-1) + 0,198.(x-1).(x-2,3)

0,662.(x-1) = 0,662.x - 0,662

0,198.(x-1).(x-2,3) = 0,198.(x^2 -3,3.x + 2,3) =

= 0,198.x^2 - 0,653.x + 0,455

Sumándolos

P(x) = 0,198.x^2 + 0,009.x - 0,007

Utilizando la Aplicación informática nos da

$$P(x) = -0{,}006 + 0{,}008.x + 0{,}198.x^2$$

Por método de Lagrange

$$m0 = \frac{0{,}20}{(1-2{,}30).(1-3{,}50)} = \frac{0{,}20}{3{,}25} = 0{,}062$$

$$m1 = \frac{1{,}06}{(2{,}30-1).(2.30-3{,}50)} = \frac{1{,}06}{(-1{,}56)} = -0{,}679$$

$$m2 = \frac{2{,}45}{(3{,}50-1).(3{,}50-2{,}30)} = \frac{2{,}45}{3} = 0{,}817$$

Multiplico factores en x

$(x-2{,}3).(x-3{,}5) = x^2 -5{,}8x +8{,}05$ ->

$0{,}062.(x^2 -5{,}8x +8{,}05) = 0{,}062.x^2 - 0{,}360.x + 0{,}499$

$(x-1).(x-3{,}5) = x^2 -4{,}5x +3{,}5$ ->

$-0{,}679.(x^2 -4{,}5x +3{,}5) = -0{,}679.x^2 + 3{,}056.x - 2{,}377$

$(x-1).(x-2{,}3) = x^2 -3{,}3.x + 2{,}3$ ->

$0{,}817.(x^2 -3{,}3.x + 2{,}3) = 0{,}817.x^2 -2{,}696.x + 1{,}879$

Sumándolos

$$P(x) = 0{,}2.x^2 + 0.x + 0{,}001$$

Vemos que existe alguna diferencia con el obtenido por el método de aproximaciones, pero estas diferencias se producen en los términos de menor grado al dominante (el mayor).

NOTA: Han sido comprobados reiteradamente.

4.- Obtener el polinomio P(x) cuya gráfica pase por los siguientes puntos

q1(-3; -1), q2(-2; 0), q3(0; 1), q4(1; 2), q5(2; 1), q6(3; 0)

Sol.: Resuélvalo el alumno.

5.- Dato: Los de la siguiente tabla

x	0	2	4	6
y	0	2	2	0

Sol.: P_0 = y0 = 0 -> a0 = 0

P_1(x) = a0 + a1.(x-0)

2 = P1(2) = 0 + a1.(2 -0) ->

a1 = 1

P_2(x) = 0 + 1.(x-0) + a2.(x-0).(x-2)

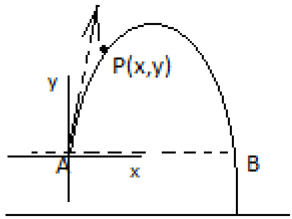
2 = P2(4) = 0 + 1.4 + a2.4.2 -> 2 = 4 + 8.a2

a2 = $\frac{-2}{8} = \frac{-1}{4} = -0{,}25$

Por tanto, si nos quedamos en grado dos como corresponde a 'parábola'

P(x) = 0 + x + $\frac{-1}{4}$.x.(x-2) = $\frac{-1}{4} . x^2 + \frac{3}{2} . x$

Comprueba que se satisface la tabla dada



$$$$oOo$$$

PROBLEMAS: Parte II

De Programación lineal:

1.-

Datos: Sistema de inecuaciones (desigualdades)

$$\begin{cases} 3x-2y+4<0 \\ 5x+2y-6<0 \\ x>0 \\ y>0 \end{cases}$$

Obtener: a)

Tomando el sistema de igualdades (rectas en el plano)

$$\begin{cases} r1: 3x-2y+4=0 \\ r2: 5x+2y-6=0 \\ r3: x=0 \\ r4: y=0 \end{cases}$$

Los puntos de corte de las rectas dos a dos

b) Los vértices del polígono de puntos factibles

Sol.: a) Obtenemos los puntos de corte de las rectas entre sí (aproximando a las milésimas)

(0’25 ; 2’375), (0 ; 2), (-1’334 ; 0),
(0 ; 3), (1’199 ; 0), (0 ; 0)

b) Vértices:

P0(0’25 ; 2’375), P1(0 ; 2),
P2(1’199 ; 0), P3(0 ; 0)

2.- Datos: Sistema de inecuaciones (desigualdades)

$$\begin{cases} 3x-2y+4<0 \\ 5x+2y-6<0 \\ x>0 \\ y>0 \end{cases}$$

Función objetivo: F(x, y) = 3x + 5y + 8

Obtener: a)
Tomando el sistema de igualdades (rectas en el plano)

$$\begin{cases} r1: 3x-2y+4=0 \\ r2: 5x+2y-6=0 \\ r3: x=0 \\ r4: y=0 \end{cases}$$

Los vértices del polígono de puntos factibles

b) El Punto P(x0, y0) (o vértice) donde F(x, y) toma el valor máximo.

c) El Punto P(x1, y1) (o vértice) donde F(x, y) toma el valor mínimo.

Sol.: a) Vértices:
P0(0’25 ; 2’375), P1(0 ; 2),
P2(1’199 ; 0), P3(0 ; 0)

b) Valor máximo en
P0(0’25 ; 2’375), F(P0) = 20’625

c) Valor mínimo en
P3(0 ; 0), F(x1, y1) = 8

3.- Sistema de restricciones

$$\begin{cases} x - y < 1 \\ x + y > 3 \\ 2x + 3y < 22 \\ -x + 4y < 22 \\ x > 0, y > 0 \end{cases}$$

A) Función objetivo: z = 3x -2y + 6
B) Función objetivo: z = 3x + 2y -6

Sol.:

Vértices: P1(2; 1), P2(5; 4), P3(2; 6), P4(-2; 5)

Caso A: Máximo en P2, Mínimo en P4

F(P1) = 10, F(P2) = 13, F(P3) = 0, F(P4) = -10

Caso B: Máximo en P2, Mínimo en P4

F(P1) = 2, F(P2) = 17, F(P3) = 12, F(P4) = -2

De Interpolación:

1.- Datos:

(xo, yo), (x1, y1), (x2, y2), (x3, y3), (x4, y4)

(15, 30), (16, 40), (20, 50), (30, 60), (40, 70)

En general, para n+1 puntos obtendremos un polinomio de grado n.

