



UNIVERSIDAD
DE GRANADA

Técnicas Cuantitativas I

Temario de la asignatura

Ismael Sallami Moreno

Recursos Ingeniería Informática y Ade

Licencia

Este trabajo está bajo una Licencia Creative Commons BY-NC-ND 4.0.

Permisos: Se permite compartir, copiar y redistribuir el material en cualquier medio o formato.

Condiciones: Es necesario dar crédito adecuado, proporcionar un enlace a la licencia e indicar si se han realizado cambios. No se permite usar el material con fines comerciales ni distribuir material modificado.



Técnicas Cuantitativas I

Ismael Sallami Moreno

<https://elblogdeismael.github.io>

Índice general

1	Variables estadísticas unidimensionales	9
1.1	Conceptos previos	9
1.2	Tablas estadísticas	9
1.3	Momentos	10
1.4	Medidas de posición	10
1.5	Medidas de dispersión	11
1.6	Medidas de forma	12
1.7	Medidas de concentración	12
1.8	Ejercicios	13
2	Variables estadísticas bidimensionales	15
2.1	Representación	15
2.2	Distribuciones marginales y condicionadas	15
2.3	Independencia	16
2.4	Covarianza	16
2.5	Coefficiente de correlación lineal	16
2.6	Regresión por mínimos cuadrados	17
2.7	Ejercicios	18
3	Números índices	19
3.1	Tasas de variación	19
3.2	Índices elementales	19
3.3	Índices sintéticos	20
3.4	Enlace de series con distinta base	21
3.5	Deflación	21
3.6	Participación de un grupo	21

3.7	Índices encadenados	22
3.8	Ejercicios	22
4	Análisis descriptivo de series cronológicas	25
4.1	Definición	25
4.2	Componentes	25
4.3	Tendencia secular	26
4.4	Variación estacional	27
4.5	Desestacionalización	27
4.6	Predicción	27
4.7	Ejercicios	28
5	Probabilidad	29
5.1	Experimentos y sucesos	29
5.2	Definición y asignación	29
5.3	Probabilidad condicionada	30
5.4	Independencia	30
5.5	Probabilidad total y Bayes	31
5.6	Ejercicios	32
6	Variables aleatorias y distribuciones de probabilidad	33
6.1	Concepto	33
6.2	Función de distribución	33
6.3	Valor esperado	34
6.4	Momentos y otras características	35
6.5	Variables aleatorias bidimensionales	35
6.6	Ejercicios	36
7	Distribuciones discretas de probabilidad	39
7.1	Uniforme discreta	39
7.2	Binomial	39
7.3	Poisson	40
7.4	Hipergeométrica	40
7.5	Geométrica	41

7.6	Cómo se elige	41
7.7	Ejercicios	42
8	Temario práctico	43
8.1	De dónde salen los datos	43
8.2	Práctica 1. Análisis descriptivo unidimensional	43
8.3	Práctica 2. Análisis descriptivo bidimensional	44
8.4	Práctica 3. Números índices y deflación	44
8.5	Práctica 4. Series cronológicas	44
8.6	Práctica 5. Probabilidad y distribuciones	45
8.7	Sobre la memoria	45

Variables estadísticas unidimensionales

Tema 1 del programa. Tablas y representaciones, momentos, y las medidas de posición, dispersión, forma y concentración.

1.1 Conceptos previos

Término	Qué es
Población	conjunto de elementos objeto de estudio
Muestra	subconjunto observado
Individuo	cada elemento
Carácter	la propiedad que se observa
Modalidades	los valores que puede tomar

Tipo de carácter	Modalidades	Ejemplo
Cualitativo nominal	sin orden	sector de actividad
Cualitativo ordinal	con orden	nivel de satisfacción
Cuantitativo discreto	valores aislados	número de empleados
Cuantitativo continuo	cualquier valor de un intervalo	facturación

El tipo decide qué medidas tienen sentido. Calcular la media de un carácter nominal codificado con números es el error más común del tema, y produce cifras que no significan nada.

