Material Docente

ESTADÍSTICA EMPRESARIAL.

Apuntes y ejercicios desarrollados en clase.

Tarancón Morán, Miguel Ángel.
Quintana Rojo, Consolación.
Rojas Peña, Daniel.

2023

Material Docente.

ESTADÍSTICA EMPRESARIAL.

Este manual ha sido realizado por los profesores del área de Estadística Empresarial de la Facultad de Derecho y Ciencias Sociales de Ciudad Real. Universidad de Castilla-La Mancha:

MIGUEL ÁNGEL TARANCÓN MORÁN

Catedrático de Universidad.
Personal Docente e Investigador de la Facultad de Derecho y Ciencias Sociales de Ciudad Real. Universidad de Castilla-La Mancha.
Departamento de Economía Aplicada I.

CONSOLACIÓN QUINTANA ROJO

Doctora en Economía y Empresa.
Personal Docente e Investigador de la Facultad de Derecho y Ciencias Sociales de Ciudad Real. Universidad de Castilla-La Mancha.
Departamento de Economía Aplicada I.

DANIEL ROJAS PEÑA

Profesor Asociado de la Facultad de Derecho y Ciencias Sociales de Ciudad Real.
Universidad de Castilla-La Mancha.
Departamento de Economía Aplicada I.

ÁREA DE ESTADÍSTICA EMPRESARIAL. DEPARTAMENTO DE ECONOMÍA APLICADA I. FACULTAD DE DERECHO Y CCSS DE CIUDAD REAL (UCLM).

Estadística Empresarial. Material Docente.

INTRODUCCIÓN

Presentación

Este documento es una recopilación de las diapositivas y ejercicios realizados en clase de la asignatura obligatoria Estadística Empresarial impartida en el primer curso del Grado en Administración y Dirección de Empresas y Doble Grado en Derecho y Administración y Dirección de Empresas de la Facultad de Derecho y Ciencias Sociales de Ciudad Real. Universidad de Castilla-La Mancha.

ÁREA DE ESTADÍSTICA EMPRESARIAL. DEPARTAMENTO DE ECONOMÍA APLICADA I. FACULTAD DE DERECHO Y CCSS DE CIUDAD REAL (UCLM).

TEORÍA

BLOQUE TEÓRICO

En este apartado encontrarás el material teórico usado para impartir la docencia en esta asignatura.

ÁREA DE ESTADÍSTICA EMPRESARIAL. DEPARTAMENTO DE ECONOMÍA APLICADA I. FACULTAD DE DERECHO Y CCSS DE CIUDAD REAL (UCLM).

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 1.1

Estadística Económica y Empresarial

Contenido

1. Qué es la Estadística.
2. Conceptos básicos.
3. Distribuciones de frecuencias.
4. Representaciones gráficas.

- Ciencia Instrumental.

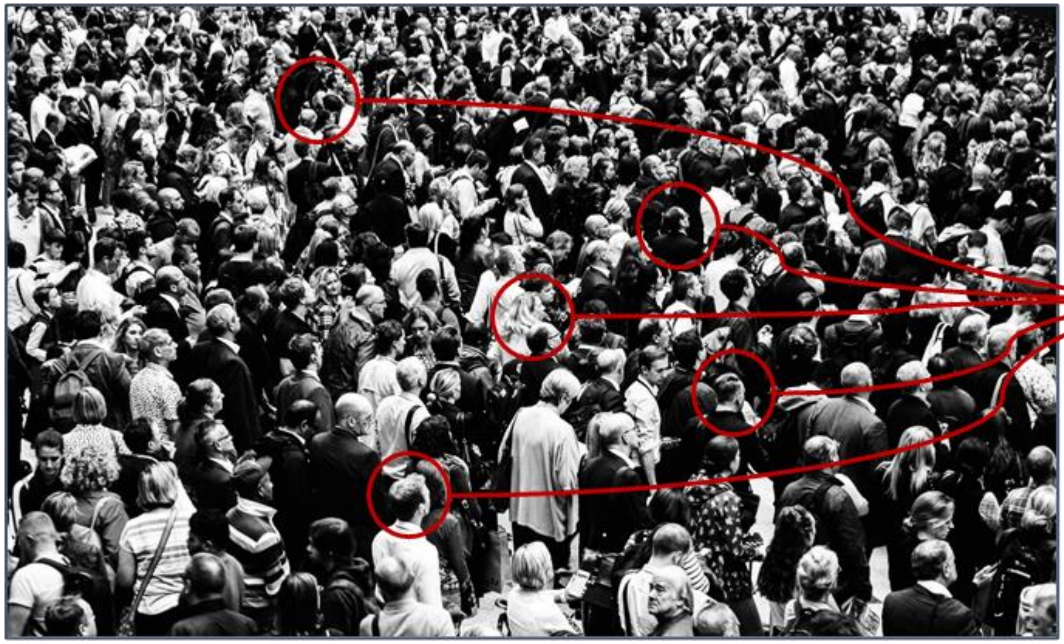
- Objeto:

  1. Recogida y análisis de datos que describan la realidad.
  2. Modelizar el comportamiento de tal realidad cuando existe incertidumbre.

- ¿Para qué?

  Toma de **decisiones**.

2.1. Conceptos básicos: población y muestra.



Población: conjunto de casos que poseen una determinada **característica**.
Ej.: personas asalariadas en España.
Cada elemento: individuo o caso.

Muestra: subconjunto de la población elegido en términos de **representatividad**.

2.2. Conceptos básicos: método estadístico.

Queremos estudiar una **característica** que afecta a todos los casos de una población.
Seleccionamos una **muestra** representativa de la población.
Medimos la característica en la muestra.
Extendemos las conclusiones a toda la población.

Método Inferencial: De lo particular a lo general.

2.3. Conceptos básicos: Estadística Descriptiva.

Objetivo:

Conocer el comportamiento de un grupo de casos (que pueden ser una muestra o una población) de acuerdo a una o varias características.

Pasos:

1. Medir las características de interés en los casos: obtención de datos.
2. Organizar los datos. Presentarlos de modo sistemático.
3. Resumir los datos mediante diferentes medidas: extracción de información.

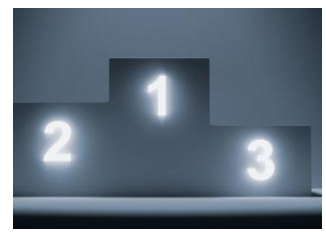
2.4. Conceptos básicos: variables, atributos, escalas.

Variables.
Características que afectan a los casos y que se concretan en valores numéricos: salario percibido.

Escala métrica.

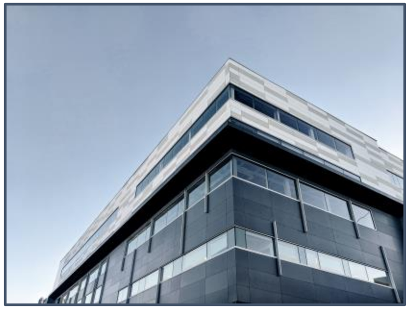
Atributos (variables cualitativas, factores).

Características que afectan a los casos y que NO se concretan en valores numéricos: nivel de estudios, sector productivo al que se pertenece, color de los ojos.



Escala ordinal.

Escala nominal.



Empresa con plantilla de 49 trabajadores.

1. Medir características / recopilar datos sobre casos (encuesta, bases de datos…).
2(a). Organizar datos (tabular).

Apellidos	Nombre	Salario (en cientos de euros)	Nivel de Estudios	Departamento
López Ruiz	Andrés	8	Básicos	Almacén
Martínez Díaz	Ángela	8	Medios	Almacén
Alonso Rodríguez	Miguel	9	Básicos	Almacén
Mascaraque Marqués	Luis	9	Básicos	Almacén
García Santacruz	María	9	Medios	Gestión Interna
León Ramírez	Lourdes	10	Medios	Gestión Interna
Burgos Maestre	Carlos	10	Medios	Gestión Interna
Martín Pescador	Adriana	10	Medios	Gestión Interna
Cabello Rubio	Ricardo	10	Medios	Gestión Proveedores
López Trujillo	Juan José	10	Medios	Gestión Proveedores
Guzmán Valdepeñas	Daniel	10	Medios	Gestión Proveedores
Ledesma Cáceres	Pedro	12	Básicos	Almacén
Madrid Colchonero	Ana	12	Básicos	Almacén
Ruiz Menéndez	Isabel	12	Básicos	Almacén
Malasaña Teruel	Manuel	12	Medios	Gestión Proveedores
Huesca Gutiérrez	Isidro	12	Medios	Gestión Proveedores
Lorenzo Carrascal	Carla	12	Medios	Gestión Proveedores
Buenavista Toledo	Fernando	12	Universitarios	Gestión Clientes
Cañada Mayor	Belén	12	Universitarios	Marketing
Fernández Sempere	Carla	15	Básicos	Almacén
Alegría Sánchez	Andrea	15	Básicos	Almacén
Pastor Céspedes	Pablo	15	Medios	Gestión Interna
Cuenca García	Celia	15	Medios	Gestión Interna
Ruidera Lagos	Alba	15	Medios	Gestión Interna
Mercante Arias	Nicolás	15	Medios	Gestión Proveedores
Ruiz Ruiz	David	15	Medios	Gestión Proveedores
Pérez Castor	Elena	15	Medios	Gestión Proveedores
Huertas Paredes	Victoria	15	Medios	Gestión Proveedores
Miño Portugal	Antonio	15	Universitarios	Gestión Clientes
Bernal Saavedra	Tomás	15	Universitarios	Gestión Clientes
Solana Galán	Bartolomé	20	Universitarios	Almacén
Navas Laguna	Irene	20	Universitarios	Gestión Proveedores
Rodero Carril	Guadalupe	20	Universitarios	Gestión Proveedores
López Megía	Ignacio	20	Universitarios	Gestión Proveedores
Montes Laderas	Ernesto	20	Universitarios	Marketing
Llanos Toboso	Abel	20	Universitarios	Gestión Clientes
Blanco Atienza	Nieves	20	Universitarios	Gestión Clientes
Martos Vilches	Carolina	20	Universitarios	Gestión Clientes
Bilbao Lugo	Luisa	22	Universitarios	Gestión Clientes
Mediodía Gracia	Alberto	22	Universitarios	Marketing
Hernández Ginés	Paula	22	Universitarios	Almacén
Fuentes Ojeda	Sergio	22	Universitarios	Gestión Proveedores
Sevilla Fernández	Estrella	22	Universitarios	Gestión Interna
Ruiz Martín	Alicia	22	Universitarios	Coordinación
Canales Fernández	Marta	25	Universitarios	Dirección
Robles Encina	Alfonso	25	Universitarios	Dirección
Olivares Andújar	Mar	25	Universitarios	Dirección
Pérez Tejedor	Martín	25	Universitarios	Dirección
Segovia Menéndez	Blanca	30	Universitarios	Dirección

3. Frecuencias

Empresa con plantilla de 49 trabajadores.

2(b). Sistematizar los datos: frecuencias.

- x_i: valores que puede tomar la variable (categorías o niveles, si se trata de atributo).
- n_i: frecuencias absolutas.
- $N = \sum_i n_i$: frecuencia total.
- $N_i = \sum_{j=1}^{i} n_i$: frecuencias absolutas acumuladas.
- $f_i = \frac{x_i}{N}$: frecuencias relativas.
- $F_i = \sum_{j=1}^{i} f_i$: frecuencias relativas acumuladas.

¿Tienen sentido las frecuencias acumuladas en la última tabla? ¿Por qué?

Estas tablas: **distribuciones de frecuencias** (en especial, las dos primeras columnas).

x_i = Salario	n_i	N_i	f_i	F_i
8	2	2	0,0408	0,0408
9	3	5	0,0612	0,1020
10	6	11	0,1224	0,2245
12	8	19	0,1633	0,3878
15	11	30	0,2245	0,6122
20	8	38	0,1633	0,7755
22	6	44	0,1224	0,8980
25	4	48	0,0816	0,9796
30	1	49	0,0204	1,0000
N=	49		1,0000	

x_i = N. Estudios	n_i	N_i	f_i	F_i
Básicos	8	8	0,1633	0,1633
Medios	18	26	0,3673	0,5306
Universitarios	23	49	0,4694	1,0000
N=	49		1,0000	

x_i = Depart.	n_i	N_i	f_i	F_i
Almacén	11	11	0,2245	0,2245
Gestión Interna	8	19	0,1633	0,3878
Gestión Proveedores	14	33	0,2857	0,6735
Gestión Clientes	7	40	0,1429	0,8163
Marketing	3	43	0,0612	0,8776
Coordinación	1	44	0,0204	0,8980
Dirección	5	49	0,1020	1,0000
N=	49		1,0000	

3. Frecuencias

A veces, se resumen los datos por intervalos de valores: **distribuciones de frecuencias agrupadas en intervalos**.

- L_{i-1}: extremo inferior del intervalo.
- L_i: extremo superior del intervalo.
- n_i: frecuencias absolutas.
- $x_i = \frac{L_{i-1}+L_i}{2}$: marca de clase.
- $c_i = L_i - L_{i-1}$: Amplitud de intervalo.
- $d_i = \frac{n_i}{c_i}$: densidad de intervalo.

Salvo el primero, los intervalos son: $(L_{i-1}, L_i]$

x_i = Salario	n_i	N_i	f_i	F_i
8	2	2	0,0408	0,0408
9	3	5	0,0612	0,1020
10	6	11	0,1224	0,2245
12	8	19	0,1633	0,3878
15	11	30	0,2245	0,6122
20	8	38	0,1633	0,7755
22	6	44	0,1224	0,8980
25	4	48	0,0816	0,9796
30	1	49	0,0204	1,0000
N=	49		1,0000	

L_{i-1}	L_i	n_i	x_i	c_i	d_i
0	10	11	5	10	1,1000
10	20	27	15	10	2,7000
20	30	11	25	10	1,1000
N=		49			

4.1. Representaciones gráficas: Atributos

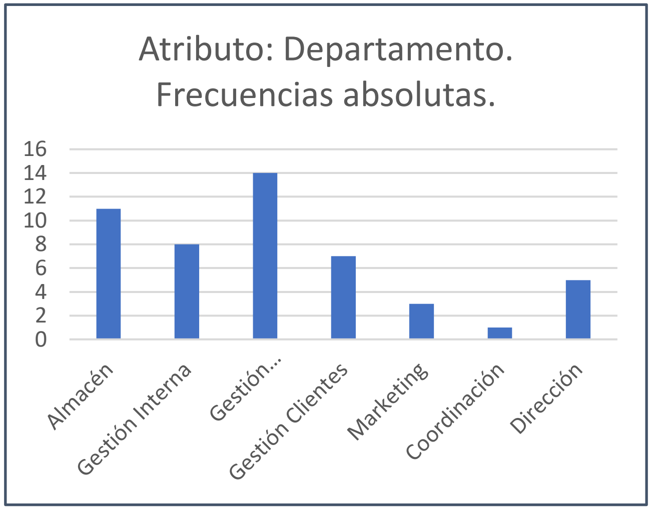


Gráfico de barras

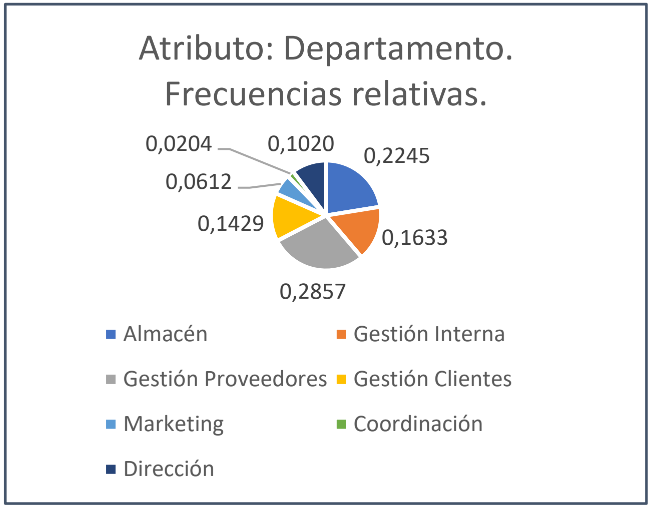


Gráfico de sectores

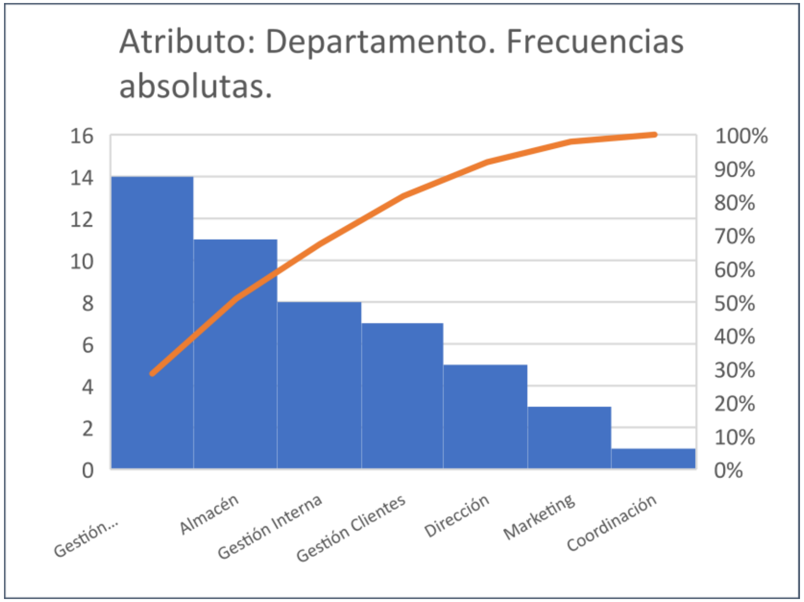


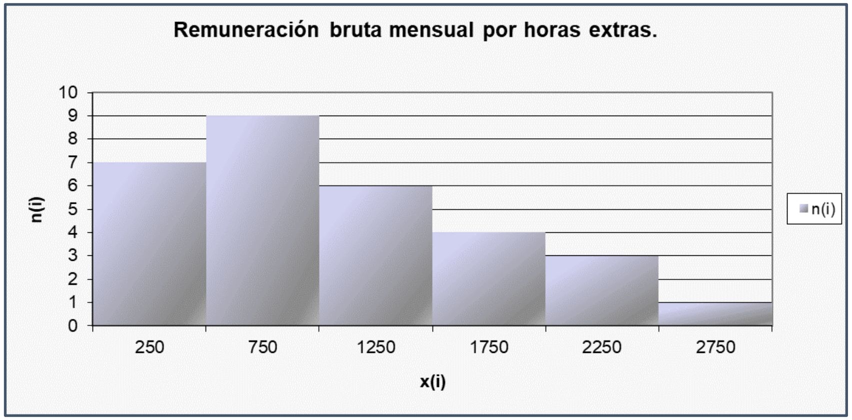
Diagrama de Pareto

4.2. Representaciones gráficas: Histograma

Variables agrupadas en intervalos

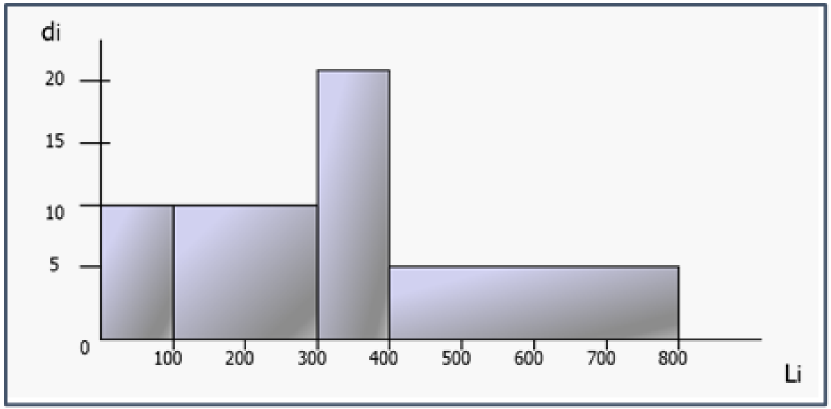
Intervalos de igual amplitud

	n(i)	c(i)
0-500	7	500
500-1000	9	500
1000-1500	6	500
1500-2000	4	500
2000-2500	3	500
2500-3000	1	500
N	30	



Intervalos de distinta amplitud

L(i)-L(i-1)	n(i)	c(i)	d(i)
0-100	1000	100	10
100-300	2000	200	10
300-400	2000	100	20
400-800	2000	400	5




¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 1.2

Variable estadística unidimensional

1. Características de una distribución de frecuencias.
2. Medidas de posición.
3. Momentos respecto al origen y respecto a la media.
4. Medidas de dispersión.
5. Medidas de forma.

1. Características de una distribución de frecuencias.

- Medidas: instrumentos matemáticos que extraen y sintetizan la información contenida en una distribución de frecuencias.

- Muchas de ellas son aplicables al caso de atributos; pero vamos a centrarnos en variables.

x_i = Salario	n_i	N_i	f_i	F_i
8	2	2	0,0408	0,0408
9	3	5	0,0612	0,1020
10	6	11	0,1224	0,2245
12	8	19	0,1633	0,3878
15	11	30	0,2245	0,6122
20	8	38	0,1633	0,7755
22	6	44	0,1224	0,8980
25	4	48	0,0816	0,9796
30	1	49	0,0204	1,0000
N=	49		1,0000	

1. Características de una distribución de frecuencias.

- **Medidas de posición.**
  - Posición central.
    - Media.
    - Mediana.
    - Moda.
  - Cuantiles.
    - Cuartiles.
    - Deciles.
    - Percentiles.
- **Medidas de dispersión.**
  - Absolutas.
    - Varianza.
    - Desviación típica.
  - Relativas.
    - Coeficiente de dispersión.
- **Medidas de forma.**
  - Asimetría.
    - Coeficiente de asimetría de Fisher.
  - Apuntamiento o curtosis.
    - Coeficiente de apuntamiento de Fisher.

2.1. Medidas de posición: media.

- Medida de posición central.
- Solo aplicable a variables.
- Existen 3 **tipos** de medias:
  - **Aritmética:** es la más utilizada y conocida.
  - **Geométrica:** sirve para promediar variaciones acumulativas (finanzas).
  - **Armónica:** sirve para promediar rendimientos, productividades (cocientes entre magnitudes).

x_i = Salario	n_i	N_i	f_i	F_i	$x_i \cdot n_i$
8	2	2	0,0408	0,0408	16
9	3	5	0,0612	0,1020	27
10	6	11	0,1224	0,2245	60
12	8	19	0,1633	0,3878	96
15	11	30	0,2245	0,6122	165
20	8	38	0,1633	0,7755	160
22	6	44	0,1224	0,8980	132
25	4	48	0,0816	0,9796	100
30	1	49	0,0204	1,0000	30
N=	49		1,0000	Suma =	786
				$\bar{x}$=	16,0408

- Media aritmética.
  - $\bar{x} = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + \cdots + x_h \cdot n_h}{N} = \frac{\sum_{i=1}^{h} x_i \cdot n_i}{N}$
  - Ventajas: única, siempre calculable y utiliza todos los valores de la distribución.
  - Inconveniente: queda invalidada ante la existencia de valores extremos (*outliers*).

16,0408 es el valor representativo de la distribución de frecuencias (**centro de gravedad de la distribución**): si todos los trabajadores ganaran 16,0408 cientos de euros, el montante global de los salarios sería igualmente 16,0408 x 49 = 786 cientos de euros.

2.1. Medidas de posición: media.

- ¿Qué ocurre cuando la distribución de frecuencias está agrupada en intervalos?

- En lugar de los valores de la variable (ya no están especificados) hay que utilizar las **marcas de clase**.

- Esto da lugar a un **error de agrupamiento**.

x_i = Salario	n_i	N_i	f_i	F_i	$x_i \cdot n_i$
8	2	2	0,0408	0,0408	16
9	3	5	0,0612	0,1020	27
10	6	11	0,1224	0,2245	60
12	8	19	0,1633	0,3878	96
15	11	30	0,2245	0,6122	165
20	8	38	0,1633	0,7755	160
22	6	44	0,1224	0,8980	132
25	4	48	0,0816	0,9796	100
30	1	49	0,0204	1,0000	30
N=	49		1,0000	Suma =	786
				$\bar{x}$=	16,0408

L_{i-1}	L_i	x_i	n_i	c_i	d_i	$x_i \cdot n_i$
0	10	5	11	10	1,1000	55
10	20	15	27	10	2,7000	405
20	30	25	11	10	1,1000	275
		N=	49		Suma =	735
					$\bar{x}$=	15,0000

15,00 cientos de euros es el valor representativo de la distribución de frecuencias (**centro de gravedad de la distribución**): pero en este caso, la media por la frecuencia total no coincide con la suma de los salarios: 15,00 x 49 = 73,50 cientos de euros, cuando vimos que si no se agrupan las observaciones, la suma de salarios es 78,60 cientos de euros: error de agrupamiento.

2.2. Medidas de posición: mediana.

- Medida de posición central.
- Aplicable a variables y atributos en escala ordinal.
- Valor de la variable que toma el caso que divide a la distribución de frecuencias en **dos grupos con el mismo número de casos**.
  - Si N es impar, la mediana es el valor que toma el caso en la posición $\frac{N+1}{2}$
  - Si N es par, puede haber dos medianas: los valores que toman los casos en las posiciones $\frac{N}{2}$ y $\frac{N}{2}+1$. En ese caso, se toma como única mediana el promedio de ambas.
- Buena medida de posición central ante la existencia de valores extremos (*outliers*).

x_i = Salario	n_i	N_i
8	2	2
9	3	5
10	6	11
12	8	19
15	11	30
20	8	38
22	6	44
25	4	48
30	1	49
N=	49	Me = 15

En este caso **N es impar** (49), por lo que habrá un valor mediano, correspondiente al caso que ocupa la posición: (49+1)/2=25. Ayudándonos de la frecuencia absoluta acumulada, vemos que el valor de la variable del caso que ocupa la posición 25 es **15,00 cientos de euros**.

2.2. Medidas de posición: mediana.

- ¿Qué ocurre cuando la distribución de frecuencias está agrupada en intervalos?
- Se procede del mismo modo para localizar el intervalo mediano
  - Si N es impar, el intervalo mediano es aquel en el que toma valor el caso en la posición $\frac{N+1}{2}$
  - Si N es par, puede haber dos intervalos medianos: aquellos en los que toman valores los casos en las posiciones $\frac{N}{2}$ y $\frac{N}{2}+1$.
- Dentro de los intervalos medianos, las medianas se calcularán con las fórmulas:
  - N impar: $Me = L_{i-1} + \frac{\frac{N+1}{2} - N_{i-1}}{n_i} c_i$
  - N par: $Me = L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} c_i$ y

    $Me = L_{i-1} + \frac{\frac{N}{2} + 1 - N_{i-1}}{n_i} c_i$

L_{i-1}	L_i	x_i	n_i	N_i	c_i
0	10	5	11	11	10
10	20	15	27	38	10
20	30	25	11	49	10
		N=	49		Me=15,1852

$$Me = L_{i-1} + \frac{\frac{N+1}{2} - N_{i-1}}{n_i} c_i$$

$$= 10 + \frac{\frac{49+1}{2} - 11}{27} \cdot 10 = 15{,}1852$$

En el caso de que **N sea par**, se hará lo mismo con las fórmulas correspondientes a los casos $\frac{N}{2}$ y $\frac{N}{2}+1$, obteniéndose dos medianas, que se promedian para obtener una única mediana.

- Medida de posición central.
- Aplicable a variables y atributos en escala ordinal y nominal.
- Valor de la variable más repetido en la distribución.
- Mejor medida de posición central para atributos en escala nominal.

x_i = Salario	n_i	N_i
8	2	2
9	3	5
10	6	11
12	8	19
15	11	30
20	8	38
22	6	44
25	4	48
30	1	49
N=	49	Mo = 15

- ¿Qué ocurre cuando la distribución de frecuencias está agrupada en intervalos?
- Depende de si los intervalos tienen o no la misma amplitud.

  - Si todos los intervalos tienen la misma amplitud, se buscará aquel que tenga una mayor frecuencia absoluta (intervalo modal), y se obtendrá la moda con la fórmula:

$$Mo = L_{i-1} + \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} \cdot c_i$$

  - Si NO todos los intervalos tienen la misma amplitud, se buscará aquel que tenga una mayor densidad de frecuencia (intervalo modal), y se obtendrá la moda con la fórmula:

$$Mo = L_{i-1} + \frac{d_i - d_{i-1}}{(d_i - d_{i-1}) + (d_i - d_{i+1})} \cdot c_i$$

L_{i-1}	L_i	x_i	n_i	c_i	d_i	Mo
0	10	5	11	10	1,1000	
10	20	15	27	10	2,7000	15,0000
20	30	25	11	10	1,1000	
		N=	49			

$$Mo = L_{i-1} + \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} \cdot c_i$$

$$= 10 + \frac{27 - 11}{(27 - 11) + (27 - 11)} \cdot 10 = 15$$

2.4. Medidas de posición: cuantiles.

- Medidas de **posición no-central**.

- Generalización de la mediana.

- Valores de la variable que dividen a la distribución en un determinado número de grupos, de manera que todos contienen el mismo número de casos.

- Más habituales:
  - Cuartiles.
  - Deciles.
  - Percentiles.

3. Momentos respecto al origen y respecto a la media.

- Los “momentos” son expresiones matemáticas que, respecto a un origen arbitrario O_t , se definen como:

$$M_r = \frac{\sum_{i=1}^{h}(x_i - O_t)^r \cdot n_i}{N}$$

- Si $O_t = 0$: momentos respecto al origen:

$$a_r = \frac{\sum_{i=1}^{h} {x_i}^r \cdot n_i}{N} \qquad ¿a_1?$$

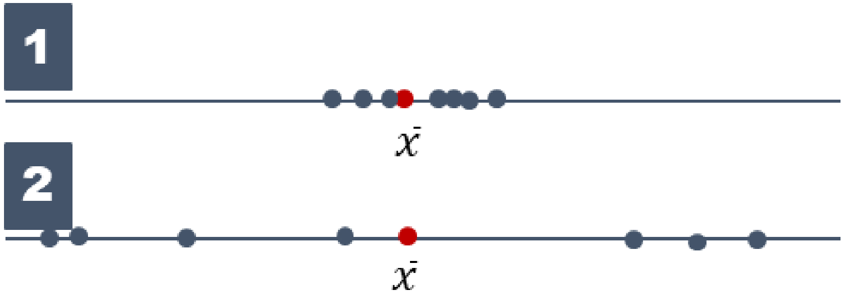
- Si $O_t = \bar{x}$: momentos respecto a la media:

$$m_r = \frac{\sum_{i=1}^{h}(x_i - \bar{x})^r \cdot n_i}{N}$$

- Propiedad:

$$m_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} a_1^k a_{r-k}$$

- Supongamos 2 distribuciones de frecuencias de una misma variable, que tienen la misma media:



- ¿En qué distribución la media es más representativa de los valores que toman los casos o individuos?

- Cuantificar esta representatividad de las medidas de posición central: **medidas de dispersión.**

- Si la distribución está agrupada en intervalos: utilizar marcas de clase.

- **Varianza:**

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot n_i}{N}$$

  - Media de la distancia al cuadrado de los valores que toman los casos con respecto a la media.

  - $S^2 = m_2 = a_2 - a_1^2 = \frac{\sum_{i=1}^{N} x_i^2 n_i}{N} - (\bar{x})^2$

- **Desviación típica:**

$$S = +\sqrt{S^2}$$

- Varianza y desviación típica tienen dos inconvenientes:
  - **No están acotadas** superiormente.
  - Medidas **absolutas**.

4.2. Medidas de dispersión: Varianza y Desviación Típica.

- **Varianza:**

$$S^2 = m_2 = a_2 - a_1^2 = \frac{\sum_{i=1}^{N} x_i^2 n_i}{N} - (\bar{x})^2 = 287{,}7959 - (16{,}0408)^2 =$$
$30{,}4881$ cientos de euros al cuadrado.

- **Desviación típica:**

$$S = +\sqrt{S^2} = \sqrt{30{,}4881} = 5{,}5216$$
cientos de euros.

- ¿Unidades de ambas medidas?

- ¿La varianza / desviación típica son suficientemente pequeñas para afirmar que la media aritmética es representativa de la distribución de frecuencias? ¿Por qué?

x_i	n_i	$x_i \cdot n_i$	$x_i^2 \cdot n_i$	$x_i^3 \cdot n_i$	$x_i^4 \cdot n_i$
8	2	16	128	1024	8192
9	3	27	243	2187	19683
10	6	60	600	6000	60000
12	8	96	1152	13824	165888
15	11	165	2475	37125	556875
20	8	160	3200	64000	1280000
22	6	132	2904	63888	1405536
25	4	100	2500	62500	1562500
30	1	30	900	27000	810000
N=	49	786	14102	277548	5868674
		a_1	a_2	a_3	a_4
		$\bar{x}$= 16,0408	287,7959	5664,2449	119768,8571

- ¿Cuál de estas dos empresas tiene un salario medio más representativo?

x_i: salario (cientos de euros)

x_i	n_i	$x_i \cdot n_i$	$x_i^2 \cdot n_i$	$x_i^3 \cdot n_i$	$x_i^4 \cdot n_i$
8	2	16	128	1024	8192
9	3	27	243	2187	19683
10	6	60	600	6000	60000
12	8	96	1152	13824	165888
15	11	165	2475	37125	556875
20	8	160	3200	64000	1280000
22	6	132	2904	63888	1405536
25	4	100	2500	62500	1562500
30	1	30	900	27000	810000
N=	49	786	14102	277548	5868674
		a_1	a_2	a_3	a_4
		$\bar{x}$= 16,0408	287,7959	5664,2449	119768,8571
$S^2 = m_2 =$	30,4881				
$S =$	5,5216				

y_i: salario (cientos de euros)

y_i	n_i	$y_i \cdot n_i$	$y_i^2 \cdot n_i$	$y_i^3 \cdot n_i$	$y_i^4 \cdot n_i$
8	3	24	192	1536	12288
9	4	36	324	2916	26244
10	5	50	500	5000	50000
12	7	84	1008	12096	145152
15	10	150	2250	33750	506250
20	7	140	2800	56000	1120000
22	5	110	2420	53240	1171280
25	4	100	2500	62500	1562500
30	4	120	3600	108000	3240000
N=	49	814	15594	335038	7833714
		a_1	a_2	a_3	a_4
		$\bar{y}$= 16,6122	318,2449	6837,5102	159871,7143
$S^2 = m_2 =$	42,2782				
$S =$	6,5022				

4.3. Medidas de dispersión: Coeficiente de dispersión.

- ¿Y si en la primera empresa se paga en florines húngaros (1 euro = 236 florines)?

x_i: salario (miles de florines)

x_i	n_i	$x_i \cdot n_i$	$x_i^2 \cdot n_i$	$x_i^3 \cdot n_i$	$x_i^4 \cdot n_i$
188,8	2	377,6	71290,9	13459718,1	2541194785,6
212,4	3	637,2	135341,3	28746487,9	6105754024,0
236	6	1416	334176,0	78865536,0	18612266496,0
283,2	8	2265,6	641617,9	181706194,9	51459194408,1
354	11	3894	1378476,0	487980504,0	172745098416,0
472	8	3776	1782272,0	841232384,0	397061685248,0
519,2	6	3115,2	1617411,8	839760227,3	436003510028,7
590	4	2360	1392400,0	821516000,0	484694440000,0
708	1	708	501264,0	354894912,0	251265597696,0
N=	49	18549,6	7854249,9	3648161964,3	1820488741102,4
		a_1	a_2	a_3	a_4
	$\bar{x}=$	378,6	160290,8	74452285,0	37152831451,1
$S^2 = m_2 =$ 16980,6689					
$S =$ 130,3099					

y_i: salario (cientos de euros)

y_i	n_i	$y_i \cdot n_i$	$y_i^2 \cdot n_i$	$y_i^3 \cdot n_i$	$y_i^4 \cdot n_i$
8	3	24	192	1536	12288
9	4	36	324	2916	26244
10	5	50	500	5000	50000
12	7	84	1008	12096	145152
15	10	150	2250	33750	506250
20	7	140	2800	56000	1120000
22	5	110	2420	53240	1171280
25	4	100	2500	62500	1562500
30	4	120	3600	108000	3240000
N=	49	814	15594	335038	7833714
		a_1	a_2	a_3	a_4
	$\bar{y}=$	16,6122	318,2449	6837,5102	159871,7143
$S^2 = m_2 =$ 42,2782					
$S =$ 6,5022					

- ¿Cuál de estas dos empresas tiene un salario medio más representativo ahora?

4.3. Medidas de dispersión: Coeficiente de dispersión.

- La estructura salarial es la misma en la primera empresa, se pague en euros o en florines; pero cambia el resultado de la comparación: efecto de tener diferentes unidades.
- **Medidas relativas**: sin unidades.
- **Coeficiente de dispersión** de Pearson: $V = \frac{S}{\bar{x}}$
- ¿Interpretación?

x_i: salario (miles de florines)

x_i	n_i	$x_i \cdot n_i$	$x_i^2 \cdot n_i$	$x_i^3 \cdot n_i$	$x_i^4 \cdot n_i$
188,8	2	377,6	71290,9	13459718,1	2541194785,6
212,4	3	637,2	135341,3	28746487,9	6105754024,0
236	6	1416	334176,0	78865536,0	18612266496,0
283,2	8	2265,6	641617,9	181706194,9	51459194408,1
354	11	3894	1378476,0	487980504,0	172745098416,0
472	8	3776	1782272,0	841232384,0	397061685248,0
519,2	6	3115,2	1617411,8	839760227,3	436003510028,7
590	4	2360	1392400,0	821516000,0	484694440000,0
708	1	708	501264,0	354894912,0	251265597696,0
N=	49	18549,6	7854249,9	3648161964,3	1820488741102,4
		a_1	a_2	a_3	a_4
	$\bar{x}$=	378,6	160290,8	74452285,0	37152831451,1
$S^2 = m_2 =$ 16980,6689					
S = 130,3099			V= 0,3444		

y_i: salario (cientos de euros)

y_i	n_i	$y_i \cdot n_i$	$y_i^2 \cdot n_i$	$y_i^3 \cdot n_i$	$y_i^4 \cdot n_i$
8	3	24	192	1536	12288
9	4	36	324	2916	26244
10	5	50	500	5000	50000
12	7	84	1008	12096	145152
15	10	150	2250	33750	506250
20	7	140	2800	56000	1120000
22	5	110	2420	53240	1171280
25	4	100	2500	62500	1562500
30	4	120	3600	108000	3240000
N=	49	814	15594	335038	7833714
		a_1	a_2	a_3	a_4
	$\bar{y}$=	16,6122	318,2449	6837,5102	159871,7143
$S^2 = m_2 =$ 42,2782					
S = 6,5022			V= 0,3914		

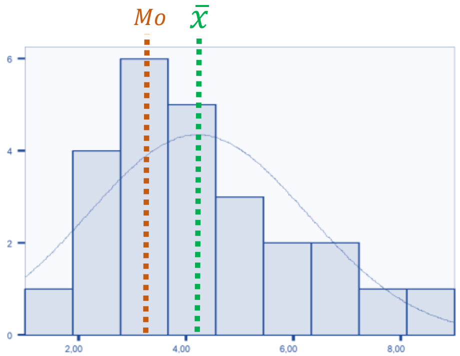
¿Cuánto valdrá *V* en la primera empresa con los salarios en cientos de euros?

- Miden el grado de deformación vertical y horizontal de la representación gráfica de una distribución de frecuencias.

- Si las frecuencias están agrupadas en intervalos, se trabaja con marcas de clase.

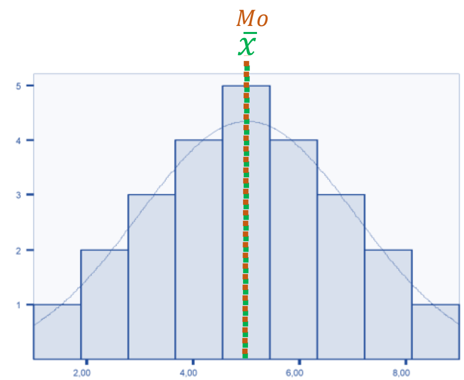
- Dos tipos:
  - **Medidas de asimetría:** deformación vertical.
  - **Medidas de apuntamiento o curtosis:** deformación horizontal.

- Miden el grado de **deformación vertical** con respecto a un "eje de simetría", que es aquel que pasa por el valor medio de la distribución.



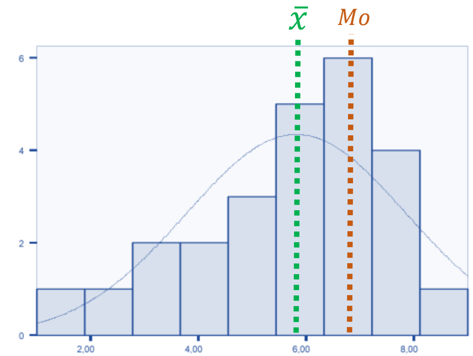
Distribución asimétrica positiva

$Mo < \bar{x}$



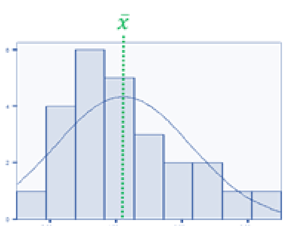
Distribución simétrica

$Mo = \bar{x}$



Distribución asimétrica negativa

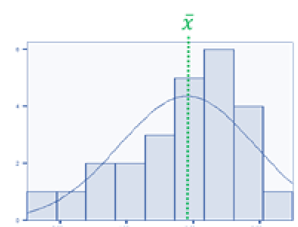
$Mo > \bar{x}$



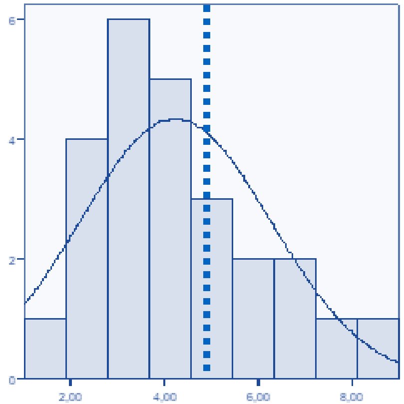
+

$$m_3 = \sum_{i=1}^{n} (x_i - \dot{x})^3 \frac{n_i}{N}$$

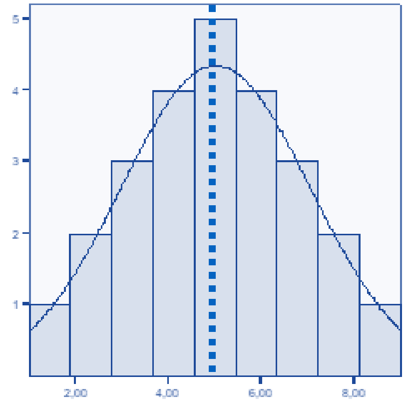
-



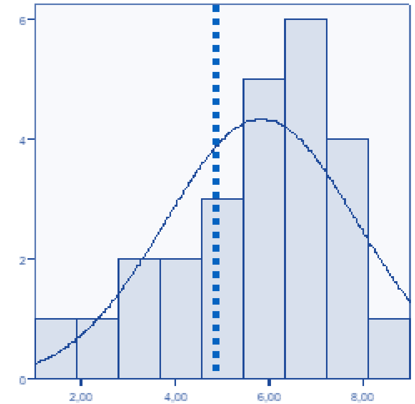
5.2. Medidas de forma: Asimetría.



Distribución
asimétrica positiva



Distribución
simétrica



Distribución
asimétrica negativa

Para poder comparar distribuciones: medida relativa.

Coeficiente de asimetría de Fisher:

$$g_1 = \frac{m_3}{S^3}$$

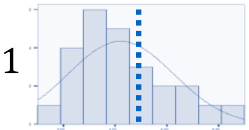
$g_1 > 0$ $\quad$ $g_1 = 0$ $\quad$ $g_1 < 0$

$g_1 = 0$ es condición necesaria pero no suficiente para la simetría

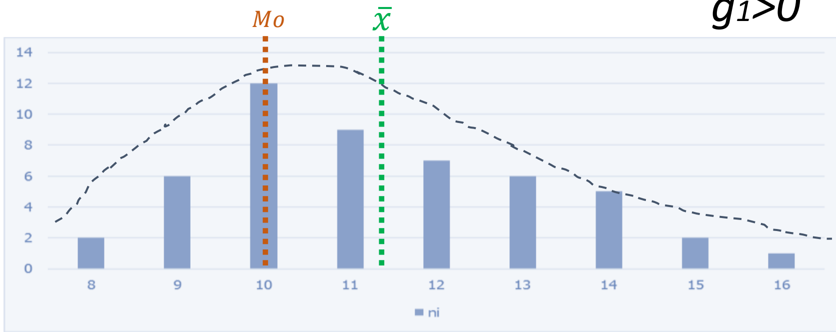
5.2. Medidas de forma: Asimetría.

x_i	n_i	$x_i \cdot n_i$	$(x_i)^2 \cdot n_i$	$(x_i)^3 \cdot n_i$
8	2	16	128	1024
9	6	54	486	4374
10	12	120	1200	12000
11	9	99	1089	11979
12	7	84	1008	12096
13	6	78	1014	13182
14	5	70	980	13720
15	2	30	450	6750
16	1	16	256	4096
N=	50	Suma = 567	Suma = 6611	Suma = 79221
		$a_1 = \bar{x} = 11{,}34$	$a_2 = 132{,}22$	$a_3 = 1584{,}42$

$$g_1 = \frac{m_3}{S^3} = \frac{a_3 - 3a_2a_1 + 2a_1^3}{\left(+\sqrt{a_2 - a_1^2}\right)^3} = 0{,}4121$$

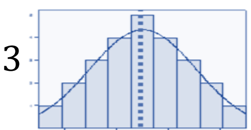


$g_1 > 0$

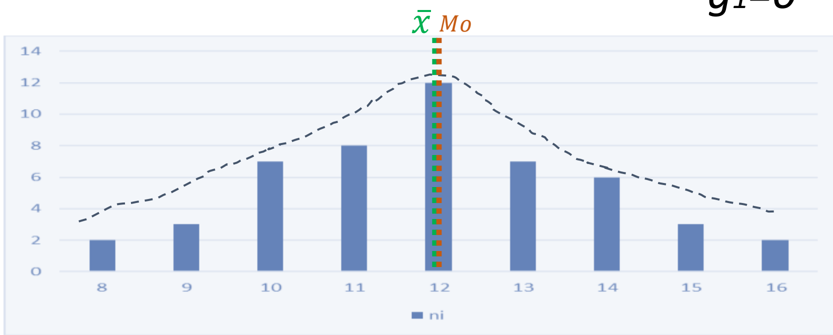


x_i	n_i	$x_i \cdot n_i$	$(x_i)^2 \cdot n_i$	$(x_i)^3 \cdot n_i$
8	2	16	128	1024
9	3	27	243	2187
10	7	70	700	7000
11	8	88	968	10648
12	12	144	1728	20736
13	7	91	1183	15379
14	6	84	1176	16464
15	3	45	675	10125
16	2	32	512	8192
N=	50	Suma = 597	Suma = 7313	Suma = 91755
		$a_1 = \bar{x} = 11{,}94$	$a_2 = 146{,}26$	$a_3 = 1835{,}10$

$$g_1 = \frac{m_3}{S^3} = \frac{a_3 - 3a_2a_1 + 2a_1^3}{\left(+\sqrt{a_2 - a_1^2}\right)^3} = 0{,}0683$$

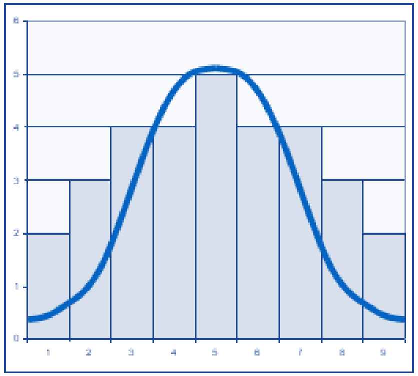


$g_1 = 0$

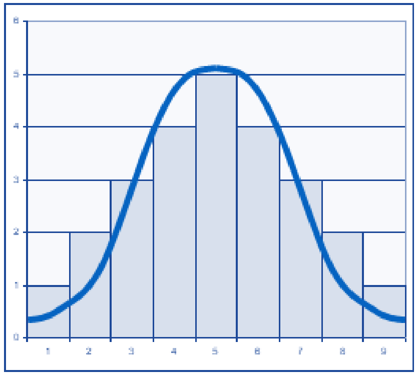


5.3. Medidas de forma: apuntamiento o curtosis.

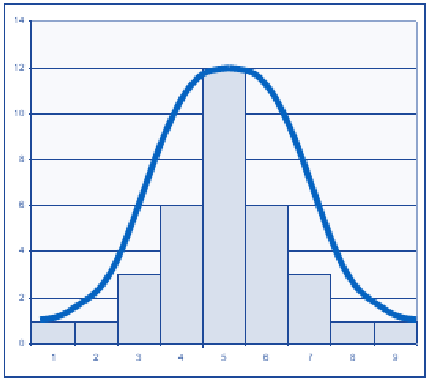
- Miden el grado de deformación horizontal con respecto a una distribución "tipo", la distribución normal.
- Se supone previamente que la distribución es campaniforme, unimodal, y simétrica o con ligera asimetría.