Aplicaremos el Método parabólico progresivo.

En nuestro caso n = 4, y el polinomio esperado, de cuarto grado, viene expresado de la forma siguiente:

P(x)= a0 + a1.(x-x0) + a2.(x-x0).(x-x1) +

a3.(x-x0).(x-x1).(x-x2) + a4.(x-x0).(x-x1). .(x-x2).(x-x3) +

a5.(x-x0).(x-x1).(x-x2).(x-x3).(x-x4)

P(x0) = a0 -- > a0 = y0

y1 = P(x1) = a0 +a1.(x1-x0) -- > a1 = (y1-y0)/(x1-x0)

Llamo p1(x) = a0 + a1.(x-x0)

y2 = P(x2) = p1(x2) + a2.(x2-x0).(x2-x1), de donde

$$a2 = \frac{(y2-p1(x2))}{[(x2-x0).(x2-x1)]}$$

Llamo

p2(x) = a0 + a1.(x-x0) + a2.(x-x0).(x-x1)

y3 = P(x3) = p2(x3) + a3.(x3-x0).(x3-x1).(x3-x2)

de donde

$$a3 = \frac{(y3-p2(x3))}{[(x3-x0).(x3-x1).(x3-x2)]}$$

Llamo

p3(x) = a0 + a1.(x-x0) + a2.(x-x0).(x-x1) + a3(x-x0).(x-x1).(x-x2)

Análogamente obtenemos

$$a4 = \frac{(y4-p3(x4))}{[(x4-x0).(x4-x1).(x4-x2).(x4-x3)]}$$

Llamo

p4(x) = a0 + a1.(x-x0) + a2.(x-x0).(x-x1) +

+ a3.(x-x0).(x-x1).(x-x2) + a4.(x-x0).(x-x1).(x-x2).(x-x3)

Este polinomio es de grado 4 y es el polinomio deseado.

Tomando los valores dados (Datos) obtenemos el valor de los coeficientes.

a0 = 30,

a1 = $\frac{40-30}{16-15} = 10$

a2 = $\frac{(50-p1(20))}{[(20-15).(20-16)]} = (*)$

Obtengo

p1(20) = 30 +10.(20-15) = 30 +50 = 80

(*) = … = $\frac{-3}{2}$

a3 = $\frac{(60-p2(30))}{[(30-15).(30-16).(30-20)]} = (*)$

Obtengo

p2(30) = 30 +10.(30-15) -3/2.(30-15).(30-16) =

= 30 +150 -3/2.210 = 180 -315 = -135

(*) = $\frac{(60+135)}{[15.14.10]} = \ldots = \frac{39}{420} = \frac{13}{140}$

a4 = $\frac{(70-p3(40))}{[40-15).(40-16).(40-20).(40-30)]} = (*)$

Obtengo

p3(40) = 30 +10.(40-15) -3/2.(40-15).(40-16) +13/140.(40-15).(40-16).(40-20) =

= 280 -3/2.600 +13/140.12000 = 280 -900 + 15600/14 = -620 +7800/7 = 3460/7

$$(*) = \frac{70-\frac{3460}{7}}{[25.24.20.10]} = \frac{\frac{-2970}{7}}{120000} = \frac{-2970}{840000} = \frac{-99}{28000}$$

Resultado:

P(x) =

$$= 30 + 10(x-15) - \frac{3}{2}.(x-15).(x-16) + \frac{13}{140}.(x-15).(x-16).(x-20) -$$

$$- \frac{99}{28000}.(x-15).(x-16).(x-20).(x-30)$$

Comprobamos:

P(15) = 30

P(16) = 30 +10.1 = 40

P(16) = 40

P(20) = 30 +10.5 -3/2.5.4 = 80 -30 = 50

P(20) = 50

P(30) = 30 +10.15 -3/2.15.14 +13/140.15.14.10 =

= 180 -315 +195 = 60

P(30) = 60

P(40) = 30 +10.25 -3/2.25.24 +13/140.25.24.20 –

-99/28000.25.24.20.10 =

$= 280 -900 +13/14.1200 -99/280.1200 =$

$-620 +13/7.600 -99/28.120 = -620 +7800/7 -2970/7 =$

$-620 +4830/7 = -620 +690 = 70$

$P(40) = 70$

2.- Datos:

p0 p1 p2 p3 p4

(xo,yo), (x1,y1), (x2,y2), (x3,y3), (x4,y4)

(-3,-2), (-2,-1), (0,0), (1,0), (2,-1)

Obtendremos un polinomio de grado 4

Siguiendo el proceso anterior, salvando algunos cálculos que realizará el Alumno, obtenemos lo siguiente.

$a0 = -2$

$-1 = p1(-2)$ de donde $a1 = 1$

$p1(x) = -2 + (x+3)$

$p2(x) = p1(x) + a2.(x+3).(x+2)$

$0 = p2(0) = p1(0) + a2.3.2$, de donde $a2 = -\frac{1}{6}$

$p2(x) = p1(x) -1/6.(x+3)(x+2)$

$p3(x) = p2(x) + a3.(x+3).(x+2).x$

$0 = p3(1) = ...$, de donde $a3 = 0$

p3(x) = p1(x) + p2(x)

p4(x) = p3(x) + a4.(x+3).(x+2).x.(x-1)

-1 = p4(x) = p3(2) + a4.5.4.2.1

$$p3(2) = -\frac{1}{3}$$

Obtengo a4 = $-\frac{1}{60}$

p4(x) = -2 + (x+3) - $\frac{1}{6}$.(x+3).(x+2) – $\frac{1}{60}$.(x+3).(x+2).x.(x-1)

Hechos algunos cálculos:

$(x^2 + 5x + 6).(x^2 - x) = x^4 + 4x^3 + x^2 - 6x$

llegamos a

$p4(x) = \frac{4}{15} .x - \frac{11}{60} .x^2 - \frac{1}{15} .x^3 - \frac{1}{60} .x^4$