1.2 Tablas estadísticas

Frecuencia	Definición
Absoluta n_i	número de veces que aparece x_i
Relativa $f_i = n_i/N$	proporción
Absoluta acumulada N_i	suma de las absolutas hasta x_i
Relativa acumulada F_i	suma de las relativas hasta x_i

Con datos agrupados en intervalos hacen falta dos elementos más:

$$c_i = \frac{L_{i-1} + L_i}{2} \quad (\text{marca de clase}), \quad a_i = L_i - L_{i-1} \quad (\text{amplitud})$$

y la **densidad de frecuencia** $h_i = n_i/a_i$, que es lo que se representa en el histograma.

Nota 1.1 . Con intervalos de amplitud desigual, el histograma debe representar la **densidad** y no la frecuencia. Si se dibujan alturas iguales a n_i , un intervalo ancho parece tener más casos de los que le corresponden, y la forma de la distribución sale falseada. Es el gráfico mal hecho más frecuente en informes reales.

Representación	Para qué carácter
Diagrama de barras	discreto o cualitativo
Histograma	continuo agrupado
Polígono de frecuencias	continuo, uniendo las marcas de clase
Diagrama de sectores	cualitativo, con pocas modalidades
Diagrama acumulado	para leer percentiles

1.3 Momentos

$$a_r = \frac{1}{N} \sum_i x_i^r n_i \quad (\text{no centrados}), \quad m_r = \frac{1}{N} \sum_i (x_i - \bar{x})^r n_i \quad (\text{centrados})$$

Casos particulares: $a_1 = \bar{x}$ es la media, $m_1 = 0$ siempre y $m_2 = s^2$ es la varianza.

La relación entre unos y otros es la que ahorra cálculos:

$$m_2 = a_2 - a_1^2, \quad m_3 = a_3 - 3a_1a_2 + 2a_1^3$$

La primera es la fórmula abreviada de la varianza, y se usa constantemente.

1.4 Medidas de posición

1.4.1 Medias

Media	Fórmula	Cuándo
Aritmética	$\bar{x} = \frac{1}{N} \sum x_i n_i$	magnitudes que se suman
Geométrica	$G = \sqrt[N]{\prod x_i^{n_i}}$	tasas de variación, índices
Armónica	$H = \frac{N}{\sum n_i/x_i}$	velocidades, ratios inversos

Proposición 1.1 . Para datos positivos, $H \leq G \leq \bar{x}$, con igualdad solo si todos los valores coinciden.

Elegir la media equivocada da un resultado sistemáticamente sesgado. Para promediar

tasas de crecimiento hay que usar la geométrica: crecer un 50 % y después caer un 50 % no deja igual, deja un 25 % menos, y solo la geométrica lo refleja.

Ejemplo 1.1 . Una inversión crece un 50 % el primer año y cae un 50 % el segundo. Los factores son 1,5 y 0,5.

La media aritmética de los factores es 1, que sugeriría que no ha pasado nada. La geométrica es $\sqrt{1,5 \cdot 0,5} = 0,866$, es decir una caída media anual del 13,4 %, y es la correcta: el capital final es el 75 % del inicial, y en efecto $0,866^2 = 0,75$.

Propiedades de la media aritmética:

Propiedad	Expresión
Desviaciones	$\sum (x_i - \bar{x})n_i = 0$
Linealidad	si $y = a + bx$, entonces $\bar{y} = a + b\bar{x}$
Mínimo cuadrático	$\sum (x_i - k)^2 n_i$ es mínimo con $k = \bar{x}$
Sensibilidad	un valor extremo la arrastra

1.4.2 Medidas de posición no centrales

Medida	Definición
Moda	el valor más frecuente
Mediana	deja el 50 % de los datos a cada lado
Cuartiles Q_1, Q_2, Q_3	dividen en cuatro partes iguales
Deciles, percentiles	en diez y en cien partes

Para datos agrupados, el percentil k se interpola dentro de su intervalo:

$$P_k = L_{i-1} + \frac{\frac{kN}{100} - N_{i-1}}{n_i} a_i$$

La mediana es robusta y la media no. Con la renta de un barrio, un solo millonario mueve la media y deja la mediana intacta, y por eso los informes de renta publican siempre la mediana.