Distribución platicúrtica



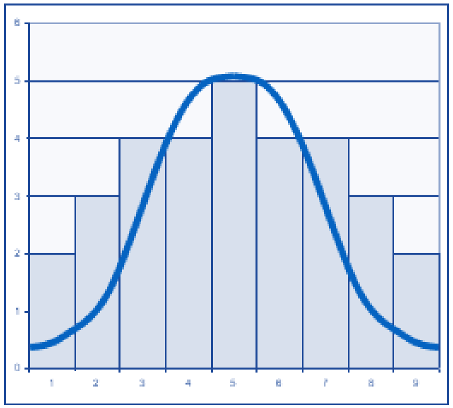
Distribución mesocúrtica



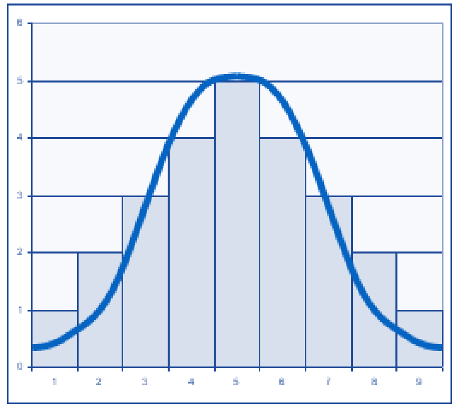
Distribución leptocúrtica

- Para encontrar una medida de apuntamiento, Fisher se da cuenta de que, **en la distribución normal** $m_4 = 3S^4$, con lo que: $\frac{m_4}{S^4} - 3 = 0$.

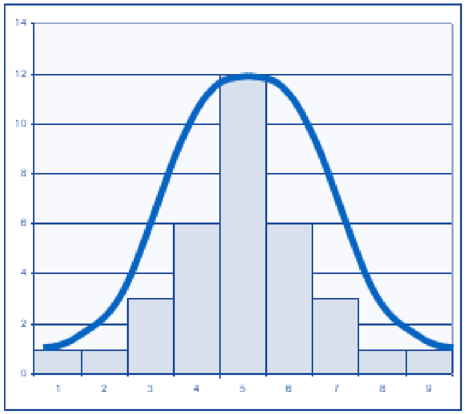
5.3. Medidas de forma: apuntamiento o curtosis.



Distribución
platicúrtica



Distribución
mesocúrtica



Distribución
leptocúrtica

Así obtiene una medida relativa, el **coeficiente de apuntamiento de Fisher:**

$$g_2 = \frac{m_4}{S^4} - 3$$

$g_2 < 0$

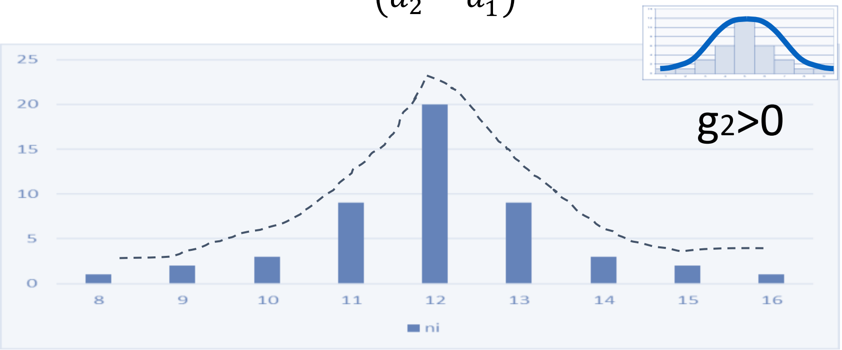
$g_2 = 0$

$g_2 > 0$

5.3. Medidas de forma: apuntamiento o curtosis.

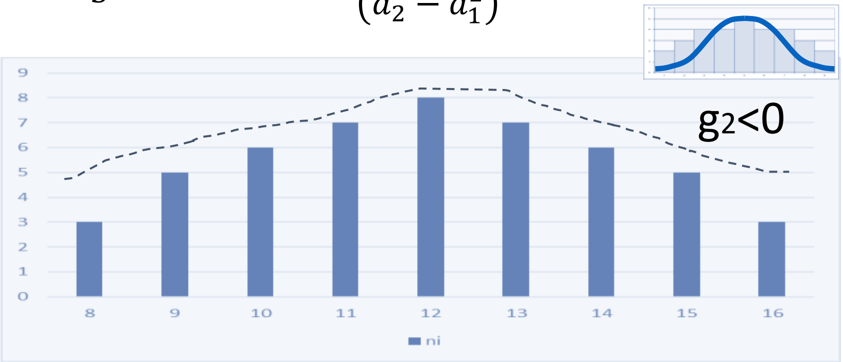
x_i	n_i	$x_i \cdot n_i$	$(x_i)^2 \cdot n_i$	$(x_i)^3 \cdot n_i$	$(x_i)^4 \cdot n_i$
8	1	8	64	512	4096
9	2	18	162	1458	13122
10	3	30	300	3000	30000
11	9	99	1089	11979	131769
12	20	240	2880	34560	414720
13	9	117	1521	19773	257049
14	3	42	588	8232	115248
15	2	30	450	6750	101250
16	1	16	256	4096	65536
N=	50	Suma = 600	Suma = 7310	Suma = 90360	Suma = 1132790
		$a_1 = \bar{x} = 12{,}00$	$a_2 = 146{,}20$	$a_3 = 1807{,}20$	$a_4 = 22655{,}80$

$$g_2 = \frac{m_4}{S^4} - 3 = \frac{a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4}{\left(a_2 - a_1^2\right)^2} - 3 = 0{,}9256$$



x_i	n_i	$x_i \cdot n_i$	$(x_i)^2 \cdot n_i$	$(x_i)^3 \cdot n_i$	$(x_i)^4 \cdot n_i$
8	3	24	192	1536	12288
9	5	45	405	3645	32805
10	6	60	600	6000	60000
11	7	77	847	9317	102487
12	8	96	1152	13824	165888
13	7	91	1183	15379	199927
14	6	84	1176	16464	230496
15	5	75	1125	16875	253125
16	3	48	768	12288	196608
N=	50	Suma = 600	Suma = 7448	Suma = 95328	Suma = 1253624
		$a_1 = \bar{x} = 12{,}00$	$a_2 = 148{,}96$	$a_3 = 1906{,}56$	$a_4 = 25072{,}48$

$$g_2 = \frac{m_4}{S^4} - 3 = \frac{a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4}{\left(a_2 - a_1^2\right)^2} - 3 = -0{,}9253$$




¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 2.1 Variable estadística bidimensional y n-dimensional

Contenido

1. Distribución bidimensional de frecuencias.
2. Independencia y dependencia estadística.
3. Momentos bidimensionales.

1. Distribución bidimensional de frecuencias.

- Tenemos 2 características que afectan a un mismo grupo de casos.

- Se podría generalizar para el caso de n características (n-dimensional).

- Uno de los temas más importantes es el de la **independencia o dependencia estadística** entre las variables que integran la variable bidimensional.

- **Dependencia estadística**: coincidencias entre los comportamientos de las variables, lo que puede llevar a pensar que dichos comportamientos están relacionados.

1. Distribución bidimensional de frecuencias.

- Tabla de correlación:

X\Y	y_1	y_2	…	y_j	…	y_k	Marg. x
x_1	n_{11}	n_{12}	…	n_{1j}	…	n_{1k}	$n_{1\bullet}$
x_2	n_{21}	n_{22}	…	n_{2j}	…	n_{2k}	$n_{2\bullet}$
…	…	…	…	…	…	…	…
x_i	n_{i1}	n_{i2}	…	n_{ij}	…	n_{ik}	$n_{i\bullet}$
…	…	…	…	…	…	…	…
x_h	n_{h1}	n_{h2}	…	n_{hj}	…	n_{hk}	$n_{h\bullet}$
Marg. y	$n_{\bullet 1}$	$n_{\bullet 2}$	…	$n_{\bullet j}$	…	$n_{\bullet k}$	N

n_{ij}: frecuencia absoluta conjunta.

$n_{i\bullet} = \sum_{j=1}^{k} n_{ij}$: frecuencia marginal de la variable X.

$n_{\bullet j} = \sum_{i=1}^{h} n_{ij}$: frecuencia marginal de la variable Y.

$$\sum_{i=1}^{h} n_{i\bullet} = \sum_{j=1}^{k} n_{\bullet j} = \sum_{i=1}^{h} \sum_{j=1}^{k} n_{ij} = N$$

1. Distribución bidimensional de frecuencias.

X\Y	y_1	y_2	…	y_j	…	y_k	Marg. x
x_1	$f_{11}=\frac{n_{11}}{N}$	$f_{12}=\frac{n_{12}}{N}$	…	$f_{1j}=\frac{n_{1j}}{N}$	…	$f_{1k}=\frac{n_{1k}}{N}$	$f_{1\bullet}=\frac{n_{1\bullet}}{N}$
x_2	$f_{21}=\frac{n_{21}}{N}$	$f_{22}=\frac{n_{22}}{N}$	…	$f_{2j}=\frac{n_{2j}}{N}$	…	$f_{2k}=\frac{n_{2k}}{N}$	$f_{2\bullet}=\frac{n_{2\bullet}}{N}$
…	…	…	…	…	…	…	…
x_i	$f_{i1}=\frac{n_{i1}}{N}$	$f_{i2}=\frac{n_{i2}}{N}$	…	$f_{ij}=\frac{n_{ij}}{N}$	…	$f_{ik}=\frac{n_{ik}}{N}$	$f_{i\bullet}=\frac{n_{i\bullet}}{N}$
…	…	…	…	…	…	…	…
x_h	$f_{h1}=\frac{n_{h1}}{N}$	$f_{h2}=\frac{n_{h2}}{N}$	…	$f_{hj}=\frac{n_{hj}}{N}$	…	$f_{hk}=\frac{n_{hk}}{N}$	$f_{h\bullet}=\frac{n_{h\bullet}}{N}$
Marg. y	$f_{\bullet 1}=\frac{n_{\bullet 1}}{N}$	$f_{\bullet 2}=\frac{n_{\bullet 2}}{N}$	…	$f_{\bullet j}=\frac{n_{\bullet j}}{N}$	…	$f_{\bullet k}=\frac{n_{\bullet k}}{N}$	1

X\Y	y_1	y_2	…	y_j	…	y_k
x_1	$f_{y=y_1/x=x_1}=\frac{n_{11}}{n_{1\bullet}}$	$f_{y=y_2/x=x_1}=\frac{n_{12}}{n_{1\bullet}}$	…	$f_{y=y_j/x=x_1}=\frac{n_{1j}}{n_{1\bullet}}$	…	$f_{y=y_k/x=x_1}=\frac{n_{1k}}{n_{1\bullet}}$
x_2	$f_{y=y_1/x=x_2}=\frac{n_{21}}{n_{2\bullet}}$	$f_{y=y_2/x=2}=\frac{n_{22}}{n_{2\bullet}}$	…	$f_{y=y_j/x=x_2}=\frac{n_{2j}}{n_{2\bullet}}$	…	$f_{y=y_k/x=x_2}=\frac{n_{2k}}{n_{2\bullet}}$
…	…	…	…	…	…	…
x_i	$f_{y=y_1/x=x_i}=\frac{n_{i1}}{n_{i\bullet}}$	$f_{y=y_2/x=x_i}=\frac{n_{i2}}{n_{i\bullet}}$	…	$f_{y=y_j/x=x_i}=\frac{n_{ij}}{n_{i\bullet}}$	…	$f_{y=y_k/x=x_i}=\frac{n_{ik}}{n_{i\bullet}}$
…	…	…	…	…	…	…
x_h	$f_{y=y_1/x=x_h}=\frac{n_{h1}}{n_{h\bullet}}$	$f_{y=y_2/x=x_1}=\frac{n_{h2}}{n_{h\bullet}}$	…	$f_{y=y_j/x=x_h}=\frac{n_{hj}}{n_{h\bullet}}$	…	$f_{y=y_k/x=x_h}=\frac{n_{hk}}{n_{h\bullet}}$

X\Y	y_1	y_2	…	y_j	…	y_k
x_1	$f_{x=x_1/y=y_1}=\frac{n_{11}}{n_{\bullet 1}}$	$f_{x=x_1/y=y_2}=\frac{n_{12}}{n_{\bullet 2}}$	…	$f_{x=x_1/y=y_j}=\frac{n_{1j}}{n_{\bullet j}}$	…	$f_{x=x_1/y=y_k}=\frac{n_{1k}}{n_{\bullet k}}$
x_2	$f_{x=x_2/y=y_1}=\frac{n_{21}}{n_{\bullet 1}}$	$f_{x=x_2/y=y_2}=\frac{n_{22}}{n_{\bullet 2}}$	…	$f_{x=x_2/y=y_j}=\frac{n_{2j}}{n_{\bullet j}}$	…	$f_{x=x_2/y=y_k}=\frac{n_{2k}}{n_{\bullet k}}$
…	…	…	…	…	…	…
x_i	$f_{x=x_i/y=y_1}=\frac{n_{i1}}{n_{\bullet 1}}$	$f_{x=x_i/y=y_2}=\frac{n_{i2}}{n_{\bullet 2}}$	…	$f_{x=x_i/y=y_j}=\frac{n_{ij}}{n_{\bullet j}}$	…	$f_{x=x_i/y=y_k}=\frac{n_{ik}}{n_{\bullet k}}$
…	…	…	…	…	…	…
x_h	$f_{x=x_h/y=y_1}=\frac{n_{h1}}{n_{\bullet 1}}$	$f_{x=x_h/y=y_2}=\frac{n_{h2}}{n_{\bullet 2}}$	…	$f_{x=x_h/y=y_j}=\frac{n_{hj}}{n_{\bullet j}}$	…	$f_{x=x_h/y=y_k}=\frac{n_{hk}}{n_{\bullet k}}$

- $f_{ij} = \frac{n_{ij}}{N}$ Frecuencia relativa conjunta.
- $f_{i\bullet} = \frac{n_{i\bullet}}{N}$ Frecuencia relativa marginal de la variable X.
- $f_{\bullet j} = \frac{n_{\bullet j}}{N}$ Frecuencia relativa marginal de la variable Y.
- $f_{x=x_i/y=y_j} = \frac{n_{ij}}{n_{\bullet j}}$ Frecuencia relativa de la variable X condicionada a Y.
- $f_{y=y_j/x=x_i} = \frac{n_{ij}}{n_{i\bullet}}$ Frecuencia relativa de la variable Y condicionada a X.

2. Independencia y dependencia estadística.

- **Independencia estadística:** los valores que toman las variables para cada individuo no están relacionados.
- **Dependencia estadística:** los valores que toman las variables para cada individuo parecen tener algún tipo de coincidencia en sus comportamientos.
- Si hay dependencia estadística, ¿los comportamientos de las variables están relacionados?
- **Condición de independencia:**
- $f_{x=x_i/y=y_j} = \frac{n_{ij}}{n_{\bullet j}} = \frac{n_{i\bullet}}{N} = f_{i\bullet} \qquad \forall \boldsymbol{i}, \boldsymbol{j}$
- $f_{y=y_j/x=x_i} = \frac{n_{ij}}{n_{i\bullet}} = \frac{n_{\bullet j}}{N} = f_{\bullet j} \qquad \forall \boldsymbol{i}, \boldsymbol{j}$
- ¿Interpretación?

Ejemplo de **independencia**:

•X="salario percibido mensual (cientos de euros)"

•Y="gasto en ocio mensual (cientos de euros)"

X/Y	1	2	3	Marg. X
20	20	40	10	70
30	30	60	15	105
40	24	48	12	84
Marg. Y	74	148	37	259

Frecuencias relativas:

X/Y	1	2	3	Marg. X
20	0,0772	0,1544	0,0386	0,2703
30	0,1158	0,2317	0,0579	0,4054
40	0,0927	0,1853	0,0463	0,3243
Marg. Y	0,2857	0,5714	0,1429	1

Frecuencias condicionadas:

X/Y	1	2	3
20	0,2857	0,5714	0,1429
30	0,2857	0,5714	0,1429
40	0,2857	0,5714	0,1429

$f_{y=y_j/x=x_i} = \frac{n_{ij}}{n_{i\bullet}} = \frac{n_{\bullet j}}{N} = f_{\bullet j}$

X/Y	1	2	3
20	0,2703	0,2703	0,2703
30	0,4054	0,4054	0,4054
40	0,3243	0,3243	0,3243

$f_{x=x_i/y=y_j} = \frac{n_{ij}}{n_{\bullet j}} = \frac{n_{i\bullet}}{N} = f_{i\bullet}$

- Condición de independencia en una expresión más manejable (se llega a lo mismo tomando como condicionante X):

$$f_{x=x_i/y=y_j} = \frac{n_{ij}}{n_{\bullet j}} = \frac{n_{i\bullet}}{N} = f_{i\bullet} \qquad \forall \boldsymbol{i}, \boldsymbol{j}$$

$$n_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{N}$$

$$\frac{n_{ij}}{N} = \frac{n_{i\bullet}}{N} \cdot \frac{n_{\bullet j}}{N}$$

$$f_{ij} = f_{i\bullet} \cdot f_{\bullet j} \qquad \forall \boldsymbol{i}, \boldsymbol{j}$$

Ejemplo de **independencia**:

- X="salario percibido mensual (cientos de euros)"
- Y="gasto en ocio mensual (cientos de euros)"

X/Y	1	2	3	Marg. X
20	20	40	10	70
30	30	60	15	105
40	24	48	12	84
Marg. Y	74	148	37	259

Frecuencias relativas conjuntas:

X/Y	1	2	3
20	0,0772	0,1544	0,0386
30	0,1158	0,2317	0,0579
40	0,0927	0,1853	0,0463

$$f_{11} = \frac{20}{259}$$

Producto de frecuencias relativas marginales:

X/Y	1	2	3	Marg. X
20	0,0772	0,1544	0,0386	0,2703
30	0,1158	0,2317	0,0579	0,4054
40	0,0927	0,1853	0,0463	0,3243
Marg. Y	0,2857	0,5714	0,1429	1

$$f_{1\bullet} \cdot f_{\bullet 1} = \frac{70}{259} \cdot \frac{74}{259}$$

En **todas** las frecuencias conjuntas se da la igualdad: independencia.

Ejemplo de **dependencia estadística**:

- X="salario percibido mensual (cientos de euros)"
- Y="gasto en ocio mensual (cientos de euros)"

X/Y	1	2	3	Marg. X
20	15	30	29	74
30	40	60	48	148
40	10	15	12	37
Marg. Y	65	105	89	259

Frecuencias relativas conjuntas:

X/Y	1	2	3
20	0,0579	0,1158	0,1120
30	0,1544	0,2317	0,1853
40	0,0386	0,0579	0,0463

$$f_{11} = \frac{15}{259}$$

Producto de frecuencias relativas marginales:

X/Y	1	2	3	Marg. X
20	0,0772	0,1544	0,0386	0,2703
30	0,1158	0,2317	0,0579	0,4054
40	0,0927	0,1853	0,0463	0,3243
Marg. Y	0,2857	0,5714	0,1429	1

$$f_{1\bullet} \cdot f_{\bullet 1} = \frac{74}{259} \cdot \frac{65}{259}$$

NO en **todas** las frecuencias conjuntas se da la igualdad: dependencia estadística.

3. Momentos bidimensionales.

- Momentos bidimensionales respecto al origen:

$$a_{rs} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} x_i^r \cdot y_j^s \cdot n_{ij}}{N}$$

- Momentos respecto al origen importantes:

$$a_{10} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} x_i^1 \cdot y_j^0 \cdot n_{ij}}{N} = \frac{\sum_{i=1}^{h} x_i^1 \cdot n_{i\bullet}}{N} = \bar{x}$$

$$a_{01} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} x_i^0 \cdot y_j^1 \cdot n_{ij}}{N} = \frac{\sum_{j=1}^{k} y_j^1 \cdot n_{\bullet j}}{N} = \bar{y}$$

$$a_{11} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} x_i \cdot y_j \cdot n_{ij}}{N}$$

- Momentos bidimensionales respecto a las medias:

$$m_{rs} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} (x_i - \bar{x})^r \cdot \left(y_j - \bar{y}\right)^s \cdot n_{ij}}{N}$$

- Momentos respecto a las medias importantes:

$$m_{20} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} (x_i - \bar{x})^2 \cdot \left(y_j - \bar{y}\right)^0 \cdot n_{ij}}{N} = \frac{\sum_{i=1}^{h} (x_i - \bar{x})^2 \cdot n_{i\bullet}}{N} = S_x^2$$

$$m_{02} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} (x_i - \bar{x})^0 \cdot \left(y_j - \bar{y}\right)^2 \cdot n_{ij}}{N} = \frac{\sum_{i=1}^{h} \left(y_j - \bar{y}\right)^2 \cdot n_{\bullet j}}{N} = S_y^2$$

$$m_{11} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{k} (x_i - \bar{x}) \cdot \left(y_j - \bar{y}\right) \cdot n_{ij}}{N} = S_{xy}$$: **Covarianza.**

- Cálculo:
  - $S_x^2 = m_{20} = a_{20} - a_{10}^2$
  - $S_y^2 = m_{02} = a_{02} - a_{01}^2$
  - $S_{xy} = m_{11} = a_{11} - a_{10} \cdot a_{01}$

- Ejemplo 1:

X/Y	1	2	3	Marg. X	$x_i * n_{i\bullet}$
20	20	40	10	70	1400
30	30	60	15	105	3150
40	24	48	12	84	3360
Marg. Y	74	148	37	259	7910
$y_j * n_{\bullet j}$	74	296	111	481	

400	1600	600
900	3600	1350
960	3840	1440
	Suma:	14690
	a_{11}=	56,7181
	a_{10}=	30,5405
	a_{01}=	1,8571
	$S_{xy}=m_{11}$=	0,0000

$$\sum_{i=1}^{h}\sum_{j=1}^{k} x_i \cdot y_j \cdot n_{ij}\ a_{11}$$

- Ejemplo 2:

X/Y	1	2	3	Marg. X	$x_i * n_{i\bullet}$
20	15	30	29	74	1480
30	40	60	48	148	4440
40	10	15	12	37	1480
Marg. Y	65	105	89	259	7400
$y_j * n_{\bullet j}$	65	210	267	542	

300	1200	1740
1200	3600	4320
400	1200	1440
	Suma:	15400
	a_{11}=	59,4595
	a_{10}=	28,5714
	a_{01}=	2,0927
	$S_{xy}=m_{11}$=	-0,3309

Propiedad de la covarianza S_{xy}: Si dos variables son independientes, su covarianza siempre vale 0. En cambio, hay casos en los que la covarianza vale 0; pero las variables no son independientes.


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 2.2

Regresión y correlación

Contenido

1. Regresión.
2. Bondad del ajuste.
3. Correlación.
4. Generalización de la regresión.

1.1. Regresión: Introducción

- Supuesto de partida: existe dependencia estadística.

- Cuestiones:

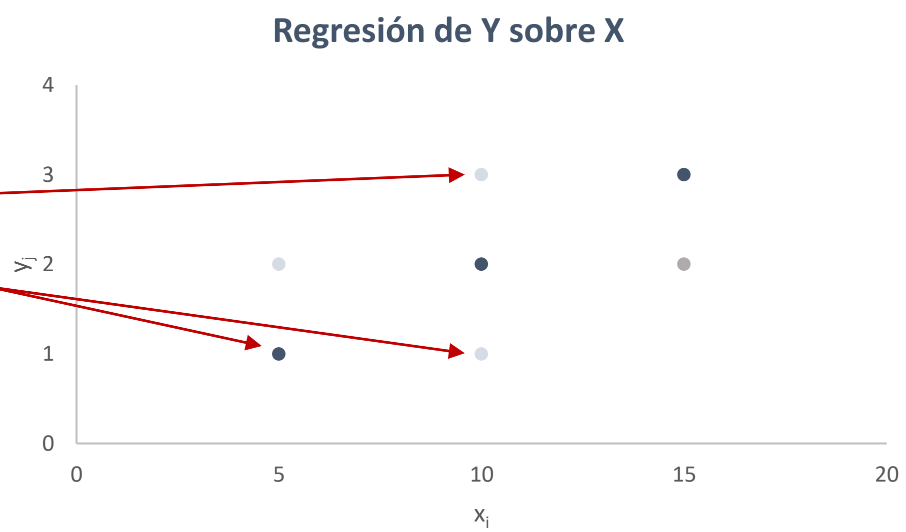
  - ¿Qué **estructura** tiene esa dependencia? (representación matemática) **Análisis de regresión**.

  - ¿Qué **intensidad** tiene? **Análisis de correlación**.

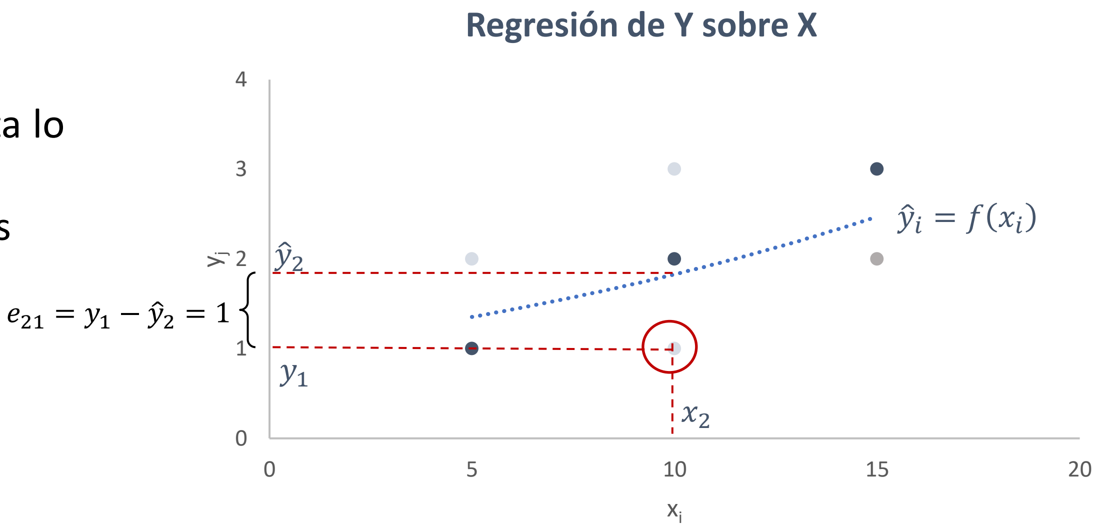
1.1. Regresión: Introducción

- Supongamos una variable explicada o dependiente *Y*; y una serie de variables explicativas X_1, X_2,…, X_n.
- El **análisis de regresión** trata de encontrar la estructura matemática que mejor explique *Y* en función de las variables explicativas.
- Regresión simple: una única variable ***X***.
- Tipos:
  - Regresión tipo I.
  - Regresión **tipo II.**

X/Y	$y_1=1$	$y_2=2$	$y_3=3$	$n_{i\bullet}$
$x_1=5$	11	1	0	12
$x_2=10$	1	12	1	14
$x_3=15$	0	2	12	14
$n_{\bullet j}$	12	15	13	40

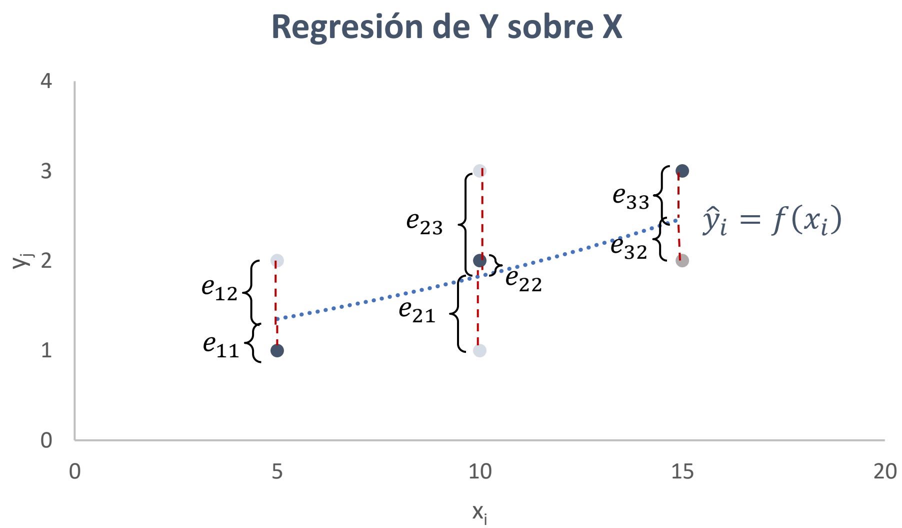


- Las frecuencias no nulas se plasman en el gráfico de dispersión como puntos.
- Puntos más oscuros: mayores frecuencias más importantes, pues implican a más casos.
- Regresión: función matemática $\hat{y}_i = f(x_i)$ cuya representación representa lo mejor posible a la nube de puntos, sobre todo a los puntos más oscuros (más casos).
- y_j: Valor real de la variable Y.
- $\hat{y}_i$: Valor propuesto por la regresión cuando $x = x_i$.
- e_{ij}: Error cometido por la regresión: residuo. $e_{ij} = y_j - \hat{y}_i \Rightarrow y_j = \hat{y}_i + e_{ij}$



1.1. Regresión: Introducción

X/Y	y_1=1	y_2=2	y_3=3	$n_{i\bullet}$
x_1=5	11	1	0	12
x_2=10	1	12	1	14
x_3=15	0	2	12	14
$n_{\bullet j}$	12	15	13	40



- La regresión $\hat{y}_i = f(x_i)$ consta de una forma funcional (tipo de función matemática) y unos parámetros estimados (valores de los parámetros que están presentes en la función)

- Ejemplo: si la forma funcional de la regresión es una parábola: $\hat{y}_i = a + b \cdot x_i + c \cdot x_i^2$, y los parámetros a estimar son: *a*, *b* y *c*.

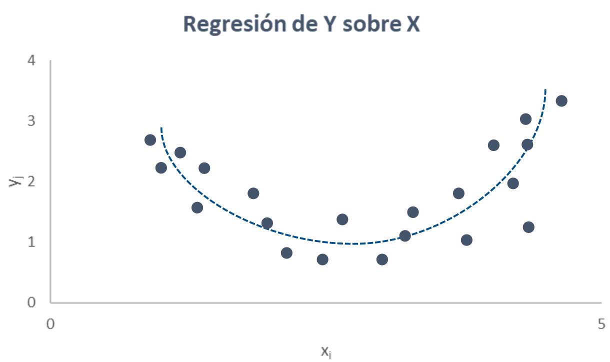
- ¿Cuál será el valor de los parámetros? Aquellos que $\min \sum_i \sum_j e_{ij}^2 \cdot n_{ij}$

- Interpretación.

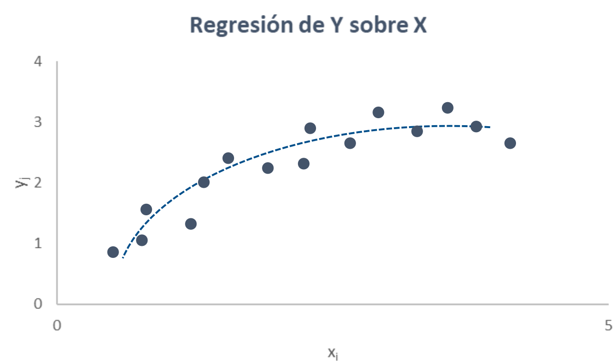
$$\min \sum_i \sum_j e_{ij}^2 \cdot n_{ij} = \min \sum_i \sum_j (y_j - \hat{y}_i)^2 \cdot n_{ij}$$

Criterio de estimación de la regresión mínimo-cuadrático

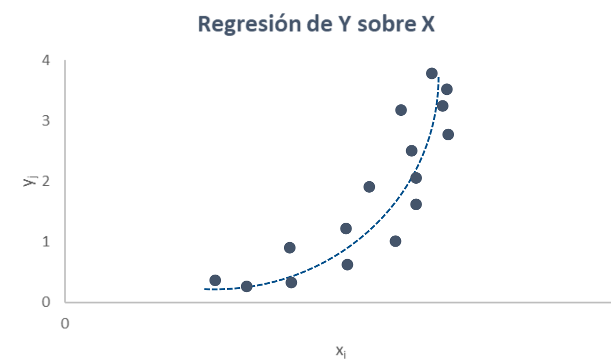
1.1. Regresión: Introducción



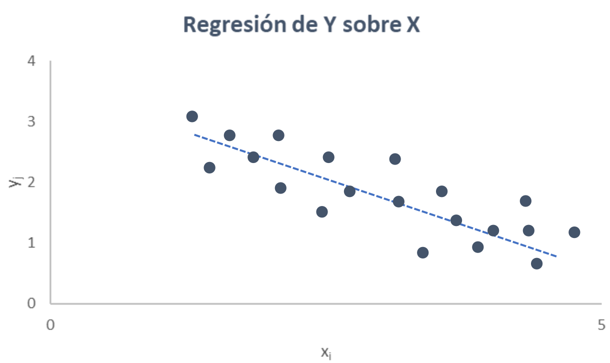
$$\hat{y}_i = a + b \cdot x_i + c \cdot x_i^2$$



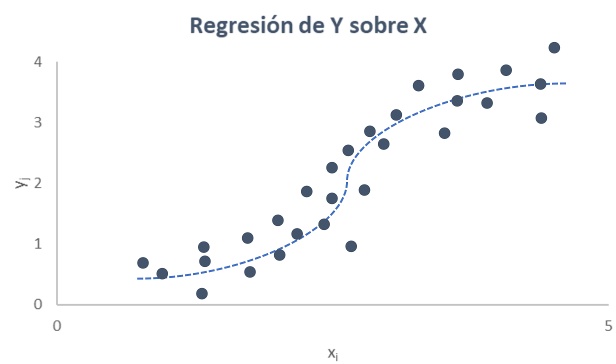
$$\hat{y}_i = a \cdot x_i^b$$



$$\hat{y}_i = a \cdot b^{x_i}$$



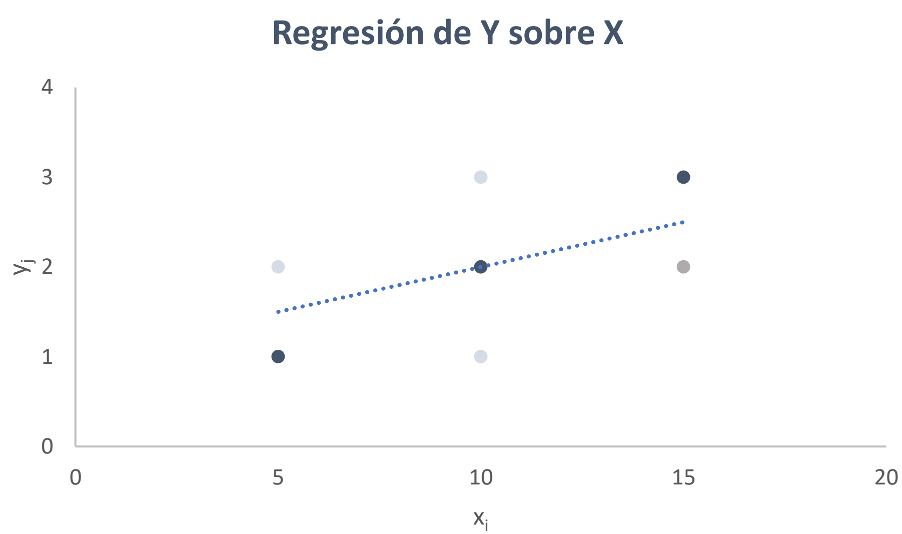
$$\hat{y}_i = a + b \cdot x_i$$



$$\hat{y}_i = \frac{1}{1 + e^{-(a + b \cdot x_i)}}$$

$$\hat{y}_i = ¿?$$

X/Y	1	2	3	$n_{i\bullet}$
5	11	1	0	12
10	1	12	1	14
15	0	2	12	14
$n_{\bullet j}$	12	15	13	40



- Regresión $\hat{y}_i = f(x_i) = a + b \cdot x_i$
- ¿Cuál será el valor de los parámetros? Aquellos que $\min \sum_i \sum_j e_{ij}^2 \cdot n_{ij}$
- En el caso lineal:

$$\min \sum_i \sum_j e_{ij}^2 \cdot n_{ij} =$$

$$\min \sum_i \sum_j (y_j - \hat{y}_i)^2 \cdot n_{ij} =$$

$$\min \sum_i \sum_j (y_j - a - bx_i)^2 \cdot n_{ij}$$

- Tras resolver el problema de minimización:

$$\hat{b} = \frac{S_{xy}}{S_x^2} \qquad \hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$$

- Así: $y_j = \hat{y}_i + e_{ij} = \hat{a} + \hat{b} \cdot x_i + e_{ij}$

1.2. Regresión: caso lineal.

X/Y	1	2	3	$n_{i\bullet}$	$x_i{*}n_{i\bullet}$	$x_i^2{*}n_{i\bullet}$
5	11	1	0	12	60	300
10	1	12	1	14	140	1400
15	0	2	12	14	210	3150
$n_{\bullet j}$	12	15	13	40	410	4850
$y_j{*}n_{\bullet j}$	12	30	39	81		
$y_j^2{*}n_{\bullet j}$	12	60	117	189		
	55	10	0			
	10	240	30			
	0	60	540			
		Suma:	945			

$$\sum_{i=1}^{h}\sum_{j=1}^{k} x_i \cdot y_j \cdot n_{ij}$$

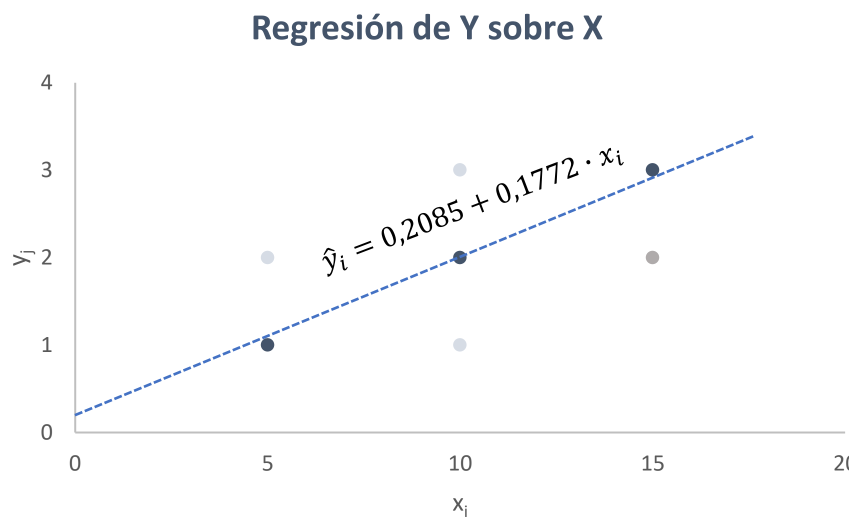
		Resultados Regresión:	
a_{11}=	23,6250	$\hat{b}$=	0,1772
$\bar{x}=a_{10}$=	10,2500	$\hat{a}$=	0,2085
$\bar{y}=a_{01}$=	2,0250		
a_{20}=	121,2500		
a_{02}=	4,7250		
$S_{xy}=m_{11}$=	2,8688		
$S_x^2=m_{20}$=	16,1875		
$S_y^2=m_{02}$=	0,6244		

$$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x} \qquad \hat{b} = \frac{S_{xy}}{S_x^2}$$

$$S_{xy} = m_{11} = a_{11} - a_{10} \cdot a_{01}$$

$$S_x^2 = m_{20} = a_{20} - a_{10}^2$$

$$S_y^2 = m_{02} = a_{02} - a_{01}^2$$



$$\hat{y}_i = \hat{a} + \hat{b} \cdot x_i = 0{,}2085 + 0{,}1772 \cdot x_i$$

Predicción / Previsión:

¿Cuánto se prevé que valdrá Y si $x_i = 20$?

$\hat{y}_i = 0{,}2085 + 0{,}1772 \cdot 20 = 3{,}7529$

2.1. Bondad del ajuste: Coeficiente de determinación general

- ¿Representa bien la regresión la realidad? **Bondad del ajuste**.

- **Coeficiente de determinación**:

  - Para un caso: $y_j = \hat{y}_i + e_{ij}$
  - Restando la media de Y: $y_j - \bar{y} = \hat{y}_i - \bar{y} + e_{ij}$
  - Elevando al cuadrado: $\left(y_j - \bar{y}\right)^2 = \left((\hat{y}_i - \bar{y}) + e_{ij}\right)^2$
  - Desarrollando: $\left(y_j - \bar{y}\right)^2 = (\hat{y}_i - \bar{y})^2 + e_{ij}^2 - 2 \cdot (\hat{y}_i - \bar{y}) \cdot e_{ij}$
  - Sumando para todos los casos:

$$\sum_i \sum_j \left(y_j - \bar{y}\right)^2 \cdot n_{ij} = \sum_i \sum_j (\hat{y}_i - \bar{y})^2 \cdot n_{ij} + \sum_i \sum_j e_{ij}^2 \cdot n_{ij} - 2 \cdot \underbrace{\sum_i \sum_j (\hat{y}_i - \bar{y}) \cdot e_{ij} \cdot n_{ij}}_{\text{0 en el caso mínimo-cuadrático}}$$

  - Dividiendo entre N:

$$\frac{\sum_j \left(y_j - \bar{y}\right)^2 \cdot n_{\bullet j}}{N} = \frac{\sum_i (\hat{y}_i - \bar{y})^2 \cdot n_{i\bullet}}{N} + \frac{\sum_i \sum_j e_{ij}^2 \cdot n_{ij}}{N}$$

$S_y^2 = S_{\hat{y}}^2 + S_e^2 \Longrightarrow S_{\hat{y}}^2 = S_y^2 - S_e^2 \Longrightarrow R^2 = \frac{S_{\hat{y}}^2}{S_y^2} = 1 - \frac{S_e^2}{S_y^2}$: **proporción de varianza** (comportamiento) **de Y recogida por la regresión**

  - S_y^2: varianza de Y; $S_{\hat{y}}^2$: varianza debida a la regresión de Y sobre X; S_e^2: varianza residual

2.2. Bondad del ajuste: Coeficiente de determinación lineal

- $0 < R^2 < 1$
- A mayor valor, mejor ajuste (representación de la realidad).
- Caso lineal: $R^2 = \frac{S_{xy}^2}{S_x^2 \cdot S_y^2}$

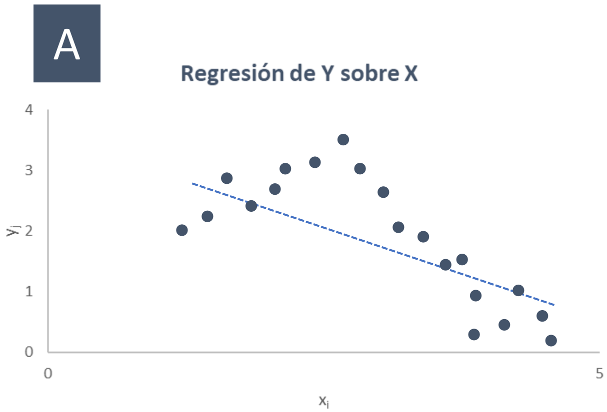
X/Y	1	2	3	$n_{i\bullet}$	$x_i{}^*n_{i\bullet}$	$x_i{}^2{}^*n_{i\bullet}$
5	11	1	0	12	60	300
10	1	12	1	14	140	1400
15	0	2	12	14	210	3150
$n_{\bullet j}$	12	15	13	40	410	4850
$y_j{}^*n_{\bullet j}$	12	30	39	81		
$y_j{}^2{}^*n_{\bullet j}$	12	60	117	189		
	55	10	0			
	10	240	30			
	0	60	540			
		Suma:	945	Resultados Regresión:		
		a_{11}=	23,6250	$\hat{b}$=	0,1772	
		$\bar{X}=a_{10}$=	10,2500	$\hat{a}$=	0,2085	
		$\bar{y}=a_{01}$=	2,0250	Coeficiente de Determinación:		
		a_{20}=	121,2500	R^2=	0,8143	
		a_{02}=	4,7250	¿Interpretación?		
		$S_{xy}=m_{11}$=	2,8688			
		$S_x{}^2=m_{20}$=	16,1875			
		$S_y{}^2=m_{02}$=	0,6244			

- **Análisis de correlación:** Medición de la intensidad con que se relacionan estadísticamente dos variables.

- El R^2 mide la calidad de la representación de la realidad por parte de la regresión. Esto depende de:

  - Si la forma funcional es correcta (A).

  - Si la relación estadística real entre las dos variables es intensa (B).



- Vamos a **suponer que la forma funcional es correcta**. Entonces, el R^2 nos informará de la intensidad de la relación existente entre las dos variables:

Alto R^2: relación muy intensa.

Bajo R^2: relación poco intensa.

3. Correlación

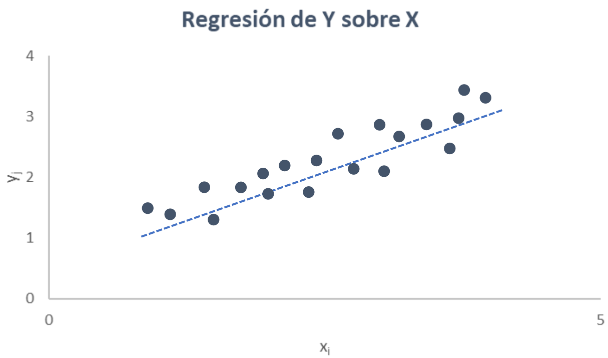
- A partir del R^2 se obtiene otra medida: el coeficiente de correlación (R): raíz de R^2 con el signo de la covarianza:

$$R = \mp\sqrt{1 - \frac{S_e^2}{S_y^2}}$$

- En el caso lineal:

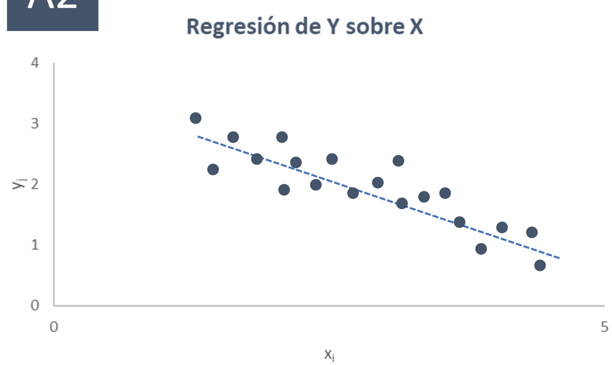
$$R = \frac{S_{xy}}{S_x \cdot S_y}$$

A1



R próximo a **+1**: relación muy intensa y en el mismo sentido.

A2



R próximo a **-1**: relación muy intensa y en el sentido opuesto.