3.- Los mismos datos que en 2, pero aplicamos el Método de Lagrange

Datos:

p0	p1	p2	p3	p4
(xo,yo),	(x1,y1),	(x2,y2),	(x3,y3),	(x4,y4)
(-3,-2),	(-2,-1),	(0,0),	(1,0),	(2,-1)

P4(x) = m0.(x+2).x.(x-1).(x-2) + m1.(x+3).x.(x-1).(x-2) +

+m2.(x+3).(x+2).(x-1).(x-2) + m3.(x+3).(x+2).x.(x-2) +

+ m4.(x+3).(x+2)x.(x-1)

En p0: -2 = P4(-3) = m0.-1.-3.-4.-5 , de donde

$$m0 = \frac{-1}{30}$$

En p1: -1 = P4(-2) = m1.1.-2.-3.-4, m1 = $\frac{1}{24}$

En p2: 0 = P4(0) = m2.3.2.-1.-2, m2 = 0

En p3: 0 = P4(1) = m3.4.3.1.-1, m3 = 0

En p4: -1= P4(2) = m4.5.4.2.1, m4 = $\frac{-1}{40}$

Cálculos para obtener P4(x) en formato normal.

Bloque de m0:

$$(x^2-3x+2).(x^2+2x)= x^4 - x^3 - 4x^2 + 4x$$

Bloque de m1:

$$(x^2-3x+2).(x^2+3x)= x^4 -7x^2 + 6x$$

Bloque de m4:

$$(x^2+5x+6).(x^2-x)= x^4 + 4x^3 + x^2 -6x$$

$$P4(x) = -\frac{1}{30}.(x^4-x^3-4x^2+4x) + \frac{1}{24}.(x^4 -7x^2+6x) -$$

$$-\frac{1}{40}.(x^4 + 4x^3 + x^2 - 6x) = ...$$

$$P4(x) = \frac{4}{15}.x - \frac{11}{60}.x^2 - \frac{1}{15}.x^3 - \frac{1}{60}.x^4$$

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

PROBLEMAS Primera parte

De Estadística:

1.- La siguiente Tabla muestra el resultado obtenido de una población de 60 hombres, a quienes se les ha consultado y tomado nota de: Su altura X, su peso Y. (Observa que son datos agrupados en intervalos)

Estudiar:

a) La distribución marginal de X, y de Y
b) Media y varianza marginales de X, y de Y

Sol.:

Y \ X	[1,55 – 1,65)	[1,65 – 1,75)	[1,75 – 1,85)
[55 – 65)	3	2	—
[65 – 75)	6	10	4
[75 – 85)	4	11	5
[85 – 95)	1	6	8

Sol.: a) Distribución marginal de X:

x_i	[1,55 – 1,65)	[1,65 – 1,75)	[1,75 – 1,85)
Frecuencia absoluta	14	29	17
Frecuencia relativa	14/60	29/60	17/60

Nota: Para cada valor x_i, fi. = $\sum_{j=1,\ldots} fij$, por ejemplo: f1. = 3+6+4+1 = 14

N = $\sum_{i=1,\ldots} fi.$ = 60

Distribución marginal de Y:

y_i	Frecuencia absoluta	Frecuencia relativa
[55 – 65)	5	5/60
[65 – 75)	20	20/60
[75 – 85)	20	20/60
[85 – 95)	15	15/60

Nota: Para cada valor y_i, f.j = $\sum_{i=1,\ldots} fij$, por ejemplo: f.1 = 3+2 = 5, f.2 = 6+10+4 = 20

N = $\sum_{j=1,\ldots} f.j$ = 60

b)Media y Varianza de X:

Recuerda que hemos de tomar la 'marca de clase' de cada intervalo:

$$m_x = \frac{14.1,60+29.1,70+17.1,80}{60} = 1,705$$

$$S_x^2 = \frac{14.1,60^2+29.1,70^2+17.1,80^2}{60} - 1,705^2 = 0,005$$

Media y Varianza de Y:

Recuerda que hemos de tomar la 'marca de clase' de cada intervalo:

$$m_y = \frac{5.60+20.70+20.80+15.90}{60} = 77{,}5$$

$$S_y^{\,2} = \frac{5.60^2+20.70^2+20.80^2+15.90^2}{60} - 77{,}5^2 = 85{,}41$$

2.- Tengo la Tabla de valores de una variable bidimensional (X,Y): (Son datos No agrupados)

x_i	y_i
1	3
2	4
4	3
5	4
6	5

Obtener las rectas de regresión: De 'y' en función de 'x', de 'x' en función de 'y'. También el coeficiente de correlación lineal, y calificar el tipo de correlación entre X e Y.

Sol.: En primer lugar completo la siguiente tabla

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	3	3	1	9
2	4	8	4	16
4	3	12	16	9
5	4	20	25	16
6	5	30	36	25
18	19	73	82	75

Observa que N = 5

Con ayuda de la tabla obtenida calculamos los estadísticos marginales y la covarianza:

m_x = 18/5 = 3,60 ; m_y = 19/5 = 3,80

$S_x^2 = \frac{82}{5} - 3{,}60^2 = 3{,}44$; $S_y^2 = \frac{75}{5} - 3{,}80^2 = 0{,}56$

Covarianza: $S_{xy} = \frac{73}{5}$ -3,60.3,80 = 0,92

Coefi. de correlación: $r = \frac{S_{xy}}{S_x . S_y} = \frac{0{,}92}{3{,}44.0{,}56} = 0{,}74$

Rectas de regresión:

FÓRMULAS: $y - m_y = \frac{S_{xy}}{S_x^2} . (x - m_x)$,

$x - m_x = \frac{S_{xy}}{S_y^2} . \left(y - m_y\right)$

En nuestro caso: $\frac{S_{xy}}{S_x^2} = \frac{0{,}92}{3{,}44} = 0{,}2674$, $\frac{S_{xy}}{S_x^2} . m_x =$

$= 0{,}2674.3{,}60 = 0{,}9626$ -> y = 0,2674.x +2,8374

$\frac{S_{xy}}{S_y^2} = \frac{0{,}92}{0{,}56} = 1{,}6429$, $\frac{S_{xy}}{S_y^2} . m_y = 1{,}6429.3{,}80 = 6{,}2430$ ->

x = 1,6429.y -2,6430

La correlación es directa por ser r > 0. Sí existe correlación entre X e Y, si bien es moderada o aleatoria.