1.5 Medidas de dispersión

Medida	Fórmula	Unidades
Recorrido	$x_{\text{máx}} - x_{\text{mín}}$	las de la variable
Recorrido intercuartílico	$Q_3 - Q_1$	las de la variable
Varianza	$s^2 = \frac{1}{N} \sum (x_i - \bar{x})^2 n_i$	al cuadrado
Desviación típica	$s = \sqrt{s^2}$	las de la variable
Coefficiente de variación	$CV = s/\bar{x}$	adimensional

La última fila es la que permite comparar dispersiones entre variables distintas. Una desviación típica de 100 es enorme si la media es 200 y despreciable si la media es 100 000, y el coeficiente de variación lo dice en un número.

Nota 1.2 . El coeficiente de variación **no sirve si la media está cerca de cero**, porque se dispara, ni con variables que toman valores negativos, porque pierde el signo su sentido. En esos casos hay que quedarse con la desviación típica y decir respecto de qué.

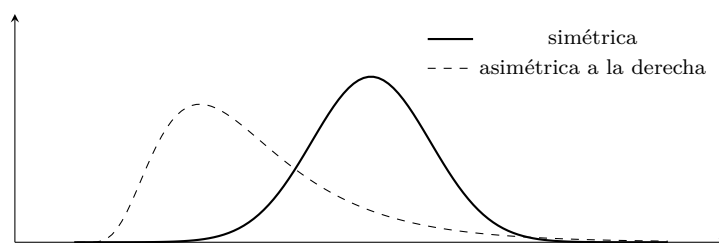
Propiedades de la varianza:

Propiedad	Expresión
Traslación	$\text{Var}(x + a) = \text{Var}(x)$
Escala	$\text{Var}(bx) = b^2 \text{Var}(x)$
Fórmula abreviada	$s^2 = \overline{x^2} - \bar{x}^2$
No negatividad	$s^2 \geq 0$, con igualdad si todos los valores son iguales

1.6 Medidas de forma

Medida	Fórmula	Interpretación
Asimetría de Fisher	$g_1 = m_3/s^3$	> 0 cola a la derecha, < 0 a la izquierda
Curtosis de Fisher	$g_2 = m_4/s^4 - 3$	> 0 más apuntada que la normal

El -3 de la curtosis está para que la distribución normal dé exactamente 0, y así el signo se lea como comparación con ella.

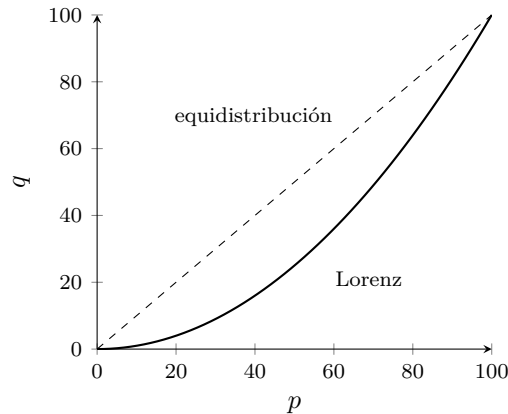


La **asimetría a la derecha** es la forma típica de las variables económicas: rentas, patrimonios, tamaños de empresa. Y es la razón de que en ellas la media supere a la mediana.

1.7 Medidas de concentración

Miden cómo se reparte el total de una magnitud entre los individuos. En un reparto perfectamente equitativo, el 20 % de la población acumula el 20 % del total.

Elemento	Definición
p_i	porcentaje acumulado de individuos
q_i	porcentaje acumulado del total de la magnitud
Curva de Lorenz	la poligonal que une los puntos (p_i, q_i)
Diagonal	$q = p$: equidistribución perfecta



Definición 1.1 (Índice de Gini).

$$IG = \frac{\sum_{i=1}^{k-1} (p_i - q_i)}{\sum_{i=1}^{k-1} p_i}$$

Vale 0 en la equidistribución perfecta y 1 en la concentración total.

Geométricamente, el Gini es el doble del área entre la diagonal y la curva de Lorenz. Cuanto más se hunde la curva, mayor es la desigualdad.

La **mediala** es el valor que divide la magnitud total en dos mitades: por debajo de ella se acumula el 50 % del total. Comparada con la mediana, cuantifica la concentración: en un reparto equitativo coinciden, y cuanto más se separan, más desigual es la distribución.