- $-1 < R < +1$
- A mayor valor absoluto, mayor relación estadística entre X e Y.
- Caso lineal: $R = \frac{S_{xy}}{S_x \cdot S_y}$

X/Y	1	2	3	Marg. X	$x_i{*}n_{i\bullet}$	$x_i^2{*}n_{i\bullet}$
5	11	1	0	12	60	300
10	1	12	1	14	140	1400
15	0	2	12	14	210	3150
Marg. Y	12	15	13	40	410	4850
$y_j{*}n_{\bullet j}$	12	30	39	81		
$y_j^2{*}n_{\bullet j}$	12	60	117	189		
	55	10	0			
	10	240	30			
	0	60	540			
		Suma:	945	Resultados Regresión:		
		a_{11}=	23,6250	$\hat{b}$=	0,1772	
		$\bar{x}=a_{10}$=	10,2500	$\hat{a}$=	0,2085	
		$\bar{y}=a_{01}$=	2,0250	Coeficiente de Determinación:		
		a_{20}=	121,2500	R^2=	0,8143	
		a_{02}=	4,7250	R=	0,9024	
		$S_{xy}=m_{11}$=	2,8688			
		$S_x^2=m_{20}$=	16,1875			
		$S_y^2=m_{02}$=	0,6244			

¿Interpretación?

4. Generalización de la regresión

- **Casos no-lineales:** a veces se plantean entre variables relaciones estadísticas que no son de tipo lineal. Algunas de ellas, no obstante, son linealizables, y se pueden estimar previa transformación de los valores de las variables (caso exponencial o potencial).

- **Regresión múltiple:** existen varias variables explicativas. Estimación: álgebra lineal.

¡Muchas gracias!

Tema 3.0.

Números Índice

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

1. Concepto de número índice.
2. Números índice simples.
3. Propiedades de los números índice.
4. Números índice complejos.
5. Enlace y cambios de base.
6. Deflactación de series económicas.

1. Concepto de número índice

- Tabla de precios:

Período	Precio bien A (euros)	Precio bien B (dólares)	Precio bien C (euros)
0	10	5	1100
1	15	12	1250
2	20	28	1500
3	28	30	1800

- Cuestiones que se pueden plantear:
  - Comparación de la evolución de los precios de cada bien: números índice simples.
  - Medida de la evolución global o conjunta de los precios de todos los 3 bienes: números índice complejos.

- **Número índice:** medida estadística de los cambios que se producen en una magnitud **simple** o **compleja** a lo largo del tiempo o del espacio.

- Períodos a tener en cuenta:
  - **Período base:** momento del tiempo en el que el valor de la magnitud se toma como referencia.
  - **Período actual:** momento del tiempo en el que el valor de la magnitud se compara con el valor que tenía en el período base.
- Según el período base cambie o no, los números índice pueden ser **de base fija** o **encadenados**.

2.1. Números índices simples: de base fija

- De **base fija**:

Período	Precio Producto A (euros/unidad)	Cantidad Vendida Producto A (unidades)	Ventas Producto A (euros)
0	30	1200	36000
1	33	1500	49500
2	45	1800	81000
3	54	2100	113400

- Queremos saber la evolución que han seguido estas 3 magnitudes (precio, cantidad y valor) a lo largo del tiempo, comparando el valor en cada período actual con el que tenían en el período base:

$$p_0^t = \frac{p_t}{p_0} \qquad q_0^t = \frac{q_t}{q_0} \qquad V_0^t = \frac{p_t q_t}{p_0 q_0} = p_0^t \cdot q_0^t$$

Período	Índice de precios	Índice de cantidades	Índice de valor
0	1,0000	1,0000	1,0000
1	1,1000	1,2500	1,3750
2	1,5000	1,5000	2,2500
3	1,8000	1,7500	3,1500

- $p_0^3 = \frac{p_{i3}}{p_{i0}} = \frac{54}{30} = 1{,}8$
- El precio se ha incrementado del período base (0) al período actual (3) en $(1{,}8 - 1) \cdot 100 = 80\%$

2.2. Números índices simples: encadenados

- **Encadenados:**

Período	Precio Producto A (euros/unidad)	Cantidad Vendida Producto A (unidades)	Ventas Producto A (euros)
0	30	1200	36000
1	33	1500	49500
2	45	1800	81000
3	54	2100	113400

- Queremos saber la evolución que han seguido estas 3 magnitudes (precio, cantidad y valor) a lo largo del tiempo, comparando el valor en cada período actual con el que tenían en el período base (el inmediatamente anterior):

$$p_{t-1}^{t} = \frac{p_t}{p_{t-1}} \qquad q_{t-1}^{t} = \frac{q_t}{q_{t-1}} \qquad V_{t-1}^{t} = \frac{p_t q_t}{p_{t-1} q_{t-1}} = p_{t-1}^{t} \cdot q_{t-1}^{t}$$

Período	Índice de precios	Índice de cantidades	Índice de valor
0	-	-	-
1	1,1000	1,2500	1,3750
2	1,3636	1,2000	1,6364
3	1,2000	1,1667	1,4000

- $p_2^3 = \frac{p_{i3}}{p_{i2}} = \frac{54}{45} = 1{,}2$
- El precio se ha incrementado del período base (0) al período actual (3) en $(1{,}2 - 1) \cdot 100 = 20\%$

2.3. Números índices simples: presentación

- Las instituciones suelen presentar los números índice multiplicados por 100. En los de **base fija** esto implica que, en el **período base**, el número siempre vale 100:

  - Base Fija:

Período	Índice de precios	Índice de cantidades	Índice de valor
0	100,00	100,00	100,00
1	110,00	125,00	137,50
2	150,00	150,00	225,00
3	180,00	175,00	315,00

  - Encadenados:

Período	Índice de precios	Índice de cantidades	Índice de valor
0	-	-	-
1	110,00	125,00	137,50
2	136,36	120,00	163,64
3	120,00	116,67	140,00

3. Propiedades de los números índice.

Propiedades que es deseable que verifiquen los números índice (a continuación se utilizan de precios, podrían ser de cantidades o valor):

- **Unicidad:** Todo número índice debe existir. O sea, el denominador NUNCA debe ser nulo.

- **Identidad:** Si se hace coincidir el período actual con el base, el número índice debe valer 1 (o 100 si se presentan multiplicados por 100):

$$p_t^t = \frac{p_t}{p_t} = 1$$

- **Inversión o reversión cronológica:** Si se intercambian los períodos base y actual, el número índice resultante debe ser el inverso del original:

$$p_b^a = \frac{p_a}{p_b} = z; \qquad p_a^b = \frac{p_b}{p_a} = \frac{1}{z}$$

- **Proporcionalidad constante independiente de la base:** el cociente entre dos números índice con la misma base y períodos actuales *a* y *b* respectivamente, será igual a un número índice de base *b* y período actual *a*.

$$\frac{p_h^a}{p_h^b} = \frac{\frac{p_a}{p_h}}{\frac{p_b}{p_h}} = \frac{p_a}{p_b} = p_b^a$$

4. Números índice complejos.

- Normalmente no estamos interesados en estudiar cantidades, precios o valores de bienes o servicios individuales, sino de grupos de éstos, de una forma global: números índice complejos.

- Un número índice complejo es un **promedio ponderado** de los números **índice simples** de cada una de las N magnitudes.

- Para el caso de un índice de precios de varios bienes:

$$IC_0^t = \sum_{i=1}^{N} p_{i0}^t \cdot \frac{w_i}{\sum_{i=1}^{N} w_i} = p_{i=1,0}^t \cdot \frac{w_1}{\sum_{i=1}^{N} w_i} + p_{i=2,0}^t \cdot \frac{w_2}{\sum_{i=1}^{N} w_i} + \cdots + p_{i=N,0}^t \cdot \frac{w_N}{\sum_{i=1}^{N} w_i}$$

- $\frac{w_i}{\sum_{i=1}^{N} w_i}$ son las ponderaciones.
  Pueden variar entre 0 y 1, tal que $\sum_{i=1}^{N} \frac{w_i}{\sum_{i=1}^{N} w_i} = 1$.

- Según como se calculen las ponderaciones w_i se obtendrán distintos tipos de números índice complejos.

- **Índices de precios:**

  - **Laspeyres:** $L_p = \sum_{i=1}^{N} p_{i0}^{t} \cdot \frac{w_i}{\sum_{i=1}^{N} w_i} = \sum_{i=1}^{N} \frac{p_{it}}{p_{i0}} \cdot \frac{p_{i0} \cdot q_{i0}}{\sum_{i=1}^{N} p_{i0} \cdot q_{i0}} = \frac{\sum_{i=1}^{N} p_{it} \cdot q_{i0}}{\sum_{i=1}^{N} p_{i0} \cdot q_{i0}}$

  - **Paasche:** $P_p = \sum_{i=1}^{N} p_{i0}^{t} \cdot \frac{w_i}{\sum_{i=1}^{N} w_i} = \sum_{i=1}^{N} \frac{p_{it}}{p_{i0}} \cdot \frac{p_{i0} \cdot q_{it}}{\sum_{i=1}^{N} p_{i0} \cdot q_{it}} = \frac{\sum_{i=1}^{N} p_{it} \cdot q_{it}}{\sum_{i=1}^{N} p_{i0} \cdot q_{it}}$

  - **Fisher:** $F_p = \sqrt{L_p P_p}$

- **Índices de cantidades:**

  - **Laspeyres:** $L_q = \sum_{i=1}^{N} q_{i0}^{t} \cdot \frac{w_i}{\sum_{i=1}^{N} w_i} = \sum_{i=1}^{N} \frac{q_{it}}{q_{i0}} \cdot \frac{q_{i0} \cdot p_{i0}}{\sum_{i=1}^{N} q_{i0} \cdot p_{i0}} = \frac{\sum_{i=1}^{N} q_{it} \cdot p_{i0}}{\sum_{i=1}^{N} q_{i0} \cdot p_{i0}}$

  - **Paasche:** $P_q = \sum_{i=1}^{N} q_{i0}^{t} \cdot \frac{w_i}{\sum_{i=1}^{N} w_i} = \sum_{i=1}^{N} \frac{q_{it}}{q_{i0}} \cdot \frac{q_{i0} \cdot p_{it}}{\sum_{i=1}^{N} q_{i0} \cdot p_{it}} = \frac{\sum_{i=1}^{N} q_{it} \cdot p_{it}}{\sum_{i=1}^{N} q_{i0} \cdot p_{it}}$

  - **Fisher:** $F_q = \sqrt{L_q P_q}$

4. Números índice complejos.

- Ejemplo **precios**: índices de precios de los 2 productos comercializados por una empresa. Calcular los números índices para t=3, teniendo de base el período 0.

Período	Precio Producto A (euros/unidad)	Cantidad Producto A (unidades)	Precio Producto B (euros/unidad)	Cantidad Producto B (unidades)
0	5	6000	12	2000
1	6	5500	15	3500
2	7	7000	20	4000
3	10	7500	20	5000

$$L_p = \frac{\sum_{i=1}^{N} p_{it} \cdot q_{i0}}{\sum_{i=1}^{N} p_{i0} \cdot q_{i0}} = \frac{10 \cdot 6000 + 20 \cdot 2000}{5 \cdot 6000 + 12 \cdot 2000} = 1{,}8519 \quad \rightarrow 185{,}19 \; si \; 0 = 100$$

$$P_p = \frac{\sum_{i=1}^{N} p_{it} \cdot q_{it}}{\sum_{i=1}^{N} p_{i0} \cdot q_{it}} = \frac{10 \cdot 7500 + 20 \cdot 5000}{5 \cdot 7500 + 12 \cdot 5000} = 1{,}7949 \rightarrow 179{,}49 \; si \; 0 = 100$$

$$F_p = \sqrt{L_p \cdot P_p} = \sqrt{1{,}8519 \cdot 1{,}7949} = 1{,}8231 \rightarrow 182{,}31 \; si \; 0 = 100$$

4. Números índice complejos.

- Ejemplo **cantidades**: índices de cantidades de los 2 productos comercializados por una empresa. Calcular los números índices para t=3, teniendo de base el período 0.

Período	Precio Producto A (euros/unidad)	Cantidad Producto A (unidades)	Precio Producto B (euros/unidad)	Cantidad Producto B (unidades)
0	5	6000	12	2000
1	6	5500	15	3500
2	7	7000	20	4000
3	10	7500	20	5000

$$L_q = \frac{\sum_{i=1}^{N} q_{it} \cdot p_{i0}}{\sum_{i=1}^{N} q_{i0} \cdot p_{i0}} = \frac{7500 \cdot 5 + 5000 \cdot 12}{6000 \cdot 5 + 2000 \cdot 12} = 1{,}8056 \quad \rightarrow 180{,}56\ si\ 0 = 100$$

$$P_q = \frac{\sum_{i=1}^{N} q_{it} \cdot p_{it}}{\sum_{i=1}^{N} q_{i0} \cdot p_{it}} = \frac{7500 \cdot 10 + 5000 \cdot 20}{6000 \cdot 10 + 2000 \cdot 20} = 1{,}7500 \rightarrow 175{,}00\ si\ 0 = 100$$

$$F_q = \sqrt{L_q \cdot P_q} = \sqrt{1{,}8056 \cdot 1{,}7500} = 1{,}7776 \rightarrow 177{,}76\ si\ 0 = 100$$

5. Enlace y cambios de base.

- Los números índice con base fija tienen un inconveniente: la **elección del período base**, ya que este, conforme avanzamos en el tiempo, puede dejar de representar una situación comparable con las situaciones futuras.

- Soluciones:

  - Trabajar con números índice encadenados.
  - Proceder cada cierto tiempo a realizar un **cambio de base**.

- Para realizar un cambio de base, se utiliza la propiedad de **proporcionalidad constante independiente de la base:**

  - Base antigua: período 0.
  - Base nueva: período h.
  - **Enlace técnico**: $p_h^t = \frac{p_0^t}{p_0^h}$ (ejemplo con índice de precios).

5. Enlace y cambios de base.

- **Ejemplo:** Se tiene el índice de precios correspondiente a un conjunto de bienes de consumo, con base fija en el período 0. En cierto momento, se decide cambiar la base al período 4. Elabore la serie enlazada del número índice con base en período 4.

Período	Índice precios base = 0	Índice precios base = 4	Índice enlazado base = 4
0	1,0000	-	0,4545
1	1,3000	-	0,5909
2	1,6000	-	0,7273
3	1,9000	-	0,8636
4	2,2000	1,0000	1,0000
5	-	1,3000	1,3000
6	-	1,8000	1,8000
7	-	2,4000	2,4000
8	-	2,5000	2,5000

- $p_4^0 = \frac{p_0^0}{p_0^4} = \frac{1{,}0000}{2{,}2000} = 0{,}4545$
- $p_4^1 = \frac{p_0^1}{p_0^4} = \frac{1{,}3000}{2{,}2000} = 0{,}5909$
- $p_4^2 = \frac{p_0^2}{p_0^4} = \frac{1{,}6000}{2{,}2000} = 0{,}7273$
- $p_4^3 = \frac{p_0^3}{p_0^4} = \frac{1{,}9000}{2{,}2000} = 0{,}8636$

$$p_h^t = \frac{p_0^t}{p_0^h}$$

6. Deflactación de series económicas.

- Somos los habitantes de una pequeña isla/Estado. Queremos medir su PIB (cantidad de bienes y servicios producidos en la isla) en un período (año) 0:

Bienes y Servicios	Cantidad (q_0)		Precio año 0 (p_0)		Valor año 0 (q_0 X p_o)	
Coco	2500	Kg.	2	euros/Kg.	5000	euros
Vino	800	litros	6	euros/litro	4800	euros
Cestas de mimbre	600	unidades	20	euros/unidad	12000	euros
Excursiones en grupo	34	excursiones contratadas	800	euros/contrato	27200	euros
Valor de la producción año 0	¿?	¿?			49000	euros

6. Deflactación de series económicas.

- En un período (año) posterior t, nos planteamos medir de nuevo su PIB, para ver si ahora se producen más bienes y servicios que el año 0:

Bienes y Servicios	Cantidad (q_0)		Precio año 0 (p_0)		Valor año 0 (q_0 X p_0)	
Coco	2500	Kg.	2	euros/Kg.	5000	euros
Vino	800	litros	6	euros/litro	4800	euros
Cestas de mimbre	600	unidades	20	euros/unidad	12000	euros
Excursiones en grupo	34	excursiones contratadas	800	euros/contrato	27200	euros
Valor de la producción año 0	¿?	¿?			49000	euros

Bienes y Servicios	Cantidad (q_t)		Precio año t (p_t)		Valor año t (q_t X p_t)	
Coco	2000	Kg.	4	euros/Kg.	8000	euros
Vino	800	litros	8	euros/litro	6400	euros
Cestas de mimbre	800	unidades	15	euros/unidad	12000	euros
Excursiones en grupo	25	excursiones contratadas	1000	euros/contrato	25000	euros
Valor de la producción año t	¿?	¿?			51400	euros

$$Tasa\ de\ crecimiento\ de\ la\ producción = \frac{(51400 - 49000)}{49000} \cdot 100 = +4{,}90\%$$

- Pero los precios no son magnitudes inamovibles. Cambian. Esto puede dar una idea falsa de la evolución en la cantidad de bienes y servicios producidos. Con precios del año 0:

Bienes y Servicios	Cantidad (q_0)		Precio año 0 (p_0)		Valor año 0 (q_0 X p_0)	
Coco	2500	Kg.	2	euros/Kg.	5000	euros
Vino	800	litros	6	euros/litro	4800	euros
Cestas de mimbre	600	unidades	20	euros/unidad	12000	euros
Excursiones en grupo	34	excursiones contratadas	800	euros/contrato	27200	euros
Valor de la producción año 0	¿?	¿?			49000	euros

Bienes y Servicios	Cantidad (q_t)		Precio año 0 (p_0)		Valor año t a precios de año 0 (q_t X p_0)	
Coco	2000	Kg.	2	euros/Kg.	4000	euros
Vino	800	litros	6	euros/litro	4800	euros
Cestas de mimbre	800	unidades	20	euros/unidad	16000	euros
Excursiones en grupo	25	excursiones contratadas	800	euros/contrato	20000	euros
Valor de la producción año t	¿?	¿?			44800	euros

$$Tasa\ de\ crecimiento\ de\ la\ producción = \frac{(44800 - 49000)}{49000} \cdot 100 = -8{,}57\%$$

6. Deflactación de series económicas.

- Previo al análisis hemos de homogeneizar las series económicas de modo que las magnitudes en cada período estén valoradas según el mismo sistema de precios, para evitar el efecto distorsionador de su cambio.

  - Si las variables económicas vienen valoradas según el precio que rige en cada período: expresadas a **precios corrientes o en términos nominales**.

  - Si las variables económicas vienen valoradas según el precio de un mismo período: expresadas a **precios constantes o en términos reales**.

- Usualmente las variables económicas se publican en términos nominales.

- El paso de términos nominales a términos reales: **deflactación**.

6. Deflactación de series económicas.

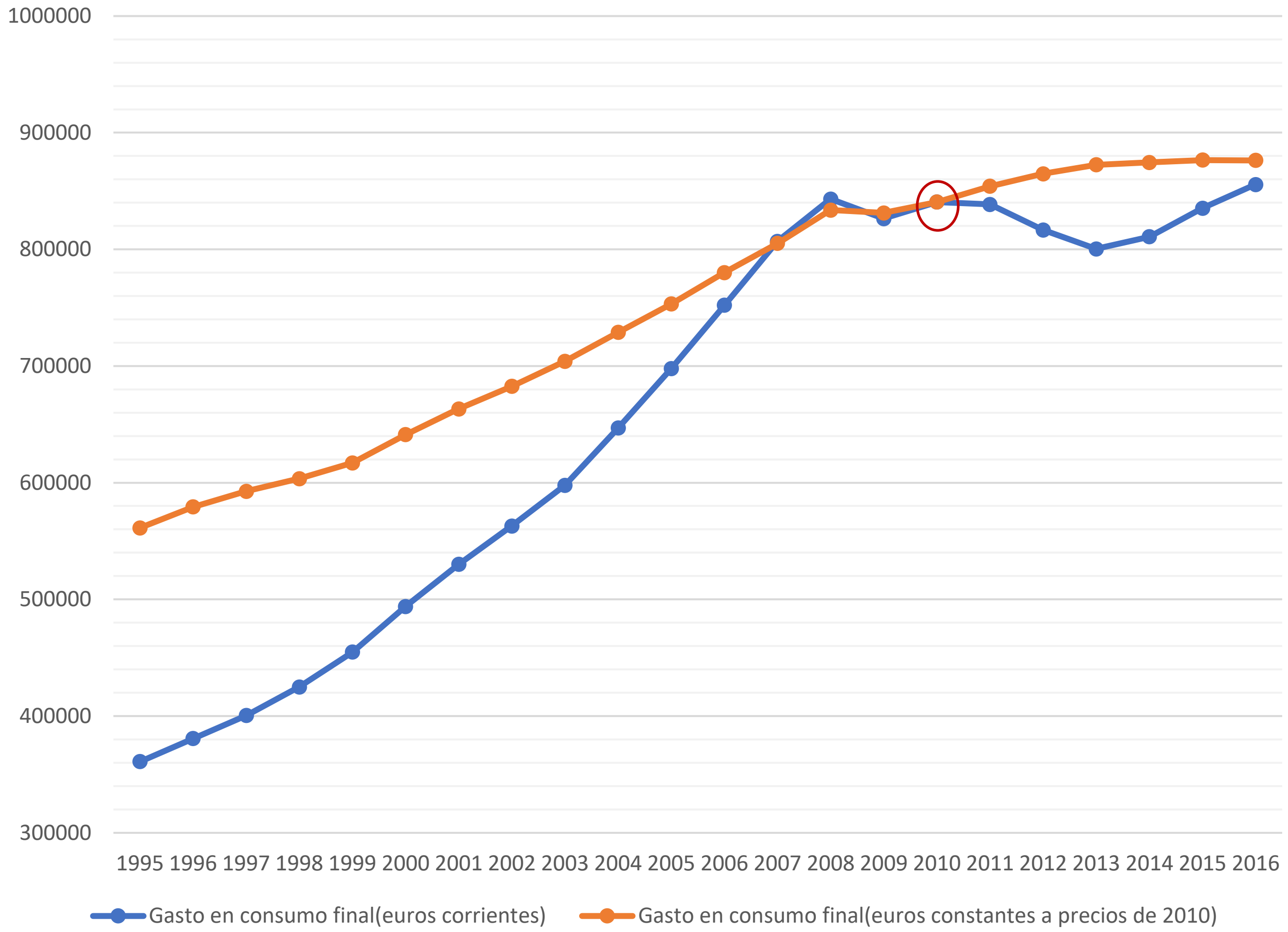
- Para **deflactar** una variable (serie) económica en términos nominales:
  - Se elige un número índice de precios adecuado, que se llama **deflactor**.
  - El valor de la variable para cada período se divide por el valor del deflactor en ese período.
  - Así, se consigue que la variable económica quede expresada en términos constantes, a precios del año base del deflactor.

- **Ejemplo**:

Período	Gasto en consumo final (precios corrientes)	Deflactor Gasto en Consumo (base 2010)	Gasto en consumo final (precios constantes de 2010)
1995	361029	64,32	561311,22
1996	380680	65,72	579240,16
1997	400460	67,57	592628,54
1998	424940	70,42	603442,33
1999	454763	73,70	617028,09
2000	493903	77,03	641155,33
2001	530079	79,92	663288,34
2002	562925	82,47	682581,41
2003	597730	84,90	704048,08
2004	646942	88,76	728837,54
2005	697774	92,65	753166,88
2006	752142	96,41	780113,04
2007	806882	100,21	805152,25
2008	843079	101,12	833712,38
2009	826392	99,42	831254,00
2010	840492	100,00	840492,00
2011	838574	98,19	854012,42
2012	816642	94,44	864752,97
2013	800381	91,74	872483,86
2014	810728	92,71	874517,17
2015	835258	95,29	876584,64
2016	855613	97,64	876313,23

$$GCF_{año\ t\ a\ precios\ del\ año\ 2010} = \frac{GCF_{año\ t\ a\ precios\ año\ t}}{Deflactor_t} \cdot 100$$

- **Ejemplo (2)**:



Millones de euros.
Fuente: INE.


¡Muchas gracias!

Tema 4.1.

Introducción a la probabilidad

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

1. Conceptos básicos.
2. Concepto de probabilidad.
3. Probabilidad condicional.
4. Independencia de sucesos.

1.1. Conceptos básicos: suceso aleatorio

- **Experimento:** cualquier acción que puede dar lugar a resultados identificables u observables.
  - **Determinista:** ante las mismas condiciones de partida, se obtiene el mismo resultado.
  - **Aleatorio:** ante las mismas condiciones, el resultado puede ser diferente y no se sabe cuál de los resultados posibles ocurrirá (sujeto a incertidumbre).

- **Suceso elemental o básico, s:** cada uno de los resultados más simples que se pueden obtener al realizar el experimento aleatorio.

- **Espacio muestral, E:** conjunto de todos los sucesos elementales que se pueden dar al realizar el experimento aleatorio.

1.1. Conceptos básicos: suceso aleatorio

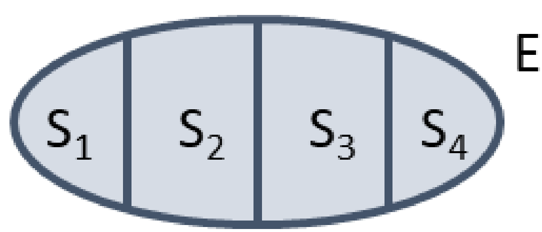
- **Suceso aleatorio S_i:** Subconjunto del Espacio muestral. **Ocurre** si el experimento da lugar a uno de los resultados básicos que lo constituyen.

- **Ejemplo 1:** experimento aleatorio → lanzar un dado una vez.
  - Sucesos elementales: que salga 1, 2, 3, 4, 5 o 6.
  - Espacio muestral: E ={1, 2, 3, 4, 5, 6}
  - Suceso 1: que salga par → S_1={2, 4, 6}
  - S_1 ocurre si al tirar el dado sale 2, 4 o 6.

- **Ejemplo 2:** experimento aleatorio → lanzar una moneda al aire 2 veces.
  - Sucesos elementales: (c, c), (c, +), (+, c), (+, +).
  - Espacio muestral: E ={(c, c), (c, +), (+, c), (+, +)}
  - Suceso: que salgan los dos lados → S_1= {(c, +), (+, c)}
  - S_1 ocurre si al hacer las tiradas sale, (c, +) o (+, c).

- **Operaciones con sucesos.**
  - **Unión de sucesos:** $\bigcup_{i=1}^{n} \boldsymbol{S_i}$ suceso compuesto por los sucesos elementales comunes y no comunes ("o").
  - **Intersección de sucesos:** $\bigcap_{i=1}^{n} \boldsymbol{S_i}$ suceso compuesto por los sucesos elementales comunes ("y").

- Si $\bigcap_{i=1}^{n} \boldsymbol{S_i} = \emptyset$ (suceso imposible), los sucesos son disjuntos.
- Si los sucesos son disjuntos 2 a 2 ($S_i \cap S_j = \emptyset\ \forall i, j$) y $\bigcup_{i=1}^{n} \boldsymbol{S_i} = \boldsymbol{E}$, los sucesos constituyen una partición del espacio muestral E.



1.2. Conceptos básicos: operaciones con sucesos

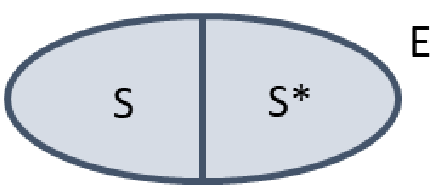
- **Ejemplo:** experimento aleatorio → lanzar un dado una vez.
  - Suceso 1: número par S_1={2, 4, 6}
  - Suceso 2: número impar S_2={1, 3, 5}
  - Suceso 3: Número mayor que 4 S_3={5, 6}

Determinar:

- $S_1 \cap S_2$
- $S_1 \cup S_2$
- $S_1 \cap S_3$
- $S_1 \cup S_3$
- $S_2 \cap S_3$
- $S_2 \cup S_3$

¿Hay alguna agrupación de estos sucesos que constituya una partición del espacio muestral?

- **Propiedades de las operaciones con sucesos.**
  - **Conmutativa:** $S_1 \cup S_2 = S_2 \cup S_1$
    $S_1 \cap S_2 = S_2 \cap S_1$
  - **Asociativa:** $S_1 \cup (S_2 \cup S_3) = (S_1 \cup S_2) \cup S_3$
    $S_1 \cap (S_2 \cap S_3) = (S_1 \cap S_2) \cap S_3$
  - **Distributiva:** $S_1 \cap (S_2 \cup S_3) = (S_1 \cap S_2) \cup (S_1 \cap S_3)$
    $S_1 \cup (S_2 \cap S_3) = (S_1 \cup S_2) \cap (S_1 \cup S_3)$
- **Sucesos complementarios:** Dos sucesos, S y S*, son complementarios si ambos constituyen una partición del espacio muestral.



2. Concepto de probabilidad

- **Probabilidad:** medida de la certidumbre de que ocurra un suceso (def. intuitiva).

- **Definición clásica (Laplace, 1812):** Al diseñar el experimento aleatorio, cociente entre el número de casos (sucesos elementales) favorables al suceso y el número total de casos, siempre que sean igualmente probables (regla de Laplace).

$$P(A) = \frac{número\ de\ casos\ favorables\ a\ A}{número\ de\ casos\ posibles}$$

  - Inconveniente 1: se supone casos equiprobables.
  - Inconveniente 2: Se supone un nº de casos posibles finito.

- **Definición frecuentista (Von Mises, 1919):**

$$P(A) = \lim_{N\to\infty} \frac{número\ de\ veces\ favorables\ a\ A}{número\ de\ veces\ que\ se\ repite\ el\ experimento}$$

2. Concepto de probabilidad

- **Definición axiomática de Kolmogorov (1933):** proporciona un modelo matemático a la Teoría de la probabilidad, aunque no ofrece un método práctico de obtención de probabilidades, para lo cual se apuesta por la definición frecuentista.
  - **Colección de sucesos Ω:** conjunto de sucesos que pueden generarse a partir del espacio muestral E, mediante las operaciones de unión, intersección y complementariedad.
  - **Espacio probabilizable:** Par (E, Ω).

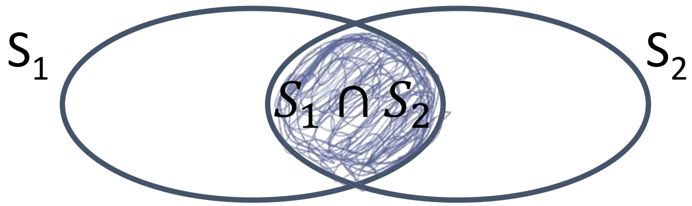
- **Axiomas:**
  1. Si S es un elemento de la colección Ω, existe un número no negativo llamado probabilidad del suceso S.
  2. $P(E) = 1$.
  3. Dada una sucesión numerable de sucesos disjuntos dos a dos, se verifica que: $P(\bigcup_{i=1}^{\infty} S_i) = \sum_{i=1}^{\infty} P(S_i)$.

- **Espacio de probabilidad:** terna (E, Ω, P).

2. Concepto de probabilidad

- **Teoremas derivados de los axiomas de Kolmogorov (1933):**

  1. $P(\emptyset) = 0$.
  2. Dados n sucesos disjuntos dos a dos $P(\bigcup_{i=1}^{n} S_i) = \sum_{i=1}^{n} P(S_i)$.
  3. $P(S_1 \cup S_2) = P(S_1)$+ $P(S_2)$- $P(S_1 \cap S_2)$.



  4. Si $S_1 \subset S \Longrightarrow P(S_1) \leq P(S)$.
  5. $P(S) \leq 1$
  6. $P(S^*) = 1 - P(S)$

- **Ejemplo:** experimento aleatorio → lanzar un dado una vez.
  - Suceso 1: número par S_1={2, 4, 6}
  - Suceso 2: Número mayor que 4 S_2={5, 6}
  - $P(S_1 \cup S_2) = P(S_1)$+ $P(S_2)$- $P(S_1 \cap S_2) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = 0{,}66$

3. Probabilidad condicional

- **Probabilidad condicional:** Dados 2 sucesos S_1 y S_2, S_2 estará condicionado por S_1 si la probabilidad de que ocurra S_2 depende de si ha ocurrido S_1. La probabilidad condicional será:

$$P(S_2 \mid S_1) = \frac{P(S_1 \cap S_2)}{P(S_1)} \Rightarrow P(S_1 \cap S_2) = P(S_1) \cdot P(S_2 \mid S_1)$$

- **Ejemplo:** En una urna hay 4 bolas negras y 3 bolas blancas. Hacemos 2 extracciones. Si en la 1ª extracción sale blanca, devolvemos la bola a la urna y metemos 2 bolas blancas adicionales. Si sale negra, solo devolvemos la bola a la urna.

  - ¿Qué probabilidad hay de que en la segunda extracción saquemos una blanca (2B) y que anteriormente hayamos sacado una blanca (1B)?

$$P(2B \mid 1B) = \frac{5}{9} = 0{,}5555 \Rightarrow P(2B \cap 1B) = P(1B) \cdot P(2B \mid 1B) = \frac{3}{7} \cdot \frac{5}{9} = 0{,}2380$$

  - ¿Qué probabilidad hay de que en la segunda extracción saquemos una blanca (2B) y que anteriormente hayamos sacado una negra (1N)?

$$P(2B \mid 1N) = \frac{3}{7} = 0{,}4286 \Rightarrow P(2B \cap 1N) = P(1N) \cdot P(2B \mid 1N) = \frac{4}{7} \cdot \frac{3}{7} = 0{,}2449$$

3. Probabilidad condicional

- **Teorema (Regla de la multiplicación):** Dados n sucesos $S_1, S_2, \ldots, S_n$ se verifica que:

$$P\left(\bigcap_{i=1}^{n} S_i\right) = P(S_1) \cdot P(S_2 \mid S_1) \cdot P(S_3 \mid S_1 \cap S_2) \cdots P(S_n \mid S_1 \cap S_2 \ldots \cap S_{n-1})$$

- **Ejemplo:** En una urna hay 4 bolas negras y 3 bolas blancas. Hacemos 3 extracciones. Si en una extracción sale blanca, devolvemos la bola a la urna y metemos 2 bolas blancas adicionales. ¿Qué probabilidad hay de sacar 3 blancas seguidas?

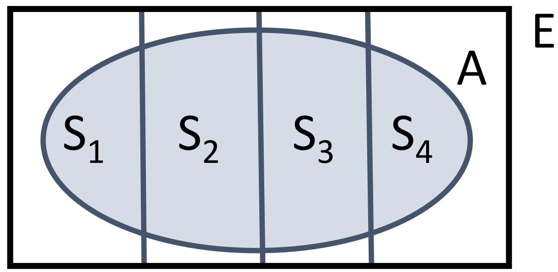
$$P(3\ B\ seguidas) = P(1B) \cdot P(2B \mid 1B) \cdot P(3B \mid 1B \cap 2B) =$$

$$= \frac{3}{7} \cdot \frac{5}{9} \cdot \frac{7}{11} = 0{,}1515$$

3. Probabilidad condicional

- **Teorema de la Probabilidad Total:** Dados un suceso A y n sucesos $S_1, S_2, \ldots, S_n$ que forman una partición del espacio muestral E, tal que $A \cap S_i \neq \emptyset \ \forall i$, $P(A) > 0$ y $P(S_i) > 0$:

$$P(A) = \sum_{i=1}^{n} P(A \mid S_i) \cdot P(S_i)$$



- **Ejemplo:** En una empresa que produce chips tomamos 5 lotes de producto, cada uno compuesto de 500 chips. Hay dos tipos de lotes. Los del tipo 1 (S_1) tienen 480 chips OK y 20 defectuosos. Los del tipo 2 (S_2) tienen 450 chips OK y 50 defectuosos. Tenemos 3 lotes tipo 1 y 2 lotes tipo 2. Si se toma uno de los 5 lotes al azar y se saca de este un chip, ¿qué probabilidad hay de que sea defectuoso?

$$P(D) = P(D \mid S_1) \cdot P(S_1) + P(D \mid S_2) \cdot P(S_2) = \frac{20}{500} \cdot \frac{3}{5} + \frac{50}{500} \cdot \frac{2}{5} = 0{,}0694$$

3. Probabilidad condicional

- **Teorema de Bayes:** Dados un suceso A y n sucesos $S_1, S_2, \ldots, S_n$ que forman una partición del espacio muestral E, tal que $A \cap S_i \neq \emptyset \; \forall i$, $P(A) > 0$ y $P(S_i) > 0$; y siendo conocidas las probabilidades $P(A/S_i)$:

$$P(S_i \mid A) = \frac{P(A \mid S_i) \cdot P(S_i)}{\sum_{i=1}^{n} P(A \mid S_i) \cdot P(S_i)}$$

- **Ejemplo:** En una empresa que produce chips tomamos 5 lotes de producto, cada uno compuesto de 500 chips. Hay dos tipos de lotes. Los del tipo 1 (S_1) tienen 480 chips OK y 20 defectuosos. Los del tipo 2 (S_2) tienen 450 chips OK y 50 defectuosos. Tenemos 3 lotes tipo 1 y 2 lotes tipo 2. Si se toma uno de los 5 lotes al azar y se saca de este un chip, y es defectuoso, ¿qué probabilidad hay de que ese chip pertenezca a un lote del tipo 1?

$$P(S_1 \mid D) = \frac{P(D \mid S_1) \cdot P(S_1)}{\sum_{i=1}^{n} P(D \mid S_i) \cdot P(S_i)} = \frac{\frac{20}{500} \cdot \frac{3}{5}}{\frac{20}{500} \cdot \frac{3}{5} + \frac{50}{500} \cdot \frac{2}{5}} = 0{,}375$$

4. Independencia de sucesos

- **Suceso S_2 condicionado por el suceso S_1:** $P(S_2) \neq P(S_2 \mid S_1)$ con $P(S_1) > 0$.

  - **Ejemplo:** Un experimento consiste en extraer 2 bolas de una urna al azar, sucesivamente, y **sin** reemplazamiento. En la urna hay 6 bolas blancas y 4 negras. ¿Cuál es la probabilidad de que la segunda bola sea blanca?

    Probabilidad de extraer de la urna 2ª bola blanca (T. Prob. Total): $P(2B) = P(2B \mid 1B) \cdot P(1B) + P(2B/1N) \cdot P(1N) = \frac{5}{9} \cdot \frac{6}{10} + \frac{6}{9} \cdot \frac{4}{10} = 0{,}6$. Pero en el caso de la 2ª extracción, si la 1ª fue blanca: $P(2B/1B) = 0{,}55$ y si fue negra $P(2B \mid 1N) = 0{,}67$.

- **Sucesos S_1 y S_2 independientes:** $P(S_2) = P(S_2 \mid S_1)$ con $P(S_1) > 0$.

  - **Ejemplo:** Mismo experimento; pero las extracciones se realizan **con** reemplazamiento. ¿Cuál es la probabilidad de que la segunda bola sea blanca?

    Probabilidad de extraer de la urna 2ª bola blanca (T. Prob. Total): $P(2B) = \frac{6}{10} = 0{,}6$.

    En la 2ª extracción, si la 1ª fue blanca $P(2B \mid 1B) = \frac{6}{10}$ y si la 2ª fue negra $P(2B \mid 1N) = \frac{6}{10}$.

4. Independencia de sucesos

- **Sucesos S_1 y S_2 independientes:** $P(S_2) = P(S_2 \mid S_1)$ con $P(S_1) > 0$.

  - De lo anterior se deduce que:

$$P(S_2) = P(S_2 \mid S_1) = \frac{P(S_1 \cap S_2)}{P(S_1)} \Longrightarrow P(S_1 \cap S_2) = P(S_1) \cdot P(S_2)$$

  - **Ejemplo:** en el ejemplo anterior, la probabilidad de sacar 2 blancas, si es un experimento **con** reemplazamiento:

$$P(1B \cap 2B) = P(1B) \cdot P(2B) = \frac{6}{10} \cdot \frac{6}{10} = 0{,}36$$

- El concepto de independencia de sucesos es generalizable a la ocurrencia de más de dos sucesos. Por ejemplo, S_1, S_2, …, S_n son independientes entre sí si:

$$P(S_1 \cap S_2 \cap \cdots \cap S_n) = P(S_1) \cdot P(S_2) \cdots P(S_n)$$


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 4.2.

Variables aleatorias unidimensionales y bidimensionales

Contenido

1. Variable aleatoria. Definición. Función de distribución.
2. Variables aleatorias discretas. Función de cuantía.
3. Variables aleatorias continuas. Función de densidad.
4. Variables aleatorias bidimensionales.
5. Probabilidades condicionales.
6. Independencia de variables aleatorias.

1.1. Variable aleatoria: definición

- **Variable aleatoria:** Dado el espacio de probabilidad (E, Ω, P), se define la variable aleatoria ξ como la aplicación $\xi : E \to \mathbb{R}$

- Los sucesos pasan a convertirse en números reales, y las probabilidades de ocurrencia de los primeros en las probabilidades asociadas a los valores que puede tomar la variable.

- Es decir, una variable aleatoria ξ es una cantidad variable cuyos valores dependen del azar y para la cual existe una **distribución de probabilidad**. Queda definida al conocer:
  - Su **campo de variación**.
  - **Conjunto de probabilidades** con que toma valores ese campo.

- **Ejemplo:**
  - Experimento: lanzar tres veces una moneda al aire y contar el número de caras.
  - ξ="Número de caras"

$$E \dashrightarrow \mathbb{R} \qquad P$$

$E \longrightarrow$	$\mathbb{R}$	P
{+,+,+} $\longrightarrow$	0	$P(\xi = 0) = \frac{1}{8}$
{c,+,+}; {+,c,+}; {+,+,c} $\longrightarrow$	1	$P(\xi = 1) = \frac{3}{8}$
{c,c,+}; {c,+,c}; {+,c,c} $\longrightarrow$	2	$P(\xi = 2) = \frac{3}{8}$
{c,c,c} $\longrightarrow$	3	$P(\xi = 3) = \frac{1}{8}$

Campo de variación

Distribución de probabilidad

- **Función de distribución:** Es una de las formas de expresar la probabilidad asociada a una variable aleatoria ξ.

$$F(x) = P(\xi \leq x)\ \forall x \in \mathbb{R}\ y\ \xi \in S \text{ con } S\text{: campo de variación}$$

- **Propiedades:**

1. $\forall x \in \mathbb{R}\ \ 0 \leq F(x) \leq 1$
2. $\lim\limits_{x \longrightarrow +\infty} F(x) = 1$ y $\lim\limits_{x \longrightarrow -\infty} F(x) = 0$
3. $x_1 < x_2 \rightarrow F(x_1) \leq F(x_2)$ (función no decreciente).
4. $\lim\limits_{\substack{h \longrightarrow 0 \\ h > 0}} F(x+h) = F(x)$ (continua por la derecha).

2. Variable aleatoria discreta: Función de cuantía.

- **Variable aleatoria ξ de tipo discreto:** cuando su campo de variación se compone de un número finito o infinito numerable de puntos, existiendo entre ellos masa discreta (puntos de salto).
  - **Ejemplo:** en el experimento del lanzamiento de un dado, la variable aleatoria ξ=“puntuación alcanzada” es discreta.

- **Función de probabilidad o cuantía:** es la masa de probabilidad asignada a un punto perteneciente al campo de variación de ξ de tipo discreto:

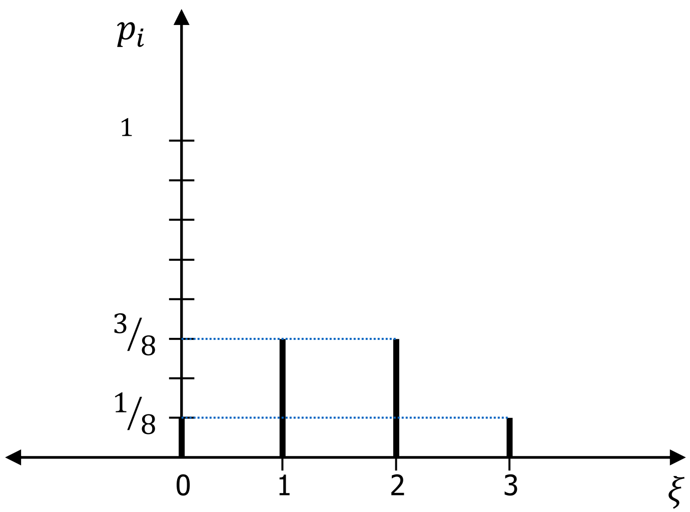
$$P(\xi = x_i) = p_i \quad \text{con } p_i > 0 \quad \text{y} \quad \sum_{i=1}^{\infty} p_i = 1$$

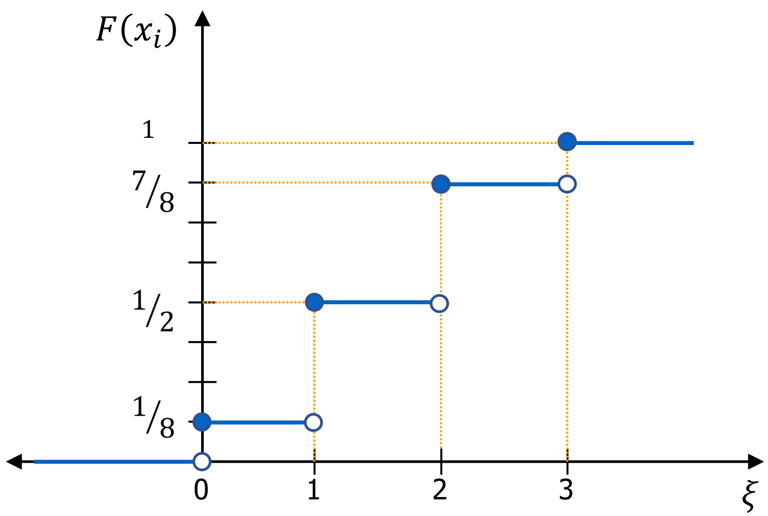
- **Función de distribución de una variable discreta:**

$$F(x) = P(\xi \leq x) = \sum_{x_i \leq x} p_i$$

- **Ejemplo:** Experimento: lanzar 3 veces una moneda. ξ="número de caras".

Espacio muestral	ξ	*F. de cuantía*	**F. de distribución**
E	x_i	$P(\xi = x_i) = p_i$	$F(x_i) = P(\xi \leq x_i)$
{+,+,+}	1	$P(\xi = 0) = \frac{1}{8}$	$F(0) = \frac{1}{8}$
{c,+,+}; {+,c,+}; {+,+,c}	2	$P(\xi = 1) = \frac{3}{8}$	$F(1) = \frac{1}{2}$
{c,c,+}; {c,+,c}; {+,c,c}	3	$P(\xi = 2) = \frac{3}{8}$	$F(2) = \frac{7}{8}$
{c,c,c}	4	$P(\xi = 3) = 1/8$	$F(3) = 1$





3. Variables aleatorias continuas: Función de densidad.

- **Variable aleatoria ξ de tipo continuo:** el número de posibles valores que puede tomar es infinito no numerable.

  - **Ejemplo:** en el experimento de pesar una pieza de una máquina con una báscula de precisión, la variable aleatoria ξ="peso de la pieza" es continua (salvo restricciones técnicas).

- La **probabilidad concentrada en un punto** de una variable aleatoria continua es **0**.

  - Entre 2 puntos del campo de variación de ξ existen infinitos puntos intermedios pertenecientes a dicho campo. Así, la probabilidad asignada a un punto concreto debe ser nula, solo pueden calcularse probabilidades de intervalos pertenecientes al campo de variación.

- Su **función de distribución** es continua y su primera derivada existe y también es continua.