NOTA:
Cuando r > 0 y próxima a 1 se dice que 'existe fuerte correlación' y la llamamos 'funcional'. Si r es próxima a cero 'no existe casi correlación, y decimos que son casi aleatoriamente independientes'.

3.- La siguiente Tabla contiene los valores de una v.a. bidimensional al consultar dos constantes biológicas de 12 individuos dentro de una población (estadística):

X	65	50	55	65	55	70	65	70	55	70	50	50
Y	85	79	76	90	85	87	94	98	81	95	76	74

Contestar a las siguientes cuestiones:

a) ¿Qué valor podemos predecir para Y cuando X tome el valor 80?
b) ¿Qué valor podemos predecir par X cuando Y tome el valor 100?
c) Explica qué tipo de dependencia existe entre X e Y.

Sol.: Necesito las rectas de regresión. Haciendo cálculos obtengo:

$m_x = 60$, $m_y = 85$,

$s_x^2 = 62{,}5$, $s_y^2 = 59{,}5$, $s_{xy} = 53{,}75$,

a) Recta $y = 0{,}86.x + 33{,}40$

$x = 80 \to y = 0{,}86.80 + 33{,}40 = 102{,}20$

b) Recta $x = 0{,}90.y – 16{,}80$

$y = 100 \to x = 0{,}90.100 -16{,}80 = 73{,}20$

c) r = 0,88 -> dependencia aleatoria fuerte

4.- En la siguiente Tabla presentamos los datos obtenidos de un cultivo de bacterias después de un tiempo transcurrido representado por X:

X Número de horas:	0	1	2	3	4	5
Y Núm. bacterias por unidad de V.:	12	19	23	34	56	62

Contesta a las siguientes cuestiones:

a) ¿Cuántas horas pasarán hasta tener 100 bacterias?
b) ¿Cuántas esperamos tener cuando han transcurrido 6 h.?
c) Calcula el coeficiente de correlación (lineal) e interprétalo?

Sol.: El alumno comprobará los siguientes resultados.

Hechos los cálculos obtenemos las rectas de regresión:

$y = 10{,}68.x + 7{,}61$;

$x = 0{,}088.y – 0{,}55$

x = 6 h -> cultivadas y = 71,69 bacterias

y = 100 bacterias -> transcurridas x = 8,25 h

Otra forma: 100 = 10,68.x +7,61 -> x = 8,65 h.

r = 0,97 -> la dependencia es (aleatoriamente)

muy fuerte.

De Variables aleatorias y Distribuciones:

1.- Sea X una v.a. discreta cuya función de probabilidad viene dada en la tabla:

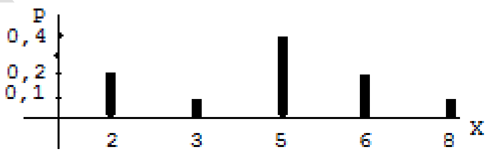
x	2	3	5	6	8
p	0,2	0,1	0,4	0,2	0,1

a) Determina la f. de distribución F(x)
b) Representa la f. de probabilidad: $p_i = P(x_i)$
c) Representa la función F(x)

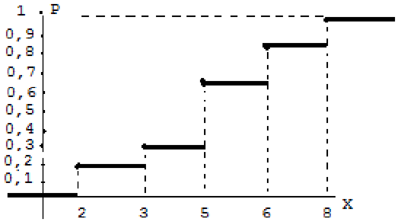
Sol.:

a) $$F(x) = \begin{cases} 0 & si \quad x < 2 \\ 0{,}2 & si \quad 2 \leq x < 3 \\ 0{,}3 & si \quad 3 \leq x < 5 \\ 0{,}7 & si \quad 5 \leq x < 6 \\ 0{,}9 & si \quad 6 \leq x < 8 \\ 1 & si \quad 8 \leq x \end{cases}$$

b)



c)



2.- La v.a. discrete X representa la suma de puntos en el lanzamiento de dos dados. Expresa los siguientes sucesos:

Nota: Represento por x, y los sucesos elementales al lanzar un dado.

S1 = $\{(x,y); x+y=3\}$, S2 = $\{(x,y); x+y=1\}$

S3 = $\{(x,y); 4 \leq x+y < 7\}$, S4 = $\{(x,y); x+y > 12\}$

S5 = $\{(x,y); 2 < x+y \leq 12\}$, S6 = $\{(x,y); x+y=6\}$

S7 = $\{(x,y); x+y=12\}$

Sol.: S1 = $\{(1,2),(2,1)\}$, S2 = $\emptyset$

S3 = $\left\{\begin{matrix}(1,3),(1,4),(1,5),(2,2),(2,3),(2,4),(3,1),(3,2),(3,3),(4,1),(4,2),\\ (5,1)\end{matrix}\right\}$

S4 = $\emptyset$,

S5 = $\left\{\begin{matrix}(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\ldots\ldots..\\ continuar\acute{a}\ el\ alumno\end{matrix}\right\}$

S6 = $\{(1,5),(2,4),(3,3),(4,2),(5,1)\}$, S7 = $\{(6,6)\}$

3.- Calcula la probabilidad de cada uno de los sucesos S_i del número 2.

Sol.: Casos posibles del experimento (lanzar dos dados): 6.6 = 36

P(S1) = 2/36, P(S2) = 0, P(S3) = 12/36,

P(S4) = 0, P(S5) = 35/36, P(S6) = 5/36,

P(S7) = 1/36

4.- En medicina se acepta que de los enfermos de Hepatitis la tercera parte sana. Tomamos al azar cinco enfermos de hepatitis. Se pide:

a) Probabilidad de que se curen dos

b) Probabilidad de que sane al menos dos

c) Probabilidad de que los sanados sean menos de cuatro.