1.8 Ejercicios

Ejercicio 1.8.1. Los salarios mensuales de seis empleados son 1200, 1300, 1400, 1500, 1600 y 9000. Calcular media, mediana y coeficiente de variación, y decir cuál describe mejor.

Solución 1.8.1. La media es $16\,000/6 = 2666,7$ y la mediana $(1400 + 1500)/2 = 1450$. La desviación típica es 2829,4, así que $CV = 1,06$.

La mediana describe mucho mejor: cinco de los seis empleados cobran menos de la media, y el sexto la arrastra él solo. Un coeficiente de variación mayor que 1 es ya la señal de que la media no representa a nadie.

Ejercicio 1.8.2. Una acción sube un 20 % un año y baja un 20 % al siguiente. ¿Cuál es la rentabilidad media anual?

Solución 1.8.2. Los factores son 1,2 y 0,8, cuyo producto es 0,96: se ha perdido un 4 %. La media

geométrica es $\sqrt{0,96} = 0,9798$, es decir una pérdida media anual del 2,02 %. La media aritmética de los porcentajes daría 0, que es falso.

Ejercicio 1.8.3. En una distribución, la media es 50 y la desviación típica 10. Se aplica la transformación $y = 3x - 20$. Hallar la media y la desviación típica de y .

Solución 1.8.3. $\bar{y} = 3 \cdot 50 - 20 = 130$ por linealidad. Y $\text{Var}(y) = 3^2 \text{Var}(x) = 900$, así que $s_y = 30$: la traslación no afecta a la dispersión y el factor de escala la multiplica en valor absoluto. El coeficiente de variación pasa de 0,20 a 0,23, porque la traslación sí cambia la media.

La estadística descriptiva unidimensional está desarrollada en Hernández Bastida, 2007, Castillo Manrique y Guijarro Garvi, 2006 y Martín Pliego, 2004, con problemas resueltos en Hermoso Gutiérrez y Hernández Bastida, 2000 y Amor Pulido, 2016.

Variables estadísticas bidimensionales

Tema 2 del programa. Representaciones en dos columnas y en tablas de contingencia, distribuciones marginales y condicionadas, independencia, covarianza, coeficiente de correlación y recta de regresión.

2.1 Representación

Formato	Cuándo se usa
Dos columnas	pocos datos, sin repeticiones
Tabla de contingencia	muchos datos, con repeticiones o agrupados

En una tabla de contingencia, n_{ij} es el número de individuos con $X = x_i$ e $Y = y_j$, y las **frecuencias marginales** son las sumas por filas y por columnas:

$$n_{i.} = \sum_j n_{ij}, \quad n_{.j} = \sum_i n_{ij}, \quad N = \sum_{i,j} n_{ij}$$

2.2 Distribuciones marginales y condicionadas

Distribución	Qué describe
Marginal de X	el comportamiento de X ignorando Y
Marginal de Y	ídem al revés
Condicionada $X Y = y_j$	el comportamiento de X dentro de una columna

$$f_{i|j} = \frac{n_{ij}}{n_{.j}}$$

Las condicionadas son las que permiten comparar grupos: la renta media de los hombres frente a la de las mujeres, o la nota media según el grupo de prácticas.

Nota 2.1 . Comparar marginales cuando lo que interesa son las condicionadas produce la **paradoja de Simpson**: una tendencia que aparece en cada subgrupo puede invertirse al agregar. Ocurre cuando el tamaño de los grupos está desequilibrado, y es la razón de que un dato global pueda contradecir todos los datos parciales sin que ninguno sea falso.

2.3 Independencia

Definición 2.1 (Independencia estadística). X e Y son independientes si

$$f_{ij} = f_{i\cdot} f_{\cdot j} \quad \text{para todo par } (i, j)$$

equivalentemente, si todas las distribuciones condicionadas de X coinciden entre sí.

La condición debe cumplirse en todas las casillas. Que se cumpla en algunas no dice nada, y comprobarla exige recorrer la tabla entera.