3. Variables aleatorias continuas: Función de densidad.

- **Función de densidad de una variable continua:** es otra función que expresa la distribución de probabilidad de una variable continua. Se define como la derivada de la función de distribución de probabilidad. **NO** es una probabilidad.

$$f(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x)$$

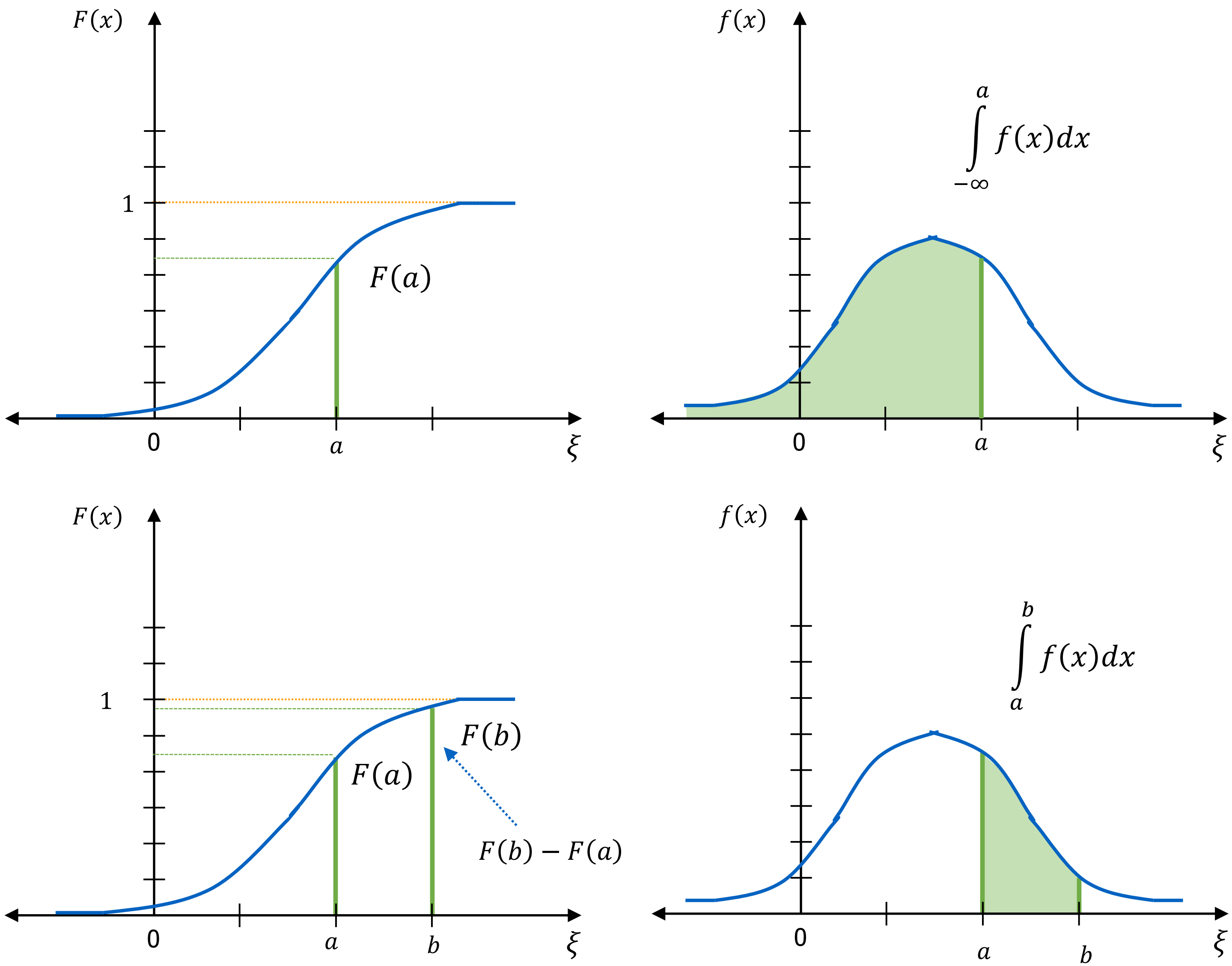
- $f(x) \cdot d(x)$ llamada **probabilidad elemental** se interpreta como la probabilidad asignada a un **intervalo infinitesimal** de puntos pertenecientes al campo de variación.

- **Condiciones básicas:**
  1. $f(x) \geq 0 \ \forall x$
  2. $P(E) = P(-\infty < \xi < \infty) = F(\infty) - F(-\infty) = \int_{-\infty}^{\infty} f(x) \cdot d(x) = 1$

- **Cálculo de probabilidades con variables aleatorias continuas:**

$P(\xi \le a)$

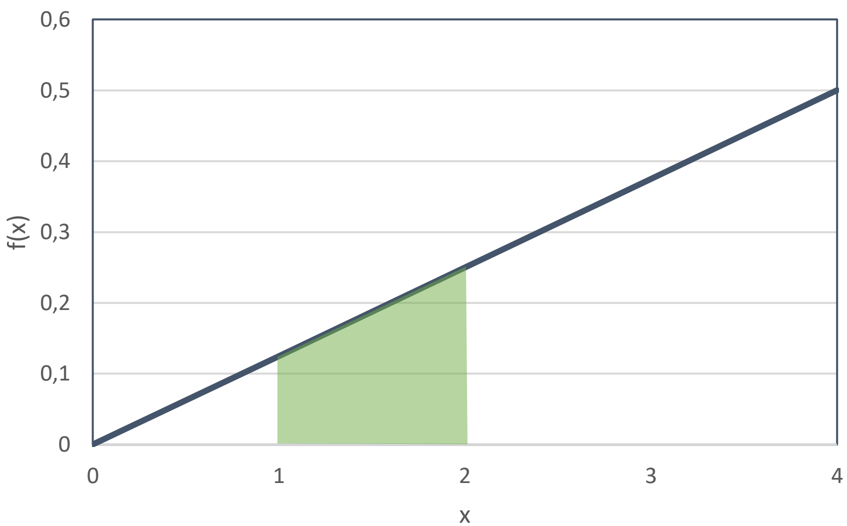
$P(a < \xi \le b)$



- **Ejemplo:** Dada ξ con función de densidad:

$$f(x) = \begin{cases} \frac{1}{8} \cdot x & si \quad 0 < x < 4 \\ 0 & en\,el\,resto \end{cases}$$

Hallar $P(1 < \xi < 2)$.



$$P(1 < \xi < 2) = \int_1^2 f(x) \cdot dx =$$

$$\int_1^2 \frac{1}{8} \cdot x \cdot dx = \frac{1}{8} \cdot \int_1^2 x \cdot dx =$$

$$\frac{1}{8} \cdot \left[\frac{x^2}{2}\right]_1^2 = \frac{1}{8} \cdot \left(2 - \frac{1}{2}\right) = \frac{3}{16}$$

4. Variables aleatorias bidimensionales

- **Variable aleatoria bidimensional:** vector $(\xi;\eta)$ cuyas componentes son dos variables aleatorias definidas en el espacio probabilizado (E,Ω,P).

- **Función de distribución conjunta:** Dada $(\xi;\eta)$ cuyo campo de variación es el recinto formado por $-\infty < x < \infty$ y $-\infty < y < \infty$:

$$F(x;y) = P(\xi \leq x;\eta \leq y)$$

- **Caso discreto:** una v. a. bidimensional es de tipo discreto si sus componentes son variables de tipo discreto.
    - Función de cuantía conjunta: $p_{ij} = P\left(\xi = x_i;\eta = y_j\right)$ tal que $\sum_{i,j} p_{ij} = 1$
    - Función de cuantía marginal de ξ: $p_{i\bullet} = P(\xi = x_i) = \sum_j p_{ij}$
    - Función de cuantía marginal de η: $p_{\bullet j} = P\left(\eta = y_j\right) = \sum_i p_{ij}$

- **Ejemplo** caso discreto:

$\xi \vert \eta$	1,2	2,3	7,8	9,6	$p_{i\bullet}$
1	0,1	0,1	0,2	0,3	0,7
2	0,05	0,05	0,025	0,025	0,15
3	0,05	0,025	0,05	0,025	0,15
$p_{\bullet j}$	0,2	0,175	0,275	0,35	1

- $p_{1;1} = P(\xi = 1 \cap \eta = 1{,}2) = 0{,}1$
- $p_{1\bullet} = P(\xi = 1) = \sum_{j=1}^{4} p_{1j} = 0{,}1 + 0{,}1 + 0{,}2 + 0{,}3 = 0{,}7$
- $p_{\bullet 1} = P(\eta = 1{,}2) = \sum_{i=1}^{3} p_{i1} = 0{,}1 + 0{,}05 + 0{,}05 = 0{,}2$

- **Caso continuo:** una v. a. bidimensional es de tipo continuo si sus componentes son variables de tipo continuo.

  - Función de densidad conjunta:

$$f(x;y)=\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{P(x<\xi\leq x+\Delta x;y<\eta\leq y+\Delta y)}{\Delta x\cdot\Delta y}$$

Derivada mixta doble de la función de distribución conjunta de ξ y η. NO es una probabilidad.

- **Condiciones básicas:**

1. $f(x;y)\geq 0$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x;y)\cdot dy\cdot dx=1$

- Función de densidad marginal de ξ:

$$f_{\xi}(x) = \int_{-\infty}^{\infty} f(x;y) \cdot dy$$

- Función de densidad marginal de η:

$$f_{\eta}(y) = \int_{-\infty}^{\infty} f(x;y) \cdot dx$$

- **Ejemplo** caso continuo:

  - Función de **distribución conjunta** de la variable $(\xi;\eta)$:

$$F(x;y)=\frac{x^2\cdot y+6\cdot x\cdot y^2}{7} \text{ con } 0\leq\xi\leq 1 \text{ y } 0\leq\eta\leq 1$$

  - Función de **densidad conjunta**:

$$\frac{\partial F(x;y)}{\partial x}=\frac{1}{7}\cdot(2x\cdot y+6\cdot y^2)$$

$$f(x;y)=\frac{\partial^2 F(x;y)}{\partial x\partial y}=\frac{\partial\left(\frac{2\cdot x\cdot y+6\cdot y^2}{7}\right)}{\partial y}=\frac{2\cdot x+12\cdot y}{7}=$$

$$\frac{2}{7}(x+6\cdot y) \text{ con } 0\leq\xi\leq 1 \text{ y } 0\leq\eta\leq 1$$

- **Ejemplo** caso continuo:
  - Función de **densidad conjunta**:

$$f(x;y) = \frac{2}{7}(x + 6 \cdot y) \text{ con } 0 \le \xi \le 1 \text{ y } 0 \le \eta \le 1$$

  - Función de **densidad marginal de $\boldsymbol{\xi}$**:

$$f_\xi(x) = \int_{-\infty}^{\infty} f(x;y) \cdot dy = \int_0^1 \frac{2}{7}(x + 6 \cdot y) \cdot dy =$$

$$\frac{2}{7}\left(x \cdot \int_0^1 dy + 6 \int_0^1 y \cdot dy\right) = \frac{2}{7}\left(x \cdot [y]_0^1 + 6\left[\frac{y^2}{2}\right]_0^1\right) =$$

$$= \frac{2}{7}(x + 3) \text{ con } 0 \le \xi \le 1$$

  - Función de **densidad marginal de $\boldsymbol{\eta}$**:

$$f_\eta(y) = \int_{-\infty}^{\infty} f(x;y) \cdot dx = \int_0^1 \frac{2}{7}(x + 6 \cdot y) \cdot dx =$$

$$\frac{2}{7}\left(\int_0^1 x \cdot dx + 6 \cdot y \int_0^1 dx\right) = \frac{2}{7}\left(\left[\frac{x^2}{2}\right]_0^1 + 6 \cdot y \cdot [x]_0^1\right) =$$

$$\frac{2}{7}\left(\frac{1}{2} + 6 \cdot y\right) \text{ con } 0 \le \eta \le 1$$

5. Probabilidades condicionales

- Recordando la **probabilidad condicional:** Dados 2 sucesos S_1 y S_2, S_2 estará condicionado por S_1 si la probabilidad de que ocurra S_2 depende de si ha ocurrido S_1. La probabilidad condicional será:

$$P(S_2 \mid S_1) = \frac{P(S_1 \cap S_2)}{P(S_1)} \Rightarrow P(S_1 \cap S_2) = P(S_1) \cdot P(S_2 \mid S_1)$$

- Ahora, en lugar de sucesos se trabajará con valores de las variables aleatorias que forman la variable bidimensional $(\xi; \eta)$. **Según sean variables discretas o continuas, se aplicarán funciones de probabilidad o cuantía, o funciones de densidad**.

5. Probabilidades condicionales

- **Ejemplo** variable discreta bidimensional $(\xi;\eta)$:

$\xi\|\eta$	1,2	2,3	7,8	9,6	$p_{i\bullet}$
1	0,1	0,1	0,2	0,3	0,7
2	0,05	0,05	0,025	0,025	0,15
3	0,05	0,025	0,05	0,025	0,15
$p_{\bullet j}$	0,2	0,175	0,275	0,35	1

$$P(\xi = 2 \mid \eta < 7{,}8) = \frac{P(\xi = 2 \cap \eta < 7{,}8)}{P(\eta < 7{,}8)} = \frac{0{,}05 + 0{,}05}{0{,}2 + 0{,}175} = 0{,}266$$

$$P(\eta \geq 9{,}6 \mid \xi < 2) = \frac{P(\eta \geq 9{,}6 \cap \xi < 2)}{P(\xi < 2)} = \frac{0{,}3}{0{,}7} = 0{,}428$$

$$P(\xi < 2 \mid \eta = 1{,}2) = \frac{P(\xi < 2 \cap \eta = 1{,}2)}{P(\eta = 1{,}2)} = \frac{0{,}1}{0{,}2} = 0{,}5$$

$$P(\xi \leq 2 \mid \eta \leq 2{,}3) = \frac{P(\xi \leq 2 \cap \eta \leq 2{,}3)}{P(\eta \leq 2{,}3)} = \frac{0{,}1 + 0{,}1 + 0{,}05 + 0{,}05}{0{,}2 + 0{,}175} = 0{,}8$$

- **Ejemplo** variable continua: Dada la función de densidad conjunta de la variable bidimensional $(\xi;\eta)$:

$$f(x)=\frac{2}{7}\cdot(x+6y)\text{ con }0\leq\xi\leq 1;\,0\leq\eta\leq 1$$

Calcular
$P(0{,}6<\eta\leq 0{,}7\mid 0{,}4<\xi\leq 0{,}5)$

Vimos en ejemplos anteriores que: $f_{\xi}(x)=\frac{2}{7}\cdot(x+3)$

$$P(0{,}6<\eta\leq 0{,}7\mid 0{,}4<\xi\leq 0{,}5)=$$

$$=\frac{\int_{0,4}^{0,5}\int_{0,6}^{0,7}\frac{2}{7}\cdot(x+6y)\cdot dy\cdot dx}{\int_{0,4}^{0,5}\frac{2}{7}\cdot(x+3)\cdot dx}=\frac{0{,}0124}{0{,}098}=0{,}126$$

6. Independencia de variables aleatorias

- **Caso general: sucesos independientes:** $P(S_2) = P(S_2 \mid S_1)\ con\ P(S_1) > 0$. De donde se deduce que:

$$P(S_2 \mid S_1) = P(S_2) = \frac{P(S_1 \cap S_2)}{P(S_1)} \Rightarrow P(S_1 \cap S_2) = P(S_1) \cdot P(S_2)$$

- Caso de **variables aleatorias discretas**: las componentes de la variable aleatoria bidimensional $(\xi;\eta)$ son independientes si:

$$p_{ij} = p_{i\bullet} \cdot p_{\bullet j}$$

- Caso de **variables aleatorias continuas**: las componentes de la variable aleatoria bidimensional $(\xi;\eta)$ son independientes si:

$$f(x;y) = f_{\xi}(x) \cdot f_{\eta}(y)$$

¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 4.3.

Características de las variables aleatorias

1. Momentos de una variable aleatoria.
2. Esperanza matemática.
3. Varianza y desviación típica.
5. Momentos bidimensionales.

1. Momentos de una variable aleatoria

- Momentos **respecto al origen** de una variable aleatoria ξ:
  - General: $\alpha_r = E(\xi^r)$
  - Variable de tipo **discreto**: $\alpha_r = \sum_i x_i^r \cdot p_i$
  - Variable de tipo **continuo**: $\alpha_r = \int_{-\infty}^{+\infty} x^r \cdot f(x) \cdot dx$

- Momentos **respecto a la media** de una variable aleatoria ξ:
  - General: $\mu_r = E(\xi - \mu)^r$ con $\mu = \alpha_1$
  - Variable de tipo **discreto**: $\mu_r = \sum_i (x_i - \mu)^r \cdot p_i$
  - Variable de tipo **continuo**: $\mu_r = \int_{-\infty}^{+\infty} (x - \mu)^r \cdot f(x) \cdot dx$

- Relación entre momentos respecto a la media y momentos respecto al origen:

$$\mu_r = \sum_{k=0}^{r} (-1)^k \cdot \binom{r}{k} \cdot \mu^k \cdot \alpha_{r-k}$$

- Es el valor central entorno al cual se concentra una distribución de probabilidad.
- Matemáticamente es el **momento respecto al origen de orden 1** de la variable aleatoria ξ:
  - General: $\mu = \alpha_1 = E(\xi)$
  - Variable de tipo **discreto**: $\mu = \alpha_1 = E(\xi) = \sum_i x_i \cdot p_i$
  - Variable de tipo **continuo**: $\mu = \alpha_1 = E(\xi) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx$

- **Ejemplos:**
  - Caso discreto:

x_i	p_i
1	0,2
2	0,3
3	0,5

$$\mu = E(\xi) = 1 \cdot 0{,}2 + 2 \cdot 0{,}3 + 3 \cdot 0{,}5 = 2{,}3$$

  - Caso continuo: ξ con $f(x) = 4 \cdot x^3;\ 0 \leq x \leq 1$

$$\mu = E(\xi) = \int_0^1 x \cdot 4 \cdot x^3\, dx = 4 \cdot \int_0^1 x^4 dx = 4 \cdot \left[\frac{x^5}{5}\right]_0^1 = \frac{4}{5}$$

2. Esperanza matemática

- **Propiedades:**
  - $E(C) = C$ con C una constante.
  - $E(\xi_1 \pm \xi_2 \pm \cdots \pm \xi_n) = E(\xi_1) \pm E(\xi_2) \pm \cdots \pm E(\xi_n)$
  - $E(\xi_1 \cdot \xi_2 \cdot \cdots \cdot \xi_n) = E(\xi_1) \cdot E(\xi_2) \cdot \cdots \cdot E(\xi_n)$ si y solo si las variables son independientes.
  - Si $E(\xi) = E(\xi_1) = \mu$ entonces $E(\xi - \mu) = 0$.
  - $E(A + B \cdot \xi) = A + B \cdot E(\xi)$ con A, B constantes.

3. Varianza

- Medida de **dispersión** de una variable aleatoria.

$$Var(\xi) = \sigma^2 = E\big(\xi - E(\xi)\big)^2 = E(\xi - \mu)^2$$

- Matemáticamente, es el momento respecto a la media de orden 2:

$$Var(\xi) = \sigma^2 = \mu_2 = E\left(\xi^2\right) - [E(\xi)]^2 = \alpha_2 - \alpha_1^2$$

- La raíz positiva de la varianza es la **desviación típica *σ***:

$$\sigma = +\sqrt{\sigma^2}$$

- **Ejemplos:**

  - Caso discreto:

x_i	p_i
1	0,2
2	0,3
3	0,5

$$E(\xi) = \mu = \alpha_1 = 1 \cdot 0{,}2 + 2 \cdot 0{,}3 + 3 \cdot 0{,}5 = 2{,}3$$

$$E\left(\xi^2\right) = \alpha_2 = 1^2 \cdot 0{,}2 + 2^2 \cdot 0{,}3 + 3^2 \cdot 0{,}5 = 5{,}9$$

$$Var(\xi) = \sigma^2 = \mu_2 = \alpha_2 - \alpha_1^2 = 5{,}9 - (2{,}3)^2 = 0{,}61$$

  - Caso continuo: ξ con $f(x) = 4 \cdot x^3;\ 0 \leq x \leq 1$

$$E(\xi) = \mu = \alpha_1 = \int_0^1 x \cdot 4 \cdot x^3\, dx = 4 \cdot \int_0^1 x^4 dx = 4 \cdot \left[\frac{x^5}{5}\right]_0^1 = \frac{4}{5}$$

$$E\left(\xi^2\right) = \alpha_2 = \int_0^1 x^2 \cdot 4 \cdot x^3\, dx = 4 \cdot \int_0^1 x^5 dx = 4 \cdot \left[\frac{x^6}{6}\right]_0^1 = \frac{4}{6} = \frac{2}{3}$$

$$\sigma^2 = Var(\xi) = \mu_2 = \alpha_2 - \alpha_1^2 = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = 0{,}02$$

3. Varianza

- **Propiedades:**

  - $Var(\xi) = \sigma^2 \geq 0.$

  - $Var(C) = 0$

  - $Var(\xi_1 \pm \xi_2 \pm \cdots \pm \xi_n) = Var(\xi_1) + Var(\xi_2) + \cdots + Var(\xi_n)$ si y solo si las variables son independientes.

  - $Var(A + B \cdot \xi) = B^2 \cdot Var(\xi)$ con *A*, *B* constantes.

5. Momentos bidimensionales

- Momentos **respecto al origen** de las variables ξ y η:
  - General: $\alpha_{rs} = E(\xi^r \cdot \eta^s)$
  - Variables de tipo **discreto**: $\alpha_{rs} = \sum_i \sum_j x_i^r \cdot y_j^s \cdot p_{ij}$
  - Variables de tipo **continuo**:

$$\alpha_{rs} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^r \cdot y^s \cdot f(x; y) \cdot dx \cdot dy$$

- Momentos **respecto a las medias** de de las variables ξ y η :
- General: $\mu_{rs} = E\left[\left(\xi - E(\xi)\right)^r \cdot \left(\eta - E(\eta)\right)^s\right]$
- Variables de tipo **discreto**:

$$\mu_{rs} = \sum_i \sum_j \left(x_i - E(\xi)\right)^r \cdot \left(y_j - E(\eta)\right)^s \cdot p_{ij}$$

- Variables de tipo **continuo**:

$$\mu_{rs} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(x - E(\xi)\right)^r \cdot \left(y \quad - E(\eta)\right)^s \cdot f(x; y) \cdot dy \cdot dx$$

5. Momentos bidimensionales

- Momentos **respecto al origen** más comunes:
  - $\alpha_{10} = E(\xi)$
  - $\alpha_{01} = E(\eta)$

- Momentos **respecto a las medias** más comunes:
  - $\mu_{20} = Var(\xi) = \alpha_{20} - \alpha_{10}^2$
  - $\mu_{02} = Var(\eta) = \alpha_{02} - \alpha_{01}^2$
  - Covarianza: $\mu_{11} = Cov(\xi;\eta) = \sigma_{\xi\eta} = \alpha_{11} - \alpha_{10} \cdot \alpha_{01}$

- Si dos variables aleatorias son **independientes**, su covarianza es **nula**.


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 5.1.

Modelos de distribución de probabilidad discretos univariantes.

Contenido

1. Introducción

- Las variables aleatorias quedan definidas mediante su campo de variación y su distribución de probabilidad.

- En muchos fenómenos aleatorios o experimentos, la **distribución de probabilidad** de las variables aleatorias que los representan se rigen según ciertos **modelos matemáticos**.

- En este tema se mostrarán los principales modelos de probabilidad univariantes en el caso de **variables discretas**.

- $B(1;p)$ (también se llama binomial 1;p)

- **Experimento:** un experimento que solo puede presentar dos resultados excluyentes: éxito o fracaso.
- **Variable cuya probabilidad modeliza:** ξ que toma valor 1 si el resultado ha sido éxito y 0 si ha sido fracaso.
- **Parámetros:**
  - ***p:*** probabilidad de obtener el resultado de éxito.
  - ***q:*** probabilidad de obtener el resultado de fracaso.
- **Función de probabilidad o cuantía:**
  - $\boldsymbol{P(\xi = x) = p^{x} \cdot q^{1-x}}$

  con $q = 1 - p, x = 0; 1.$
- **Características:**
  - $\mu = E(\xi) = p$
  - $\sigma^2 = Var(\xi) = p \cdot q$

- $B(1;p)$ (también se llama binomial 1;p)

- **Ejemplo:** Se tiene una moneda trucada tal que la probabilidad de que al lanzarla salga cara es de p = 0,6. Probabilidad de que al lanzar la moneda se obtenga cruz.

- **Variable cuya probabilidad modeliza:** ξ ="sale cara o cruz". Hemos elegido que el éxito sea que sale cara (ξ = 1).

$$\xi \to B(1;p)$$

- $P(\xi = 0) = p^x \cdot q^{1-x} = 0{,}6^0 \cdot 0{,}4^{1-0} = 0{,}4$

- $B(n;p)$

- **Experimento:** repetir **n veces** un experimento de **Bernoulli**. Cada repetición es independiente e idénticamente distribuida al resto.
- **Variable cuya probabilidad modeliza:** ξ que cuenta en **número de éxitos** que ha habido en las n repeticiones del experimento. $\xi \rightarrow B(n;p)$
- **Parámetros:**
  - ***n:*** número de veces que se repite el experimento.
  - ***p:*** probabilidad de obtener el resultado de éxito.
  - ***q:*** probabilidad de obtener el resultado de fracaso.
- **Función de probabilidad o cuantía:**
  - $\boldsymbol{P(\xi = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}} = \frac{n!}{x! \cdot (n-x)!} \cdot p^x \cdot q^{n-x}$

    con $q = 1 - p$, y $x = 0, 1, 2, 3, \ldots.$

- $B(n;p)$

- **Características:**
  - $\mu = E(\xi) = n \cdot p$
  - $\sigma^2 = Var(\xi) = n \cdot p \cdot q$

- **Propiedad:**
  - Aditiva o reproductiva: si $\eta = \xi_1 + \xi_2 + \cdots + \xi_m$ con $\xi_j \to B(n_j;p)$ independiente del resto, entonces se cumple $\eta \to B(n_1 + n_2 + \cdots + n_m;p)$

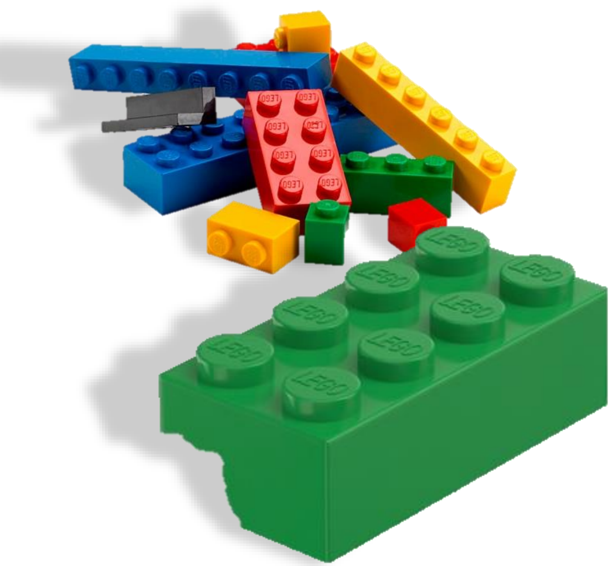
3. Distribución Binomial

- $B(n;p)$

- **Ejemplo:** Suponiendo que la probabilidad de que un estudiante acabe el grado en ADE es de 0,4 calcular, para un grupo de 5 estudiantes:
  - Probabilidad de que ninguno obtenga el grado:
    Variable cuya probabilidad modeliza: ξ =“número de aprobados”. $\xi \rightarrow B(5;0{,}4)$

    $$P(\xi=0)=\binom{n}{x}\cdot p^x\cdot q^{n-x}$$
    $$=\frac{n!}{x!\cdot(n-x)!}\cdot p^x\cdot q^{n-x}=\binom{5}{0}\cdot 0,4^0\cdot 0,6^5$$
    $$=1\cdot 1\cdot 0,6^5=0{,}0777$$
  - Probabilidad de que al menos dos lo obtengan:
    $$P(\xi\geq 2)=1-P(\xi<2)$$
    $$=1-[P(\xi=0)+P(\xi=1)]=$$
    $$=1-\left[\binom{5}{0}\cdot 0,4^0\cdot 0,6^5+\binom{5}{1}\cdot 0,4^1\cdot 0,6^4\right]$$
    $$=1-\left(0,6^5+5\cdot 0{,}4\cdot 0,6^4\right)=0{,}663$$

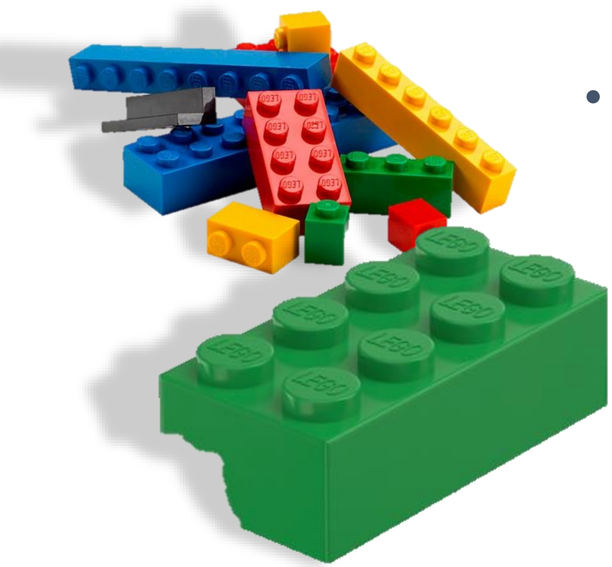
- $Po(\lambda)$



- **Experimento:** repetir **n veces** un experimento de **Bernoulli**. Cada repetición es independiente e idénticamente distribuida al resto. La **probabilidad de obtener un número de éxitos elevado decrece rápidamente**.
- **Variable cuya probabilidad modeliza:** ξ que cuenta el **número de éxitos** en las n repeticiones del experimento. $\xi \rightarrow Po(\lambda)$
- **Parámetros:**
  - λ**:** orden de la distribución. $\lambda = n \cdot p$
- **Función de probabilidad o cuantía:**
  - $\boldsymbol{P(\xi = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}}$

  con $\lambda > 0$, y $x = 0, 1, 2, 3, \ldots$
- **Características:**
  - $\mu = E(\xi) = \lambda$
  - $\sigma^2 = Var(\xi) = \lambda$

- $Po(\lambda)$



- **Propiedad:**
  - Aditiva o reproductiva: si $\eta = \xi_1 + \xi_2 + \cdots + \xi_m$ con $\xi_j \to Po(\lambda_j)$ independiente del resto, entonces se cumple que $\eta \to Po(\lambda_1 + \lambda_2 + \cdots + \lambda_m)$

- **Ejemplo:** Una máquina fabrica un tipo de piezas con una probabilidad de que una salga defectuosa de p = 0,001. Si tenemos un lote de 2000 piezas:
  - Probabilidad de que salgan en el lote 3 defectuosas: Se podría calcular con la binomial; pero es más fácil con la Poisson.

    Variable cuya probabilidad modeliza: ξ =“número de piezas defectuosas”. $\xi \to B(2000; 0{,}001) \to Po(2)$

$$P(\xi = 3) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-2}2^3}{3!} = 0{,}1804$$

- $BN(r;p)$

- **Experimento:** repetir un experimento de **Bernoulli**. Cada repetición es independiente e idénticamente distribuida al resto.

- **Variable cuya probabilidad modeliza:** ξ que cuenta en **número de fracasos** (resultado B) **antes de que se dé el r-simo éxito** (resultado A). $\xi \rightarrow BN(r;p)$

- **Parámetros:**
  - ***r:*** número de éxitos que se han de dar.
  - ***p:*** probabilidad de obtener el resultado de éxito.
  - ***q:*** probabilidad de obtener el resultado de fracaso.

- **Función de probabilidad o cuantía:**
  - $\boldsymbol{P(\xi = x) = \binom{x+r-1}{x} \cdot p^r \cdot q^x} = \frac{(x+r-1)!}{x! \cdot (r-1)!} \cdot p^r \cdot q^x$

  con $q = 1 - p$, y $x = 0, 1, 2, 3, \ldots$.

- $BN(r;p)$

- **Características:**
  - $\mu = E(\xi) = r \cdot \frac{q}{p}$
  - $\sigma^2 = Var(\xi) = r \cdot \frac{q}{p^2}$

- **Propiedad:**
  - Aditiva o reproductiva: si $\eta = \xi_1 + \xi_2 + \cdots + \xi_m$ con $\xi_j \to BN(r_j;p)$ independiente del resto, entonces se cumple que $\eta \to BN(r_1 + r_2 + \cdots + r_m;p)$

- **Nota:**
  - $\xi \to BN(1;p)$ también se llama distribución geométrica: $\xi \to G(p)$

5. Distribución binomial negativa

- $BN(r; p)$

- **Ejemplo:** Dos equipos de baloncesto A y B disputan la final de liga al mejor de 7 partidos. El factor campo no influye, y el equipo A tiene una probabilidad de ganar un partido de 0,6. Probabilidad de que A gane la final en 5 partidos.
  - Esto se dará cuando gane 4 partidos y B gane 1. Por tanto, $r = 4$ y $x = 1$.

    Variable cuya probabilidad modeliza: ξ ="número de partidos ganados por B antes de que A gane su 4º partido". $\xi \to BN(4; 0{,}6)$

$$P(\xi = 1) = \binom{1+4-1}{1} \cdot 0{,}6^4 \cdot 0{,}4^1$$

$$= \frac{(1+4-1)!}{1! \cdot (4-1)!} \cdot 0{,}6^4 \cdot 0{,}4^1 = 4 \cdot 0{,}6^4 \cdot 0{,}4$$

$$= 0{,}2074$$

6. Distribución hipergeométrica

- $H(N; n; p)$

- **Experimento:** De una muestra de *N* elementos, hay N_1 de tipo *A* y $N - N_1$ de tipo *B*. Se extraen de la muestra al azar, y **sin reemplazamiento** (no hay independencia de sucesos), n elementos.

- **Variable cuya probabilidad modeliza:** ξ cuenta el **número de elementos tipo A** que hay en los n extraídos. $\xi \to H(N; n; p)$

- **Parámetros:**
  - ***N:*** número de elementos de la muestra.
  - ***n:*** número de elementos de la muestra extraídos.
  - $\boldsymbol{p = \frac{N_1}{N}}$***:*** proporción de elementos tipo A en la muestra.

- **Función de probabilidad o cuantía:**
  - $$\boldsymbol{P(\xi = x) = \frac{\binom{N_1}{x} \cdot \binom{N-N_1}{n-x}}{\binom{N}{n}}} = \frac{\frac{N_1!}{x! \cdot (N_1 - x)!} \cdot \frac{(N-N_1)!}{(n-x)! \cdot [(N-N_1)-(n-x)]!}}{\frac{N!}{n! \cdot (N-n)!}}$$

    con $x = 0, 1, 2, 3, \ldots, n$

6. Distribución hipergeométrica

- $H(N; n; p)$

- **Características:**
  - $\mu = E(\xi) = n \cdot p$
  - $\sigma^2 = Var(\xi) = n \cdot p \cdot q \cdot \frac{N-n}{N-1}$

- **Propiedad:**
  - $\xi \rightarrow H(N; n; p) \rightarrow B(n; p)$ siempre que $\mathrm{p} = \frac{\boldsymbol{N_1}}{\boldsymbol{N}}$ sea lo suficientemente pequeño (en la práctica, $p < 0{,}1$)

- $H(N; n; p)$

- **Ejemplo:** Una comunidad autónoma tiene 50 representantes en la convención de un partido político. De ellos, 30 apoyan al candidato A y 20 al B. Si se seleccionan al azar 5 representantes, cuál es la probabilidad de que de ellos solo 1 apoye a A.

Variable cuya probabilidad modeliza: ξ ="número de partidarios de A". $\xi \to H\left(50; 5; \frac{30}{50}\right)$

$$P(\xi = 1) = \frac{\binom{30}{1} \cdot \binom{20}{4}}{\binom{50}{5}} = \frac{\frac{30!}{1! \cdot 29!} \cdot \frac{20!}{4! \cdot 16!}}{\frac{50!}{5! \cdot 45!}} = 0{,}0686$$


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

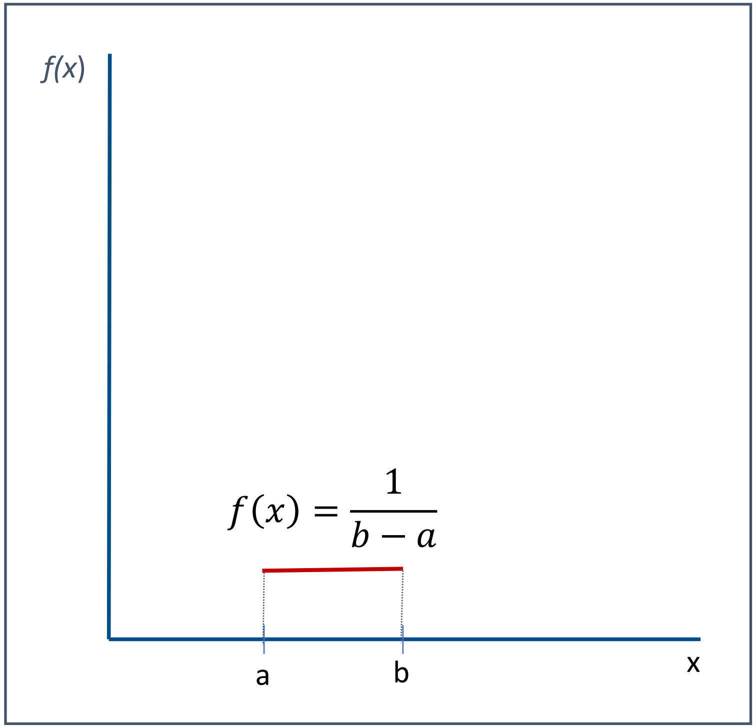
Tema 5.2.

Modelos de distribución de probabilidad continuos univariantes.

1. Distribución uniforme.
2. Distribución exponencial.
3. Distribución normal.

1. Distribución uniforme

- $U(a;b)$

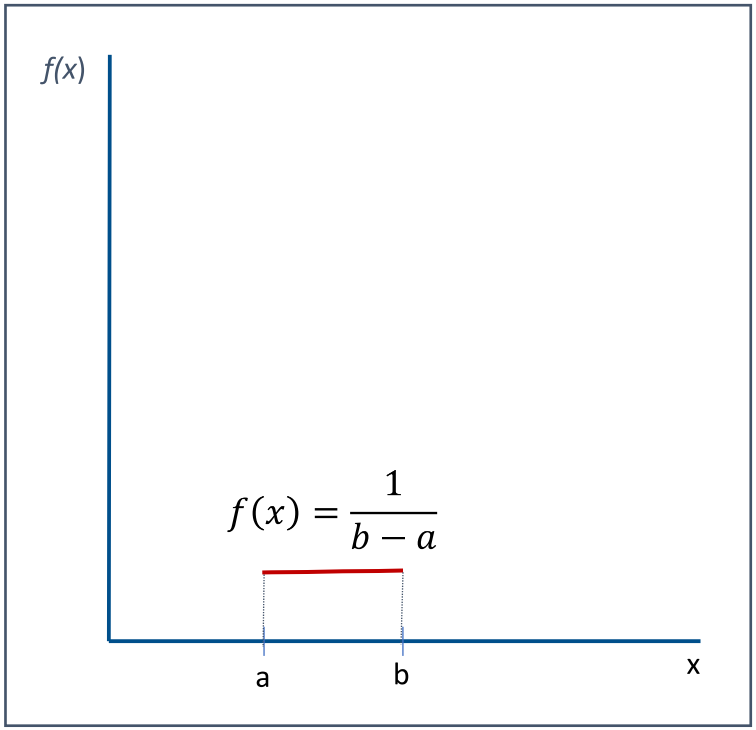


- **Parámetros:**
  - ***a:*** extremo inferior del campo de variación de la variable ξ.
  - ***b:*** extremo superior del campo de variación de la variable ξ.

- **Función de densidad:**
  - $f(x) = \frac{1}{b-a}$ con $a \leq x \leq b$.

- **Características:**
  - $\mu = E(\xi) = \frac{a+b}{2}$
  - $\sigma^2 = Var(\xi) = E(\xi - \mu)^2 = \frac{(b-a)^2}{12}$

- $U(a;b)$



- **Ejemplo:** La variable ξ ="volumen anual de ventas" de un almacén se distribuye uniformemente entre 380 y 1200 miles de euros.
  - Probabilidad de que las ventas sean superiores a 1000 miles de euros.

$$\xi \to U(380; 1200)$$

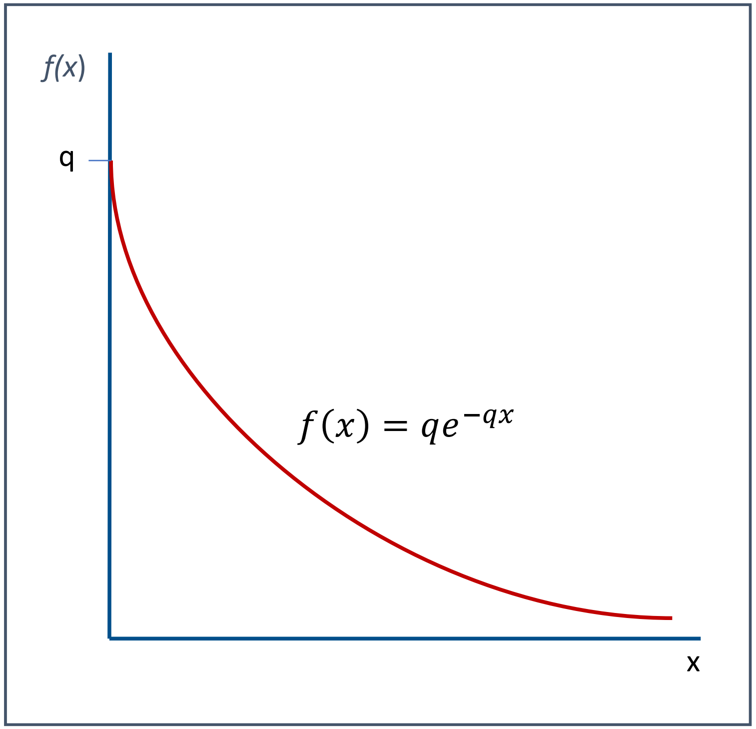
$$P(\xi > 1000)$$

$$= \int_{1000}^{1200} \frac{1}{(1200-380)} dx = \frac{200}{820} = 0{,}2439$$

  - Ventas esperadas en un año.

$$E(\xi) = \frac{a+b}{2} = \frac{380+1200}{2} = 790$$

- $Exp(q)$



- **Parámetros:**
  - ***q:*** parámetro característico de la función de densidad de la variable ξ.

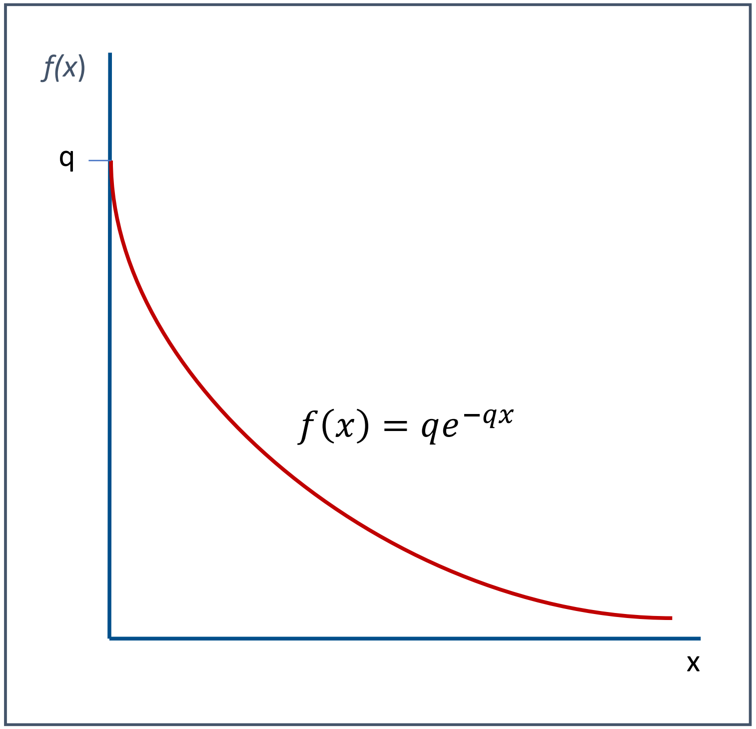
- **Función de densidad:**
  - $f(x) = qe^{-qx}$ con $x > 0$ y $q \geq 0$.

- **Características:**
  - $\mu = E(\xi) = \frac{1}{q}$
  - $\sigma^2 = Var(\xi) = E(\xi - \mu)^2 = \frac{1}{q^2}$

- **Propiedad (pérdida de memoria):**

  Si $\xi \rightarrow Exp(q)$ y tenemos dos números positivos *a* y *b*; $P(\xi > a + b \,/\, \xi > b) = P(\xi > a)$

- $Exp(q)$



- **Ejemplo:** El tiempo en horas que se tarda en reparar una máquina sigue una distribución *Exp(4)*. ¿Cuál es la probabilidad de que una avería tarde más de 1 hora en ser reparada?

  - Variable modelizada: ξ = "tiempo de reparación" $\rightarrow Exp(4)$

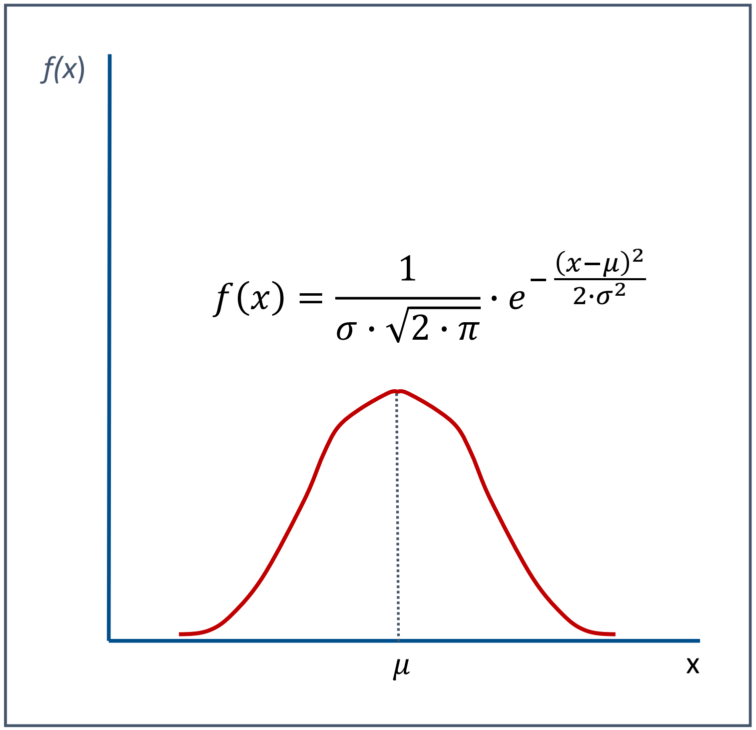
  - Probabilidad:

$$P(\xi > 1) = 1 - P(\xi \leq 1) = 1 - \int_0^1 4 \cdot e^{-4x} dx =$$

$$= 1 - 4 \cdot \int_1^\infty e^{-4x} dx = 1 - 4 \cdot \frac{1}{-4} \cdot [e^{-4x}]_0^1 =$$

$$= 1 + e^{-4} - e^0 = \frac{1}{e^4} = 0{,}0183$$

- $N(\mu;\sigma)$



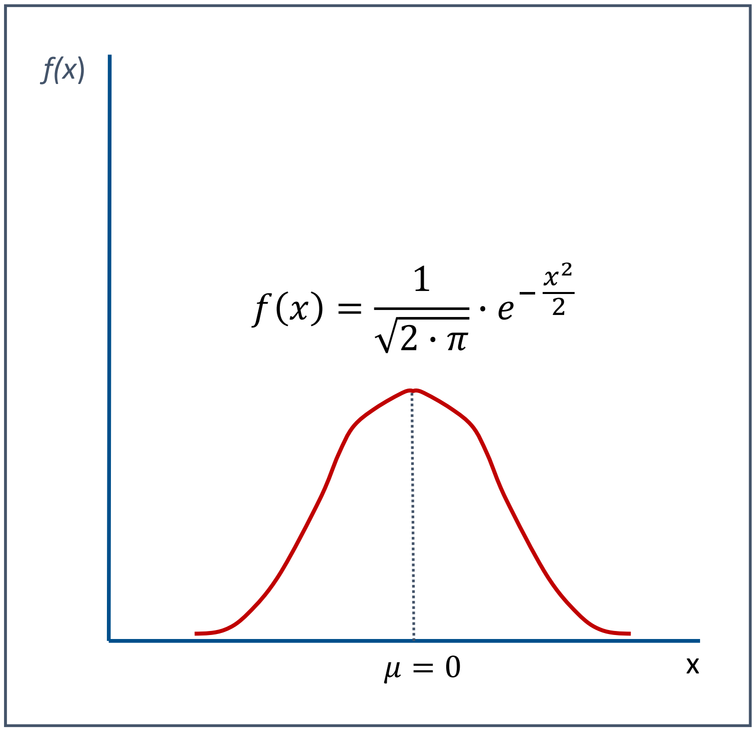
- **Parámetros:**
  - μ**:** media o esperanza matemática de la variable ξ.
  - σ**:** desviación típica de la variable ξ.
- **Función de densidad:**
  - $f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}$ con $-\infty \leq x \leq \infty$.
- **Características:**
  - $E(\xi) = \mu$
  - $Var(\xi) = E(\xi - \mu)^2 = \sigma^2$
- **Propiedad (aditiva o reproductiva):**

  Si tenemos n variables $\xi_j \rightarrow N(\mu_j;\sigma_j)$ independientes entre sí; la variable $\eta = a \pm b_1 \cdot \xi_1 \pm b_2 \cdot \xi_2 \pm \cdots \pm b_n \cdot \xi_n$, con no todos los b_j nulos, seguirá una distribución:

$$\eta \rightarrow N\left(a \pm b_1 \cdot \mu_1 \pm b_2 \cdot \mu_2 \pm \cdots \pm b_n \cdot \mu_n; \sqrt{b_1^2 \cdot \sigma_1^2 + b_2^2 \cdot \sigma_2^2 + \cdots + b_n^2 \cdot \sigma_n^2}\right)$$

3. Distribución normal

- $N(0;1)$ Normal 0,1; tipificada o estandarizada.