Sol.: a) p = 1/3, q = 1-p = 2/3, n = 5

$$p(X=2) = \binom{5}{2} . \left(\frac{1}{3}\right)^2 . \left(\frac{2}{3}\right)^3 = \cdots = 0{,}3292$$

b) $p(X \geq 2) = 1 - p(X<2) = 1 - [p(X=0) + p(X=1)] =$

$$p(X=0) = \binom{5}{0} . \left(\frac{1}{3}\right)^0 . \left(\frac{2}{3}\right)^5 = 1.1.0{,}1317 = 0{,}1317$$

$$p(X=1) = \binom{5}{1} . \left(\frac{1}{3}\right)^1 . \left(\frac{2}{3}\right)^4 = 5.0{,}0658 = 0{,}3292$$

p(X >= 2) = 1 – [0,1317 + 0,3292] = 0,5391

c)p(X< 4) = 1 – p(X >= 4) =

= 1 – [p(X = 4) + p(X= 5)]

p(X= 4) = $\binom{5}{4}.\left(\frac{1}{3}\right)^4.\left(\frac{2}{3}\right)^1 = 5.0{,}00823 = 0{,}0412$

p(X= 5) = $\binom{5}{5}.\left(\frac{1}{3}\right)^5.\left(\frac{2}{3}\right)^0 = 1.0{,}00412.1 = 0{,}0041$

p(X< 4) = 1 – $[0{,}0412 + 0{,}0041] = 0{,}9547$

Variable Continua:

5.- Tengo una v.a. continua, X, cuya función de densidad f(x) viene definida del siguiente modo:

$$f(x) = \begin{cases} 0, & para\ x \le 0 \\ \cos(x), & para\ 0 < x \le \frac{\pi}{2} \\ 0, & para\ x > \frac{\pi}{2} \end{cases}$$

Determina su función de distribución F(x)

Sol.: F(x) queda definida por la expresión

$$F(x) = \int_{-\infty}^{x} f(t).dt$$

Teniendo en cuenta los intervalos en la definición de f(x) obtengo:

$$\begin{cases} para\ x \le 0 \rightarrow F(x) = 0 \\ para\ 0 < x \le \frac{\pi}{2} \rightarrow F(x) = sen(x) \\ para\ x > \frac{\pi}{2} \rightarrow F(x) = 1 \end{cases}$$

Nota: Recuerda que: $\int_{-\infty}^{x} \cos(t)\,.\,dt = \int_{-\infty}^{0} \cos(t)\,.\,dt +$

$+\int_0^x \cos(t)\,.\,dt$ = 0 + (sen(t)$]_0^x$ = sen(x) – sen(0) =

= sen(x) ; y también que: $\int_{-\infty}^{+\infty} f(t).\,dt = 1$

6.- Tengo una v.a. continua, X, cuya función de distribución F(x) viene definida del siguiente modo:

$$F(x) = \begin{cases} 0, & para\ \ x \leq 0 \\ \text{sen}(x), & para\ 0 < x \leq \frac{\pi}{2} \\ 1, & para\ x > \frac{\pi}{2} \end{cases}$$

Halla su función de densidad f(x)

Sol.: Por definición, f(x) es una función tal que la función F(x) queda definida por

F(x) = $\int_{-\infty}^{x} f(t).\,dt$, y de modo que se cumplan las siguientes condiciones

$$F(x) = \begin{cases} Es\ creciente \\ Es\ continua\ por\ la\ derecha \\ \int_{-\infty}^{+\infty} f(t).\,dt\ = 1 \end{cases}$$

De F(x) = $\int_{-\infty}^{x} f(t).\,dt$, deducimos que f(x) = F'(x)

$$f(x) = \begin{cases} 0, & para\ \ x \leq 0 \\ \cos(x), & para\ 0 < x \leq \frac{\pi}{2} \\ 0\ , & para\ x > \frac{\pi}{2} \end{cases}$$

7.- Tengo una v.a. continua cuya f. de densidad es de la forma

$$f(x) = \begin{cases} 0, \; para\, x \leq 1 \\ \frac{1}{8}.x + a, \; ,para\; 1 < x < 5 \\ 0, \; para\;\; x \geq 5 \end{cases}$$

a) Determina el valor de 'a' para que f(x) sea una f. de densidad
b) Determina su función de distribución F(x) asociada
c) Represéntalas
d) Calcula la media de esta distribución
e) Calcula la desviación típica

Sol.: a) Tiene que cumplir $1 = \int_1^5 \left(\frac{1}{8}.x + a\right).dx =$

$= (\frac{1}{16}.x^2 + ax]_1^5 = \left(\frac{25}{16} + 5a\right) - \left(\frac{1}{16} + a\right) = \frac{24}{16} + 4a = \frac{3}{2} + 4a$ ->

2 = 3 +8.a, a = $\frac{-1}{8}$ -> $f(x)= \begin{cases} 0, \; para\, x \leq 1 \\ \frac{1}{8}.(x-1), \; ,para\; 1 < x < 5 \\ 0, \; para\;\; x \geq 5 \end{cases}$

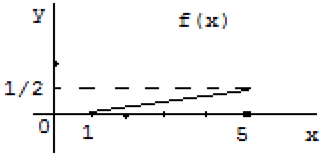
b) F(x) = $\int_{-\infty}^{x} \frac{1}{8}.(t-1).dt = \int_{-\infty}^{1} \frac{1}{8}.(t-1).dt \; +$

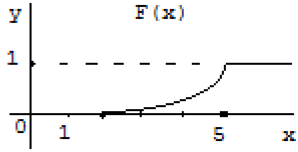
$+ \int_1^x \frac{1}{8}.(t-1).dt \; = \; 0 + \frac{1}{8}.(\frac{1}{2}.t^2 - t]_1^x =$

$= \frac{1}{8}.\left[\left(\frac{x^2}{2} - x\right) - \left(\frac{1}{2} - 1\right)\right] = \frac{1}{8}.\left[\frac{x^2}{2} - x + \frac{1}{2}\right]$;

Observa: F(5) = …. = 1/8 .8 = 1;

c)Representación:





d)La media:

$$m = E[X] = \int_1^5 x.\frac{1}{8}.(x-1).dx = \ldots = \frac{1}{8}.(\frac{x^3}{3} - \frac{x^2}{2}]_1^5 =$$

$$= \cdots = \quad 3{,}67$$

e)Desviación típica:

Varianza: $s^2 = E[(X-m)^2] =$

$$= \int_1^5 (x - 3{,}67)^2.\frac{1}{8}.(x-1).dx = \quad = \ldots\ldots$$

Continuará el alumno (Véase el siguiente)