2.4 Covarianza

$$s_{xy} = \frac{1}{N} \sum_{i,j} (x_i - \bar{x})(y_j - \bar{y}) n_{ij} = \overline{xy} - \bar{x} \bar{y}$$

Signo	Interpretación
$s_{xy} > 0$	relación directa: crecen juntas
$s_{xy} < 0$	relación inversa
$s_{xy} = 0$	incorreladas , que no es lo mismo que independientes

Proposición 2.1 . Si X e Y son independientes, $s_{xy} = 0$. El recíproco es falso.

Ejemplo 2.1 . Sea X con valores $-2, -1, 0, 1, 2$ e $Y = X^2$. La covarianza es 0 por simetría, y sin embargo Y está determinada completamente por X . La covarianza solo detecta relación **lineal**, y aquí la relación es cuadrática.

La covarianza tiene un problema práctico: **sus unidades son el producto de las dos variables**, así que su valor no se puede interpretar por sí solo ni comparar entre pares de variables distintas. De ahí el coeficiente siguiente.

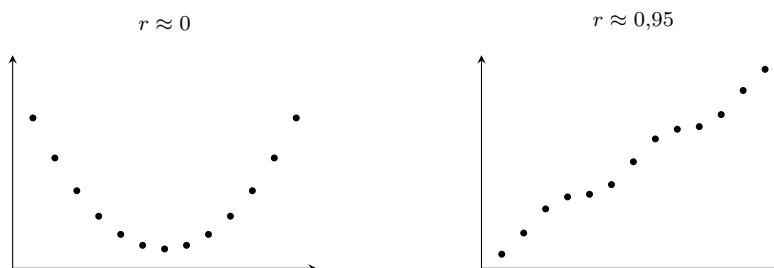
2.5 Coeficiente de correlación lineal

$$r = \frac{s_{xy}}{s_x s_y}, \quad -1 \leq r \leq 1$$

$ r $	Relación lineal
1	perfecta: todos los puntos en una recta
cercano a 1	fuerte
en torno a 0,5	moderada
cercano a 0	débil o inexistente

Es adimensional e invariante frente a cambios de escala y de origen, lo que permite compararlo entre estudios distintos.

Nota 2.2 . Dos advertencias que valen para toda la asignatura. La primera: **correlación no es causalidad**; puede haber una tercera variable que explique las dos, o ser coincidencia. La segunda: r mide *solo* relación lineal, así que $r \approx 0$ es compatible con una dependencia perfecta pero curva. Antes de interpretar r hay que mirar el diagrama de dispersión.



2.6 Regresión por mínimos cuadrados

El objetivo: la recta $y = a + bx$ que mejor aproxima la nube de puntos, minimizando la suma de los cuadrados de los residuos.

$$\min_{a,b} \sum_i (y_i - (a + bx_i))^2$$

Derivando e igualando a cero salen las **ecuaciones normales**, y de ellas:

$$b = \frac{s_{xy}}{s_x^2}, \quad a = \bar{y} - b\bar{x}$$

Propiedad	Consecuencia
Pasa por $(\bar{x}, \bar{y})$	el centro de gravedad de la nube
Los residuos suman cero	la recta no está sesgada
b tiene el signo de s_{xy}	la pendiente indica el sentido de la relación

Hay dos rectas de regresión distintas, la de Y sobre X y la de X sobre Y , y no coinciden salvo que $|r| = 1$. Minimizar errores en direcciones distintas, así que usar una para predecir en el sentido contrario es un error.

2.6.1 Bondad del ajuste

$$R^2 = r^2 = \frac{\text{varianza explicada}}{\text{varianza total}}$$

Se interpreta como el porcentaje de la variabilidad de Y que la recta explica. Con $R^2 = 0,81$, la recta da cuenta del 81 % de la variación.

Ejemplo 2.2 . Para cinco observaciones se tiene $\bar{x} = 3$, $\bar{y} = 7$, $s_x^2 = 2$, $s_y^2 = 8$ y $s_{xy} = 3,6$.

La pendiente es $b = 3,6/2 = 1,8$ y la ordenada $a = 7 - 1,8 \cdot 3 = 1,6$, así que la recta es $y = 1,6 + 1,8x$.

El coeficiente de correlación es $r = 3,6/\sqrt{2 \cdot 8} = 3,6/4 = 0,9$, y $R^2 = 0,81$: el ajuste explica el 81 % de la variabilidad.