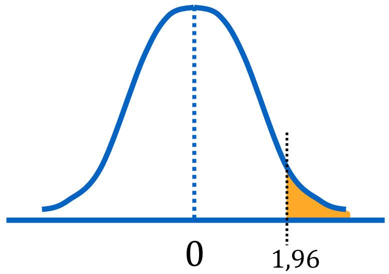


- **Parámetros:**
  - $\mu = 0$**:** media o esperanza matemática de la variable ξ.
  - $\sigma = 1$**:** desviación típica de la variable ξ.
- **Función de densidad:**
  - $f(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{x^2}{2}}$ con $-\infty \leq x \leq \infty$.
- **Características:**
  - $E(\xi) = 0$
  - $Var(\xi) = E(\xi - \mu)^2 = 1$

- Una variable aleatoria $\xi \rightarrow N(0;1)$ cuenta con la ventaja de **tener tabulada su distribución de probabilidad**, por lo que se pueden hallar probabilidades consultando la tabla sin necesidad de recurrir a integrar la función de densidad o a calcular la función de distribución.

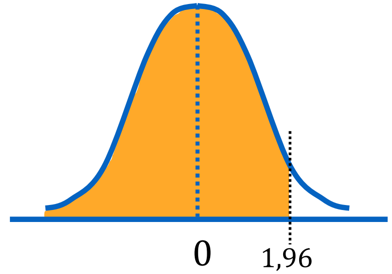
$N(0;1)$ Normal 0,1; tipificada o estandarizada. Ejemplos de uso de tabla de la distribución.

1



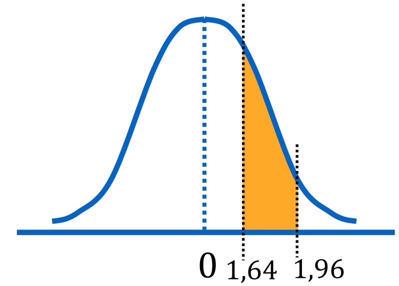
$P(\xi \geq 1{,}96) = 0{,}025$

2



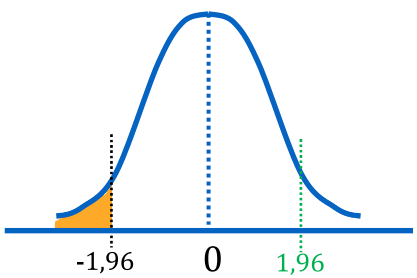
$P(\xi \leq 1{,}96) = 1 - P(\xi \geq 196) =$
$=1-0{,}025=0{,}975$

3



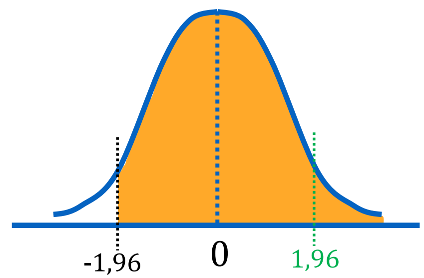
$P(1{,}64 \leq \xi \leq 1{,}96) = P(\xi \geq 1{,}64) - P(\xi \geq 1{,}96) =$
$=0{,}0505 - 0{,}025 = 0{,}0255$

4



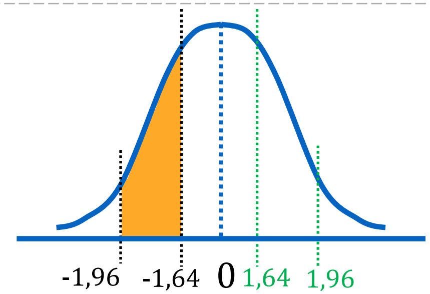
$P(\xi \leq -1{,}96) = P(\xi \geq 1{,}96) = 0{,}025$

5



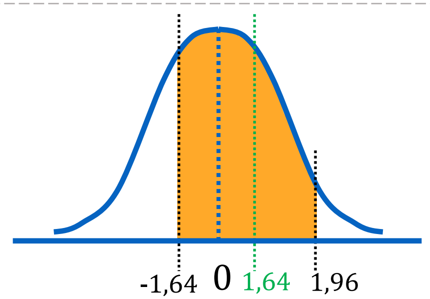
$P(\xi \geq -1{,}96) = 1 - P(\xi \geq 1{,}96) =$
$= 1 - 0{,}025 = 0{,}975$

6



$P(-1{,}96 \leq \xi \leq -1{,}64) = P(1{,}64 \leq \xi \leq 1{,}96) =$
$= P(\xi \geq 1{,}64) - P(\xi \geq 1{,}96) =$
$= 0{,}0505 - 0{,}025 = 0{,}0255$

7



$P(-1{,}64 \leq \xi \leq 1{,}96) =$
$= 1 - P(\xi \geq 1{,}64) - P(\xi \geq 1{,}96) =$
$= 1 - 0{,}0505 - 0{,}025 = 0{,}9245$

3. Distribución normal

- Si la variable sigue una distribución $\xi \rightarrow N(\mu;\sigma)$, para hallar probabilidades previamente se transformará en una variable $\xi^* \rightarrow N(0;1)$ (**tipificación** o estandarización):

$$\xi^* = \frac{\xi - \mu}{\sigma}$$

- **Ejemplo:** Se hace un *casting* para una película. Un actor no puede optar al papel con una altura inferior a 180 cm. La altura de los candidatos sigue una distribución normal de media 185 cm. y desviación típica 10 cm. ¿Cuál es la probabilidad de que un aspirante escogido al azar quede fuera del *casting* por motivo de su altura?

  - $\xi = \text{"}altura\ de\ los\ candidatos\text{"} \rightarrow N(185;10)$

$$P(\xi < 180) = P\left(\frac{\xi - 185}{10} < \frac{180 - 185}{10}\right) =$$
$$P(\xi^* < -0{,}5) = P(\xi^* > 0{,}5) = 0{,}3085$$

¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

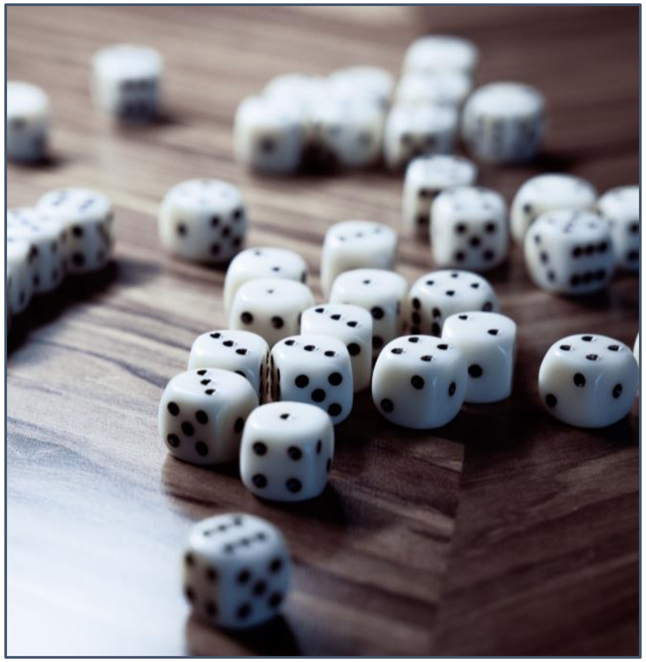
Tema 5.3.

Convergencia y Teoremas Límite.

Contenido

1. Teorema Central del Límite.
2. Esquema de aproximaciones.

- **Enunciado:**



Una sucesión de variables aleatorias $\{\xi_n\}$ verifica el Teorema Central del Límite si la variable aleatoria $\eta = \xi_1 + \xi_2 + \cdots + \xi_n$ cumple con la convergencia en distribución:

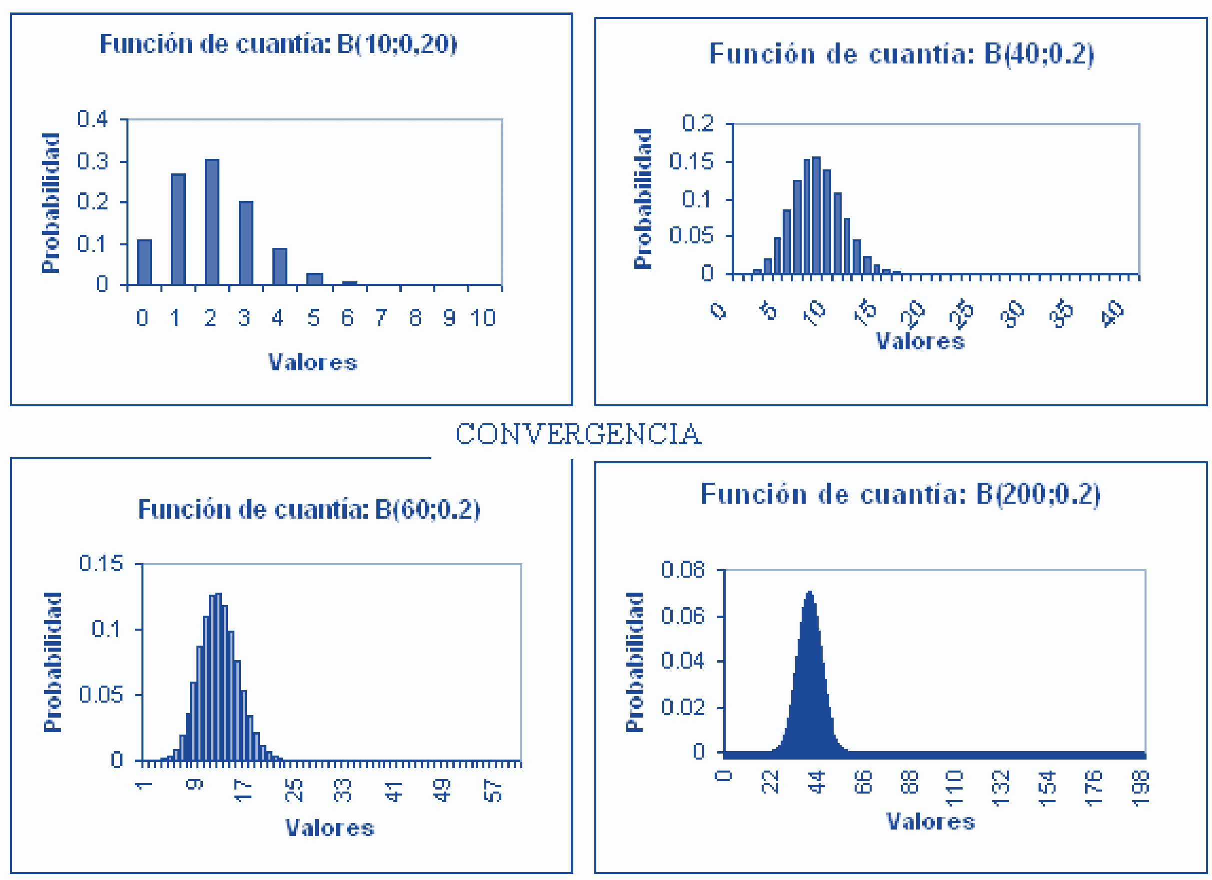
$$\frac{\eta - E(\eta)}{\sqrt{Var(\eta)}} \xrightarrow{d} N(0;1) \text{ siendo la varianza finita.}$$

- **Lindeberg-Lévy** proponen como condición para que la sucesión de variables cumplan con el TCL que las variables sean independientes e idénticamente distribuidas, con $E(\xi_j) = \mu$ y $Var(\xi_j) = \sigma^2, \forall j$. Así:

$$\frac{\eta - n \cdot \mu}{\sqrt{n \cdot \sigma^2}} \xrightarrow{d} N(0;1)$$

- **Moivre-Laplace** proponen con anterioridad a Lindeberg-Lévy un caso particular de cumplimiento del TCL. Si las variables $\xi_j \to B(1;p)\ \forall j$ entonces $\eta = \xi_1 + \xi_2 + \cdots + \xi_n \to B(n;p)$ y por tanto:

$$\frac{\eta - n\cdot p}{\sqrt{n\cdot p\cdot q}} \xrightarrow{d} N(0;1)$$



$$\begin{tikzcd}
 & Po(n \cdot p) \arrow[r, "n \geq 100"] & \boldsymbol{N(n \cdot p; \sqrt{n \cdot p})} \\
B(n;p) \arrow[ur, dotted, "{n \cdot p \leq 5,\ p \leq 0{,}05}"] \arrow[rr, "n \geq 100"'] & & \boldsymbol{N(n \cdot p; \sqrt{n \cdot p \cdot q})}
\end{tikzcd}$$

- **Ejemplo:**

Una fábrica de medicamentos realiza pruebas clínicas con 300 nuevos fármacos potenciales. Cerca del 20% de las sustancias que alcanzan esta etapa reciben finalmente la aprobación para su venta. ¿Cuál es la probabilidad de que se aprueben al menos 45 de los 300 medicamentos?

- ξ =“número de medicamentos aprobados”.

$$\xi \to B(300;0{,}2) \overset{d}{\to} N\left(300 \cdot 0{,}2; \sqrt{300 \cdot 0{,}2 \cdot 0{,}8}\right)$$

$$P(\xi > 45) = P\left(\frac{\xi - 60}{\sqrt{48}} > \frac{45 - 60}{\sqrt{48}}\right) = P(\xi^* > -2{,}16) =$$

$$= 1 - P(\xi^* \geq 2{,}16) = 0{,}9846$$

con $\xi^* \to N(0;1)$


¡Muchas gracias!

PRÁCTICA

BLOQUE PRÁCTICO

En este apartado encontrarás los ejercicios a realizar, así como las soluciones a los mismos.

ÁREA DE ESTADÍSTICA EMPRESARIAL. DEPARTAMENTO DE ECONOMÍA APLICADA I. FACULTAD DE DERECHO Y CCSS DE CIUDAD REAL (UCLM).

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 1.2

Variable estadística unidimensional - Práctica

Ej. 1 *Dada la distribución de frecuencias correspondiente a los ingresos mensuales de las familias (en miles de euros) de una determinada población, calcule:*

Ingreso Mensual	Num. Familias
0 - 0,9	5
0,9 – 1,2	15
1,2 – 1,5	17
1,5 – 3	12
3 – 6	2

L_{i-1}	L_i	x_i	n_i
0	0,9	0,45	5
0,9	1,2	1,05	15
1,2	1,5	1,35	17
1,5	3	2,25	12
3	6	4,5	2
		N=	51

Ej. 1 *Dada la distribución de frecuencias correspondiente a los ingresos mensuales de las familias (en miles de euros) de una determinada población, calcule:*

a) El ingreso mensual medio.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$
0	0,9	0,45	5	5	0,9	5,5556	2,25
0,9	1,2	1,05	15	20	0,3	50,0000	15,75
1,2	1,5	1,35	17	37	0,3	56,6667	22,95
1,5	3	2,25	12	49	1,5	8,0000	27
3	6	4,5	2	51	3	0,6667	9
		N=	51				76,95

$$\bar{x} = \frac{\sum_{i=1}^{h} x_i \cdot n_i}{N} = \frac{76,95}{51} = 1,5088\, miles\ de\ euros$$

Ej. 1

Dada la distribución de frecuencias correspondiente a los ingresos mensuales de las familias (en miles de euros) de una determinada población, calcule:

b) El ingreso mensual mediano.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$
0	0,9	0,45	5	5	0,9	5,5556	2,25
0,9	1,2	1,05	15	20	0,3	50,0000	15,75
1,2	1,5	1,35	17	37	0,3	56,6667	22,95
1,5	3	2,25	12	49	1,5	8,0000	27
3	6	4,5	2	51	3	0,6667	9
		N=	51				76,95

- $N = 51\ (impar)$
- $\frac{N+1}{2} = \frac{52}{2} = 26 \ \rightarrow\ Intervalo\ mediano\text{: } (1{,}2\ ;\ 1{,}5]$
- $Me = L_{i-1} + \frac{\frac{N+1}{2} - N_{i-1}}{n_i} c_i = 1{,}2 + \frac{26-20}{17} \cdot 0{,}3 = 1{,}3059\ miles\ de\ euros$

Ej. 1

Dada la distribución de frecuencias correspondiente a los ingresos mensuales de las familias (en miles de euros) de una determinada población, calcule:

c) El ingreso más frecuente.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$
0	0,9	0,45	5	5	0,9	5,5556	2,25
0,9	1,2	1,05	15	20	0,3	50,0000	15,75
1,2	1,5	1,35	17	37	0,3	56,6667	22,95
1,5	3	2,25	12	49	1,5	8,0000	27
3	6	4,5	2	51	3	0,6667	9
		N=	51				76,95

- $Intervalos\ de\ distinta\ amplitud$
- $Intervalo\ modal\!:\ (1{,}2\ ;\ 1{,}5]$
- $Mo = L_{i-1} + \dfrac{d_i - d_{i-1}}{(d_i - d_{i-1}) + (d_i - d_{i+1})} \cdot c_i =$

$$= 1{,}2 + \frac{56{,}6667-50}{(56{,}6667-50)+(56{,}6667-8)} \cdot 0{,}3 = 1{,}2361\ miles\ de\ euros$$

Ej. 2

Dada la distribución de frecuencias correspondiente a la variable X = "miles de euros de beneficio" de las empresas de un determinado sector productivo, calcule:

Beneficio	N. empresas
De 0 a 5	6
De 5 a 10	8
De 10 a 15	9
De 15 a 20	5
De 20 a 25	3

L_{i-1}	L_i	x_i	n_i
0	5	2,5	6
5	10	7,5	8
10	15	12,5	9
15	20	17,5	5
20	25	22,5	3
		N=	31

Ej. 2

Dada la distribución de frecuencias correspondiente a la variable X = "miles de euros de beneficio" de las empresas de un determinado sector productivo, calcule:

a) El valor medio de la distribución.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	5	2,5	6	6	5	1,2000	15	37,5
5	10	7,5	8	14	5	1,6000	60	450
10	15	12,5	9	23	5	1,8000	112,5	1406,25
15	20	17,5	5	28	5	1,0000	87,5	1531,25
20	25	22,5	3	31	5	0,6000	67,5	1518,75
		N=	31				342,5	4943,75

$$\bar{x} = \frac{\sum_{i=1}^{h} x_i \cdot n_i}{N} = \frac{342,5}{31} = \frac{342,5}{31} = 11,0484 \; miles\ de\ euros$$

Ej. 2

Dada la distribución de frecuencias correspondiente a la variable X = "miles de euros de beneficio" de las empresas de un determinado sector productivo, calcule:

b) La mediana de la distribución.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	5	2,5	6	6	5	1,2000	15	37,5
5	10	7,5	8	14	5	1,6000	60	450
10	15	12,5	9	23	5	1,8000	112,5	1406,25
15	20	17,5	5	28	5	1,0000	87,5	1531,25
20	25	22,5	3	31	5	0,6000	67,5	1518,75
		N=	31				342,5	4943,75

- $N = 31\ (impar)$
- $\frac{N+1}{2} = \frac{32}{2} = 16 \ \rightarrow \ Intervalo\ mediano\text{: } (10\ ;15]$
- $Me = L_{i-1} + \frac{\frac{N+1}{2} - N_{i-1}}{n_i} c_i = 10 + \frac{16-14}{9} \cdot 5 = 11{,}1111\ miles\ de\ euros$

Ej. 2

Dada la distribución de frecuencias correspondiente a la variable X = "miles de euros de beneficio" de las empresas de un determinado sector productivo, calcule:

c) El valor modal de la distribución.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	5	2,5	6	6	5	1,2000	15	37,5
5	10	7,5	8	14	5	1,6000	60	450
10	15	12,5	9	23	5	1,8000	112,5	1406,25
15	20	17,5	5	28	5	1,0000	87,5	1531,25
20	25	22,5	3	31	5	0,6000	67,5	1518,75
		N=	31				342,5	4943,75

- $Intervalos\ de\ misma\ amplitud$
- $Intervalo\ modal\text{:}\ (10\ ;15]$
- $Mo = L_{i-1} + \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} \cdot c_i =$

$$= 10 + \frac{9-8}{(9-8)+(9-5)} \cdot 5 = 11\ miles\ de\ euros$$

Ej. 2 *Dada la distribución de frecuencias correspondiente a la variable X = "miles de euros de beneficio" de las empresas de un determinado sector productivo, calcule:*

d) La varianza y la desviación típica de la distribución.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	5	2,5	6	6	5	1,2000	15	37,5
5	10	7,5	8	14	5	1,6000	60	450
10	15	12,5	9	23	5	1,8000	112,5	1406,25
15	20	17,5	5	28	5	1,0000	87,5	1531,25
20	25	22,5	3	31	5	0,6000	67,5	1518,75
		N=	31				342,5	4943,75

- $S^2 = m_2 = a_2 - a_1^2 = \frac{\sum_{i=1}^{N} x_i^2 n_i}{N} - (\bar{x})^2 =$

$$= \frac{4943,75}{31} - (11,0484)^2 = 37,4089 \; miles\ de\ euros\ cuadrado$$

- $S = +\sqrt{S^2} = \sqrt{37,4089} = 6,116 \; miles\ de\ euros$

Ej. 3

Dada la distribución de frecuencias correspondiente a la variable X = "miles de dólares de beneficio" de las empresas de un determinado sector productivo, calcule:

Beneficio	N. empresas
De 0 a 6	4
De 6 a 12	4
De 12 a 18	9
De 18 a 24	2
De 24 a 30	2

L_{i-1}	L_i	x_i	n_i
0	6	3	4
6	12	9	4
12	18	15	9
18	24	21	2
24	30	27	2
		N=	21

Ej. 3 *Dada la distribución de frecuencias correspondiente a la variable $X =$ "miles de dólares de beneficio" de las empresas de un determinado sector productivo, calcule:*

a) El valor medio de la distribución.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	6	3	4	4	6	0,6667	12	36
6	12	9	4	8	6	0,6667	36	324
12	18	15	9	17	6	1,5000	135	2025
18	24	21	2	19	6	0,3333	42	882
24	30	27	2	21	6	0,3333	54	1458
		N=	21				279	4725

$$\bar{x} = \frac{\sum_{i=1}^{h} x_i \cdot n_i}{N} = \frac{279}{21} = 13{,}2857\ miles\ de\ dólares$$

Ej. 3

Dada la distribución de frecuencias correspondiente a la variable $X =$ "miles de dólares de beneficio" de las empresas de un determinado sector productivo, calcule:

b) La mediana de la distribución.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	6	3	4	4	6	0,6667	12	36
6	12	9	4	8	6	0,6667	36	324
12	18	15	9	17	6	1,5000	135	2025
18	24	21	2	19	6	0,3333	42	882
24	30	27	2	21	6	0,3333	54	1458
		N=	21				279	4725

- $N = 21\ (impar)$
- $\frac{\mathrm{N}+1}{2} = \frac{22}{2} = 11 \ \rightarrow\ Intervalo\ mediano:\ (12\ ;18]$
- $Me = L_{i-1} + \frac{\frac{N+1}{2} - N_{i-1}}{n_i} c_i = 12 + \frac{11-8}{9} \cdot 6 = 14\ miles\ de\ dólares$

Ej. 3

Dada la distribución de frecuencias correspondiente a la variable X = "miles de dólares de beneficio" de las empresas de un determinado sector productivo, calcule:

c) El valor modal de la distribución.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	6	3	4	4	6	0,6667	12	36
6	12	9	4	8	6	0,6667	36	324
12	18	15	9	17	6	1,5000	135	2025
18	24	21	2	19	6	0,3333	42	882
24	30	27	2	21	6	0,3333	54	1458
		N=	21				279	4725

- $Intervalos\ de\ misma\ amplitud$
- $Intervalo\ modal\text{:}\ (12\ ;18]$
- $Mo = L_{i-1} + \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} \cdot c_i =$

$$= 12 + \frac{9-4}{(9-4)+(9-2)} \cdot 6 = 14{,}5\ miles\ de\ dólares$$

Ej. 3

Dada la distribución de frecuencias correspondiente a la variable X = "miles de dólares de beneficio" de las empresas de un determinado sector productivo, calcule:

d) La varianza y la desviación típica.

L_{i-1}	L_i	x_i	n_i	N_i	c_i	d_i	$x_i \cdot n_i$	$x^2_i \cdot n_i$
0	6	3	4	4	6	0,6667	12	36
6	12	9	4	8	6	0,6667	36	324
12	18	15	9	17	6	1,5000	135	2025
18	24	21	2	19	6	0,3333	42	882
24	30	27	2	21	6	0,3333	54	1458
		N=	21				279	4725

- $S^2 = m_2 = a_2 - a_1^2 = \frac{\sum_{i=1}^{N} x_i^2 n_i}{N} - (\bar{x})^2 =$

$= \frac{4725}{21} - (13{,}2857)^2 = 48{,}49\ miles\ de\ dólares\ cuadrado$

- $S = +\sqrt{S^2} = \sqrt{48{,}49} = 6{,}9635\ miles\ de\ dólares$

Ej. 4 ***Teniendo en cuenta las variables de los dos ejercicios anteriores, ¿cuál de ellas tiene una media más representativa de la distribución de frecuencias?***

$$Coeficiente\ de\ dispersión\ de\ Pearson:\ \ V=\frac{S}{\bar{x}}$$

- $Primera\ empresa\ (ejercicio\ 2, euros):$

$$V=\frac{S}{\bar{x}}=\frac{6{,}1163}{11{,}0484}=0{,}5536$$

- $Segunda\ empresa\ (ejercicio\ 3, dólares):$

$$V=\frac{S}{\bar{x}}=\frac{6{,}9635}{13{,}2857}=0{,}5241$$

$$Menor\ coeficiente\ de\ dispersión\ \rightarrow Media\ más\ representativa$$

Ej. 5 ***Tres empresas que compiten en un mismo sector registran los niveles de morosidad (en meses) recogidos en las siguientes distribuciones de frecuencias. Calcule las correspondientes medidas de asimetría y curtosis y compárelas entre sí.***

Empresa A

Meses de retraso en los pagos (período de morosidad)	Número de Clientes
0 a 3	145
3 a 6	160
6 a 9	210
9 a 12	155
12 a 15	142

Empresa B

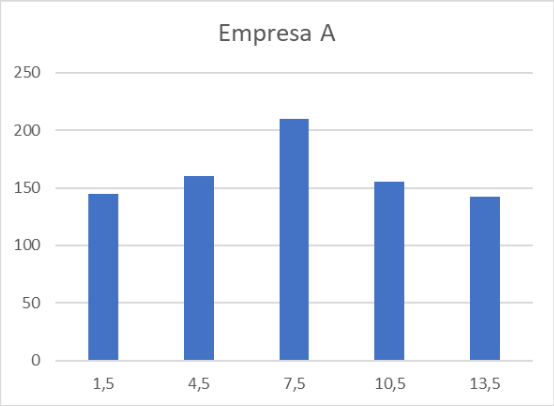
Meses de retraso en los pagos (período de morosidad)	Número de Clientes
0 a 3	121
3 a 6	235
6 a 9	811
9 a 12	240
12 a 15	125

Empresa C

Meses de retraso en los pagos (período de morosidad)	Número de Clientes
0 a 3	100
3 a 6	125
6 a 9	325
9 a 12	126
12 a 15	99

Ej. 5 *Tres empresas que compiten en un mismo sector registran los niveles de morosidad (en meses) recogidos en las siguientes distribuciones de frecuencias. Calcule las correspondientes medidas de asimetría y curtosis y compárelas entre sí.*

L_{i-1}	L_i	x_i	n_i	$x_i * n_i$	$x^2_i \cdot n_i$	$x^3_i \cdot n_i$	$x^4_i \cdot n_i$
0	3	1,5	145	217,50	326,25	489,38	734,06
3	6	4,5	160	720,00	3240,00	14580,00	65610,00
6	9	7,5	210	1575,00	11812,50	88593,75	664453,13
9	12	10,5	155	1627,50	17088,75	179431,88	1884034,69
12	15	13,5	142	1917,00	25879,50	349373,25	4716538,88
		N=	812	6057,00	58347,00	632468,25	7331370,75
				a_1	a_2	a_3	a_4
				7,4594	71,8559	778,9018	9028,7817



Empresa A

Asimetría

$$g_1 = \frac{m_3}{S^3} = \frac{a_3 - 3a_2a_1 + 2a_1^3}{\left(+\sqrt{a_2 - a_1^2}\right)^3} = \frac{778{,}9018 - 3 \cdot 71{,}8559 \cdot 7{,}4594 + 2 \cdot 7{,}4594^3}{\left(+\sqrt{71{,}8559 - 7{,}4594^2}\right)^3} = 0{,}0155; \mathit{simétrica}$$

Curtosis

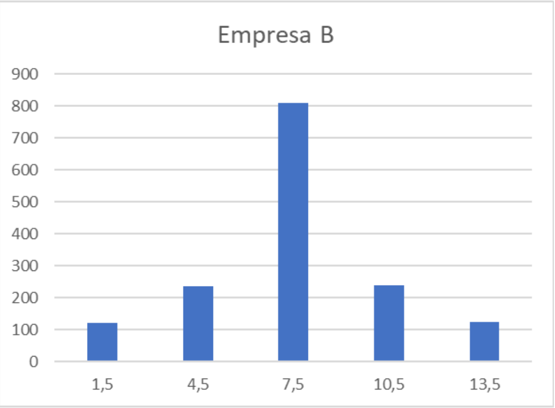
$$g_2 = \frac{m_4}{S^4} - 3 = \frac{a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4}{\left(a_2 - a_1^2\right)^2} - 3$$

$$= \frac{9028{,}7817 - 4 \cdot 778{,}9018 \cdot 7{,}4594 + 6 \cdot 71{,}8559 \cdot 7{,}4594^2 - 3 \cdot 7{,}4594^4}{(71{,}8559 - 7{,}4594^2)^2} - 3$$

$$= -1{,}1380; \mathit{platicúrtica}$$

Ej. 5 *Tres empresas que compiten en un mismo sector registran los niveles de morosidad (en meses) recogidos en las siguientes distribuciones de frecuencias. Calcule las correspondientes medidas de asimetría y curtosis y compárelas entre sí.*

L_{i-1}	L_i	x_i	n_i	$x_i * n_i$	$x^2_i \cdot n_i$	$x^3_i \cdot n_i$	$x^4_i \cdot n_i$
0	3	1,5	121	181,50	272,25	408,38	612,56
3	6	4,5	235	1057,50	4758,75	21414,38	96364,69
6	9	7,5	811	6082,50	45618,75	342140,63	2566054,69
9	12	10,5	240	2520,00	26460,00	277830,00	2917215,00
12	15	13,5	125	1687,50	22781,25	307546,88	4151882,81
		N=	1532	11529,00	99891,00	949340,25	9732129,75
				a_1	a_2	a_3	a_4
				7,5255	65,2030	619,6738	6352,5651



Empresa B

Asimetría

$$g_1 = \frac{m_3}{S^3} = \frac{a_3 - 3a_2a_1 + 2a_1^3}{\left(+\sqrt{a_2 - a_1^2}\right)^3} = \frac{619{,}6738 - 3 \cdot 65{,}2030 \cdot 7{,}5255 + 2 \cdot 7{,}5255^3}{\left(+\sqrt{65{,}2030 - 7{,}5255^2}\right)^3} = -0{,}0001; \textit{simétrica}$$

Curtosis

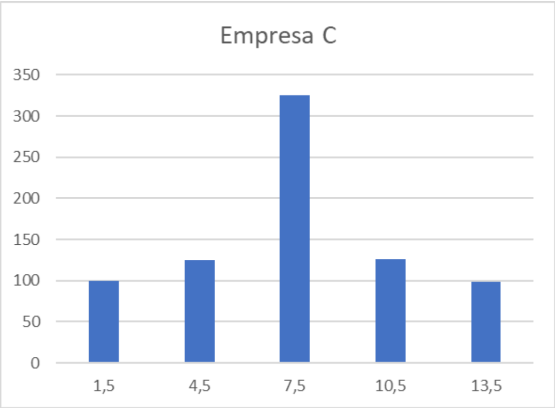
$$g_2 = \frac{m_4}{S^4} - 3 = \frac{a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4}{\left(a_2 - a_1^2\right)^2} - 3$$

$$= \frac{6352{,}5651 - 4 \cdot 619{,}6738 \cdot 7{,}5255 + 6 \cdot 65{,}2030 \cdot 7{,}5255^2 - 3 \cdot 7{,}5255^4}{(65{,}2030 - 7{,}5255^2)^2} - 3$$

$$= 0{,}1746; \textit{leptocúrtica}$$

Ej. 5 *Tres empresas que compiten en un mismo sector registran los niveles de morosidad (en meses) recogidos en las siguientes distribuciones de frecuencias. Calcule las correspondientes medidas de asimetría y curtosis y compárelas entre sí.*

L_{i-1}	L_i	x_i	n_i	$x_i * n_i$	$x_i^2 \cdot n_i$	$x_i^3 \cdot n_i$	$x_i^4 \cdot n_i$
0	3	1,5	100	150,00	225,00	337,50	506,25
3	6	4,5	125	562,50	2531,25	11390,63	51257,81
6	9	7,5	325	2437,50	18281,25	137109,38	1028320,31
9	12	10,5	126	1323,00	13891,50	145860,75	1531537,88
12	15	13,5	99	1336,50	18042,75	243577,13	3288291,19
		N=	775	5809,50	52971,75	538275,38	5899913,44
				a_1	a_2	a_3	a_4
				7,4961	68,3506	694,5489	7612,7915



Empresa C

Asimetría

$$g_1 = \frac{m_3}{S^3} = \frac{a_3 - 3a_2a_1 + 2a_1^3}{\left(+\sqrt{a_2 - a_1^2}\right)^3} = \frac{694{,}5489 - 3 \cdot 68{,}3506 \cdot 7{,}4961 + 2 \cdot 7{,}4961^3}{\left(+\sqrt{68{,}3506 - 7{,}4961^2}\right)^3} = -0{,}0024; \textit{simétrica}$$

Curtosis

$$g_2 = \frac{m_4}{S^4} - 3 = \frac{a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4}{\left(a_2 - a_1^2\right)^2} - 3$$

$$= \frac{7612{,}7915 - 4 \cdot 694{,}5489 \cdot 7{,}4961 + 6 \cdot 68{,}3506 \cdot 7{,}4961^2 - 3 \cdot 7{,}4961^4}{(68{,}3506 - 7{,}4961^2)^2} - 3$$

$$= -0{,}5715; \textit{platicúrtica}$$


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 2.1.

Variable estadística bidimensional y n-dimensional

Práctica

Ej. 1

A continuación, se muestra una tabla de correlaciones con las empresas que operan en un determinado sector productivo, agrupadas según las variables "porcentaje de beneficios anuales destinados a I+D" (X) y "cifra de ventas en millones de euros" (Y). Se pide:

X \ Y	(0-10)	(10-25)	(25-50)	(50-100)
(0-5)	6	6	4	0
(5-10)	2	2	3	1
(10-20)	1	2	1	1

X/Y	5	17,5	37,5	75	ni.
2,5	6	6	4	0	16
7,5	2	2	3	1	8
15	1	2	1	1	5
n.j	9	10	8	2	29

Ej. 1

A continuación, se muestra una tabla de correlaciones con las empresas que operan en un determinado sector productivo, agrupadas según las variables "porcentaje de beneficios anuales destinados a I+D" (X) y "cifra de ventas en millones de euros" (Y). Se pide:

a) Determine si ambas variables son estadísticamente independientes.

X/Y	5	17,5	37,5	75	ni.
2,5	6	6	4	0	16
7,5	2	2	3	1	8
15	1	2	1	1	5
n.j	9	10	8	2	29

Frecuencias relativas conjuntas

X/Y	5	17,5	37,5	75
2,5	0,2069	0,2069	0,1379	0,0000
7,5	0,0690	0,0690	0,1034	0,0345
15	0,0345	0,0690	0,0345	0,0345

Producto de frecuencias relativas marginales

X/Y	5	17,5	37,5	75	ni.
2,5	0,1712	0,1902	0,1522	0,0380	0,5517
7,5	0,0856	0,0951	0,0761	0,0190	0,2759
15	0,0535	0,0595	0,0476	0,0119	0,1724
n.j	0,3103	0,3448	0,2759	0,0690	1

Diferencia

X/Y	5	17,5	37,5	75
2,5	0,0357	0,0166	-0,0143	-0,0380
7,5	-0,0166	-0,0262	0,0273	0,0155
15	-0,0190	0,0095	-0,0131	0,0226

Conclusión: *como no todas las celdas de esta tabla son* $0 \rightarrow$ *dependencia estadística.*

Ej. 1 *A continuación, se muestra una tabla de correlaciones con las empresas que operan en un determinado sector productivo, agrupadas según las variables "porcentaje de beneficios anuales destinados a I+D" (X) y "cifra de ventas en millones de euros" (Y). Se pide:*

b) Calcule la covarianza de ambas variables.

X/Y	5	17,5	37,5	75	ni.	xi*ni.
2,5	6	6	4	0	16	40
7,5	2	2	3	1	8	60
15	1	2	1	1	5	75
n.j	9	10	8	2	29	175
yj*n.j	45	175	300	150	670	

$$\boxed{S_{xy} = m_{11} = a_{11} - a_{10} \cdot a_{01}}$$

75	262,5	375	0
75	262,5	843,75	562,5
75	525	562,5	1125
	Suma:	4743,75	

$$a_{11} = \frac{4743,75}{29} = 163,5776$$

$$a_{10} = \frac{175}{29} = 6,0345$$

$$a_{01} = \frac{670}{29} = 23,1034$$

$$S_{xy} = m_{11} = a_{11} - a_{10} \cdot a_{01} = 163,5776 - 6,0345 \cdot 23,1034 = 24,1602$$

Ej. 1

A continuación, se muestra una tabla de correlaciones con las empresas que operan en un determinado sector productivo, agrupadas según las variables "porcentaje de beneficios anuales destinados a I+D" (X) y "cifra de ventas en millones de euros" (Y). Se pide:

c) ¿En cuál de las variables la media aritmética es más representativa del total de empresas?

$a_{11} = 105{,}3879$

$a_{10} = 6{,}0345$

$a_{01} = 23{,}1034$

X/Y	5	17,5	37,5	75	ni.	xi*ni.	x^2i*ni.
2,5	6	6	4	0	16	40	100
7,5	2	2	3	1	8	60	450
15	1	2	1	1	5	75	1125
n.j	9	10	8	2	29	175	1675
yj*n.j	45	175	300	150	670		
y^2j*n.j	225	3062,5	11250	11250	25787,5		

$$Coeficiente\ de\ dispersión\ de\ Pearson:\ V = \frac{S}{\bar{x}} \qquad S^2 = m_2 = a_2 - a_1^2$$

<u>*Variable X*</u>

$$a_{20} = \frac{1675}{29} = 57{,}7586$$

$$S_x^2 = 57{,}7586 - (6{,}0345)^2 = 21{,}3436$$

$$S_x = \sqrt{21{,}3436} = 4{,}6199$$

$$V_x = \frac{S}{\bar{x}} = \frac{4{,}6199}{6{,}0345} = \boxed{0{,}7656}$$

<u>*Variable Y*</u>

$$a_{02} = \frac{25787{,}5}{29} = 889{,}2241$$

$$S_y^2 = 889{,}2241 - (23{,}1034)^2 = 355{,}4548$$

$$S_y = \sqrt{355{,}4548} = 18{,}8535$$

$$V_y = \frac{S}{\bar{y}} = \frac{18{,}8535}{23{,}1034} = 0{,}8160$$

Ej. 2

Una franquicia de restaurantes está estudiando la relación que existe entre el número de mesas de sus salones (X) y el beneficio antes de impuestos semanal generado en miles de euros (Y). La franquicia cuenta con 25 establecimientos, y la tabla de correlación de ambas variables es la siguiente:

X/Y	2	3	4
10	4	2	4
20	3	2	2
30	2	2	4

X/Y	2	3	4	ni.
10	4	2	4	10
20	3	2	2	7
30	2	2	4	8
n.j	9	6	10	25

Ej. 2

Una franquicia de restaurantes está estudiando la relación entre el número de mesas de sus salones (X) y el beneficio antes de impuestos generado en miles de euros (Y). La franquicia cuenta con 25 establecimientos, y la tabla de correlación de ambas variables es la siguiente:

a) Calcule la covarianza entre ambas variables.

X/Y	2	3	4	ni.	xi*ni.
10	4	2	4	10	100
20	3	2	2	7	140
30	2	2	4	8	240
n.j	9	6	10	25	480
yj*n.j	18	18	40	76	

$$\boxed{S_{xy} = m_{11} = a_{11} - a_{10} \cdot a_{01}}$$

80	60	160
120	120	160
120	180	480
	Suma:	1480

$$\mathrm{a}_{11} = \frac{1480}{25} = 59{,}20$$

$$a_{10} = \frac{480}{25} = 19{,}20$$

$$a_{01} = \frac{76}{25} = 3{,}040$$

$$\mathrm{S_{xy} = m_{11} = a_{11} - a_{10} \cdot a_{01} = 59{,}2 - 19{,}2 \cdot 3{,}04 = 0{,}8320}$$

Ej. 2

Una franquicia de restaurantes está estudiando la relación entre el número de mesas de sus salones (X) y el beneficio antes de impuestos generado en miles de euros (Y). La franquicia cuenta con 25 establecimientos, y la tabla de correlación de ambas variables es la siguiente:

b) Determine si ambas variables son estadísticamente independientes.

$$\mathrm{S_{xy}} = 0{,}8320$$

$$NO\ son\ estadísticamente\ independientes$$
$$ya\ que\ S_{xy}\ es\ distinta\ de\ 0.$$


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 2.2.

Regresión y correlación

Práctica

Ej. 1 ***A continuación, se muestra una tabla de correlación con las variables "número de hectáreas cultivadas" (X) y "tasa de rentabilidad de la explotación agraria (en %)" (Y) de las explotaciones agrarias de una determinada comarca:***

X \ Y	(4-6)	(6-8)	(8-10)
(0-20)	7	3	0
(20-40)	1	11	0
(40-60)	0	2	8

Para cuantificar la relación existente entre ambas variables, se propone la especificación del modelo de regresión lineal: $\hat{y}_i = a + b \cdot x_i$

X \ Y	5	7	9	ni.
10	7	3	0	10
30	1	11	0	12
50	0	2	8	10
n.j	8	16	8	32

Ej. 1

A continuación, se muestra una tabla de correlación con las variables "número de hectáreas cultivadas" (X) y "tasa de rentabilidad de la explotación agraria (en %)" (Y) de las explotaciones agrarias de una determinada comarca. Se pide:

a) Estime los parámetros correspondientes al modelo de regresión lineal

X \ Y	5	7	9	ni.	xi*ni.	x^2i*ni.
10	7	3	0	10	100	1000
30	1	11	0	12	360	10800
50	0	2	8	10	500	25000
n.j	8	16	8	32	960	36800
yj*n.j	40	112	72	224		
y^2j*n.j	200	784	648	1632		

350	210	0
150	2310	0
0	700	3600
	Suma:	7320

$$\hat{b} = \frac{S_{xy}}{S_x^2} = \frac{18{,}75}{250} = 0{,}075$$

$$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x} = 7{,}0 - 0{,}075 \cdot 30{,}0 = 4{,}75$$

$$\boldsymbol{\hat{y}_i = \hat{a} + \hat{b} \cdot x_i = 4{,}75 + 0{,}075 \cdot x_i}$$

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$S_{xy} = a_{11} - a_{10} \cdot a_{01}$

$S_x^2 = a_{20} - a_{10}^2$

$S_y^2 = a_{02} - a_{01}^2$

Ej. 1 *A continuación, se muestra una tabla de correlación con las variables "número de hectáreas cultivadas" (X) y "tasa de rentabilidad de la explotación agraria (en %)" (Y) de las explotaciones agrarias de una determinada comarca. Se pide:*

b) Comente la bondad del ajuste de la recta de regresión.

X \ Y	5	7	9	ni.	xi*ni.	x^2i*ni.
10	7	3	0	10	100	1000
30	1	11	0	12	360	10800
50	0	2	8	10	500	25000
n.j	8	16	8	32	960	36800
yj*n.j	40	112	72	224		
y^2j*n.j	200	784	648	1632		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$$\hat{b} = 0{,}075$$
$$\hat{a} = 4{,}75$$

$$\hat{y}_i = 4{,}75 + 0{,}075 \cdot x_i$$

$$R^2 = \frac{S_{xy}^2}{S_x^2 \cdot S_y^2} = \frac{(18{,}75)^2}{250 \cdot 2} = \mathbf{0,7031}$$

$El\ 70{,}3\%\ de\ la\ varianza\ o\ comportamiento\ de\ la\ variable\ Y\ viene\ recogida\ por\ la\ regresión.$

Ej. 1

A continuación, se muestra una tabla de correlación con las variables "número de hectáreas cultivadas" (X) y "tasa de rentabilidad de la explotación agraria (en %)" (Y) de las explotaciones agrarias de una determinada comarca. Se pide:

c) Realice el análisis de correlación entre las variables.

X \ Y	5	7	9	ni.	xi*ni.	x^2i*ni.
10	7	3	0	10	100	1000
30	1	11	0	12	360	10800
50	0	2	8	10	500	25000
n.j	8	16	8	32	960	36800
yj*n.j	40	112	72	224		
y^2j*n.j	200	784	648	1632		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$$\hat{y}_i = 4{,}75 + 0{,}075 \cdot x_i$$

$$\hat{b} = 0{,}075$$

$$\hat{a} = 4{,}75$$

$$R^2 = 0{,}7031$$

$$R = \sqrt{R^2} = \sqrt{0{,}7031} = \mathbf{0{,}8385}$$

La relación entre ambas variables es alta y en el mismo sentido (signo +)

$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$

Ej. 1

A continuación, se muestra una tabla de correlación con las variables "número de hectáreas cultivadas" (X) y "tasa de rentabilidad de la explotación agraria (en %)" (Y) de las explotaciones agrarias de una determinada comarca. Se pide:

d) ¿Qué significado económico tienen los parámetros obtenidos?

X \ Y	5	7	9	ni.	xi*ni.	x^2i*ni.
10	7	3	0	10	100	1000
30	1	11	0	12	360	10800
50	0	2	8	10	500	25000
n.j	8	16	8	32	960	36800
yj*n.j	40	112	72	224		
y^2j*n.j	200	784	648	1632		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$$\hat{y}_i = 4{,}75 + 0{,}075 \cdot x_i$$

$$\hat{b} = 0{,}075$$
$$\hat{a} = 4{,}75$$

$$R^2 = 0{,}7031$$
$$R = 0{,}8385$$

- $\hat{b}$ → ***propensión marginal de la productividad con respecto a la extensión de las explotaciones:*** *por cada hectárea adicional de extensión, la rentabilidad se incrementa en un 0,075%.*
- $\hat{a}$ → ***productividad autónoma de las explotaciones****: si no hay ninguna hectárea cultivada, la rentabilidad es de 4,75%.*

$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$

Ej. 1

A continuación, se muestra una tabla de correlación con las variables "número de hectáreas cultivadas" (X) y "tasa de rentabilidad de la explotación agraria (en %)" (Y) de las explotaciones agrarias de una determinada comarca. Se pide:

e) Predicción de rentabilidad para unas explotaciones de 65 y 70 hectáreas.