8.- Una v.a. continua tiene f. de distribución que indicamos:

$$F(x) = \begin{cases} 0, \; para \;\; x \leq 0 \\ x^2, \; para \;\; 0 < x \leq 1 \\ 1, \; para \; x > 1 \end{cases}$$

a) Determina su f. de densidad asociada.
b) Calcula la media
c) Calcula la varianza y la desviación típica

Sol.: a) $f(x) = \begin{cases} 0, \;\; para \; x \leq 0 \\ 2x, \; para \; 0 < x \leq 1 \\ 0, \;\; para \; x > 1 \end{cases}$

b) m = E[X] = $\int_0^1 x.f(x).dx = \int_0^1 2x^2.dx = (\frac{2x^3}{3}]_0^1 = \frac{2}{3}$

c) $s^2 = E[(X-m)^2] = \int_0^1 \left(x-\frac{2}{3}\right)^2.f(x).dx =$

$= \int_0^1 2x.\left(x^2-\frac{4x}{3}+\frac{4}{9}\right).dx = \int_0^1 \left(2x^3-\frac{8}{3}.x^2+\frac{8}{9}.x\right).dx =$

$= (\frac{x^4}{2}-\frac{8x^3}{9}+\frac{8x^2}{18}]_0^1 = \left(\frac{1}{2}-\frac{8}{9}+\frac{8}{18}\right)-(0) = \frac{9-8+8}{18} = \frac{1}{18}$

$s = \sqrt{\frac{1}{9.2}} = \frac{\sqrt{2}}{6}$

9.- Tengo una v.a. continua X del tipo N(0;1). Calcula:

a) p(X>=1,32), b) p(X<=2,17),

c) p(1,52<X<=2,03)

Sol.: NOTA: Hacemos uso de la Tabla de la Normal. El alumno debe aprovecharlo para comprobar si hace un uso correcto de dichas tablas.

a) p(X>=1,32) =1 –p(X<1,32)= 1-0,9066 =0,00934

b) p(X<=2,17) = 0,9850

c) p(1,52<X<=2,03) = p(X<=2,03) –p(X<=1,52) =

= 0,9788 – 0,9357 = 0,0431

10.- En una determinada población se ha podido comprobar que la talla (la v.a. X que nos da la talla) sigue una distribución con media m = 1,75 m, y desviación típica s = 8 cm. Elegimos un individuo al azar, y deseamos saber:

Probabilidad de que su talla sea

a) Mayor que 1,80, b) Menor que 1,70,

c)Esté entre 1,70 y 1,80

Sol.: NOTA: Recuerda que en el caso de m <> 0, y s <> 1, hemos de 'Tipificarla' mediante la Fórmula:

P(X < x) = P(Y< $\frac{y-m}{s}$) , donde Y es v.a. normal tipificada, es decir, del tipo N(0;1)

a) p(X > 1,80) = p(Y > $\frac{1,80-1,75}{8}$) $= p(Y > 0,006) =$

= 1- p(Y $\leq$ 1,80) = 1-p(Y<= 0,006) =

= 1-0,5001= 0,4999

NOTA:
El valor 0,006 es tan pequeño que no figura en la entrada de la tabla. Pero tampoco sería aceptable tomarlo como igual a cero, y entonces p(Y<= 0,006) = 0,5000; he decidido tomar el valor aproximado p(Y<= 0,006) = 0,5001 (Observando el área bajo la curva de la Normal).

b) p(X< 1,70) = p(Y< $\frac{1,70-1,75}{8}$) $= p(Y < -0,006) =$

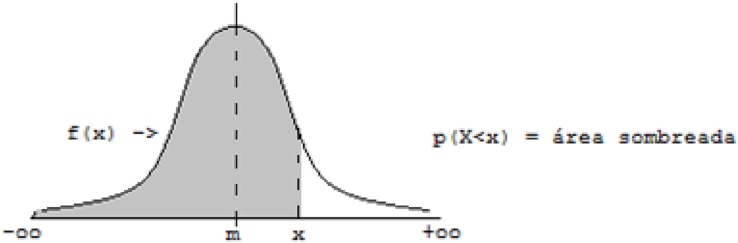
= p(Y $\geq$ 0,006) = 1 –p(Y< 0,006) = 0,4999

c) p(1,70 < X < 1,80) = p(X< 1,80) – p(X< 1,70) =

= p(Y< $\frac{1,80-1,75}{8}$) – p(Y< $\frac{1,70-1,75}{8}$) = p(Y< 0,006) –

-p(Y< -0,006) = 0,5001 – [1-p(Y< 0,006)] =

= 2.0,5001 – 1 = 1,0002 – 1 = 0,0002



p(X< x) = F(x) = $\int_{-\infty}^{+\infty} f(t).dt$

11.- El peso aceptado de una tableta de aspirina es 324 mgs. En el proceso de fabricación se acepta que los pesos de las tabletas siguen una distribución normal con desviación típica s = 10 mgs (en las tabletas). Se desea saber:

a) El tanto por ciento de tabletas con peso menor que 310 mgs.

b) Lo mismo de las que tienen peso superior a 330 mgs.

Sol.: m = 324, s = 10

a) p(X< 310) = p(Y< $\frac{310-324}{10}$) $= p(Y < -1{,}4) =$

= 1 –p(Y< 1,4) = 1 -0,9192 = 0,0808

Este resultado representa el 'tanto por uno', por lo que 8,08 es el % pedido.

b) p(X > 330) = p(Y > $\frac{330-324}{10}$) $= p(Y > 0{,}6 =$
= 1 –p(Y < 0,6) = 1 -0,7257 = 0,2743,

y por tanto 27,43 es el % pedido.