Nota 2.3 . Extrapolar fuera del rango observado no está justificado. La recta se ha ajustado con datos de un intervalo y nada garantiza que la relación siga siendo lineal fuera de él. Predecir la facturación para un tamaño de empresa diez veces mayor que cualquiera de la muestra es aritmética, no estadística.

2.7 Ejercicios

Ejercicio 2.7.1. En una tabla de contingencia, $n_{11} = 20$, $n_{12} = 30$, $n_{21} = 10$ y $n_{22} = 15$. ¿Son independientes las dos variables?

Solución 2.7.1. Las marginales son $n_{1.} = 50$, $n_{2.} = 25$, $n_{.1} = 30$, $n_{.2} = 45$ y $N = 75$. La independencia exige $n_{ij} = n_{i.}n_{.j}/N$. Para la primera casilla: $50 \cdot 30/75 = 20$, que coincide. Comprobando las otras tres: 30, 10 y 15, todas coinciden. **Sí son independientes.**

Ejercicio 2.7.2. Un estudio encuentra $r = 0,85$ entre el número de bomberos enviados a un incendio y los daños causados. ¿Significa que los bomberos causan daños?

Solución 2.7.2. No. Hay una tercera variable, la magnitud del incendio, que explica las dos: a mayor incendio, más bomberos y más daños. Es un caso de correlación espuria por variable oculta. Para separar el efecto habría que comparar incendios de tamaño similar, es decir condicionar por esa tercera variable.

Ejercicio 2.7.3. Con $\bar{x} = 10$, $\bar{y} = 50$, $s_x = 2$, $s_y = 6$ y $r = 0,8$, hallar la recta de regresión de Y sobre X y predecir y para $x = 12$.

Solución 2.7.3. $s_{xy} = r s_x s_y = 0,8 \cdot 2 \cdot 6 = 9,6$, así que $b = 9,6/4 = 2,4$ y $a = 50 - 2,4 \cdot 10 = 26$. La recta es $y = 26 + 2,4x$, y para $x = 12$ predice $y = 54,8$.

Conviene añadir que $R^2 = 0,64$: el modelo explica el 64 % de la variabilidad, así que la predicción tiene un margen de error apreciable.

El análisis bidimensional está desarrollado en Castillo Manrique y Guijarro Garvi, 2006, Martín Pliego, 2004 y Newbold et al., 2013, con problemas resueltos en Hermoso Gutiérrez y Hernández Bastida, 2000 y Arnaldos García et al., 2003.

Números índices

Tema 3 del programa. Tasas de variación, índices elementales y sintéticos, índices de precios, cantidades y valor, enlace de series, deflación, participación de un grupo y los índices encadenados, con el IPC como ejemplo.

3.1 Tasas de variación

Tasa	Fórmula
Variación absoluta	$x_t - x_{t-1}$
Tasa de variación	$\frac{x_t - x_{t-1}}{x_{t-1}}$
Tasa acumulada en k periodos	$\frac{x_t}{x_{t-k}} - 1$
Tasa media anual	$\sqrt[k]{\frac{x_t}{x_{t-k}}} - 1$

Las tasas no se suman, se componen. Dos subidas consecutivas del 10 % no dan un 20 % sino un 21 %, porque la segunda se aplica sobre lo ya aumentado. La tasa media es geométrica, no aritmética, y es el mismo resultado del tema 1.

$$(1 + t_1)(1 + t_2) \cdots (1 + t_k) = 1 + T$$

3.2 Índices elementales

$$I_{t/0} = \frac{x_t}{x_0} \times 100$$

El periodo 0 es la **base**, que vale 100 por construcción. Un índice de 118 significa un aumento del 18 % respecto de la base.

Propiedad	Enunciado
Identidad	$I_{0/0} = 100$
Inversión	$I_{t/0} \cdot I_{0/t} = 10\,000$
Circular	$I_{t/0} = I_{t/s} \cdot I_{s/0}/100$
Proporcionalidad	multiplicar la serie por k no cambia el índice

Nota 3.1 . Un índice **no se puede sumar ni promediar directamente** entre periodos: es un cociente, no una magnitud. Y la variación entre dos índices no es su diferencia sino su cociente. De 120 a 132 la subida no es de 12 puntos porcentuales de crecimiento, es del 10 %.