X \ Y	5	7	9	ni.	xi*ni.	x^2i*ni.
10	7	3	0	10	100	1000
30	1	11	0	12	360	10800
50	0	2	8	10	500	25000
n.j	8	16	8	32	960	36800
yj*n.j	40	112	72	224		
y^2j*n.j	200	784	648	1632		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$$\hat{y}_i = 4{,}75 + 0{,}075 \cdot x_i$$

$$\hat{b} = 0{,}075$$
$$\hat{a} = 4{,}75$$

$$R^2 = 0{,}7031$$
$$R = 0{,}8385$$

- $x = 65 \rightarrow \hat{y}_i = 4{,}75 + 0{,}075 \cdot (65) = \mathbf{9{,}6250}\%$
- $x = 70 \rightarrow \hat{y}_i = 4{,}75 + 0{,}075 \cdot (70) = \mathbf{10{,}0}\%$

$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$

Ej. 2

Una consultora está estudiando, para cierto sector productivo, la relación existente entre la inversión en I+D que realizan las empresas (variable X, en miles de euros), y el valor de su cifra de negocio (variable Y, en cientos de miles de euros). Las frecuencias obtenidas se recogen en la siguiente tabla de correlación:

X \ Y	(0-4)	(4-8)	(8-12)	(12-16)
(0-10)	10	1	0	0
(10-20)	2	11	0	0
(20-30)	1	2	8	1
(30-40)	0	0	2	6
(40-50)	0	0	0	2

Para cuantificar la relación existente entre ambas variables, se propone la especificación del modelo de regresión lineal: $\hat{y}_i = a + b \cdot x_i$

X \ Y	2	6	10	14	ni.
5	10	1	0	0	11
15	2	11	0	0	13
25	1	2	8	1	12
35	0	0	2	6	8
45	0	0	0	2	2
n.j	13	14	10	9	46

Ej. 2

Una consultora está estudiando, para cierto sector productivo, la relación existente entre la inversión en I+D que realizan las empresas (variable X, en miles de euros), y el valor de su cifra de negocio (variable Y, en cientos de miles de euros). Se pide:

a) Estime los parámetros correspondientes al modelo de regresión lineal

X \ Y	2	6	10	14	ni.	xi*ni.	x^2i*ni.
5	10	1	0	0	11	55	275
15	2	11	0	0	13	195	2925
25	1	2	8	1	12	300	7500
35	0	0	2	6	8	280	9800
45	0	0	0	2	2	90	4050
n.j	13	14	10	9	46	920	24550
yj*n.j	26	84	100	126	336		
y^2j*n.j	52	504	1000	1764	3320		

100	30	0	0
60	990	0	0
50	300	2000	350
0	0	700	2940
0	0	0	1260
		Suma:	8780

$$\hat{b} = \frac{S_{xy}}{S_x^2} = \frac{44{,}7826}{133{,}6957} = 0{,}3350$$

$$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x} = 7{,}3043 - 0{,}3350 \cdot 20{,}0 = 0{,}6052$$

$$\boldsymbol{\hat{y}_i = \hat{a} + \hat{b} \cdot x_i = 0{,}6052 + 0{,}3350 \cdot x_i}$$

$a_{11}=$	190,8696	
$\bar{X}=a_{10}=$	20,0000	
$\bar{y}=a_{01}=$	7,3043	
$a_{20}=$	533,6957	
$a_{02}=$	72,1739	
$S_{xy}=m_{11}=$	44,7826	$S_{xy} = a_{11} - a_{10} \cdot a_{01}$
$S_x^2=m_{20}=$	133,6957	$S_x^2 = a_{20} - a_{10}^2$
$S_y^2=m_{02}=$	18,8204	$S_y^2 = a_{02} - a_{01}^2$

Ej. 2 *Una consultora está estudiando, para cierto sector productivo, la relación existente entre la inversión en I+D que realizan las empresas (variable X, en miles de euros), y el valor de su cifra de negocio (variable Y, en cientos de miles de euros). Se pide:*

b) Comente la bondad del ajuste de la recta de regresión.

X \ Y	2	6	10	14	ni.	xi*ni.	x^2i*ni.
5	10	1	0	0	11	55	275
15	2	11	0	0	13	195	2925
25	1	2	8	1	12	300	7500
35	0	0	2	6	8	280	9800
45	0	0	0	2	2	90	4050
n.j	13	14	10	9	46	920	24550
yj*n.j	26	84	100	126	336		
y^2j*n.j	52	504	1000	1764	3320		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x{}^2=m_{20}=$	250,0000
$S_y{}^2=m_{02}=$	2,0000

$$\hat{b} = 0{,}3350$$

$$\hat{a} = 0{,}6052$$

$$\hat{y}_i = 0{,}6052 + 0{,}3350 \cdot x_i$$

$$R^2 = \frac{S_{xy}^2}{S_x^2 \cdot S_y^2} = \frac{(44{,}7826)^2}{133{,}6957 \cdot 18{,}8204} = \mathbf{0,7970}$$

El 79,7% de la varianza o comportamiento de la variable Y viene recogida por la regresión.

Ej. 2

Una consultora está estudiando, para cierto sector productivo, la relación existente entre la inversión en I+D que realizan las empresas (variable X, en miles de euros), y el valor de su cifra de negocio (variable Y, en cientos de miles de euros). Se pide:

c) Realice el análisis de correlación entre las variables.

X \ Y	2	6	10	14	ni.	xi*ni.	x^2i*ni.
5	10	1	0	0	11	55	275
15	2	11	0	0	13	195	2925
25	1	2	8	1	12	300	7500
35	0	0	2	6	8	280	9800
45	0	0	0	2	2	90	4050
n.j	13	14	10	9	46	920	24550
yj*n.j	26	84	100	126	336		
y^2j*n.j	52	504	1000	1764	3320		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$$\hat{b} = 0{,}3350$$

$$\hat{a} = 0{,}6052$$

$$\hat{y}_i = 0{,}6052 + 0{,}3350 \cdot x_i$$

$$R^2 = 0{,}7970$$

$$R = \sqrt{R^2} = \sqrt{0{,}7970} = \mathbf{0{,}8928}$$

La relación entre ambas variables es alta y en el mismo sentido (signo +)

Ej. 2 *Una consultora está estudiando, para cierto sector productivo, la relación existente entre la inversión en I+D que realizan las empresas (variable X, en miles de euros), y el valor de su cifra de negocio (variable Y, en cientos de miles de euros). Se pide:*

d) ¿Qué significado económico tienen los parámetros obtenidos?

X \ Y	2	6	10	14	ni.	xi*ni.	x^2i*ni.
5	10	1	0	0	11	55	275
15	2	11	0	0	13	195	2925
25	1	2	8	1	12	300	7500
35	0	0	2	6	8	280	9800
45	0	0	0	2	2	90	4050
n.j	13	14	10	9	46	920	24550
yj*n.j	26	84	100	126	336		
y^2j*n.j	52	504	1000	1764	3320		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$$\hat{b} = 0{,}3350$$

$$\hat{a} = 0{,}6052$$

$$\hat{y}_i = 0{,}6052 + 0{,}3350 \cdot x_i$$

$$R^2 = 0{,}7970$$

$$R = 0{,}8928$$

- ***$\hat{b}$ → propensión marginal de la cifra de negocios con respecto a la inversión en I+D:*** *por cada 1000 euros adicionales invertidos en I+D, la cifra de negocios se incrementa en 33500 euros.*
- ***$\hat{a}$ → cifra de negocio autónoma de las empresas****: si no hay ninguna inversión en I+D, la empresa tiene una cifra de negocio de 60520 euros.*

$$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$$

Ej. 2

Una consultora está estudiando, para cierto sector productivo, la relación existente entre la inversión en I+D que realizan las empresas (variable X, en miles de euros), y el valor de su cifra de negocio (variable Y, en cientos de miles de euros). Se pide:

e) Predicción de cifra de ventas para una inversión en I+D de 56.000 euros.

X \ Y	2	6	10	14	ni.	xi*ni.	x^2i*ni.
5	10	1	0	0	11	55	275
15	2	11	0	0	13	195	2925
25	1	2	8	1	12	300	7500
35	0	0	2	6	8	280	9800
45	0	0	0	2	2	90	4050
n.j	13	14	10	9	46	920	24550
yj*n.j	26	84	100	126	336		
y^2j*n.j	52	504	1000	1764	3320		

$a_{11}=$	228,7500
$\bar{x}=a_{10}=$	30,0000
$\bar{y}=a_{01}=$	7,0000
$a_{20}=$	1150,0000
$a_{02}=$	51,0000
$S_{xy}=m_{11}=$	18,7500
$S_x^2=m_{20}=$	250,0000
$S_y^2=m_{02}=$	2,0000

$$\hat{b} = 0{,}3350$$

$$\hat{a} = 0{,}6052$$

$$\hat{y}_i = 0{,}6052 + 0{,}3350 \cdot x_i$$

$$R^2 = 0{,}7970$$

$$R = 0{,}8385$$

- $x = 56000 \rightarrow \hat{y}_i = 0{,}6052 + 0{,}3350 \cdot (56) = 19{,}3652 \; cientos\ de\ miles\ de\ euros = 1936520 \; euros$

$\hat{a} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$


¡Muchas gracias!

Tema 3.0.
Números Índice
Práctica

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Ej. 1

Una cadena de ropa comercializa 3 tipos de camisetas: línea básica, línea casual y línea premium. A lo largo de 6 años, los datos referentes al precio de las camisetas (en euros), y a las cantidades vendidas (en miles de unidades) son los siguientes:

	Línea Básica		Línea Casual		Línea Premium	
Período	Precio	Cantidad	Precio	Cantidad	Precio	Cantidad
0	12	48	15	36	30	9
1	14	49	18	38	32	12
2	15	53	20	39	35	15
3	16	55	24	41	40	19
4	18	52	28	44	50	23
5	20	60	30	46	55	28

Se pide, para todos los períodos considerados:

Ej. 1

Una cadena de ropa comercializa 3 tipos de camisetas: línea básica, línea casual y línea premium. A lo largo de 6 años, los datos referentes al precio de las camisetas (en euros), y a las cantidades vendidas (en miles de unidades) son los siguientes:

	Línea Básica		Línea Casual		Línea Premium	
Período	Precio	Cantidad	Precio	Cantidad	Precio	Cantidad
0	12	48	15	36	30	9
1	14	49	18	38	32	12
2	15	53	20	39	35	15
3	16	55	24	41	40	19
4	18	52	28	44	50	23
5	20	60	30	46	55	28

a) Números índice de precios simples para las 3 líneas, con base fija en el período 0 (0=100).

$$p_0^t = \frac{p_t}{p_0}$$

	Línea Básica	Línea Casual	Línea Premium
Período	Precio	Precio	Precio
0	100,00	100,00	100,00
1	116,67	120,00	106,67
2	125,00	133,33	116,67
3	133,33	160,00	133,33
4	150,00	186,67	166,67
5	166,67	200,00	183,33

Ej. 1

Una cadena de ropa comercializa 3 tipos de camisetas: línea básica, línea casual y línea premium. A lo largo de 6 años, los datos referentes al precio de las camisetas (en euros), y a las cantidades vendidas (en miles de unidades) son los siguientes:

	Línea Básica		Línea Casual		Línea Premium	
Período	Precio	Cantidad	Precio	Cantidad	Precio	Cantidad
0	12	48	15	36	30	9
1	14	49	18	38	32	12
2	15	53	20	39	35	15
3	16	55	24	41	40	19
4	18	52	28	44	50	23
5	20	60	30	46	55	28

b) Números índice de cantidades simples encadenados para las 3 líneas (multiplicados por 100).

$$q_{t-1}^{t} = \frac{q_t}{q_{t-1}}$$

Período	Línea Básica Cantidad	Línea Casual Cantidad	Línea Premium Cantidad
0	-	-	-
1	102,08	105,56	133,33
2	108,16	102,63	125,00
3	103,77	105,13	126,67
4	94,55	107,32	121,05
5	115,38	104,55	121,74

Ej. 1

Una cadena de ropa comercializa 3 tipos de camisetas: línea básica, línea casual y línea premium. A lo largo de 6 años, los datos referentes al precio de las camisetas (en euros), y a las cantidades vendidas (en miles de unidades) son los siguientes:

Período	Línea Básica		Línea Casual		Línea Premium	
	Precio	Cantidad	Precio	Cantidad	Precio	Cantidad
0	12	48	15	36	30	9
1	14	49	18	38	32	12
2	15	53	20	39	35	15
3	16	55	24	41	40	19
4	18	52	28	44	50	23
5	20	60	30	46	55	28

c) Números índice de precios complejos de las 3 líneas, con base fija en el período 0 (0=100), por los métodos de Laspeyres, Paasche y Fisher.

Laspeyres: $L_p = \frac{\sum_{i=1}^{N} p_{it} \cdot q_{i0}}{\sum_{i=1}^{N} p_{i0} \cdot q_{i0}}$

Paasche: $P_p = \frac{\sum_{i=1}^{N} p_{it} \cdot q_{it}}{\sum_{i=1}^{N} p_{i0} \cdot q_{it}}$

Fisher: $F_p = \sqrt{L_p P_p}$

Período	Laspeyres Precio	Paasche Precio	Fisher Precio
0	100,00	100,00	100,00
1	116,02	115,55	115,78
2	126,62	125,67	126,15
3	143,72	142,22	142,97
4	167,53	168,09	167,81
5	182,90	183,11	183,01

Ej. 1

Una cadena de ropa comercializa 3 tipos de camisetas: línea básica, línea casual y línea premium. A lo largo de 6 años, los datos referentes al precio de las camisetas (en euros), y a las cantidades vendidas (en miles de unidades) son los siguientes:

	Línea Básica		Línea Casual		Línea Premium	
Período	Precio	Cantidad	Precio	Cantidad	Precio	Cantidad
0	12	48	15	36	30	9
1	14	49	18	38	32	12
2	15	53	20	39	35	15
3	16	55	24	41	40	19
4	18	52	28	44	50	23
5	20	60	30	46	55	28

d) Números índice de cantidades complejos de las 3 líneas, con base fija en el período 0 (0=100), por los métodos de Laspeyres, Paasche y Fisher.

$$\text{Laspeyres: } L_q = \frac{\sum_{i=1}^{N} q_{it} \cdot p_{i0}}{\sum_{i=1}^{N} q_{i0} \cdot p_{i0}}$$

$$\text{Paasche: } P_q = \frac{\sum_{i=1}^{N} q_{it} \cdot p_{it}}{\sum_{i=1}^{N} q_{i0} \cdot p_{it}}$$

$$\text{Fisher: } F_q = \sqrt{L_q P_q}$$

	Laspeyres	Paasche	Fisher
Período	Cantidades	Cantidades	Cantidades
0	100,00	100,00	100,00
1	109,52	109,08	109,30
2	120,56	119,66	120,11
3	133,12	131,73	132,42
4	142,42	142,89	142,66
5	162,34	162,52	162,43

Ej. 2 *Se dispone de la serie del valor las exportaciones de bienes industriales de un determinado país (variable X, en millones de euros) desde el período (año) 0 al 12, en términos nominales:*

Período	X
0	75000
1	83000
2	88000
3	91000
4	100000
5	102000
6	108000
7	112000
8	118000
9	125000
10	136000
11	142000
12	154000

Se cuenta además con los números índice de precios de bienes industriales en base 0=100 y en base 6=100 siguientes:

Período	p^t_0
0	100
1	102
2	105
3	106
4	108
5	112
6	115

Período	p^t_6
6	100
7	108
8	112
9	114
10	118
11	120
12	122

a) Calcule la serie de tasas de crecimiento reales anuales de X desde el período 0 al 12.

Período	X	p^t_0	p^t_6	Deflactor (6=100)
0	75000	100	-	86,9565
1	83000	102	-	88,6957
2	88000	105	-	91,3043
3	91000	106	-	92,1739
4	100000	108	-	93,9130
5	102000	112	-	97,3913
6	108000	115	100	100
7	112000	-	108	108
8	118000	-	112	112
9	125000	-	114	114
10	136000	-	118	118
11	142000	-	120	120
12	154000	-	122	122

- $p_6^0 = \frac{p_0^0}{p_0^6} = \frac{100}{115} = 0{,}869565$
- $p_6^1 = \frac{p_0^1}{p_0^6} = \frac{102}{115} = 0{,}886957$
- $p_6^2 = \frac{p_0^2}{p_0^6} = \frac{105}{115} = 0{,}913043$
- $p_6^3 = \frac{p_0^3}{p_0^6} = \frac{106}{115} = 0{,}921739$
- $p_6^4 = \frac{p_0^4}{p_0^6} = \frac{108}{115} = 0{,}939130$
- $p_6^5 = \frac{p_0^5}{p_0^6} = \frac{112}{115} = 0{,}973913$
- $p_6^6 = \frac{p_0^6}{p_0^6} = \frac{115}{115} = 1$

$$p_h^t = \frac{p_0^t}{p_0^h}$$

Ej. 2 *Se dispone de la serie del valor las exportaciones de bienes industriales de un determinado país (variable X, en millones de euros) desde el período (año) 0 al 12, en términos nominales:*

Período	X
0	75000
1	83000
2	88000
3	91000
4	100000
5	102000
6	108000
7	112000
8	118000
9	125000
10	136000
11	142000
12	154000

Se cuenta además con los números índice de precios de bienes industriales en base 0=100 y en base 6=100 siguientes:

Período	p^t_0
0	100
1	102
2	105
3	106
4	108
5	112
6	115

Período	p^t_6
6	100
7	108
8	112
9	114
10	118
11	120
12	122

a) Calcule la serie de tasas de crecimiento reales anuales de X desde el período 0 al 12.

Período	X	p^t_0	p^t_6	Deflactor (6=100)	X a precios de periodo 6	% crec. anual real
0	75000	100	-	86,9565	86250,00	-
1	83000	102	-	88,6957	93578,43	8,4967
2	88000	105	-	91,3043	96380,95	2,9948
3	91000	106	-	92,1739	98726,42	2,4335
4	100000	108	-	93,9130	106481,48	7,8551
5	102000	112	-	97,3913	104732,14	-1,6429
6	108000	115	100	100	108000,00	3,1202
7	112000	-	108	108	103703,70	-3,9781
8	118000	-	112	112	105357,14	1,5944
9	125000	-	114	114	109649,12	4,0737
10	136000	-	118	118	115254,24	5,1119
11	142000	-	120	120	118333,33	2,6716
12	154000	-	122	122	126229,51	6,6728

Ej. 2 *Se dispone de la serie del valor las exportaciones de bienes industriales de un determinado país (variable X, en millones de euros) desde el período (año) 0 al 12, en términos nominales:*

Período	X
0	75000
1	83000
2	88000
3	91000
4	100000
5	102000
6	108000
7	112000
8	118000
9	125000
10	136000
11	142000
12	154000

Se cuenta además con los números índice de precios de bienes industriales en base 0=100 y en base 6=100 siguientes:

Período	p^t_0
0	100
1	102
2	105
3	106
4	108
5	112
6	115

Período	p^t_6
6	100
7	108
8	112
9	114
10	118
11	120
12	122

b) ¿Cuál fue la tasa de crecimiento real de X entre los años 0 y 12?

Período	X	p^t_0	p^t_6	Deflactor (6=100)	X a precios de periodo 6
0	75000	100	-	86,9565	86250,00
1	83000	102	-	88,6957	93578,43
2	88000	105	-	91,3043	96380,95
3	91000	106	-	92,1739	98726,42
4	100000	108	-	93,9130	106481,48
5	102000	112	-	97,3913	104732,14
6	108000	115	100	100	108000,00
7	112000	-	108	108	103703,70
8	118000	-	112	112	105357,14
9	125000	-	114	114	109649,12
10	136000	-	118	118	115254,24
11	142000	-	120	120	118333,33
12	154000	-	122	122	126229,51

Tasa de crecimiento (%)

$$= \frac{x_t - x_{t-1}}{x_{t-1}} \cdot 100 =$$

$$= \frac{126229{,}51 - 86250{,}00}{86250{,}00} \cdot 100$$

$$= 46{,}3531\%$$


¡Muchas gracias!

Tema 4.1.

Introducción a la probabilidad

Práctica

Estadística Empresarial

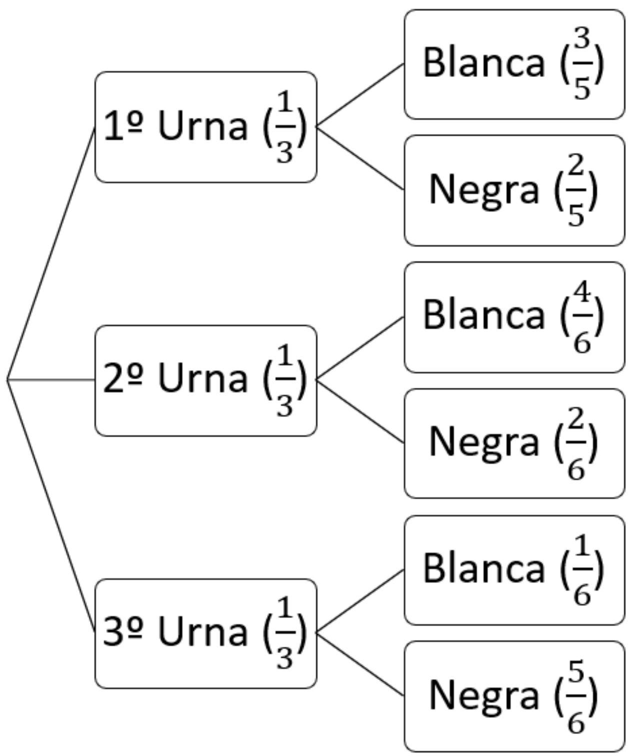
Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Ej. 1 *Sean tres urnas idénticas que contienen bolas del mismo peso y tamaño. En la primera urna hay 3 bolas blancas y 2 negras, en la segunda urna hay 4 bolas blancas y 2 negras, y en la tercera urna hay una bola blanca y 5 negras. Si se escoge al azar una urna, y de ella se extrae una bola también al azar, se pide:*

a) Probabilidad de que la bola sea blanca.

Teorema de la Probabilidad Total

$$P(Blanca) = P(1ºU) \cdot P(B \mid 1ºU) + P(2ºU) \cdot P(B \mid 2ºU) + P(3ºU) \cdot P(B \mid 3ºU)$$



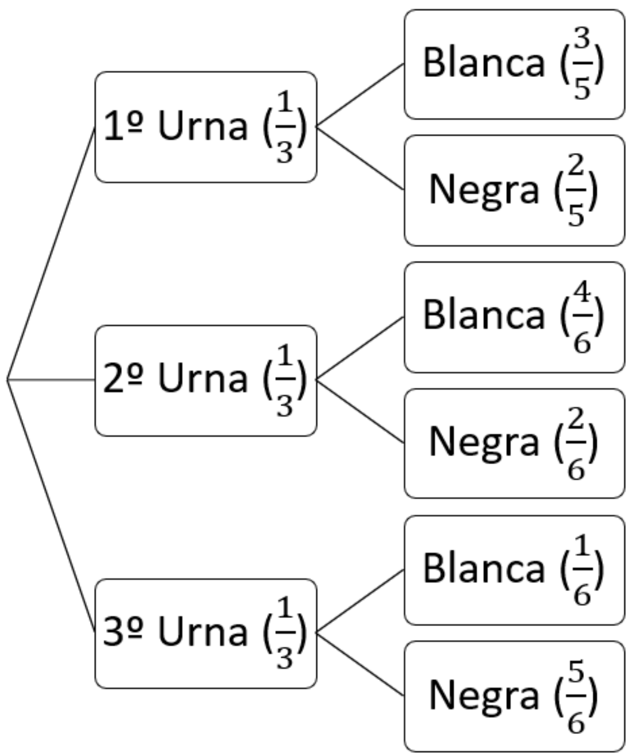
$$P(Blanca) = \frac{1}{3} \cdot \frac{3}{5} + \frac{1}{3} \cdot \frac{4}{6} \cdot + \frac{1}{3} \cdot \frac{1}{6} = 0{,}48$$

Ej. 1 *Sean tres urnas idénticas que contienen bolas del mismo peso y tamaño. En la primera urna hay 3 bolas blancas y 2 negras, en la segunda urna hay 4 bolas blancas y 2 negras, y en la tercera urna hay una bola blanca y 5 negras. Si se escoge al azar una urna, y de ella se extrae una bola también al azar, se pide:*

b) Probabilidad de que, si la bola sale negra, esta pertenezca a la segunda urna.

Teorema de Bayes

$$P(\,2^{\underline{o}}U|\,Negra) = \frac{P(N\,|\;2^{\underline{o}}U)\cdot P(2^{\underline{o}}U)}{P(N\,|\;1^{\underline{o}}U)\cdot P(1^{\underline{o}}U) + P(N\,|\;2^{\underline{o}}U)\cdot P(2^{\underline{o}}U) + P(N\,|\;3^{\underline{o}}U)\cdot P(3^{\underline{o}}U)} =$$



$$= \frac{\frac{1}{3}\cdot\frac{2}{6}}{\frac{1}{3}\cdot\frac{2}{5} + \frac{1}{3}\cdot\frac{2}{6} + \frac{1}{3}\cdot\frac{5}{6}} = 0{,}213$$

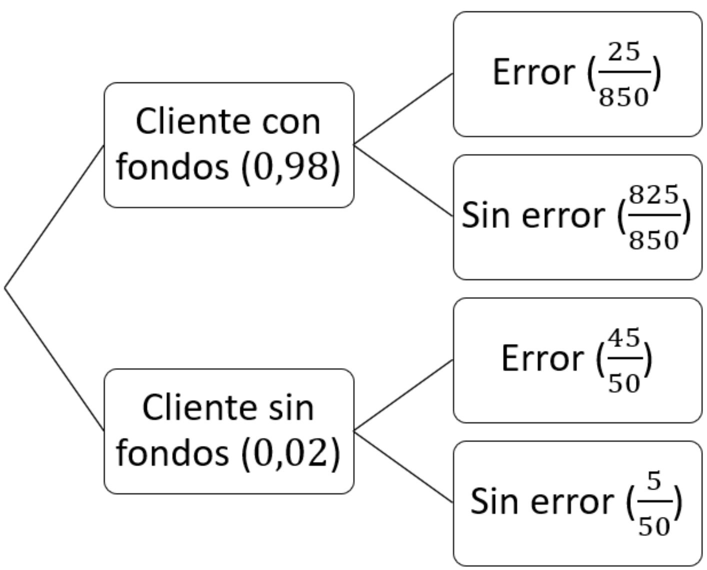
Ej. 2 *Una sucursal de una caja de ahorros dispone de la siguiente información referente a los cheques extendidos por sus clientes:*

- *De 850 clientes con fondos que extendieron un cheque, 25 cometieron algún error de forma.*
- *De 50 clientes sin fondos que extendieron un cheque, 45 cometieron algún error de forma.*
- *El 98% de los clientes tienen fondos.*

¿Cuál es la probabilidad de que, si un cliente extendió un cheque cometiendo algún error de forma, este fuera un cliente sin fondos?

Teorema de Bayes

$$P(\,Cliente\ SF|\,Error) = \frac{P(Error\,|\ Cliente\ SF)\cdot P(Cliente\ SF)}{P(Error\,|\ Cliente\ CF)\cdot P(Cliente\ CF) + P(Error\,|\ Cliente\ SF)\cdot P(Cliente\ SF)} =$$

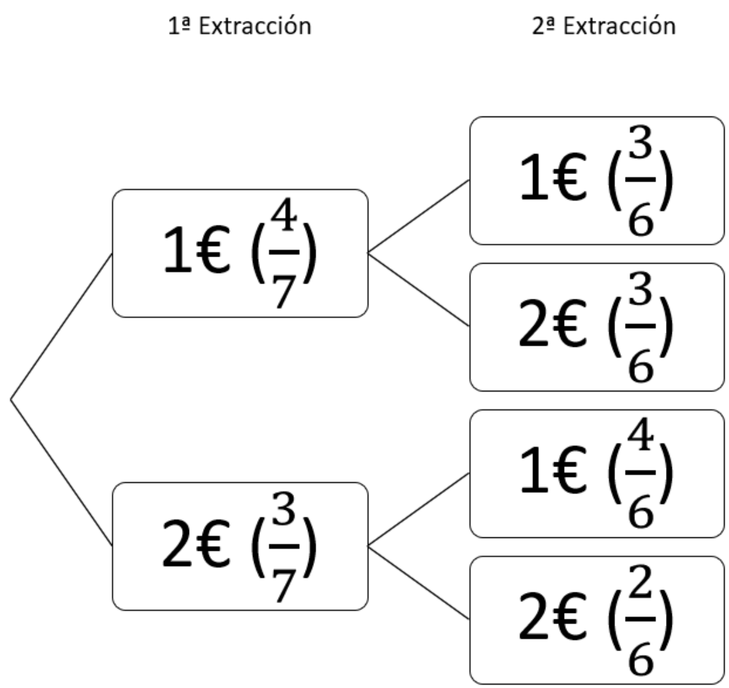


$$P(\,Cliente\ SF|\,Error) = \frac{\frac{45}{50}\cdot 0{,}02}{\frac{25}{850}\cdot 0{,}98 + \frac{45}{50}\cdot 0{,}02} = 0{,}38$$

Ej. 3

En una caja opaca hay 4 monedas de 1 euro y 3 monedas de 2 euros. Se extraen de la caja sucesivamente, y al azar, dos monedas sin devolución. Si las monedas de 1 y 2 euros tienen la misma probabilidad de salir, ¿qué probabilidad hay de obtener 4 euros al extraer las dos monedas?

Probabilidad condicional (sin reemplazamiento)

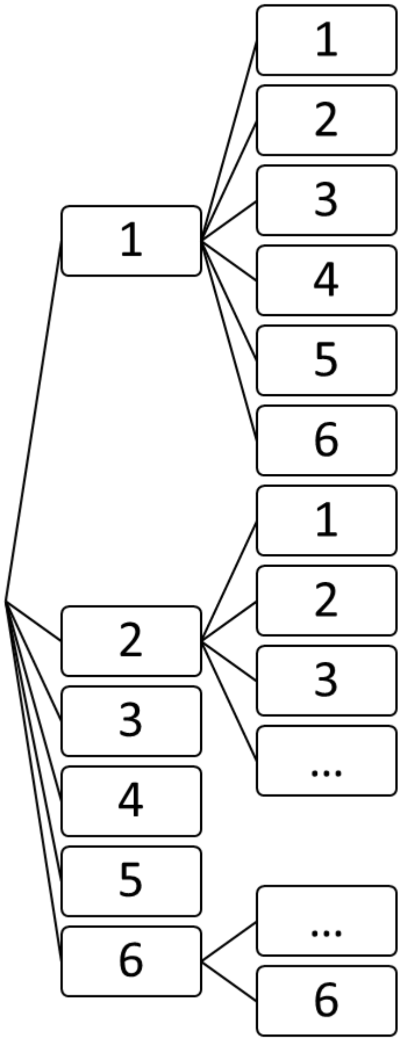


$$P(4€) = P(2€ \cap 2€) = P(2€) \cdot P(2€ \mid 2€) = \frac{3}{7} \cdot \frac{2}{6} = 0{,}14$$

Ej. 4 *Problema del Caballero de Meré. Se pide:*

a) Probabilidad de lanzar n veces dos dados no trucados y que se obtenga al menos un seis doble.

1ª Tirada 2ª Tirada



$$\mathrm{P(al\ menos\ uno)} = 1 - P(\mathrm{ninguno})$$

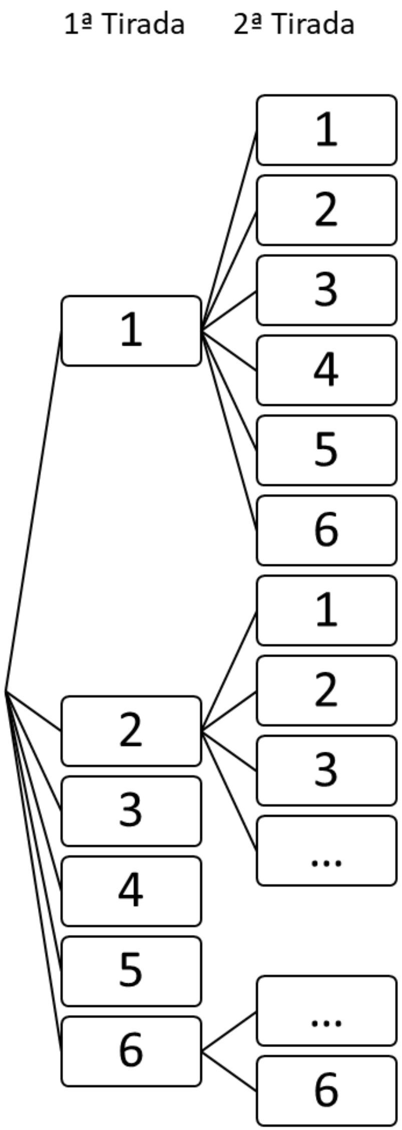
$$\mathrm{P(al\ menos\ uno)} = 1 - P(\mathrm{ninguno}) \cdot P(\mathrm{ninguno}) \cdot P(\mathrm{ninguno}) \cdot \cdots \mathrm{(n\ veces)}$$

$$\mathrm{P(al\ menos\ uno)} = 1 - P(\mathrm{ninguno})^{\mathrm{n}}$$

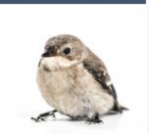
$$\mathrm{P(al\ menos\ uno)} = 1 - \left(\frac{35}{36}\right)^{\mathrm{n}}$$

Ej. 4 *Problema del Caballero de Meré. Se pide:*

b) ¿Qué es más ventajoso, apostar a obtener un seis doble en 24 lanzamientos, o hacerlo a obtener un seis en 4 lanzamientos de un único dado?

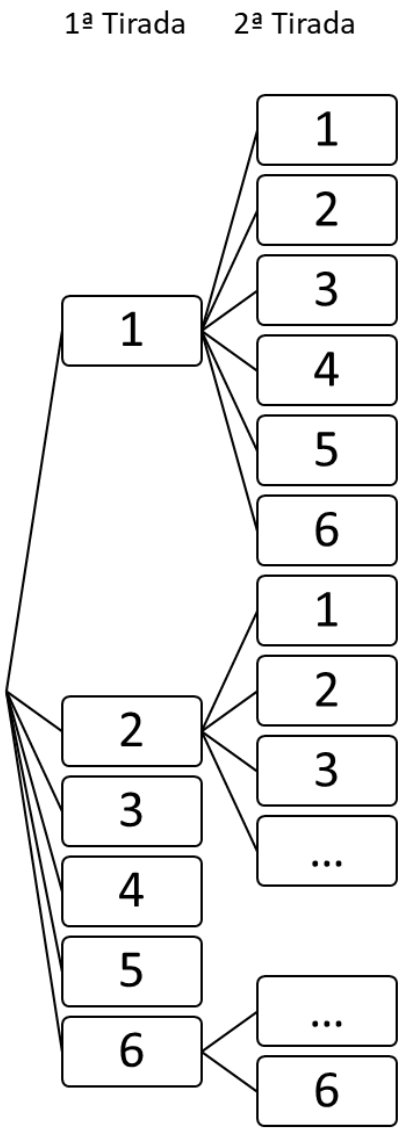


- $P(\text{al menos un seis doble}) = 1 - \left(\frac{35}{36}\right)^{24} = 0{,}4914$
- $P(\text{al menos un seis}) = 1 - \left(\frac{5}{6}\right)^{4} = 0{,}5177$



Ej. 4 *Problema del Caballero de Meré. Se pide:*

c) ¿Cuántos lanzamientos se deben de realizar para que la probabilidad de obtener un seis doble sea de 0,6?



$$P(\text{al menos uno}) = 1 - \left(\frac{35}{36}\right)^n$$

$$0{,}6 = 1 - \left(\frac{35}{36}\right)^n$$

Despejamos n:

$$\left(\frac{35}{36}\right)^n = 1 - 0{,}6$$

$$\left(\frac{35}{36}\right)^n = 0{,}4$$

$$n \cdot \mathrm{Ln}\left(\frac{35}{36}\right) = \mathrm{Ln}(0{,}4)$$

$$n = \frac{\mathrm{Ln}(0{,}4)}{\mathrm{Ln}\left(\frac{35}{36}\right)}$$

$$n = \frac{\mathrm{Ln}(0{,}4)}{\mathrm{Ln}\left(\frac{35}{36}\right)} = 32{,}53 \approx 33 \text{ tiradas}$$

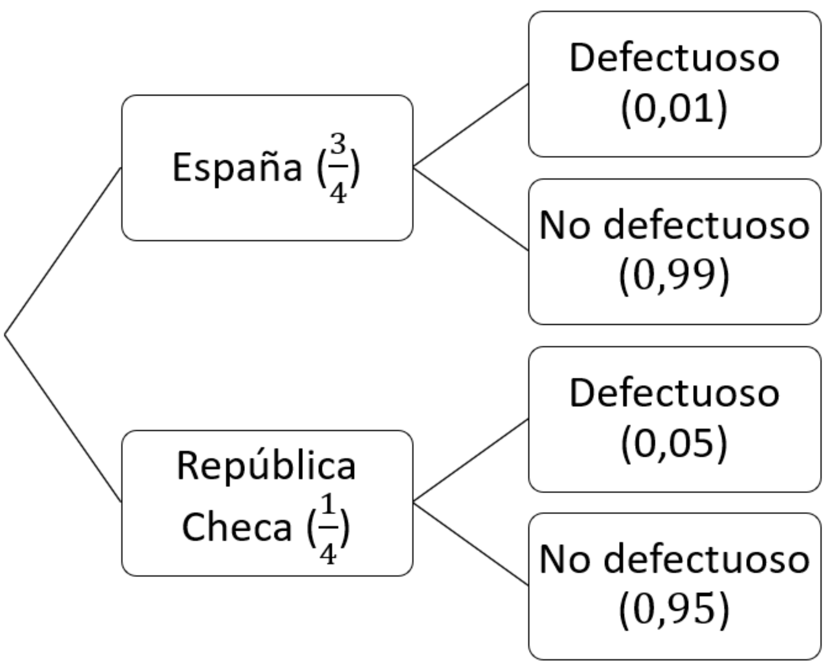
Ej. 5

Una empresa de componentes eléctricos para automóviles tiene dos factorías, una España y otra en la República Checa. De la primera factoría, un 1% de los componentes fabricados salen defectuosos; mientras que de la segunda los componentes defectuosos son el 5%. La factoría española manufactura el triple de piezas que la factoría checa. Si se elige aleatoriamente un componente de los fabricados por la empresa:

a) ¿Qué probabilidad hay de que el componente sea defectuoso?

Teorema de la Probabilidad Total

$$P(Defectuoso) = P(España) \cdot P(Def \mid España) + P(República\ Checa) \cdot P(Def \mid República\ Checa)$$



$$P(Defectuoso) = \frac{3}{4} \cdot 0{,}01 + \frac{1}{4} \cdot 0{,}05 = 0{,}02$$

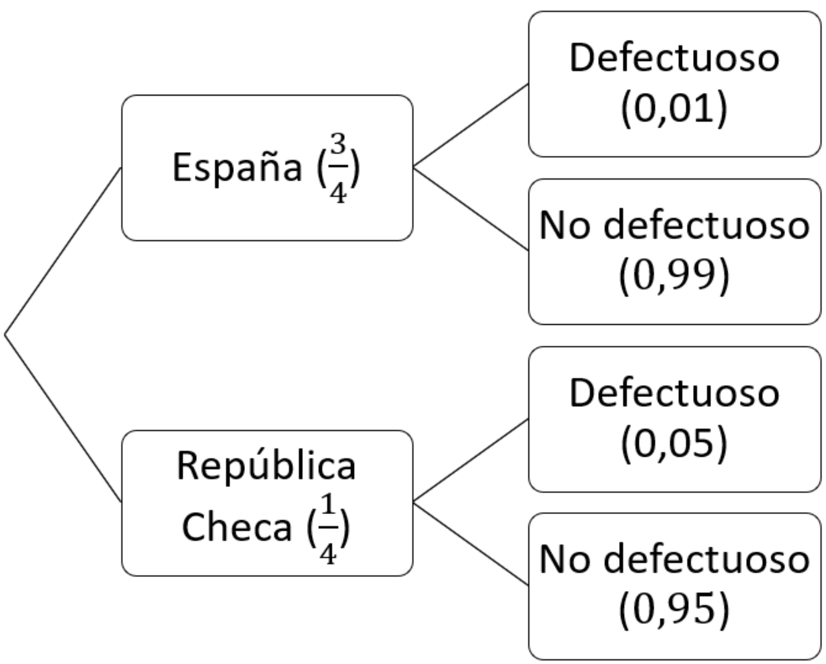
Ej. 5

Una empresa de componentes eléctricos para automóviles tiene dos factorías, una España y otra en la República Checa. De la primera factoría, un 1% de los componentes fabricados salen defectuosos; mientras que de la segunda los componentes defectuosos son el 5%. La factoría española manufactura el triple de piezas que la factoría checa. Si se elige aleatoriamente un componente de los fabricados por la empresa:

a) Si el componente defectuoso, ¿qué probabilidad hay de que proceda de la factoría española?

Teorema de Bayes

$$P(\,España|\,Defectuoso) = \frac{P(Def.|\;España)\cdot P(España)}{P(Def.|\,República\ Checa)\cdot P(República\ Checa) + P(Def.|\;España)\cdot P(España)} =$$



$$\frac{\frac{3}{4}\cdot 0{,}01}{\frac{1}{4}\cdot 0{,}05\ +\ \frac{3}{4}\cdot 0{,}01} = 0{,}375$$

Ej. 6

Un fabricante de vehículos compra cierto tipo de pieza a un proveedor A. De cada lote de piezas que recibe de ese tipo, el 5% son defectuosas (se supone que el hecho de que una pieza de un lote sea defectuosa es independiente de que el resto de piezas del lote lo sean también o no). Si cada vehículo fabricado requiere de 8 piezas, se pide:

a) ¿Qué probabilidad existe de que las 8 piezas del vehículo se encuentren en perfectas condiciones?

$$\mathrm{P(pieza\ defectuosa)} = 0{,}05 \rightarrow \mathrm{P(No\ def.)} = 0{,}95$$

$$\mathrm{P(No\ def} \cap \mathrm{No\ def} \cap \cdots . \cap \mathrm{No\ def)} = \mathrm{P(No\ def)} \cdot \mathrm{P(No\ def)} \ldots \cdot \mathrm{P(No\ def)}$$

$$\mathrm{P(8\ piezas\ No\ def.\ consecutivas)} = \mathrm{P(No\ def.)}^8 = 0{,}95^8 = 0{,}6634$$

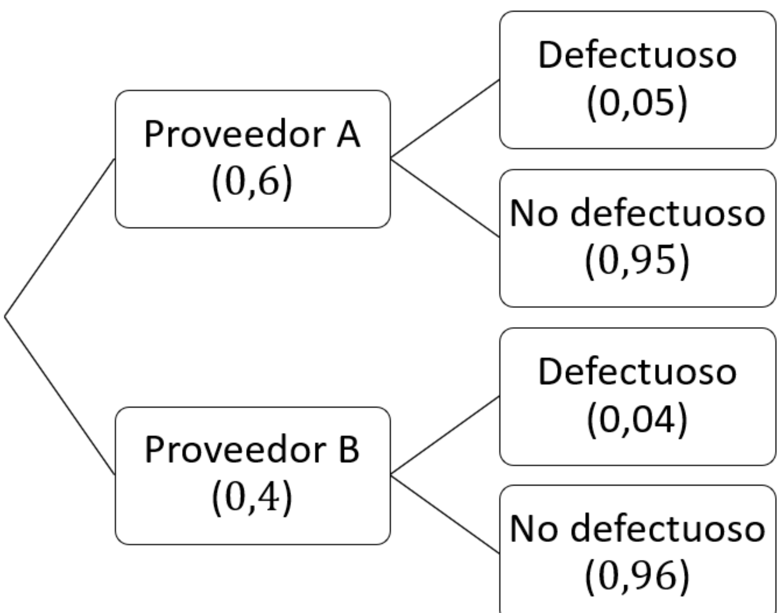
Ej. 6

Un fabricante de vehículos compra cierto tipo de pieza a un proveedor A. De cada lote de piezas que recibe de ese tipo, el 5% son defectuosas (se supone que el hecho de que una pieza de un lote sea defectuosa es independiente de que el resto de piezas del lote lo sean también o no). Si cada vehículo fabricado requiere de 8 piezas, se pide:

b) El fabricante está pensando en la opción de comprar el 40% de los lotes de la pieza anterior a un 2º proveedor B, que suministra lotes con el mismo número de piezas que el proveedor A; pero cuyo porcentaje de piezas defectuosas por lote es del 4%. Si en este caso, las piezas requeridas para fabricar un vehículo fueran elegidas al azar de entre las piezas suministradas por ambos proveedores, ¿qué probabilidad existiría de que las 8 piezas que incorpora un vehículo se encontraran en perfectas condiciones?

Teorema de la Probabilidad Total

$$P(No\ def.) = P(Prov. A) \cdot P(No\ def. \mid Prov. A) + P(Prov. B) \cdot P(No\ def. \mid Prov. B)$$



$$P(No\ def.) = 0{,}6 \cdot 0{,}95 + 0{,}4 \cdot 0{,}96 = 0{,}954$$

$$\mathrm{P}(8 \text{ piezas No def. consecutivas}) =$$
$$= \mathrm{P}(\text{No def.})^8 = 0{,}954^8 = 0{,}686$$

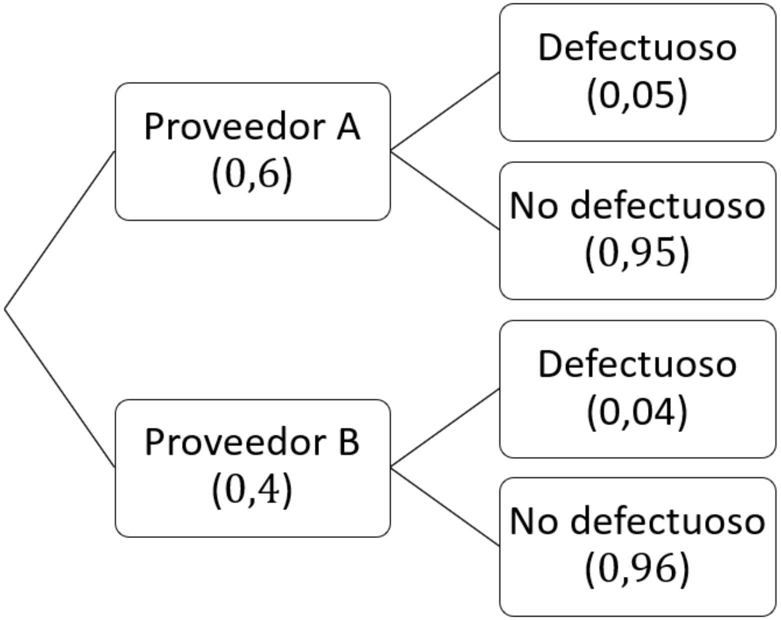
Ej. 6

Un fabricante de vehículos compra cierto tipo de pieza a un proveedor A. De cada lote de piezas que recibe de ese tipo, el 5% son defectuosas (se supone que el hecho de que una pieza de un lote sea defectuosa es independiente de que el resto de piezas del lote lo sean también o no). Si cada vehículo fabricado requiere de 8 piezas, se pide:

c) Dada la situación planteada en el apartado anterior, si seleccionamos al azar una pieza y es defectuosa, ¿qué probabilidad hay de que haya sido suministrada por el proveedor A?