12.- Tenemos 2000 individuos en un campamento (supongamos futuros soldados). Llamamos cociente intelectual al cociente entre la 'edad mental' y la 'edad real' (intelectualmente hablando). Se acepta que el citado coeficiente (en ese campamento) sigue una distribución normal con parámetros: m = 0,80, s = 0,50.
Deseamos saber:

a) Número de individuos con cociente entre 0,70 y 1,20.

b) Número con cociente superior a 1,40

Sol.: a) p(0,70 < X< 1,20) = p(X< 1,20) – p(X< 0,70) =

$= p(Y< \frac{1,20-0,80}{0,50}) - p(Y< \frac{0,70-0,80}{0,50}) =$

= p(Y< 0,80) – p(Y< -0,2) =

= 0,7881 – [1- p(Y< 0,20) = 0,7881 – (1-0,5793) =

= 0,7881 – 0,4207 = 0,3674 ; Por tanto, el

36,74 % lo cumplen, y por tanto, en total lo

cumplen $\frac{2000}{100}.36,74 = 734,8 \rightarrow 735$ individuos

b)p(X > 1,40) = 1 –p(X< 1,40) = 1 –p(Y< $\frac{1,40-0,80}{0,50}$) =

= 1 – p(Y < 1,20) = 1 – 0,8849 = 0,1151

Por tanto: 2000.0,1151 = 230,20 -- > 230

13.- Lanzo un moneda 100 veces. Calcula la probabilidad de:

a) Obtener a lo más 40 caras

b) Obtener más de 40 caras.

Sol.: El experimento nos sitúa en una ‘distribución binomial’ con p = 1/2, n = 100.

Cuando n es ‘grande’ el cálculo de $\binom{n}{k}.p^k.q^{n-k}$ resulta complicado debido a la magnitud infinitesimal de su valor. En el estudio teórico vimos que en estos caso la distribución X binomial puede ser aproximada mediante una Distribución continua cuya v.a. Z viene definida (Demostrada por Moivre) así:

$Z = \frac{X-np}{\sqrt{n.p.q}}$, donde

m = n.p (Media de de la binomial)

$s = \sqrt{n.p.q}$ (Desv. típica de la binom.)

cuya aplicación es aceptable cuando: $\begin{cases} n \geq 30 \\ 0,1 < p < 0,9 \end{cases}$,

siendo una aproximación mejor cuánto mayor es n y p más próximo a 0,5

En nuestro problema: $Z = \frac{X-100.0,5}{\sqrt{100.0,5.0,5}} = \frac{X-50}{5}$

a) $p(X < 40) = p(Z < \frac{40-50}{5}) = p(Z < -2) =$

$= 1 - p(Z < 2) = 1 - 0,9772 = 0,0228$

b) $p(X > 40) = p(Z > -2) = 1 - p(Z < -2) =$

= 1 –[1 –p(Z < 2)] = p(z < 2) = 0,9772

A) De Poisson

De Poisson: P(X= r) = $\frac{k^r}{r!}.e^{-r}$

1.- Se acepta que morir por accidente en una gran ciudad tiene probabilidad p = 0,0004. Una compañía de seguros tiene contratadas N = 10000 pólizas contra ese tipo de accidente en esta ciudad. ¿Calcula la probabilidad de que tenga que hacer frente a más de tres accidentes de este tipo?

Sol.: Llamamos ‘suceso raro’ a este tipo de suceso con probabilidad muy pequeña. Esto indica que estamos en un experimento del tipo Poisson:

$$k = n.p = 10000.0{,}0004 = 4 \quad \text{(Media)}$$

En nuestro caso r = 3

$$P(X > 3) = 1 - P(X \leq 3) = 1 - 0{,}4338 = 0{,}5665$$

(Hemos usado la tabla)

2.- En una fábrica el número de accidentes por semana sigue una ley de Poisson con $k = 2$

a) Probabilidad de que en una semana se produzca algún accidente.

b) Probabilidad de que en el transcurso de dos semanas se produzcan cuatro accidentes.

c) Probabilidad de tener 2 en una semana y otros dos en la siguiente.

d) Es lunes y ya ha habido un accidente. Probabilidad de que en esta semana no haya más de tres accidentes.

Sol.: a) P(X = 0) = $\frac{2^0}{0!}.e^{-2} = \ldots. = 0{,}1353$

que alguno: P(X > 0) = 1-P(X= 0) = 1- 0,1353 =

= 0,8647

b)k' = 2.k = 4 (por tratarse de dos semanas)

P(X = 4) = $\frac{k'^4}{4!}.e^{-4} = \cdots. = \frac{32}{3}.e^{-4}$

c)P(X= 2).P(X= 2) = …. = $(2.e^{-2})^2$

d)P(X= 0) + P(X=1) + P(X= 2) =

= 0,1353 + 0,2707 + 0,6767

De Distribuciones Hipergeométrica:

Repaso teórico:

Población finita de tamaño N cuyos individuos perteneces exclusivamente a dos clases o categorías: A, A'

Extraemos 'simultaneamente' (o sucesivamente sin reposición) n elemento al azar. Nos preguntamos por la probabilidad de que, entre estos n elementos, se presenten r de la clase A, y lógicamente n-r serán de la clase A'.

Para resolverlo necesitamos conocer un valor o parámetro: p = fracción de elementos que en la población son de clase A, y lógicamente tendremos también, de forma inmediata, la fracción q de los de clase A'. Este valor p suele venir dado en forma de %. El número de elementos de clase A, del total N, es N.p, y de clase A' serán N.q.

Los casos favorables al suceso 'que figuren r de clase A' viene dado por:

$$\binom{N.p}{r}.\binom{N.q}{n-r}$$

y el número de casos posibles es: $\binom{N}{n}$, y por tanto la probabilidad pedida es

$$\mathrm{P(X = r)} = \frac{\left(\binom{N.p}{r}.\binom{N.q}{n-r}\right)}{\binom{N}{n}}$$

Nota:
Los parámetros en esta distribución son:

N, n, p

Media y Varianza en una Distristribución Hipergeométrica:

-Media: $m_x = n.p$ (o bien E[X])

-Varianza: $S^2 = n.p.q.\frac{N-n}{N-1}$ (o bien V[X])

Observaciones:

-Si N es grande en comparación con n, el valor

$$\frac{N-n}{N-1}$$

está próximo a 1, lo cual demuestra que puede ser aproximada mediante la Binomial B(n,p), con media n.p, y desviación n.p.q ;

-Cuando N.p y N.q son grandes con relación a n (ocurre si $\frac{n}{N} <$ 0,01 y N ≥ 60) puede ser aproximada por una binomial con

$$p = \frac{N.p}{N.p+N.q}$$

-Cuando n.p ≥ 4, la v.a. definida mediante

$$Z = \frac{r-n.p}{\sqrt{n.p.q.\frac{N-n}{N-1}}}$$

se distribuye aproximadamente como una ley normal.

1.- Tengo una urna con 20 bolas: 12 blancas, 8 negras.