3.3 Índices sintéticos

Cuando hay varios bienes, hay que agregarlos con algún criterio de ponderación.

Índice	Fórmula	Ponderaciones
Laspeyres de precios	$L_p = \frac{\sum p_t q_0}{\sum p_0 q_0} \times 100$	cantidades del periodo base
Paasche de precios	$P_p = \frac{\sum p_t q_t}{\sum p_0 q_t} \times 100$	cantidades del periodo actual
Fisher	$F = \sqrt{L \cdot P}$	media geométrica de los dos
Índice de valor	$V = \frac{\sum p_t q_t}{\sum p_0 q_0} \times 100$	—

Los índices de cantidades son los mismos intercambiando el papel de p y q .

	Laspeyres	Paasche
Cesta	fija, la del periodo base	cambia cada periodo
Coste de cálculo	bajo: una sola encuesta de consumo	alto: encuesta cada periodo
Sesgo	sobrestima la inflación	la subestima
Uso real	el habitual, incluido el IPC	contabilidad nacional

Por qué Laspeyres sobrestima: al mantener fija la cesta, no recoge que los consumidores sustituyen los bienes que más se encarecen por otros. Paasche tiene el sesgo contrario porque usa la cesta ya adaptada. El índice de Fisher, media geométrica de los dos, se llama «ideal» porque compensa los dos sesgos y cumple las propiedades de inversión y circular.

Proposición 3.1 (Descomposición del valor).

$$V = \frac{L_p \cdot P_q}{100} = \frac{P_p \cdot L_q}{100}$$

El índice de valor se descompone siempre en un índice de precios de un tipo por uno de cantidades del otro.

Ejemplo 3.1 . Dos bienes, con precios y cantidades:

	p_0	q_0	p_t	q_t
A	10	100	12	90
B	20	50	22	60

$$\sum p_0 q_0 = 1000 + 1000 = 2000; \sum p_t q_0 = 1200 + 1100 = 2300; \sum p_t q_t = 1080 + 1320 = 2400; \sum p_0 q_t = 900 + 1200 = 2100.$$

$L_p = 115$, $P_p = 2400/2100 \times 100 = 114,29$ y $F = \sqrt{115 \cdot 114,29} = 114,64$.

Se cumple lo anunciado: $L_p > P_p$. Y el índice de valor es $V = 2400/2000 \times 100 = 120$.

3.4 Enlace de series con distinta base

Cuando una serie cambia de base hay que **enlazarla** para poder comparar todo el periodo. Con un año solapado en las dos bases, el coeficiente de enlace es el cociente de los dos valores en ese año.

$$I_t^{\text{nuevo}} = I_t^{\text{viejo}} \times \frac{I_s^{\text{nuevo}}}{I_s^{\text{viejo}}}$$

Ejemplo 3.2 . Una serie con base 2010 vale 130 en 2015, y la nueva serie con base 2015 vale 100 ese mismo año. Para llevar los valores viejos a la base nueva se multiplican por $100/130 = 0,769$. Así, un 110 de 2012 en base 2010 pasa a ser 84,6 en base 2015.

3.5 Deflación

Convertir una serie en unidades monetarias corrientes a unidades constantes, es decir, eliminar el efecto de los precios.

$$x_t^{\text{real}} = \frac{x_t^{\text{nominal}}}{I_t} \times 100$$

Magnitud	Qué mide
Nominal, o corriente	el valor en euros de cada año
Real, o constante	el valor a precios de un año de referencia

Es la operación con más consecuencias prácticas del tema. Una serie de salarios que crece un 3 % anual con una inflación del 4 % está **cayendo** en términos reales, y solo la deflación lo hace visible.

Ejemplo 3.3 . Un salario pasa de 24 000 a 27 000 euros en cinco años, mientras el IPC pasa de 100 a 118.

El salario real final es $27\,000/118 \times 100 = 22\,881$ euros de base. Ha subido un 12,5 % en nominal y ha **bajado** un 4,7 % en real. Esa diferencia es todo lo que hay que entender del tema.