Teorema de Bayes

$$P(\,Prov.A|\,Def) = \frac{P(Def.|\ Prov.A)\cdot P(Prov.A)}{P(Def.|\ Prov.A)\cdot P(Prov.A) + P(Def.|\ Prov.B)\cdot P(Prov.B)} =$$

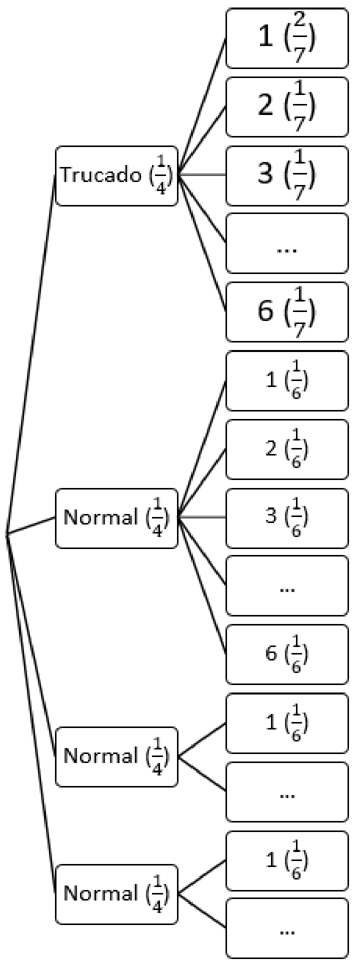


$$\frac{0{,}05\cdot 0{,}6}{0{,}04\cdot 0{,}4\ +\ 0{,}05\cdot 0{,}6} = 0{,}6522$$

Ej. 7 *Se tienen cuatro dados normales y uno trucado. En el trucado, la posibilidad de obtener un 1 es el doble a la de cualquier otra cara, sabiendo que el resto de caras si son equiprobables. Si se elige un dado al azar, se lanza y se obtiene un 1, ¿cuál es la probabilidad de que se trate del dado trucado?*

Teorema de Bayes

$$P(\,Truc|\,1) = \frac{P(1\mid Truc)\cdot P(Truc)}{P(1\mid Truc)\cdot P(Truc) + P(1\mid Normal)\cdot P(Normal) + P(1\mid Normal)\cdot P(Normal) + P(1\mid Normal)\cdot P(Normal)}$$



$$P(\,Trucado|\,1) = \frac{\frac{2}{7}\cdot\frac{1}{4}}{\frac{2}{7}\cdot\frac{1}{4} + \frac{1}{6}\cdot\frac{1}{4} + \frac{1}{6}\cdot\frac{1}{4} + \frac{1}{6}\cdot\frac{1}{4}} = 0{,}3$$

Ej. 8 *La probabilidad de que el índice IBEX-35 aumente cualquier día de apertura de los Mercados es del 0,5. Si el comportamiento del índice en un día es independiente de su comportamiento en los días anteriores, se pide:*

a) Probabilidad de que el índice aumente durante cuatro días consecutivos.

$$\mathrm{P}(\text{Aumente} \cap \text{Aumente} \cap \text{Aumente} \cap \text{Aumente}) =$$
$$= \mathrm{P}(\text{Aumente}) \cdot \mathrm{P}(\text{Aumente}) \cdot \mathrm{P}(\text{Aumente}) \cdot \mathrm{P}(\text{Aumente})$$

$$\mathrm{P}(\text{Aumente cuatro días consecutivos}) = \mathrm{P}(\text{Aumente})^4 = 0{,}5^4 = 0{,}0625$$

Ej. 8 ***La probabilidad de que el índice IBEX-35 aumente cualquier día de apertura de los Mercados es del 0,5. Si el comportamiento del índice en un día es independiente de su comportamiento en los días anteriores, se pide:***

b) Probabilidad de que el índice disminuya en al menos uno de cuatro días consecutivos observados.

$$\text{P(al menos un día disminuya)} = 1 - P(\text{todos los días aumente})$$

$$\text{P(al menos un día disminuya)} = 0{,}9375$$

$$\text{P(al menos un día disminuya)} = 1 - P(0{,}0625)$$


¡Muchas gracias!

Tema 4.2.

Variables aleatorias unidimensionales y bidimensionales

Práctica

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Ej. 1 *Sea una variable aleatoria ξ con función de densidad:*

$$f(x) = \begin{cases} 0 & si & x < 0 \\ k \cdot x^2 & si & 0 \leq x \leq 2 \\ 0 & si & x > 2 \end{cases}$$

a) Hallar k para que $f(x)$ sea verdaderamente una función de densidad.

$$\boxed{\int_{-\infty}^{\infty} f(x) \cdot dx = 1}$$

$$\int_{-\infty}^{0} 0 \cdot dx + \int_{0}^{2} k \cdot x^2 \cdot dx + \int_{2}^{\infty} 0 \cdot dx = 1$$

$$0 + k \int_{0}^{2} x^2 \cdot dx + 0 = 1$$

$$k \cdot \left[\frac{x^3}{3}\right]_0^2 = 1$$

$$k \cdot \left[\frac{2^3}{3} - \frac{0^3}{3}\right] = 1$$

$$\frac{8}{3} k = 1$$

$$k = \frac{3}{8}$$

Función de densidad completa:

$$f(x) = \begin{cases} 0 & si & x < 0 \\ \frac{3}{8} \cdot x^2 & si & 0 \leq x \leq 2 \\ 0 & si & x > 2 \end{cases}$$

Ej. 1 ***Sea una variable aleatoria ξ con función de densidad:***

$$f(x)=\begin{cases} 0 & si & x<0 \\ k\cdot x^2 & si & 0\le x\le 2 \\ 0 & si & x>2 \end{cases}$$

b) Hallar $P(1\le \xi \le 2)$.

Función de densidad completa:

$$f(x)=\begin{cases} 0 & si & x<0 \\ \frac{3}{8}\cdot x^2 & si & 0\le x\le 2 \\ 0 & si & x>2 \end{cases}$$

$$P(1\le \xi\le 2)=\int_1^2 \frac{3}{8}\cdot x^2\cdot dx=$$

$$\frac{3}{8}\cdot\int_1^2 x^2\cdot dx=\frac{3}{8}\cdot\left[\frac{x^3}{3}\right]_1^2=\frac{3}{8}\cdot\left[\frac{2^3}{3}-\frac{1^3}{3}\right]=\frac{3}{8}\cdot\left[\frac{7}{3}\right]=\frac{21}{24}=\frac{7}{8}$$

Ej. 2 ***Sea una variable aleatoria ξ con función de densidad:***

$$f(x)=\begin{cases} 0 & si \quad x<0 \\ x & si \quad 0\le x<1 \\ k-x & si \quad 1\le x<2 \\ 0 & si \quad x\ge 2 \end{cases}$$

a) Hallar k para que $f(x)$ sea verdaderamente una función de densidad.

$$\boxed{\int_{-\infty}^{\infty} f(x)\cdot dx = 1}$$

$$\int_{-\infty}^{0} 0\cdot dx+\int_{0}^{1} x\cdot dx+\int_{1}^{2} k-x\cdot dx+\int_{2}^{\infty} 0\cdot dx=1$$

$$0+\int_{0}^{1} x\cdot dx+\int_{1}^{2} k-x\cdot dx+0=1$$

$$\int_{0}^{1} x\cdot dx+\int_{1}^{2} k\cdot dx-\int_{1}^{2} x\cdot dx=1$$

$$\left[\frac{x^2}{2}\right]_0^1+k\cdot[x]_1^2-\left[\frac{x^2}{2}\right]_1^2=1$$

$$\left[\frac{1^2}{2}-\frac{0^2}{2}\right]+k\cdot[2-1]-\left[\frac{2^2}{2}-\frac{1^2}{2}\right]=1$$

$$\left[\frac{1}{2}\right]+k\cdot[1]-\left[\frac{3}{2}\right]=1$$

$$k=1-\frac{1}{2}+\frac{3}{2}=\frac{4}{2}=2$$

Función de densidad completa:

$$f(x)=\begin{cases} 0 & si \quad x<0 \\ x & si \quad 0\le x<1 \\ 2-x & si \quad 1\le x<2 \\ 0 & si \quad x\ge 2 \end{cases}$$

Ej. 2 ***Sea una variable aleatoria ξ con función de densidad:***

$$\boldsymbol{f(x)=\begin{cases}0 & si & x<0\\ x & si & 0\leq x<1\\ k-x & si & 1\leq x<2\\ 0 & si & x\geq 2\end{cases}}$$

b) Hallar $\boldsymbol{P(0\leq\xi\leq 1,5)}$.

Función de densidad completa:

$$f(x)=\begin{cases}0 & si & x<0\\ x & si & 0\leq x<1\\ 2-x & si & 1\leq x<2\\ 0 & si & x\geq 2\end{cases}$$

$$P(0\leq\xi\leq 1{,}5)=\int_0^1 x\cdot dx+\int_1^{1{,}5}2-x\cdot dx=$$

$$\int_0^1 x\cdot dx+2\int_1^{1{,}5}dx-\int_1^{1{,}5}x\cdot dx=\left[\frac{x^2}{2}\right]_0^1+2\cdot[x]_1^{1{,}5}-\left[\frac{x^2}{2}\right]_1^{1{,}5}=$$

$$\left[\frac{1^2}{2}-\frac{0^2}{2}\right]+2\cdot[1{,}5-1]-\left[\frac{1{,}5^2}{2}-\frac{1^2}{2}\right]=\left[\frac{1}{2}\right]+2\cdot[0{,}5]-\left[\frac{1{,}25}{2}\right]=0{,}875$$

Ej. 3

En una autovía hay un radar colocado en un puesto estratégico. La velocidad máxima autorizada en el punto es de 120 Km/h. El radar permite un margen de tolerancia sobre la velocidad máxima permitida del 10%. Si la velocidad a la que pasan los vehículos por ese punto queda representada mediante una variable aleatoria con función de densidad:

$$\boldsymbol{f(x) = \begin{cases} \dfrac{x}{10000} & si \quad 0 \leq x < 120 \\ \dfrac{200 - x}{10000} & si \quad 120 \leq x < 200 \\ 0 & en \quad resto \end{cases}}$$

a) ¿Qué probabilidad hay de que el conductor de un vehículo pierda puntos de su permiso de conducción al pasar por delante del radar?

$$P(\xi > 132) = 1 - P(\xi \leq 132) = 1 - \left[\int_0^{120} \frac{x}{10000} \cdot dx + \int_{120}^{132} \frac{200 - x}{10000} \cdot dx\right] =$$

$$1 - \left[\frac{1}{10000}\int_0^{120} x \cdot dx + \frac{1}{10000}\int_{120}^{132} 200 \cdot dx - \frac{1}{10000}\int_{120}^{132} x \cdot dx\right] =$$

$$1 - \frac{1}{10000} \cdot \left[\int_0^{120} x \cdot dx + \int_{120}^{132} 200 \cdot dx - \int_{120}^{132} x \cdot dx\right] =$$

$$1 - \frac{1}{10000} \cdot \left[\left[\frac{x^2}{2}\right]_0^{120} + 200 \cdot [x]_{120}^{132} - \left[\frac{x^2}{2}\right]_{120}^{132}\right] =$$

$$1 - \frac{1}{10000} \cdot \left[\left[\frac{120^2}{2} - \frac{0^2}{2}\right] + 200 \cdot [132 - 120] - \left[\frac{132^2}{2} - \frac{120^2}{2}\right]\right] = 1 - \frac{1}{10000} \cdot [8088] = 0{,}1912$$

Ej. 3

En una autovía hay un radar colocado en un puesto estratégico. La velocidad máxima autorizada en el punto es de 120 Km/h. El radar permite un margen de tolerancia sobre la velocidad máxima permitida del 10%. Si la velocidad a la que pasan los vehículos por ese punto queda representada mediante una variable aleatoria con función de densidad:

$$\boldsymbol{f(x) = \begin{cases} \dfrac{x}{10000} & si \quad 0 \leq x < 120 \\ \dfrac{200-x}{10000} & si \quad 120 \leq x < 200 \\ 0 & en \quad resto \end{cases}}$$

b) ¿Y si sabemos que circulaba a más de 120 Km./h.?

$$P(Multa \mid Circular\ a\ más\ de\ 120) = \frac{P(Multa\ (por\ circular\ a\ más\ de\ 130\ K\,m/h))}{P(Circular\ a\ más\ de\ 120)}$$

$$P(\xi > 120) = 1 - P(\xi < 120) = 1 - \int_0^{120} \frac{x}{10000} \cdot dx = 1 - \frac{1}{10000} \int_0^{120} x \cdot dx =$$

$$1 - \frac{1}{10000} \cdot \left[\left[\frac{x^2}{2} \right]_0^{120} \right] = 1 - \frac{1}{10000} \cdot \left[\frac{120^2}{2} - \frac{0^2}{2} \right] = 1 - \frac{1}{10000} \cdot \left[\frac{14400}{2} \right] = 1 - \frac{1}{10000} \cdot [7200] =$$

$$1 - 0{,}72 = 0{,}28$$

$$P(Multa \mid Circular\ a\ más\ de\ 120) = \frac{P(Multa\ (por\ circular\ a\ más\ de\ 130\ K\,m/h))}{P(Circular\ a\ más\ de\ 120)} = \frac{0{,}1912}{0{,}28} = 0{,}6829$$

Ej. 4

Una fábrica de bombillas produce dos tipos de bombillas de aspecto exterior semejante. El 70% de las bombillas son del tipo A, con una duración en años representada por la variable aleatoria ξ con función de densidad:

$$f(x) = \begin{cases} 0 & si \quad x \le 0 \\ e^{-x} & si \quad x > 0 \end{cases}$$

El 30% restante de bombillas son del tipo B, con duración en años representada por la variable η con función de densidad:

$$g(y) = \begin{cases} 0 & si \quad y \le 0 \\ 2 \cdot e^{-2y} & si \quad y > 0 \end{cases}$$

a) Probabilidad de que una bombilla de tipo A escogida al azar dure más de un año.

$$P(\xi > 1) = 1 - P(\xi \le 1) = 1 - \int_0^1 e^{-x} \cdot dx =$$

Si f(x) es un polinomio de grado uno: $\int_a^b e^{f(x)} \cdot dx = \frac{1}{f'(x)} \cdot \left[e^{f(x)}\right]_a^b$

$$= 1 - \frac{1}{-1}[e^{-x}]_0^1 = 1 + 1[e^{-1} - e^0] = 1 + \left[\frac{1}{e} - 1\right] = 1 + [0{,}3679 - 1] = 0{,}3679$$

Ej. 4

Una fábrica de bombillas produce dos tipos de bombillas de aspecto exterior semejante. El 70% de las bombillas son del tipo A, con una duración en años representada por la variable aleatoria ξ con función de densidad:

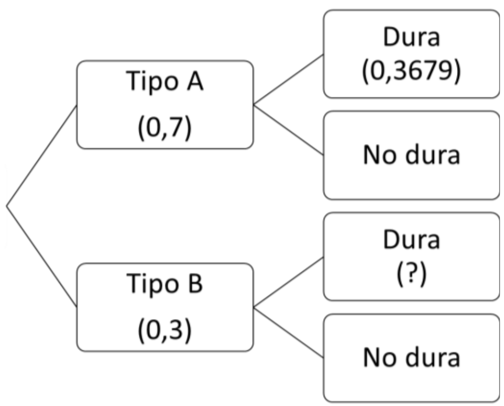
$$f(x) = \begin{cases} 0 & si \quad x \leq 0 \\ e^{-x} & si \quad x > 0 \end{cases}$$

El 30% restante de bombillas son del tipo B, con duración en años representada por la variable η con función de densidad:

$$g(y) = \begin{cases} 0 & si \quad y \leq 0 \\ 2 \cdot e^{-2y} & si \quad y > 0 \end{cases}$$

b) Si tomamos una bombilla al azar de entre toda la producción, ¿qué probabilidad tenemos de que dure más de un año?

$$P(Dure) = P(Dure \mid Tipo\ A) \cdot P(Tipo\ A) + P(Dure \mid Tipo\ B) \cdot P(Tipo\ B)$$



$$P(\eta > 1) = 1 - P(\eta \leq 1) = 1 - \int_0^1 2 \cdot e^{-2y} \cdot dy = 1 - 2 \cdot \int_0^1 e^{-2y} \cdot dy =$$

$$\boxed{\textit{Si f(x) es un polinomio de grado uno: } \int_a^b e^{f(x)} \cdot dx = \frac{1}{f'(x)} \cdot \left[e^{f(x)}\right]_a^b}$$

$$1 - \frac{2}{-2}\left[\mathrm{e}^{-2x}\right]_0^1 = 1 + 1\left[\mathrm{e}^{-1} - \mathrm{e}^0\right] = 1 + 1\left[\mathrm{e}^{-2} - \mathrm{e}^0\right]$$

$$= 1 + \left[\frac{1}{\mathrm{e}^2} - 1\right] = 1 + [0{,}1353 - 1] = 0{,}1353$$

$$P(Dure) = P(Dure \mid Tipo\ A) \cdot P(A) + P(Dure \mid Tipo\ B) \cdot P(B) = 0{,}3679 \cdot 0{,}7 + 0{,}1353 \cdot 0{,}3 = 0{,}2981$$

Ej. 4

Una fábrica de bombillas produce dos tipos de bombillas de aspecto exterior semejante. El 70% de las bombillas son del tipo A, con una duración en años representada por la variable aleatoria ξ con función de densidad:

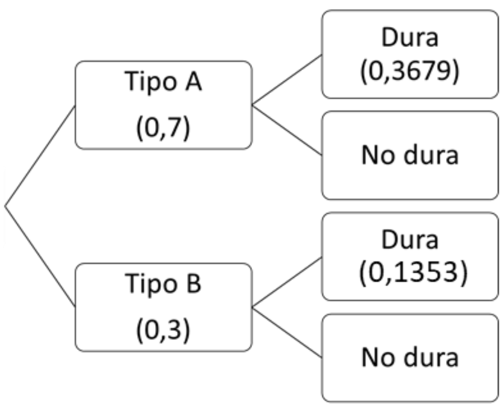
$$f(x) = \begin{cases} 0 & si \quad x \leq 0 \\ e^{-x} & si \quad x > 0 \end{cases}$$

El 30% restante de bombillas son del tipo B, con duración en años representada por la variable η con función de densidad:

$$g(y) = \begin{cases} 0 & si \quad y \leq 0 \\ 2 \cdot e^{-2y} & si \quad y > 0 \end{cases}$$

c) Y si una bombilla escogida al azar dura más de un año, ¿qué probabilidad hay de que sea del tipo A?

$$P(\,Tipo\ A|\ Dura) = \frac{P(Dure\mid Tipo\ A)\cdot P(Tipo\ A)}{PP(Dure\mid Tipo\ A)\cdot P(Tipo\ A) + P(Dure\mid Tipo\ B)\cdot P(Tipo\ B)}$$



$$= \frac{0{,}3679 \cdot 0{,}7}{0{,}3679 \cdot 0{,}7 + 0{,}1353 \cdot 0{,}3} = \frac{0{,}2575}{0{,}2981} = 0{,}8638$$

Ej. 5 *Dada la variable aleatoria bidimensional $(\xi ; \eta)$ con función de cuantía conjunta:*

$\xi \backslash \eta$	1	2	3	4	5
1	0,01	0,02	0,04	0,08	0,03
2	0,00	0,05	0,00	0,07	0,11
3	0,03	0,00	0,00	0,15	0,09
4	0,01	0,04	0,18	0,01	0,08

a) Hallar las funciones de cuantía marginales de ξ y η.

$\xi \backslash \eta$	1	2	3	4	5	$p_{i\bullet}$
1	0,01	0,02	0,04	0,08	0,03	0,18
2	0,00	0,05	0,00	0,07	0,11	0,23
3	0,03	0,00	0,00	0,15	0,09	0,27
4	0,01	0,04	0,18	0,01	0,08	0,32
$p_{\bullet j}$	0,05	0,11	0,22	0,31	0,31	1,00

Ej. 5 *Dada la variable aleatoria bidimensional (ξ ; η) con función de cuantía conjunta:*

ξ \ η	1	2	3	4	5
1	0,01	0,02	0,04	0,08	0,03
2	0,00	0,05	0,00	0,07	0,11
3	0,03	0,00	0,00	0,15	0,09
4	0,01	0,04	0,18	0,01	0,08

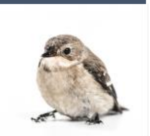
b) Calcular $P(\xi \leq 3/\eta = 2)$ *y* $P(\eta \leq 3/\xi = 2)$*.*

ξ \ η	1	2	3	4	5	$p_{i\bullet}$
1	0,01	0,02	0,04	0,08	0,03	0,18
2	0,00	0,05	0,00	0,07	0,11	0,23
3	0,03	0,00	0,00	0,15	0,09	0,27
4	0,01	0,04	0,18	0,01	0,08	0,32
$p_{\bullet j}$	0,05	0,11	0,22	0,31	0,31	1,00

$$\boxed{P(x_a \leq \xi \leq x_b | y_c \leq \eta \leq y_d) = \frac{\sum_{i=a}^{b} \sum_{j=c}^{d} p_{ij}}{\sum_{j=c}^{d} p_{\bullet j}}}$$

- $P(\xi \leq 3/\eta = 2) = P\left(\frac{\xi \leq 3 \cap \eta = 2}{\eta = 2}\right) = \frac{0{,}02+0{,}05+0{,}00}{0{,}11} = 0{,}6363$
- $P(\eta \leq 3/\xi = 2) = P\left(\frac{\eta \leq 3 \cap \xi = 2}{\xi = 2}\right) = \frac{0{,}00+0{,}05+0{,}00}{0{,}23} = 0{,}2173$

Ej. 5 *Dada la variable aleatoria bidimensional (ξ ; η) con función de cuantía conjunta:*



$\xi \backslash \eta$	1	2	3	4	5
1	0,01	0,02	0,04	0,08	0,03
2	0,00	0,05	0,00	0,07	0,11
3	0,03	0,00	0,00	0,15	0,09
4	0,01	0,04	0,18	0,01	0,08

c) ¿Son independientes ξ y η?

$\xi \backslash \eta$	1	2	3	4	5	$p_{i\bullet}$
1	0,01	0,02	0,04	0,08	0,03	0,18
2	0,00	0,05	0,00	0,07	0,11	0,23
3	0,03	0,00	0,00	0,15	0,09	0,27
4	0,01	0,04	0,18	0,01	0,08	0,32
$p_{\bullet j}$	0,05	0,11	0,22	0,31	0,31	1,00

$$\boxed{p_{ij} = p_{i\bullet} \cdot p_{\bullet j}}$$

$$\left.\begin{array}{l} \bullet \; p_{1\bullet} \cdot p_{\bullet 1} = 0{,}18 \cdot 0{,}05 = 0{,}009 \\ \bullet \; p_{11} = 0{,}01 \end{array}\right\} \; p_{1\bullet} \cdot p_{\bullet 1} \neq p_{11}$$

Ej. 6 *Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:*

$$f(x;y)=\begin{cases} k\cdot x\cdot y^2 & si \quad 0\le x<2, 0\le y<1 \\ 0 & en \quad resto \end{cases}$$

a) Hallar k para que $f(x;y)$ sea verdaderamente una función de densidad conjunta.

$$\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x;y)\cdot dy\cdot dx=1}$$

$$\int_0^2\int_0^1 k\cdot x\cdot y^2\cdot dy\cdot dx=1$$

$$\int_0^1 k\cdot x\cdot y^2\cdot dy=k\cdot x\cdot\int_0^1 y^2\cdot dy=k\cdot x\cdot\left(\left[\frac{y^3}{3}\right]_0^1\right)=k\cdot x\cdot\left(\left[\frac{1^3}{3}-\frac{0^3}{3}\right]\right)=k\cdot x\cdot\left(\frac{1}{3}\right)$$

$$\int_0^2 k\cdot x\cdot\left(\frac{1}{3}\right)dx=1$$

$$k\cdot\frac{1}{3}\cdot\left(\left[\frac{x^2}{2}\right]_0^2\right)=1$$

$$k\cdot\frac{1}{3}\cdot\left(\left[\frac{2^2}{2}-\frac{0^2}{2}\right]\right)=1$$

$$k=\frac{3}{2}$$

Ej. 6 *Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:*

$$f(x;y)=\begin{cases} k\cdot x\cdot y^2 & si \quad 0\le x<2,0\le y<1 \\ 0 & en \quad resto \end{cases}$$

b) Hallar las funciones de densidad marginales de ξ y η.

$$f(x;y)=\begin{cases} \frac{3}{2}\cdot x\cdot y^2 & si \quad 0\le x<2,0\le y<1 \\ 0 & en \quad resto \end{cases}$$

- ***Función de densidad marginal de ξ:***

$$f_\xi(x)=\int_{-\infty}^{\infty} f(x;y)\cdot dy=\int_0^1 \frac{3}{2}\cdot x\cdot y^2\cdot dy=\frac{3}{2}\cdot x\cdot\int_0^1 y^2\cdot dy=\frac{3}{2}\cdot x\cdot\left(\left[\frac{y^3}{3}\right]_0^1\right)$$
$$=\frac{3}{2}\cdot x\cdot\left(\left[\frac{1^3}{3}-\frac{0^3}{3}\right]\right)=\frac{3}{2}\cdot x\cdot\left(\left[\frac{1}{3}\right]\right)=\frac{1}{2}\cdot x$$

- ***Función de densidad marginal de η:***

$$f_\eta(x)=\int_{-\infty}^{\infty} f(x;y)\cdot dx=\int_0^2 \frac{3}{2}\cdot x\cdot y^2\cdot dx=\frac{3}{2}\cdot y^2\cdot\int_0^2 x\cdot dy=\frac{3}{2}\cdot y^2\cdot\left(\left[\frac{x^2}{2}\right]_0^2\right)$$
$$=\frac{3}{2}\cdot y^2\cdot\left(\left[\frac{2^2}{2}-\frac{0^2}{2}\right]\right)=\frac{3}{2}\cdot y^2\cdot([2])=3\cdot y^2$$

Ej. 6 *Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:*

$$\boldsymbol{f(x;y) = \begin{cases} k \cdot x \cdot y^2 & si \quad 0 \le x < 2, 0 \le y < 1 \\ 0 & en \quad resto \end{cases}}$$

c) ¿Son independientes ξ y η ?

$$f(x;y) = \begin{cases} \frac{3}{2} \cdot x \cdot y^2 & si \quad 0 \le x < 2, 0 \le y < 1 \\ 0 & en \quad resto \end{cases}$$

- ***Función de densidad marginal de ξ:***

$$f_\xi(x) = \frac{1}{2} \cdot x$$

- ***Función de densidad marginal de η:***

$$f_\eta(x) = 3 \cdot y^2$$

$$\boxed{f(x;y) = f_\xi(x) \cdot f_\eta(y)}$$

$$f_\xi(x) \cdot f_\eta(y) = \frac{1}{2} \cdot x \cdot 3 \cdot y^2 = \frac{3}{2} \cdot x \cdot y^2 \ \rightarrow son\ independientes$$

Ej. 7 *Dada la variable aleatoria bidimensional $(\xi ; \eta)$ con función de densidad conjunta:*

$$f(x;y) = \begin{cases} k \cdot x \cdot y^2 & si \quad 0 \leq y < x < 2 \\ 0 & en \quad resto \end{cases}$$

a) Hallar k para que $f(x;y)$ sea verdaderamente una función de densidad conjunta.

$$\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x;y) \cdot dy \cdot dx = 1}$$

$$\int_0^2 \int_0^x k \cdot x \cdot y^2 \cdot dy \cdot dx = 1$$

$$\int_0^x k \cdot x \cdot y^2 \cdot dy = k \cdot x \cdot \int_0^x y^2 \cdot dy = k \cdot x \cdot \left(\left[\frac{y^3}{3}\right]_0^x\right) = k \cdot x \cdot \left(\left[\frac{x^3}{3} - \frac{0^3}{3}\right]\right) = k \cdot x \cdot \left(\frac{x^3}{3}\right) = k \cdot \frac{x^4}{3}$$

$$\int_0^2 k \cdot \frac{x^4}{3} \cdot dx = 1$$

$$k \cdot \frac{1}{3} \cdot \int_0^2 x^4 \cdot dx = 1$$

$$k \cdot \frac{1}{3} \cdot \left(\left[\frac{x^5}{5}\right]_0^2\right) = 1$$

$$k \cdot \frac{1}{3} \cdot \left(\left[\frac{2^5}{5} - \frac{0^5}{5}\right]\right) = 1$$

$$k = \frac{15}{32}$$

Ej. 7 *Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:*

$$f(x;y)=\begin{cases} k\cdot x\cdot y^2 & si \quad 0\le y<x<2 \\ 0 & en \quad resto \end{cases}$$

b) Hallar las funciones de densidad marginales de ξ y η.

$$f(x;y)=\begin{cases} \frac{15}{32}\cdot x\cdot y^2 & si \quad 0\le y<x<2 \\ 0 & en \quad resto \end{cases}$$

- ***Función de densidad marginal de ξ:***

$$f_\xi(x)=\int_{-\infty}^{\infty} f(x;y)\cdot dy=\int_0^x \frac{15}{32}\cdot x\cdot y^2\cdot dy=\frac{15}{32}\cdot x\cdot\int_0^x y^2\cdot dy=\frac{15}{32}\cdot x\cdot\left(\left[\frac{y^3}{3}\right]_0^x\right)$$

$$=\frac{15}{32}\cdot x\cdot\left(\left[\frac{x^3}{3}-\frac{0^3}{3}\right]\right)=\frac{15}{32}\cdot x\cdot\left(\left[\frac{x^3}{3}\right]\right)=\frac{15}{32}\cdot\frac{x^4}{3}=\frac{15\cdot x^4}{96}=\frac{5\cdot x^4}{32}$$

- ***Función de densidad marginal de η:***

$$f_\eta(x)=\int_{-\infty}^{\infty} f(x;y)\cdot dx=\int_y^2 \frac{15}{32}\cdot x\cdot y^2\cdot dx=\frac{15}{32}\cdot y^2\cdot\int_y^2 x\cdot dy=\frac{15}{32}\cdot y^2\cdot\left(\left[\frac{x^2}{2}\right]_y^2\right)$$

$$=\frac{15}{32}\cdot y^2\cdot\left(\left[\frac{2^2}{2}-\frac{y^2}{2}\right]\right)=\frac{15}{32}\cdot y^2\cdot\left(2-\frac{y^2}{2}\right)=\frac{15}{32}\cdot\left(2\cdot y^2-\frac{y^4}{2}\right)$$

Ej. 7 *Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:*

$$\boldsymbol{f(x;y) = \begin{cases} k \cdot x \cdot y^2 & si \quad 0 \le y < x < 2 \\ 0 & en \quad resto \end{cases}}$$

c) ¿Son independientes ξ y η ?

$$f(x;y) = \begin{cases} \dfrac{15}{32} \cdot x \cdot y^2 & si \quad 0 \le y < x < 2 \\ 0 & en \quad resto \end{cases}$$

- ***Función de densidad marginal de ξ:***

$$f_\xi(x) = \frac{5 \cdot x^4}{32}$$

- ***Función de densidad marginal de η:***

$$f_\eta(x) = \frac{15}{32} \cdot \left(2 \cdot y^2 - \frac{y^4}{2}\right)$$

$$\boxed{f(x;y) = f_\xi(x) \cdot f_\eta(y)}$$

$$f_\xi(x) \cdot f_\eta(y) = \frac{5 \cdot x^4}{32} \cdot \frac{15}{32} \cdot \left(2 \cdot y^2 - \frac{y^4}{2}\right) \rightarrow NO\ son\ independientes$$

Ej. 7 ***Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:***

$$\boldsymbol{f(x;y) = \begin{cases} k \cdot x \cdot y^2 & si \quad 0 \le y < x < 2 \\ 0 & en \quad resto \end{cases}}$$

d) Calcular $\boldsymbol{P(\eta \ge 1)}$.

$$f(x;y) = \begin{cases} \dfrac{15}{32} \cdot x \cdot y^2 & si \quad 0 \le y < x < 2 \\ 0 & en \quad resto \end{cases}$$

$$\int_1^2 f_\eta(y) \cdot dy = \int_1^2 \frac{15}{32} \cdot \left(2 \cdot y^2 - \frac{y^4}{2}\right) \cdot dy = \frac{15}{32} \cdot \int_1^2 2 \cdot y^2 - \frac{y^4}{2} \cdot dy = \frac{15}{32} \cdot \left(\int_1^2 2 \cdot y^2 \cdot dy - \int_1^2 \frac{y^4}{2} \cdot dy\right) =$$

$$\frac{15}{32} \cdot \left(2 \cdot \int_1^2 y^2 \cdot dy - \frac{1}{2} \int_1^2 y^4 \cdot dy\right) = \frac{15}{32} \cdot \left(2 \cdot \left[\frac{y^3}{3}\right]_1^2 - \frac{1}{2} \cdot \left[\frac{y^5}{5}\right]_1^2\right) = \frac{15}{32} \cdot \left(2 \cdot \left[\frac{2^3}{3} - \frac{1^3}{3}\right] - \frac{1}{2} \cdot \left[\frac{2^5}{5} - \frac{1^5}{5}\right]\right) =$$

$$\frac{15}{32} \cdot \left(2 \cdot \left[\frac{7}{3}\right] - \frac{1}{2} \cdot \left[\frac{31}{5}\right]\right) = \frac{15}{32} \cdot \left(\frac{14}{3} - \frac{31}{10}\right) = \frac{15}{32} \cdot \left(\frac{140}{30} - \frac{93}{30}\right) = \frac{705}{960} = 0{,}7343$$

Ej. 8 ***Dada la variable aleatoria bidimensional $(\xi ; \eta)$ con función de densidad conjunta:***

$$f(x;y)=\begin{cases} k\cdot x\cdot y^2 & si \quad 0\le x\le y\le 2 \\ 0 & en \quad resto \end{cases}$$

a) Hallar k para que $f(x;y)$ sea verdaderamente una función de densidad conjunta.

$$\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x;y)\cdot dy\cdot dx=1}$$

$$\int_0^2\int_x^2 k\cdot x\cdot y^2\cdot dy\cdot dx=1$$

$$\int_x^2 k\cdot x\cdot y^2\cdot dy=k\cdot x\cdot\int_x^2 y^2\cdot dy=k\cdot x\cdot\left(\left[\frac{y^3}{3}\right]_x^2\right)=k\cdot x\cdot\left(\left[\frac{2^3}{3}-\frac{x^3}{3}\right]\right)=k\cdot x\cdot\left(\left[\frac{8-x^3}{3}\right]\right)=\frac{k}{3}\cdot(8\cdot x-x^4)$$

$$\int_0^2\frac{k}{3}\cdot(8\cdot x-x^4)\cdot dx=1$$

$$\frac{k}{3}\cdot\int_0^2 8\cdot x-x^4\cdot dx=1$$

$$\frac{k}{3}\cdot\left(\int_0^2 8\cdot x\cdot dx-\int_0^2 x^4\cdot dx\right)=1$$

$$\frac{k}{3}\cdot\left(8\cdot\left[\frac{x^2}{2}\right]_0^2-\left[\frac{x^5}{5}\right]_0^2\right)=1$$

$$\frac{k}{3}\cdot\left(8\cdot\left[\frac{2^2}{2}-\frac{0^2}{2}\right]-\left[\frac{2^5}{5}-\frac{0^5}{5}\right]\right)=1$$

$$\frac{k}{3}\cdot\left(8\cdot\left[\frac{4}{2}\right]-\left[\frac{32}{5}\right]\right)=1$$

$$\frac{k}{3}\cdot\left(16-\frac{32}{5}\right)=1$$

$$\frac{k}{3}\cdot\left(\frac{80}{5}-\frac{32}{5}\right)=1$$

$$\frac{48\cdot k}{15}=1$$

$$k=\frac{15}{48}=\frac{5}{16}$$

Ej. 9 ***Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:***

$$f(x;y)=\begin{cases} k\cdot\left(x^2+y^2\right) & si \quad 0<y\le x\le 1 \\ 0 & en \quad resto \end{cases}$$

a) Hallar k para que $f(x;y)$ sea verdaderamente una función de densidad conjunta.

$$\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x;y)\cdot dy\cdot dx=1}$$

$$\int_0^1\int_0^x k\cdot\left(x^2+y^2\right)\cdot dy\cdot dx=1$$

$$\int_0^x k\cdot\left(x^2+y^2\right)\cdot dy=k\cdot\int_0^x x^2+y^2\cdot dy=k\cdot\left(\int_0^x x^2\cdot dy+\int_0^x y^2\cdot dy\right)=k\cdot\left(x^2\cdot\int_0^x dy+\int_0^x y^2\cdot dy\right)$$

$$=k\cdot\left(x^2\cdot[y]_0^x+\left[\frac{y^3}{3}\right]_0^x\right)=k\cdot\left(x^2\cdot[x-0]+\left[\frac{x^3}{3}-\frac{0^2}{3}\right]\right)=k\cdot\left(x^2\cdot[x]+\left[\frac{x^3}{3}\right]\right)=k\cdot\left(x^3+\frac{x^3}{3}\right)=k\cdot\left(\frac{4x^3}{3}\right)$$

$$\int_0^1 k\cdot\left(\frac{4x^3}{3}\right)\cdot dx=1$$

$$k\cdot\frac{4}{3}\cdot\int_0^1 x^3\cdot dx=1$$

$$k\cdot\frac{4}{3}\cdot\left(\left[\frac{x^4}{4}\right]_0^1\right)=1$$

$$k\cdot\frac{4}{3}\cdot\left(\left[\frac{1^4}{4}-\frac{0^4}{4}\right]\right)=1$$

$$k\cdot\frac{4}{3}\cdot\left(\frac{1}{4}\right)=1$$

$$k\cdot\frac{1}{3}=1$$

$$k=3$$

Ej. 9 ***Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:***

$$\boldsymbol{f(x;y) = \begin{cases} k \cdot \left(x^2 + y^2\right) & si \quad 0 < y \le x \le 1 \\ 0 & en \quad resto \end{cases}}$$

b) Hallar las funciones de densidad marginales de ξ y η.

$$f(x;y) = \begin{cases} 3 \cdot \left(x^2 + y^2\right) & si \quad 0 < y \le x \le 1 \\ 0 & en \quad resto \end{cases}$$

- ***Función de densidad marginal de ξ:***

$$f_\xi(x) = \int_{-\infty}^{\infty} f(x;y) \cdot dy = \int_0^x 3 \cdot \left(x^2 + y^2\right) \cdot dy = 3 \cdot \left(\int_0^x x^2 \cdot dy + \int_0^x y^2 \cdot dy\right) = 3 \cdot \left(x^2 \cdot \int_0^x dy + \int_0^x y^2 \cdot dy\right)$$

$$= 3 \cdot \left(x^2 \cdot [y]_0^x + \left[\frac{y^3}{3}\right]_0^x\right) = 3 \cdot \left(x^2 \cdot [x - 0] + \left[\frac{x^3}{3} - \frac{0^2}{3}\right]\right) = 3 \cdot \left(x^3 + \frac{x^3}{3}\right) = 3 \cdot \left(\frac{4x^3}{3}\right) = 4x^3$$

- ***Función de densidad marginal de η:***

$$f_\eta(x) = \int_{-\infty}^{\infty} f(x;y) \cdot dx = \int_y^1 3 \cdot \left(x^2 + y^2\right) \cdot dx = 3 \cdot \int_y^1 x^2 + y^2 \cdot dx = 3 \cdot \left(\int_y^1 x^2 \cdot dx + \int_y^1 y^2 \cdot dx\right)$$

$$= 3 \cdot \left(\int_y^1 x^2 \cdot dx + y^2 \cdot \int_y^1 dx\right) = 3 \cdot \left(\left[\frac{x^3}{3}\right]_y^1 + y^2 \cdot [x]_y^1\right) = 3 \cdot \left(\left[\frac{1^3}{3} - \frac{y^3}{3}\right] + y^2 \cdot [1 - y]\right)$$

$$= 3 \cdot \left(\left[\frac{1}{3} - \frac{y^3}{3}\right] + [y^2 - y^3]\right) = \frac{3}{3} - \frac{3 \cdot y^3}{3} + 3 \cdot y^2 - 3 \cdot y^3 = 1 - y^3 + 3 \cdot y^2 - 3 \cdot y^3 = -4 \cdot y^3 + 3 \cdot y^2 + 1$$

Ej. 9 ***Dada la variable aleatoria bidimensional (ξ ; η) con función de densidad conjunta:***

$$\boldsymbol{f(x;y) = \begin{cases} k\cdot\left(x^2+y^2\right) & si \quad 0 < y \le x \le 1 \\ 0 & en \quad resto \end{cases}}$$

c) ¿Son independientes ξ y η ?

$$f(x;y) = \begin{cases} 3\cdot\left(x^2+y^2\right) & si \quad 0 < y \le x \le 1 \\ 0 & en \quad resto \end{cases}$$

- ***Función de densidad marginal de ξ:***

$$f_\xi(x) = 4x^3$$

- ***Función de densidad marginal de η:***

$$f_\eta(x) = -4\cdot y^3 + 3\cdot y^2 + 1$$

$$\boxed{f(x;y) = f_\xi(x)\cdot f_\eta(y)}$$

$$f_\xi(x)\cdot f_\eta(y) = 4x^3\cdot\left(-4\cdot y^3 + 3\cdot y^2 + 1\right)$$ → *NO son independientes*

¡Muchas gracias!

Tema 4.3.

Características de las variables aleatorias

Práctica

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Ej. 1 *Sea una variable aleatoria ξ con función de cuantía:*

x_i	p_i
-2	1/4
-1	1/8
0	1/4
1	1/8
2	1/4

y dada la transformación $\eta = 2 \cdot \xi - 2$;

a) Hallar la esperanza matemática de ξ.

$$\boldsymbol{E(\xi) = \mu = \alpha_1 = \sum_i x_i p_i =}$$

$$-2 \cdot \frac{1}{4} - 1 \cdot \frac{1}{8} + 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{4} = 0$$

Ej. 1 *Sea una variable aleatoria ξ con función de cuantía:*

x_i	p_i
-2	1/4
-1	1/8
0	1/4
1	1/8
2	1/4

y dada la transformación $\eta = 2 \cdot \xi - 2$;

b) Hallar la varianza de ξ.

$$\boldsymbol{Var(\xi) = \mu_2 = \alpha_2 - \alpha_1^2}$$

$$\alpha_1 = 0$$

$$\alpha_2 = \sum_i {x_i}^2 p_i = (-2)^2 \cdot \frac{1}{4} + (-1)^2 \cdot \frac{1}{8} + 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{8} + 2^2 \cdot \frac{1}{4} = \frac{9}{4}$$

$$Var(\xi) = \alpha_2 - \alpha_1^2 = \frac{9}{4} - (0)^2 = \frac{9}{4}$$

Ej. 1 *Sea una variable aleatoria ξ con función de cuantía:*

x_i	p_i
-2	1/4
-1	1/8
0	1/4
1	1/8
2	1/4

y dada la transformación $\eta = 2 \cdot \xi - 2$;

c) Hallar la esperanza matemática de η.

$$E(\xi) = 0$$

$$Var(\xi) = \frac{9}{4}$$

$$\boxed{\boldsymbol{E(A + B \cdot \xi) = E(A) + E(B) \cdot E(\xi) = A + B \cdot E(\xi)}}$$

$$E(\eta) = E(2 \cdot \xi - 2) = E(2 \cdot \xi) - E(2) =$$

$$E(2) \cdot E(\xi) - E(2) = 2 \cdot E(\xi) - 2 = 2 \cdot 0 - 2 = -2$$

Ej. 1 *Sea una variable aleatoria ξ con función de cuantía:*

x_i	p_i
-2	1/4
-1	1/8
0	1/4
1	1/8
2	1/4

y dada la transformación $\eta = 2 \cdot \xi - 2$;

d) Hallar la varianza de η.

$$E(\xi) = 0$$

$$Var(\xi) = \frac{9}{4}$$

$$\boxed{\boldsymbol{Var(A + B \cdot \xi) = Var(A) + Var(B) \cdot Var(\xi) = 0 + B^2 \cdot Var(\xi)}}$$

$$Var(\eta) = Var(2 \cdot \xi - 2) = Var(2 \cdot \xi) - Var(2) = Var(2 \cdot \xi) - 0 = 2^2 \cdot Var(\xi) = 2^2 \cdot \frac{9}{4} = 9$$

Ej. 2 *Sea una variable aleatoria ξ con función de densidad:*

$$f(x) = \begin{cases} 0 & si & x < 0 \\ \frac{3}{8} \cdot x^2 & si & 0 \le x \le 2 \\ 0 & si & x > 2 \end{cases}$$

a) Hallar la esperanza matemática de ξ.

$$E(\xi) = \alpha_1 = \int_{-\infty}^{+\infty} x^1 \cdot f(x) \cdot dx =$$

$$\int_0^2 x \cdot \frac{3}{8} \cdot x^2 \cdot dx = \frac{3}{8} \cdot \int_0^2 x^3 \cdot dx =$$

$$\frac{3}{8} \cdot \left[\frac{x^4}{4}\right]_0^2 = \frac{3}{8} \cdot \left[\frac{2^4}{4} - \frac{0^4}{4}\right] = \frac{3}{8} \cdot \left[\frac{16}{4}\right] = \frac{3}{2}$$

Ej. 2 *Sea una variable aleatoria ξ con función de densidad:*

$$f(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{3}{8} \cdot x^2 & si \quad 0 \leq x \leq 2 \\ 0 & si \quad x > 2 \end{cases}$$

b) Hallar la varianza de ξ.

$$Var(\xi) = \mu_2 = \alpha_2 - \alpha_1^2$$

$$\alpha_1 = \frac{3}{2}$$

$$\alpha_2 = \int_{-\infty}^{+\infty} x^2 \cdot f(x) \cdot dx = \int_0^2 x^2 \cdot \frac{3}{8} \cdot x^2 \cdot dx = \frac{3}{8} \cdot \int_0^2 x^4 \cdot dx =$$

$$\frac{3}{8} \cdot \left[\frac{x^5}{5}\right]_0^2 = \frac{3}{8} \cdot \left[\frac{2^5}{5} - \frac{0^5}{5}\right] = \frac{3}{8} \cdot \left[\frac{32}{5}\right] = \frac{96}{40} = \frac{12}{5}$$

$$Var(\xi) = \frac{12}{5} - \left(\frac{3}{2}\right)^2 = \frac{12}{5} - \frac{9}{4} = \frac{3}{20} = 0{,}15$$

Ej. 2 *Sea una variable aleatoria ξ con función de densidad:*

$$f(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{3}{8} \cdot x^2 & si \quad 0 \le x \le 2 \\ 0 & si \quad x > 2 \end{cases}$$

c) Hallar la esperanza matemática y la varianza de la transformación $\eta = 2 \cdot \xi - 3$.

$$E(\xi) = \frac{3}{2}$$

$$Var(\xi) = 0{,}15$$

- $\boldsymbol{E(\eta)}$: $\boxed{\boldsymbol{E(A + B \cdot \xi) = E(A) + E(B) \cdot E(\xi) = A + B \cdot E(\xi)}}$

$$E(\eta) = E(2 \cdot \xi - 3) = E(2 \cdot \xi) - E(3) = 2 \cdot E(\xi) - E(3) = 2 \cdot E(\xi) - 3 = 2 \cdot \frac{3}{2} - 3 = 0$$

- $\boldsymbol{Var(\eta)}$: $\boxed{\boldsymbol{Var(A + B \cdot \xi) = Var(A) + Var(B) \cdot Var(\xi) = 0 + B^2 \cdot Var(\xi)}}$

$$Var(\eta) = Var(2 \cdot \xi - 3) = Var(2 \cdot \xi) - Var(3) = 2^2 \cdot \mathrm{Var}(\xi) - 0 = 4 \cdot \frac{3}{20} = \frac{12}{20} = 0{,}6$$

Ej. 3 *Sea una variable aleatoria ξ con función de densidad:*

$$f(x)=\begin{cases}\dfrac{1}{2} & si \quad 0\le x<1\\ x-2 & si \quad 2\le x<3\\ 0 & en \quad resto\end{cases}$$

a) Hallar la esperanza matemática de ξ.