Experimentos:

A) Hago tres extracciones de una bola cada vez, devolviendo cada vez la bola antes de extraer la siguiente (tres experimentos independientes).

Estamos en el caso de la Binomial:

Probabilidad de blanca: $p = \frac{12}{20}$, Probabilidad de negra: q = $\frac{8}{20}$ = 1-p

B(20; 0,6)

Media= n.p = 3.0,6 = 1,80

Varianza = n.p.q = 3.0,6.0,4 = 0,72

B) Hago una extracción tomando las tres bolas a la vez.

Estamos en el caso de distribución

Hipergeométrica: N = 20, n = 3, p = $\frac{12}{20}$ =

= 0,6 ; q = 0,4 -- > H(20;3;0,6)

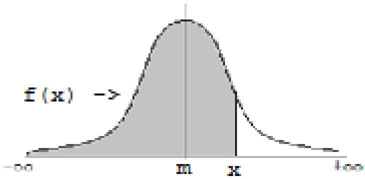
$$P(Y= r) = \frac{\binom{N.p}{r}.\binom{N.q}{n-r}}{\binom{N}{n}} = \frac{\binom{12}{r}.\binom{8}{3-r}}{\binom{20}{3}}$$

Media = n.p = 3.0,6 = 1,80

Varianza = n.p.q.$\frac{N-n}{N-1} = 3.0,6.0,4.\frac{17}{19} = 0,644$

Observa: Var[Y] = Var[X].$\frac{N-n}{N-1}$

Tabla de la Normal tipificada: N(0;1)

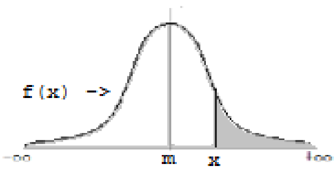


p(X<x) = área sombreada

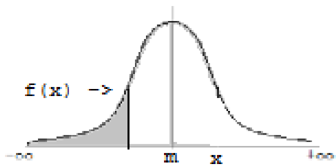
En la siguiente Tabla (pág.248) el valor obtenido es siempre

$$F(x) = p(X < x)$$

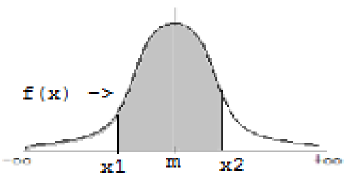
Las siguientes gráficas muestran la casuística. En las págs. 236-240 resolvemos casos concretos.



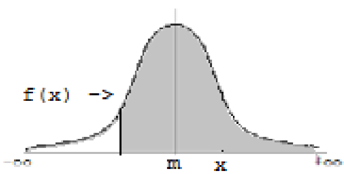
p(X > x)



p(X < x), x < 0



p(x1 < X <x2)



p(X > x), x < 0

Tabla de la normal N(0;1)

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9981	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.8988	.9988	.9989	.9989.	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

BIBLIOGRAFÍA

Análisis Estadístico Aplicado
Autor: Sixto Ríos
Paraninfo, Madrid, año: 1976

Fundamentos de Probabilidad en BioEstadística
Autor: G Alonso, J Ocaña, C.M. Cuadras
2ª Edición
Laboratorio de Cálculos Universidad de Barcelona

Exposición Intuitiva de Métodos Estadísticos
(Fundamentos y Aplicaciones a Biología, Medicina y otras Ciencias)
Autor: M. Purificación Galindo Villardón
Edita: Universidad de Salamanca, Facultad de Biología, Departamento de Ecología

Introducción a la Estadística (Documentos Didácticos)
Coordinador: Ramón Ardanuy Albajar
Edita: Instituto de Ciencias de la Educación, Universidad de Salamanca

Álgebra Binaria de Boole, y sus Aplicaciones Autor: Raoul de Palma
Traducción: Rafael Romero Mercadal
Editorial: Marcombo, S.A. de Boixareu Editores, Barcelona, año: 1973

Programación Lineal
Autor: Laureano F. Escudro
Ediciones Deusto, S.A. Año: 1976

Álgebra Lineal (incluyendo Teoría de Conjuntos),
y Problemas resueltos

Estadística, Correlación. Probabilidad, Variable aleatoria y f.de densidad, Distribuciones. Programación Lineal. …

Autor: Alberto Luzárraga
Editado por el autor, Barcelona 1968

Teoría de Conjuntos y Temas Afines (Teoría y Proble.)
Autor: Seymour Lipschutz
Editorial: Libros McGraw-Hill, 1969,México
Serie compendios SCHAUM

Teoría de Conjuntos y Grafos
Autor: … Lipchit ,
Editado: McGraw-Hill, México 1970

Elementos de Matemáticas
Autor: J. Rey Pastor, A. de Castro
Sociedad Anónima de Traductores y Autores
Madrid, 1967

Análisis Matemático, Vol.I
Autor: J. Rey Pastor, P. P1 Calleja, C.A. Trejo
Editorial Kapelusz, Buenos Aires, Octava edición 1969

NOTACIÓN y Nomenclatura. Valores:

Símbolo	Significado
*	Producto
.	Producto
^	Potencia
sqr(a)	Raíz cuadrada
rad(a)	Raíz cuadrada
rad(a;n)	Radical con índice n
rad(a;n/m)	Radical con índice n/m

∈ significa ‘pertenece a’

∞ infinito

exp(x)	Exponencial: exp(x) = e^x
exp(x;a)	Exponencial de base a>0: exp(x;a) = a^x
ln(x)	Logaritmo neperiano: y = ln(x) <--> x = e^y
log(x;a)	Logaritmo base a>0: y = log(x;a) <--> x = a^y

≅ aproximado

Δ incremento

< menor que, > mayor que, Ej.: $x < y$, $x > y$

Valores:

π = 3,1415927... (número pi, en radianes)

pi = 3,1415927... (número pi, en radianes)

e = 2,7182818... (número e, base de ln(x))

sen(0) = 0 cos(0) = 1

sen(pi/6) = $\frac{1}{2}$ cos(pi/6) = $\frac{\sqrt{3}}{2}$

sen(pi/3) = $\frac{\sqrt{3}}{2}$ cos(pi/3) = $\frac{1}{2}$

sen(pi/2) = 1, cos(pi/2) = 0