$$\boldsymbol{E(\xi)=\alpha_1=\int_{-\infty}^{+\infty} x^1\cdot f(x)\cdot dx=}$$

$$\int_0^1 x\cdot\frac{1}{2}dx+\int_2^3 x\cdot(x-2)\cdot dx=\frac{1}{2}\int_0^1 x\cdot dx+\int_2^3 x^2-2x\cdot dx$$

$$=\frac{1}{2}\int_0^1 x\,dx+\int_2^3 x^2dx-2\cdot\int_2^3 x\,dx=\frac{1}{2}\cdot\left[\frac{x^2}{2}\right]_0^1+\left[\frac{x^3}{3}\right]_2^3-2\cdot\left[\frac{x^2}{2}\right]_2^3$$

$$=\frac{1}{2}\cdot\left[\frac{1^2}{2}-\frac{0^2}{2}\right]+\left[\frac{3^3}{3}-\frac{2^3}{3}\right]-2\cdot\left[\frac{3^2}{2}-\frac{2^2}{2}\right]=\frac{1}{2}\cdot\left[\frac{1}{2}\right]+\left[\frac{27}{3}-\frac{8}{3}\right]-2\cdot\left[\frac{9}{2}-\frac{4}{2}\right]$$

$$\frac{1}{2}\cdot\left[\frac{1}{2}\right]+\left[\frac{19}{3}\right]-2\cdot\left[\frac{5}{2}\right]=\frac{1}{4}+\frac{19}{3}-5=\frac{19}{12}$$

Ej. 3 *Sea una variable aleatoria ξ con función de densidad:*

$$f(x) = \begin{cases} \frac{1}{2} & si \quad 0 \le x < 1 \\ x-2 & si \quad 2 \le x < 3 \\ 0 & en \quad resto \end{cases}$$

b) Hallar la varianza de ξ.

$$Var(\xi) = \mu_2 = \alpha_2 - \alpha_1^2$$

$$\alpha_1 = \frac{19}{12}$$

$$\alpha_2 = \int_{-\infty}^{+\infty} x^2 \cdot f(x) \cdot dx = \int_0^1 x^2 \cdot \frac{1}{2} dx + \int_2^3 x^2 \cdot (x-2) \cdot dx = \frac{1}{2}\int_0^1 x^2 \cdot dx + \int_2^3 x^3 - 2x^2 \cdot dx$$

$$= \frac{1}{2}\int_0^1 x^2\, dx + \int_2^3 x^3 dx - 2 \cdot \int_2^3 x^2\, dx = \frac{1}{2} \cdot \left[\frac{x^3}{3}\right]_0^1 + \left[\frac{x^4}{4}\right]_2^3 - 2 \cdot \left[\frac{x^3}{3}\right]_2^3$$

$$= \frac{1}{2} \cdot \left[\frac{1^3}{3} - \frac{0^3}{3}\right] + \left[\frac{3^4}{4} - \frac{2^4}{4}\right] - 2 \cdot \left[\frac{3^3}{3} - \frac{2^3}{3}\right] = \frac{1}{2} \cdot \left[\frac{1}{3}\right] + \left[\frac{65}{4}\right] - 2 \cdot \left[\frac{19}{3}\right] = \frac{1}{6} + \frac{65}{4} - \frac{38}{3} = \frac{45}{12}$$

$$Var(\xi) = \frac{45}{12} - \left(\frac{19}{12}\right)^2 = \frac{45}{12} - \frac{361}{144} = \frac{179}{144}$$

Ej. 3 *Sea una variable aleatoria ξ con función de densidad:*

$$f(x)=\begin{cases} \frac{1}{2} & si \quad 0\le x<1 \\ x-2 & si \quad 2\le x<3 \\ 0 & en \quad resto \end{cases}$$

c) Hallar la esperanza matemática y la varianza de la transformación $\eta = 3\cdot\xi - 4$.

$$E(\xi)=\frac{19}{12}$$

$$Var(\xi)=\frac{179}{144}$$

- $\boldsymbol{E(\eta)}$: $\boxed{\boldsymbol{E(A+B\cdot\xi)=E(A)+E(B)\cdot E(\xi)=A+B\cdot E(\xi)}}$

$$E(\eta)=E(3\cdot\xi-4)=E(3\cdot\xi)-E(4)=E(3)\cdot E(\xi)-E(4)=3\cdot E(\xi)-4=3\cdot\frac{19}{12}-4=\frac{3}{4}$$

- $\boldsymbol{Var(\eta)}$: $\boxed{\boldsymbol{Var(A+B\cdot\xi)=Var(A)+Var(B)\cdot Var(\xi)=0+B^2\cdot Var(\xi)}}$

$$Var(\eta)=Var(3\cdot\xi-4)=Var(3\cdot\xi)-Var(4)=3^2\cdot \mathrm{Var}(\xi)-0=9\cdot\frac{179}{144}=\frac{1611}{144}=\frac{179}{16}$$

Ej. 4 *Dada la variable aleatoria bidimensional (ξ, η) con función de densidad conjunta:*

$$f(x;y) = \begin{cases} 4 \cdot x \cdot y & si \quad 0 < x \le 1, 0 < y \le 1 \\ 0 & en \quad resto \end{cases}$$

Calcular la covarianza de ambas variables, $cov(\xi;\eta)$.

$$\boxed{\mu_{11} = Cov(\xi;\eta) = \sigma_{\xi\eta} = \alpha_{11} - \alpha_{10} \cdot \alpha_{01}}$$

- $\boldsymbol{\alpha_{10} = E(\xi)}$

$$f_\xi(x) = \int_0^1 4 \cdot x \cdot y \cdot dy = 4 \cdot x \cdot \int_0^1 y \cdot dy = 4 \cdot x \cdot \left[\frac{y^2}{2}\right]_0^1 = 4 \cdot x \cdot \left[\frac{1^2}{2} - \frac{0^2}{2}\right] = 4 \cdot x \cdot \left[\frac{1}{2}\right] = \frac{4 \cdot x}{2} = 2x$$

$$E(\xi) = \alpha_{10} = \int_0^1 x^1 \cdot 2x \cdot dx = 2 \cdot \int_0^1 x^2 \cdot dx = 2 \cdot \left[\frac{x^3}{3}\right]_0^1 = 2 \cdot \left[\frac{1^3}{3} - \frac{0^3}{3}\right] = 2 \cdot \left[\frac{1}{3}\right] = \frac{2}{3}$$

- $\boldsymbol{\alpha_{01} = E(\eta)}$

$$f_\eta(y) = \int_0^1 4 \cdot x \cdot y \cdot dx = 4 \cdot y \cdot \int_0^1 x \cdot dx = 4 \cdot y \cdot \left[\frac{x^2}{2}\right]_0^1 = 4 \cdot y \cdot \left[\frac{1^2}{2} - \frac{0^2}{2}\right] = 4 \cdot y \cdot \left[\frac{1}{2}\right] = \frac{4 \cdot y}{2} = 2y$$

$$E(\eta) = \alpha_{01} = \int_0^1 y^1 \cdot 2y \cdot dy = 2 \cdot \int_0^1 y^2 \cdot dy = 2 \cdot \left[\frac{y^3}{3}\right]_0^1 = 2 \cdot \left[\frac{1^3}{3} - \frac{0^3}{3}\right] = 2 \cdot \left[\frac{1}{3}\right] = \frac{2}{3}$$

Ej. 4 *Dada la variable aleatoria bidimensional (ξ,η) con función de densidad conjunta:*

$$f(x;y)=\begin{cases}4\cdot x\cdot y & si \quad 0<x\leq 1, 0<y\leq 1\\ 0 & en \quad resto\end{cases}$$

Calcular la covarianza de ambas variables, $cov(\xi;\eta)$.

$$\boxed{\mu_{11}=Cov(\xi;\eta)=\sigma_{\xi\eta}=\alpha_{11}-\alpha_{10}\cdot\alpha_{01}}$$

- $\boldsymbol{\alpha_{10}=E(\xi)=\frac{2}{3}}$
- $\boldsymbol{\alpha_{01}=E(\eta)=\frac{2}{3}}$
- $\boldsymbol{\alpha_{11}}=\int_0^1\int_0^1 x^1\cdot y^1\cdot(4\cdot x\cdot y)\cdot dy\cdot dx=\int_0^1\int_0^1 4\cdot x^2\cdot y^2\cdot dy\cdot dx$

$$\int_0^1 4\cdot x^2\cdot y^2\cdot dy=4\cdot x^2\cdot\int_0^1 y^2\cdot dy=4\cdot x^2\cdot\left[\frac{y^3}{3}\right]_0^1=4\cdot x^2\cdot\left[\frac{1^3}{3}-\frac{0^3}{3}\right]=4\cdot x^2\cdot\left[\frac{1}{3}\right]=\frac{4\cdot x^2}{3}$$

$$\int_0^1\frac{4\cdot x^2}{3}\cdot dx=\frac{4}{3}\int_0^1 x^2\cdot dx=\frac{4}{3}\cdot\left[\frac{x^3}{3}\right]_0^1=\frac{4}{3}\cdot\left[\frac{1^3}{3}-\frac{0^3}{3}\right]=\frac{4}{3}\cdot\left[\frac{1}{3}\right]=\frac{4}{9}$$

$$Cov(\xi;\eta)=\sigma_{\xi\eta}=\alpha_{11}-\alpha_{10}\cdot\alpha_{01}=\frac{4}{9}-\frac{2}{3}\cdot\frac{2}{3}=0$$


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 5.1.

Modelos de distribución de probabilidad discretos univariantes Práctica

Ej. 1

El gerente de un restaurante que sólo da servicio bajo reserva previa sabe por experiencia que el 20% de los clientes que hacen reserva no acuden al restaurante. Si el restaurante acepta 12 reservas; pero sólo dispone de 10 mesas, ¿qué probabilidad hay de que todos los clientes que acuden al restaurante tengan mesa?

- **Tipo de experimento:**
  - Binomial B(n;p).
- **Variable que modeliza:**
  - ξ= “Número de personas que reservan pero no acuden al restaurante”

- **Parámetros:**
  - $n = 12$
  - $p = 0{,}20$
  - $q = 0{,}80$
- **Distribución de la V.A.**
  - $\xi \to B(12\,,0'20)$

- **Función de cuantía:**
  - $P(\xi = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$

$$P(\xi \geq 2) = 1 - P(\xi < 2) = 1 - [P(\xi = 0) + P(\xi = 1)]$$

$$= 1 - [\binom{12}{0} \cdot 0,2^0 \cdot 0,8^{12-0} + \binom{12}{1} 0,2^1 \cdot 0,8^{12-1}] = 0{,}7251$$

Ej. 2

Con base en encuestas se sabe que la preferencia de los consumidores respecto a dos marcas de un bien A y B es muy pareja. Si la opción de compra entre estas marcas es independiente entre los diferentes consumidores:

• **Tipo de experimento:** ○ Binomial B(n;p). • **Variable que modeliza:** ○ ξ= "Número de personas que prefieren la marca A"	• **Parámetros:** ○ n=5 ○ p=0,5 ○ q=0,5 • **Distribución de la V.A.** ○ $\xi \rightarrow B(5\,,0'5)$	• **Función de cuantía:** ○ $P(\xi = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$

a) ¿Qué probabilidad existe de que, entre 5 personas seleccionadas al azar, no más de 1 tengan preferencia por la marca A?

$$P(\xi \leq 1) = [P(\xi = 0) + P(\xi = 1)] =$$

$$\binom{5}{0} \cdot 0,5^0 \cdot 0,5^{5-0} + \binom{5}{1} \cdot 0,5^1 \cdot 0,5^{5-1} = 1 \cdot 1 \cdot 0,5^5 + 5 \cdot 0,5 \cdot 0,5^4 = 0,1875$$

b) ¿Qué probabilidad hay de que las 5 personas prefieran la marca B?

$$P(\xi = 0) = \binom{5}{0} \cdot 0,5^0 \cdot 0,5^{5-0} = 0,0313$$

Ej. 3 *La proporción de alumnos de la Universidad con calificación media de sobresaliente es de 0,005% La calificación de un alumno es independiente de la del resto ¿Qué probabilidad hay de que, de 5000 alumnos elegidos al azar, 2 tengan esa calificación media?*

- **Tipo de experimento:**
  - Binomial B(n;p)*.
- **Variable que modeliza:**
  - ξ= "Número de alumnos con sobresaliente"

- **Parámetros:**
  - n=5000
  - p=0,005%=0,00005
  - q=0,99995
- **Distribución de la V.A.**
  - $\xi \rightarrow B(5\,,0'5) \rightarrow$
  - $\xi \rightarrow Po(0{,}25)$

- **Función de cuantía:**
  - $P(\xi = x) = \frac{e^{-\lambda}.\lambda^{x}}{x!}$

$$P(\xi = 2) = \frac{e^{-0{,}25} \cdot (0{,}25)^2}{2!} = 0{,}0243$$

Ej. 4

Una empresa de transporte tiene una flota de 5 autobuses que prestan servicio diariamente. Sabiendo que el número de servicios cancelados en un día es una variable aleatoria que se distribuye según una distribución de Poisson con esperanza matemática 1:

• **Tipo de experimento:** ○ Poisson Po(λ). • **Variable que modeliza:** ○ ξ= "Número de servicios cancelados"	• **Parámetros:** ○ n=5 ○ λ=1 • **Distribución de la V.A.** ○ $\xi \rightarrow Po(1)$	• **Función de cuantía:** ○ $P(\xi = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$

a) ¿Qué probabilidad hay de que en un día no se preste ningún servicio?

$$P(\xi = 5) = \frac{e^{-1} \cdot 1^5}{5!} = 0{,}0031$$

b) Sabiendo que el número de servicios cancelados en un día es independiente en los 7 días de una semana, ¿qué probabilidad hay de que en una semana ocurran más de dos cancelaciones?

η→ "número de cancelaciones en una semana"

$$P(\eta > 2) = 1 - P(\eta \leq 2) = 1 - [P(\eta = 0) + P(\eta = 1) + P(\eta = 2)] =$$

$$1 - \frac{e^{-7} \cdot 7^0}{0!} - \frac{e^{-7} \cdot 7^1}{1!} - \frac{e^{-7} \cdot 7^2}{2!} = 0{,}9704$$

Ej. 5 *El número de clientes que entra en una sucursal de un banco en un minuto puede representarse mediante una variable aleatoria distribuida según una distribución de Poisson. La entrada de un cliente a la sucursal es independiente del comportamiento del resto de clientes. Si el promedio de clientes es de 135 por hora:*

- **Tipo de experimento:**
  - Poisson Po(λ).
- **Variable que modeliza:**
  - ξ= "Número de clientes que entran en un minuto"

- **Parámetros:**
  - λ=(135 clientes)/hora=
  =(2,25 clientes)/minuto
- **Distribución de la V.A.**
  - $\xi \rightarrow$ Po(2,25)

- **Función de cuantía:**
  - $P(\xi = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$

a) ¿Qué probabilidad hay de que entren al menos 2 clientes en un minuto?

$$P(\xi \geq 2) = 1 - P(\xi \leq 1) = 1 - [P(\xi = 0) + P(\xi = 1)] =$$

$$= 1 - \frac{e^{-2.25} \cdot (2{,}25)^0}{0!} - \frac{e^{-2.25} \cdot (2{,}25)^1}{1!} = 0{,}6575$$

b) Sabiendo que el número de servicios cancelados en un día es independiente en los 7 días de una semana, ¿qué probabilidad hay de que en una semana ocurran más de dos cancelaciones?

$$P(\xi = 0) = \frac{e^{-2.25} \cdot (2.25)^0}{0!} = 0{,}1053$$

Ej. 6

Supóngase que en un tramo de una carretera muy transitada ocurren de manera aleatoria e independiente, de media, 2 accidentes por semana. ¿Cuál será la probabilidad de que haya un accidente una semana y 3 en la siguiente?

- **Tipo de experimento:**
  - Poisson Po(λ).
- **Variable que modeliza:**
  - ξ_1= "Número de accidentes en la semana uno"
  - ξ_2= "Número de accidentes en la semana dos"

- **Parámetros:**
  - $\lambda=2$
- **Distribución de la V.A.**
  - $\xi_1 \to Po(2)$
  - $\xi_2 \to Po(2)$

- **Función de cuantía:**
  - $P(\xi = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$

$$\boxed{P(S_1 \cap S_2) = P(S_1) \cdot P(S_2)}$$

$$P(\xi_1 = 1 \cap \xi_2 = 3) = P(\xi_1 = 1) \cdot P(\xi_2 = 3)$$

$$\to \quad P(\xi_1 = 1) = \frac{e^{-2} \cdot 2^1}{1!} = 0{,}2707$$

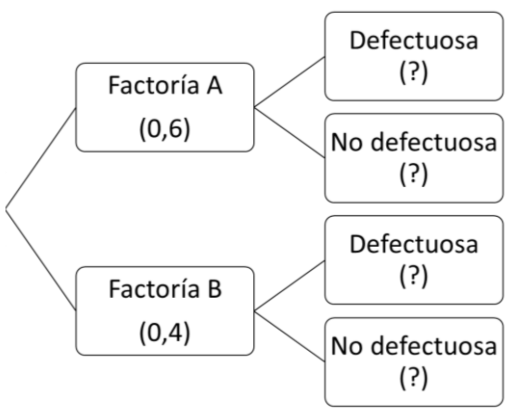
$$\to \quad P(\xi_2 = 3) = \frac{e^{-2} \cdot 2^3}{3!} = 0{,}1804$$

$$P(\xi_1 = 1 \cap \xi_2 = 3) = P(\xi_1 = 1) \cdot P(\xi_2 = 3) = 0{,}2707 \cdot 0{,}1804 = 0{,}0488$$

Ej. 7 ***La producción de un determinado bien por parte de una empresa depende de sus dos factorías. La primera de ellas (A), que fabrica el 60% de la producción de la empresa, tiene un promedio diario de elaboración de unidades defectuosas de 5. La segunda (B), por su parte, tiene un promedio diario de elaboración de unidades defectuosas de 2. Los procesos de fabricación de ambas factorías, así como la elaboración de unidades defectuosas, son independientes. Si llega al almacén la producción del día de una de las factorías, y en dicha producción detectamos la existencia de 3 unidades defectuosas, ¿qué probabilidad existe de que se trate de la producción de la factoría A?***

- **Tipo de experimento:**
  - Poisson Po(λ).
- **Variable que modeliza:**
  - ξ_A = *"Número de unidades defectuosas en un día en factoría A"*
  - ξ_B = *"Número de unidades defectuosas en un día en factoría B"*
- **Parámetros:**
  - $\lambda_A=5$
  - $\lambda_B=2$
- **Distribución de la V.A.**
  - $\xi_A \rightarrow \text{Po}(5)$
  - $\xi_B \rightarrow \text{Po}(2)$
- **Función de cuantía:**
  - $P(\xi = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$

$$P(\text{Factoría } A \mid 3\ defectuosas) = \frac{P(3\ Def \mid Factoría\ A) \cdot P(Factoría\ A)}{P(3\ Def \mid Factoría\ B) \cdot P(Factoría\ B) + P(3\ Def \mid Factoría\ A) \cdot P(Factoría\ A)}$$



$$\rightarrow P(\xi_A = 3) = \frac{e^{-5} \cdot 5^3}{3!} = 0{,}1403$$

$$\rightarrow P(\xi_B = 3) = \frac{e^{-2} \cdot 2^3}{3!} = 0{,}1804$$

$$= \frac{0{,}1403 \cdot 0{,}6}{0{,}1403 \cdot 0{,}6 + 0{,}1804 \cdot 0{,}4} = 0{,}5385$$

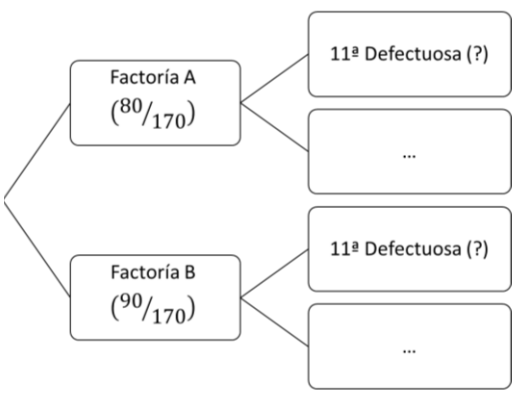
Ej. 8 *La producción de un determinado producto por parte de una empresa depende de sus dos factorías. La primera de ellas (A) tiene una probabilidad de fabricar una unidad defectuosa estimada en 0,005, mientras que la segunda (B) tiene una probabilidad de que una unidad sea defectuosa de 0,002, siendo independientes los procesos de fabricación en ambas factorías. Las producciones diarias de cada factoría son 80 y 90 unidades, respectivamente. Si en una de las factorías, al iniciar la producción diaria, se produce la unidad número 11 como primera defectuosa, ¿qué probabilidad existe de que dicha unidad defectuosa provenga de la factoría A?*

- **Tipo de experimento:**
  - Binomial Negativa BN(1;p), Geométrica G(p).
- **Variable que modeliza:**
  - ξ_A = *"Número de unidades válidas (fracasos) antes de la primera defectuosa (éxito) en un día en factoría A"*
  - ξ_B = *"Número de unidades válidas (fracasos) antes de la primera defectuosa (éxito) en un día en factoría B"*

- **Parámetros:**
  - p_A=0,005
  - p_B=0,002
- **Distribución de la V.A.**
  - $\xi_A \rightarrow$ G(0,005)
  - $\xi_B \rightarrow$ G(0,002)

- **Función de cuantía:**
  - $P(\xi = x) = \binom{x+r-1}{x} \cdot p^r \cdot q^x$

$$P(Factoría\ A|11^{a}\ Def) = \frac{P(11^{a}\ Def|\ Factoría\ A) \cdot P(Factoría\ A)}{P(11^{a}\ Def|\ Factoría\ B) \cdot P(Factoría\ B) + P(11^{a}\ Def|\ Factoría\ A) \cdot P(Factoría\ A)} =$$



$$\rightarrow \quad P(11^{a}\ Def\ A) = P(\xi_A = 10) = \binom{10+1-1}{10} \cdot (0{,}995)^{10} \cdot (0{,}005)^1 = 0{,}0047$$

$$\rightarrow \quad P(11^{a}\ Def\ B) = P(\xi_B = 10) = \binom{10+1-1}{10} \cdot (0{,}998)^{10} \cdot (0{,}002)^1 = 0{,}0020$$

$$= \frac{0{,}0047 \cdot \frac{80}{170}}{0{,}0047 \cdot \frac{80}{170} + 0{,}0020 \cdot \frac{90}{170}} = 0{,}6831$$

Ej. 9 *En una empresa se dispone de un sistema de control de calidad del producto manufacturado. El producto se clasifica, en dicho control, en defectuoso o no-defectuoso, siendo el estado de las unidades producidas independiente unas de otras. Sabiendo que la probabilidad de que una unidad de producto sea defectuosa es del 7%; ¿qué probabilidad hay de que la décima unidad controlada sea la tercera defectuosa?*

• **Tipo de experimento:** ○ Binomial Negativa BN(r;p). • **Variable que modeliza:** ○ ξ= "Número de unidades válidas (fracasos) antes de la tercera defectuosa (éxito)".	• **Parámetros:** ○ r=3 ○ p=0,07 ○ q=0,93 • **Distribución de la V.A.** ○ $\xi \to BN(3; 0'07)$	• **Función de cuantía:** ○ $P(\xi = x) = \binom{x+r-1}{x} \cdot p^r \cdot q^x$

$$P(\xi = 7) = \binom{7+3-1}{7} \cdot (0{,}07)^3 \cdot (0{,}93)^7 = 0{,}0074$$

Ej. 10 *Un fabricante de automóviles planea adquirir 40 motores de un proveedor. Para hacer efectiva la adquisición se seleccionan 8 motores al azar. Se efectuará la compra si ninguno de esos 8 motores está defectuoso. Si el lote de 40 motores contiene 2 defectuosos, ¿qué probabilidad hay de realizar la compra?*

- **Tipo de experimento:**
  - Hipergeométrica H(N;n; p).
- **Variable que modeliza:**
  - ξ= "Número de motores defectuosos en la muestra seleccionada".

- **Parámetros:**
  - N=40
  - n=8
  - N_1=2
  - p= $\frac{N_1}{N}$=$\frac{2}{40}$=0,05
- **Distribución de la V.A.**
  - $\xi \rightarrow H(40; 8;\ 0'05)$

- **Función de cuantía:**
  - $P(\xi = x) = \frac{\binom{N_1}{x} \cdot \binom{N-N_1}{n-x}}{\binom{N}{n}}$

$$P(\xi = 0) = \frac{\binom{2}{0} \cdot \binom{38}{8}}{\binom{40}{8}} = 0{,}6359$$

Ej. 11

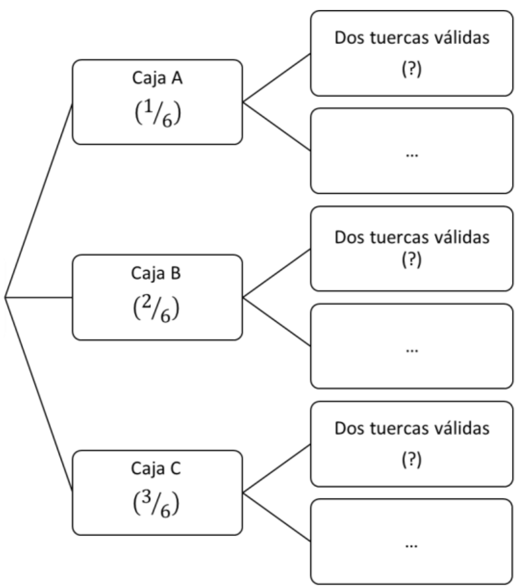
Se tienen 6 cajas que contienen, cada una, doce tuercas. En una de las cajas (tipo A), hay 8 tuercas utilizables y 4 defectuosas. En dos cajas (tipo B), hay 6 tuercas utilizables y 6 defectuosas. En el resto de cajas (tipo C), hay 4 tuercas utilizables y 8 defectuosas. Se elige una caja al azar y de ella se extraen 3 tuercas sin reemplazamiento. Se obtienen 2 tuercas útiles y una defectuosa. ¿Qué probabilidad hay de que hayamos extraído las tuercas de una caja del tipo B?

- **Tipo de experimento:**
  - Hipergeométrica H(N;n; p).
- **Variable que modeliza:**
  - ξ_A = "Número de tuercas válidas de la caja del tipo A".
  - ξ_B = "Número de tuercas válidas de la caja del tipo B".
  - ξ_C = "Número de tuercas válidas de la caja del tipo C".

- **Parámetros:**
  - Tipo A (1 caja)
    - $N_A=12$
    - $n=3$
    - $N_{A1}=8$
    - $p=\frac{N_{A1}}{N_A}=\frac{8}{12}=0{,}6667$
  - Tipo B (2 cajas)
    - $N_B=12$
    - $n=3$
    - $N_{B1}=6$
    - $p=\frac{N_{B1}}{N_B}=\frac{6}{12}=0{,}5$
  - Tipo C (3 cajas)
    - $N_C=12$
    - $n=3$
    - $N_{C1}=4$
    - $p=\frac{N_{C1}}{N_C}=\frac{4}{12}=0{,}3333$
- **Distribución de la V.A.**
  - $\xi \to H(12;3;0'6667)$
  - $\xi \to H(12;3;0'5)$
  - $\xi \to H(12;4;0'333)$

- **Función de cuantía:**
  - $P(\xi = x) = \frac{\binom{N_1}{x}\cdot\binom{N-N_1}{n-x}}{\binom{N}{n}}$

$$P(2\,Val|Caja\ B) = \frac{P(2\,Val|Caja\ B)\cdot P(Caja\ B)}{P(2\,Val|Caja\ A)\cdot P(Caja\ A) + P(2\,Val|Caja\ B)\cdot P(Caja\ B) + P(2\,Val|Caja\ C)\cdot P(Caja\ C)}$$



$$\to P(\xi_A=2)=\frac{\binom{8}{2}\cdot\binom{4}{1}}{\binom{12}{3}} = 0{,}509$$

$$\to P(\xi_B=2)=\frac{\binom{6}{2}\cdot\binom{6}{1}}{\binom{12}{3}} = 0{,}409$$

$$= \frac{0{,}409\cdot\frac{2}{6}}{0{,}509\cdot\frac{1}{6}+0{,}409\cdot\frac{2}{6}+0{,}218\cdot\frac{3}{6}} = 0{,}4128$$

$$\to P(\xi_C=2)=\frac{\binom{4}{2}\cdot\binom{8}{1}}{\binom{12}{3}} = 0{,}218$$


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales.
Universidad de Castilla – La Mancha.

Tema 5.2.

Modelos de distribución de probabilidad continuos univariantes Práctica

Ej. 1 *La concentración de cierto compuesto químico en el agua de un río se encuentra distribuida de manera uniforme en el intervalo de 4 a 20 partes por millón. Si se considera tóxica una concentración igual o superior a 15 partes por millón, ¿cuál es la probabilidad de que al analizarse una muestra de agua dé como resultado que esta no es potable?*

• **Variable que modeliza:**	• **Parámetros:**	• **Distribución de la V.A.**	• **Función de densidad:**
○ ξ= "Concentración de contaminante en el agua"	○ $a = 4$ ○ $b = 20$	○ $\xi \rightarrow U(4; 20)$	○ $f(x) = \frac{1}{b-a}$

$$P(\xi \geq 15) = \int_{15}^{20} \frac{1}{20-4} \cdot dx = \frac{1}{20-4} \cdot \int_{15}^{20} \mathrm{dx} =$$

$$= \frac{1}{16} \cdot [\mathrm{x}]_{15}^{20} = \frac{1}{16} \cdot [20 - 15] = \frac{5}{16} = 0{,}3125$$

Ej. 2

Una fábrica produce velas de cera cuyas duraciones encendidas (en horas) siguen una distribución exponencial de parámetro q = 1/4. Se escoge una vela de las producidas al azar y se enciende. ¿Cuál es la probabilidad de que permanezca más de 5 horas encendida?

• **Variable que modeliza:**	• **Parámetros:**	• **Distribución de la V.A.**	• **Función de densidad:**
○ ξ= "Horas de duración de la vela"	○ $q = 1/4$	○ $\xi \rightarrow Exp\left(\frac{1}{4}\right)$	○ $f(x) = qe^{-qx}$

$$P(\xi > 5) = \int_5^{\infty} \frac{1}{4} \cdot e^{-\frac{1}{4}x} dx = \int_5^{\infty} \frac{1}{4} \cdot e^{-\frac{1}{4}x} dx = \frac{1}{4} \cdot \int_5^{\infty} e^{-\frac{1}{4}x} dx = \frac{1}{4} \cdot \frac{1}{-\frac{1}{4}} \cdot \left[e^{-\frac{1}{4}x}\right]_5^{\infty} =$$

$$= -1 \cdot \left[e^{-\frac{1}{4}\cdot\infty} - e^{-\frac{1}{4}\cdot 5}\right] = -1 \cdot \left[0 - e^{-\frac{5}{4}}\right] = -1 \cdot [-0{,}2865] = 0{,}2865$$

Ej. 3 *Una fábrica produce piezas cuyos diámetros se distribuyen en probabilidad según un modelo N(5 ; 0,001). Para que una pieza sirva, el diámetro debe estar comprendido entre 4,998 cm. y 5,002 cm. Si es menor se desecha, si es superior se recicla*

• **Variable que modeliza:** ○ ξ= "Cm de diámetro de una pieza"	• **Parámetros:** ○ $\mu = 5$ ○ $\sigma = 0{,}001$	• **Distribución de la V.A.** ○ $\xi \to N(5; 0{,}001)$

a) ¿Qué porcentaje de piezas probablemente servirá?

$$P(4{,}998 \leq \xi \leq 5{,}002) = P\left(\frac{4{,}998-5}{0{,}001} \leq \frac{\xi-5}{0{,}001} \leq \frac{5{,}002-5}{0{,}001}\right) = P(-2 \leq \xi^* \leq 2) =$$

$$= 1 - P(\xi^* \geq 2) - P(\xi^* \geq 2) = 1 - 0{,}0228 - 0{,}0228 = 0{,}9544 = 95{,}44\%$$

b) ¿Qué porcentaje probablemente será desechado?

$$P(\xi < 4{,}998) = P\left(\frac{\xi-5}{0{,}001} < \frac{4{,}998-5}{0{,}001}\right) = P(\xi^* < -2) = P(\xi^* > 2) = 0{,}0228 = 2{,}28\%$$

c) ¿Qué porcentaje probablemente será reciclado?

$$P(\xi > 5{,}002) = P\left(\frac{\xi-5}{0{,}001} > \frac{5{,}002-5}{0{,}001}\right) = P(\xi^* > 2) = 0{,}0228 = 2{,}28\%$$

Ej. 4 *La demanda mensual de cierto producto sigue una distribución de probabilidad N(200;40). Si el almacén no se repone a lo largo del mes, ¿qué cantidad mínima de existencias debe permanecer en stock al principio de mes para que la probabilidad de agotar las existencias sea 0,025?*

• **Variable que modeliza:** ○ ξ= "Cm de diámetro de una pieza"	• **Parámetros:** ○ $\mu = 200$ ○ $\sigma = 40$	• **Distribución de la V.A.** ○ $\xi \rightarrow N(200; 40)$

$$P(\xi \geq x) = 0{,}025$$

Tipificamos:

$$P\left(\frac{\xi - 200}{40} \geq \frac{x - 200}{40}\right) = P\left(\xi^* \geq \frac{x - 200}{40}\right) = 0{,}025$$

Hacemos la conversión:

$$\frac{x - 200}{40} = 1{,}96$$

Despejando:

$$x = 1{,}96 \cdot 40 + 200 = 278{,}4$$

Ej. 5 *La duración en horas de un tipo de baterías recargables sigue una distribución de probabilidad N(240;10). Si elegimos dos baterías independientemente y al azar, ¿qué probabilidad hay de que la duración conjunta de ambas baterías sea inferior a 490 horas?*

• **Variable que modeliza:**	• **Parámetros:**	• **Distribución de la V.A.**
○ ξ_1 = "Duración en horas de la batería 1" ○ ξ_2 = "Duración en horas de la batería 2"	○ $\mu = 240$ ○ $\sigma = 10$	○ $\xi_1 \to N(240; 10)$ ○ $\xi_2 \to N(240; 10)$

$$\eta = \xi_1 + \xi_2 = \text{"duración conjunta de ambas baterías"}$$

$$\eta \to N\left(240 + 240; \sqrt{10^2 + 10^2}\right) = N(480; 14{,}14)$$

$$P(\eta < 490) = P\left(\frac{\eta - 480}{14{,}14} < \frac{490 - 480}{14{,}14}\right) =$$

$$P(\eta^* < 0{,}71) = 1 - P(\eta^* \geq 0{,}71) = 1 - 0{,}2389 = 0{,}7611$$

Ej. 6 *Sea ξ_1 con distribución N(40;9) el número de viajeros que salen de Madrid en el AVE de las 7:15 h. Sea ξ_2 con distribución N(30;12) el número de viajeros que se bajan en Ciudad Real. Sea ξ_3 con distribución N(100;20) el número de viajeros que suben en Ciudad Real con destino a Puertollano. Suponiendo que las tres variables son independientes entre sí, ¿qué probabilidad hay de que el número de viajeros que parten desde Ciudad Real a Puertollano sea menor que 70?*

• **Variable que modeliza:** ○ ξ_1= "Número de viajeros que salen de Madrid a las 7:15". ○ ξ_2= "Número de viajeros que bajan en Ciudad Real". ○ ξ_3= "Número de viajeros que suben en Ciudad Real con destino Puertollano".	• **Distribución de la V.A.** ○ $\xi_1 \to N(40;9)$ ○ $\xi_2 \to N(30;12)$ ○ $\xi_3 \to N(100;20)$

$$\eta = \xi_1 - \xi_2 + \xi_3 = \text{"número final de viajeros"}$$

$$\eta \to N\left(40 - 30 + 100; \sqrt{9^2 + 12^2 + 20^2}\right) = N(110;25)$$

$$P(\eta < 70) = P\left(\frac{\eta - 110}{25} < \frac{70 - 110}{25}\right) = P(\eta^* < -1{,}6) = P(\eta^* > 1{,}6) = 0{,}0548$$

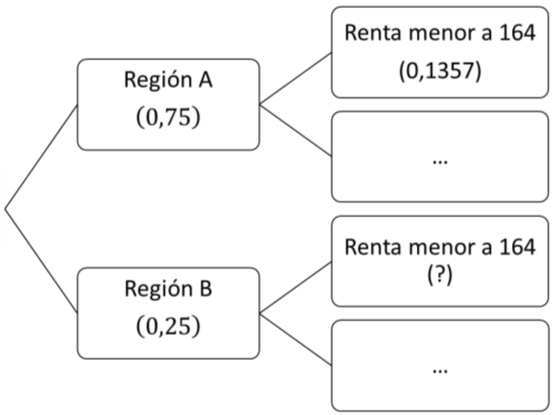
Ej. 7 ***La duración en horas de un tipo de baterías recargables sigue una distribución de probabilidad N(240;10). Si elegimos dos baterías independientemente y al azar, ¿qué probabilidad hay de que la duración conjunta de ambas baterías sea inferior a 490 horas?***

• **Variable que modeliza:** ○ ξ_A= "Renta anual de la población de la región A". ○ ξ_B= "Renta anual de la población de la región B".	• **Distribución de la V.A.** ○ $\xi_A \to N(175;10)$ ○ $\xi_B \to N(170;10)$

a) Si se elige al azar un individuo de la región A, ¿qué probabilidad hay de que su renta sea inferior a 164?

$$P(\xi_A < 164) = P\left(\frac{\xi_A - 175}{10} < \frac{164-175}{10}\right) = P(\xi_A^* < -1{,}1) = P(\xi_A^* > 1{,}1) = 0{,}1357$$

b) Si se elige al azar un individuo del país, ¿qué probabilidad hay de que su renta sea inferior a 164?



$$P(\xi_B < 164) = P\left(\frac{\xi_B - 170}{10} < \frac{164-170}{10}\right) = P(\xi_B^* < -0{,}6) = P(\xi_B^* > 0{,}6) = 0{,}2743$$

$$P(Ren < 164) =$$

$$P(Región\ A) \cdot P(Ren < 164 \mid Región\ A) + P(Región\ B) \cdot P(Ren < 164 \mid Región\ B)$$

$$= 0{,}75 \cdot 0{,}1357 + 0{,}25 \cdot 0{,}2743 = 0{,}1703$$

c) Si la renta de un individuo del país es inferior a 164, ¿qué probabilidad existe de que se trate de un individuo de la región A?

$$P(\,Región\ A \mid Ren < 164) = \frac{P(Ren < 164 \mid\ Región\ A) \cdot P(Región\ A)}{P(Ren < 164 \mid\ Región\ A) \cdot P(Región\ A) + P(Ren < 164 \mid\ Región\ B) \cdot P(Región\ B)} =$$

$$P(\,Región\ A \mid Renta < 164) = \frac{0{,}1357 \cdot 0{,}75}{0{,}75 \cdot 0{,}1357 + 0{,}25 \cdot 0{,}2743} = 0{,}5974$$


¡Muchas gracias!

Estadística Empresarial

Facultad de Derecho y Ciencias Sociales. Universidad de Castilla – La Mancha.

Tema 5.3.

Convergencia y Teoremas Límite

Práctica

Ej. 1

Una máquina tiene una fiabilidad para fabricar un determinado tipo de pieza del 92%. ¿Cuál será la probabilidad de que la máquina fabrique más de 90 piezas correctamente de entre un total de 100?

- **Variable aleatoria:**
  - ξ= "Concentración de contaminante en el agua"
- **Parámetros:**
  - a = 4
  - b = 20

- **Aproximación:**

$$\text{Como } n \geq 100, \longrightarrow B(100; 0{,}92) \xrightarrow{d} N\left(100 \cdot 0{,}92; \sqrt{100 \cdot 0{,}92 \cdot 0{,}08}\right) = (92; 2{,}7129)$$

$$P(\xi > 90) = P\left(\frac{\xi - 92}{2{,}7129} > \frac{90 - 92}{2{,}7129}\right) = P(\xi^* > -0{,}7372) =$$

$$= 1 - P(\xi^* \geq 0{,}7372) = 1 - 0{,}2296 = 0{,}7704$$

Ej. 2 *La duración en horas de unas baterías recargables tiene una distribución N(240;10). Si elegimos 100 baterías independientemente y al azar, ¿qué probabilidad hay de que más de 80 de ellas duren menos de 250 horas?*

• **Variable aleatoria:** ○ ξ= "Duración en horas de una batería i"	• **Parámetros:** ○ $\mu_i = 240$ ○ $\sigma_i = 10$

Probabilidad de que una de estas baterías dure menos de 250 horas:

$$(\xi_i < 250) = P\left(\frac{\xi_i - 240}{10} < \frac{250 - 240}{10}\right) = P({\xi_i}^* < 1) = 1 - P\left(\xi_j^* \geq 1\right) = 1 - 0{,}1587 = 0{,}8413$$

$$\eta = \text{"Número de baterías con duración menor a 250 horas"} = \xi_1 + \xi_2 + \cdots + \xi_n \longrightarrow B(100; 0{,}8413)$$

- **Aproximación:**

$$\xi \longrightarrow B(100; 0{,}8413) \xrightarrow{d} N\left(100 \cdot 0{,}8413; \sqrt{100 \cdot 0{,}8413 \cdot (1 - 0{,}8413)}\right) = N(84{,}13; 3{,}6540)$$

$$P(\xi > 80) = P\left(\frac{\xi - 84{,}13}{3{,}6540} > \frac{80 - 84{,}13}{3{,}6540}\right) = P(\xi^* > -1{,}1303) =$$

$$= 1 - P(\xi^* \geq 1{,}1303) = 1 - 0{,}1292 = 0{,}8708$$

Ej. 3

En una fábrica se sabe que la probabilidad de que salgan X artículos defectuosos en un día concreto es $Po(4)$, siendo independiente el número de artículos defectuosos en un día respecto al de cualquier otro día. ¿Cuál será la probabilidad de que en 100 días el número total de artículos defectuosos sea menor que 410?

• **Variable aleatoria:** ○ ξ= "Número de artículos defectuosos en el día j"	• **Parámetros:** ○ $\lambda = 4$

$$\eta \to \mathrm{Po}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) = \eta \to \mathrm{Po}(4 + 4 + 4 + \cdots + 4) = \eta \to \mathrm{Po}(100 \cdot 4) = \mathrm{Po}(400)$$

- **Aproximación:**

$$\text{Como } n \geq 100, \longrightarrow \mathrm{Po}(400) \xrightarrow{d} N\left(100 \cdot 4; \sqrt{100 \cdot 4}\right) = (400; 20)$$

$$P(\eta < 410) = P\left(\frac{\eta - 410}{20} < \frac{410 - 400}{20}\right) = P(\eta^* < 0{,}5) =$$

$$= 1 - P(\eta^* \geq 0{,}5) = 1 - 0{,}3085 = 0{,}6915$$

Ej. 4

Una cadena hotelera posee un hotel en una determinada localidad costera. Por experiencia, los directivos saben que el 60% de las reservas realizadas con al menos seis meses de antelación relativas a la ocupación de las habitaciones durante el mes de junio son confirmadas, mientras que el 40% restante acaban anulándose, siendo las confirmaciones independientes unas de otras. En esta ocasión, las reservas realizadas con al menos 6 meses de antelación ascienden a 216. ¿Qué probabilidad hay de que, en total, se confirmen al menos 140 reservas de entre todas esas?

• **Variable aleatoria:** ○ ξ= "Número de confirmaciones correspondientes a las reservas realizadas con al menos 6 meses de antelación"	• **Parámetros:** ○ p = 0,60 ○ q = 0,40 ○ n = 216

- **Aproximación:**

$$\text{Como } n\geq 100, \longrightarrow B(216;0{,}60)\xrightarrow{d} N\left(216\cdot 0{,}60;\sqrt{216\cdot 0{,}60\cdot 0{,}40}\right)=(129{,}6;7{,}2)$$

$$P(\xi \geq 140) = P\left(\frac{\xi - 129{,}6}{7{,}2} \geq \frac{140-129{,}6}{7{,}2}\right) = P(\xi^* \geq 1{,}44) = 0{,}0749$$


¡Muchas gracias!

Estadística Empresarial. Material Docente.

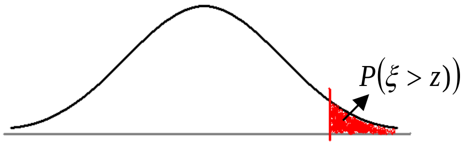
MATERIAL ADICIONAL

BLOQUE PRÁCTICO

En este apartado encontrarás recursos adicionales como las tablas y los principales estadísticos utilizados y sus distribuciones.

ÁREA DE ESTADÍSTICA EMPRESARIAL. DEPARTAMENTO DE ECONOMÍA APLICADA I. FACULTAD DE DERECHO Y CCSS DE CIUDAD REAL (UCLM).

DISTRIBUCIÓN NORMAL ESTÁNDAR N(0;1)



Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010

PRINCIPALES MODELOS DE PROBABILIDAD

Distribución	Función cuantía / densidad	Parámetros	Media	Varianza
$\boldsymbol{B(1;p)}$	$\boldsymbol{P(\xi = x) = p^x \cdot q^{1-x}}$ $q = 1-p; x = 0; 1$	• ***p:*** probabilidad de obtener el resultado de éxito. • ***q:*** probabilidad de obtener el resultado de fracaso.	p	$p \cdot q$
$\boldsymbol{B(n;p)}$	$\boldsymbol{P(\xi = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}}$ $q = 1-p; x = 0,\ 1,\ 2,\ 3,\ \dots$	• ***n:*** número de veces que se repite el experimento. • ***p:*** probabilidad de obtener el resultado de éxito. • ***q:*** probabilidad de obtener el resultado de fracaso.	$n \cdot p$	$n \cdot p \cdot q$
$\boldsymbol{Po(\lambda)}$	$\boldsymbol{P(\xi = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}}$ $\lambda > 0; x = 0,\ 1,\ 2,\ 3,\ \dots$	• $\boldsymbol{\lambda}$***:*** orden de la distribución. $\lambda = n \cdot p$.	λ	λ
$\boldsymbol{BN(r;p)}$	$\boldsymbol{P(\xi = x) = \binom{x+r-1}{x} \cdot p^r \cdot q^x}$ $x = 0,\ 1,\ 2,\ 3,\ \dots$	• ***r:*** número de éxitos que se han de dar. • ***p:*** probabilidad de obtener el resultado de éxito. • ***q:*** probabilidad de obtener el resultado de fracaso.	$r \cdot \frac{q}{p}$	$r \cdot \frac{q}{p^2}$
$\boldsymbol{H(N;n;p)}$	$\boldsymbol{P(\xi = x) = \frac{\binom{N_1}{x} \cdot \binom{N-N_1}{n-x}}{\binom{N}{n}}}$ $x = 0,\ 1,\ 2,\ 3,\ \dots,\ n$	• ***N:*** número de elementos de la muestra. • ***n:*** número de elementos de la muestra extraídos. • $\boldsymbol{p = \frac{N_1}{N}}$***:*** proporción de elementos tipo A en la muestra.	$n \cdot p$	$n \cdot p \cdot q \cdot \frac{N-n}{N-1}$
$\boldsymbol{U(a;b)}$	$\boldsymbol{f(x) = \frac{1}{b-a}}$ $a \leq x \leq b$	• ***a:*** extremo inferior del campo de variación de la variable ξ. • ***b:*** extremo superior del campo de variación de la variable ξ.	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$\boldsymbol{Exp(q)}$	$\boldsymbol{f(x) = qe^{-qx}}$ $x > 0;\ q \geq 0$	• ***q:*** parámetro característico de la función de densidad de la variable ξ.	$\frac{1}{q}$	$\frac{1}{q^2}$
$\boldsymbol{N(\mu;\sigma)}$	$\boldsymbol{f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}}$ $-\infty \leq x \leq \infty$	• $\boldsymbol{\mu}$***:*** media o esperanza matemática de la variable ξ. • $\boldsymbol{\sigma}$***:*** desviación típica de la variable ξ.	μ	σ^2

$$\boldsymbol{m_3 = a_3 - 3a_2a_1 + 2a_1^3}$$

$$\boldsymbol{m_4 = a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4}$$

Material docente.
Estadística Empresarial.

FIN

ÁNIMO.

contacto

Miguel Ángel Tarancón Morán
miguelangel.tarancon@uclm.es

Consolación Quintana Rojo
consolacion.quintana@uclm.es

Daniel Rojas Peña
daniel.rojas@uclm.es

ÁREA DE ESTADÍSTICA EMPRESARIAL. DEPARTAMENTO DE ECONOMÍA APLICADA I. FACULTAD DE DERECHO Y CCSS DE CIUDAD REAL (UCLM).