3.6 Participación de un grupo

Un índice general se descompone en la aportación de cada grupo de productos:

$$I_{\text{general}} = \sum_g w_g I_g$$

con w_g el peso del grupo. La **contribución** de un grupo a la variación total es su peso por su variación:

$$\text{contribuci3n}_g = w_g (I_g - 100)$$

Es lo que permite decir que «la subida del IPC se debe sobre todo a la energí», y distingue dos cosas que se confunden: un grupo puede subir muchísimo y aportar poco si su peso es pequeño.

Ejemplo 3.4 . La energí sube un 30 % con un peso del 8 %, y los alimentos un 4 % con un peso del 20 %.

Contribuciones: $0,08 \cdot 30 = 2,4$ puntos frente a $0,20 \cdot 4 = 0,8$ puntos. La energí sube casi ocho veces más y aporta tres veces más, no ocho: el peso amortigua.

3.7 Índices encadenados

En vez de una base fija durante años, se calcula cada periodo respecto del anterior y se multiplican los eslabones:

$$I_{t/0} = \prod_{s=1}^t \frac{I_{s/s-1}}{100} \times 100$$

Ventaja	Inconveniente
La cesta se actualiza cada año	pierde la propiedad de agregaci3n
Recoge la aparici3n de productos nuevos	no es aditivo entre componentes
Reduce el sesgo de Laspeyres	dificulta comparar periodos lejanos

3.7.1 El IPC

El índice de precios de consumo espaol es un **Laspeyres encadenado** con base 2021 y actualizaci3n anual de ponderaciones, calculado por el INE a partir de una cesta de varios cientos de artículos agrupados en doce divisiones.

Concepto	Definici3n
Inflaci3n interanual	variaci3n del IPC respecto del mismo mes del ańo anterior
Inflaci3n mensual	respecto del mes anterior
Inflaci3n subyacente	excluye alimentos no elaborados y energí

La subyacente existe porque los dos componentes excluidos son muy volátiles y responden a factores externos, así que la general oscila por razones que no reflejan la tendencia de fondo de los precios.

3.8 Ejercicios

Ejercicio 3.8.1. Una serie sube un 8 % un ańo y baja un 5 % al siguiente. ¿Cuál es la variaci3n acumulada y cuál la media anual?

Solución 3.8.1. Acumulada: $1,08 \cdot 0,95 = 1,026$, es decir un 2,6 %. La media anual es $\sqrt{1,026} = 1,0129$, o sea un 1,29 %. Restar los porcentajes daría un 3 %, y la media aritmética un 1,5 %: los dos están mal.

Ejercicio 3.8.2. El PIB nominal crece un 6 % y el deflactor un 4 %. ¿Cuánto crece el PIB real?

Solución 3.8.2. $1,06/1,04 = 1,0192$: un 1,92 %. La aproximación por diferencia, $6 - 4 = 2$, es buena porque las tasas son pequeñas, pero deja de serlo con inflaciones altas: con un nominal del 50 % y un deflactor del 40 %, la diferencia daría 10 % y el valor correcto es $1,5/1,4 - 1 = 0,071$, es decir un 7,1 %.

Ejercicio 3.8.3. Con los datos del ejemplo de la sección de índices sintéticos, comprobar la descomposición del índice de valor.

Solución 3.8.3. $L_p = 115$ y el índice de cantidades de Paasche es $P_q = \sum p_t q_t / \sum p_t q_0 \times 100 = 2400/2300 \times 100 = 104,35$.

Su producto dividido por 100 da $115 \cdot 104,35/100 = 120$, que es exactamente el índice de valor. La identidad se cumple, y sirve como comprobación de que los cuatro sumatorios están bien calculados.

Los números índices están desarrollados en Martín Pliego, 2004 y Newbold et al., 2013, con problemas resueltos en Hermoso Gutiérrez y Hernández Bastida, 2000 y Amor Pulido, 2016.

Análisis descriptivo de series cronológicas

Tema 4 del programa. Definición y representación, componentes y modelos, tendencia secular, variación estacional, desestacionalización y predicción.

4.1 Definición

Definición 4.1 (Serie cronológica). Sucesión de observaciones de una variable ordenadas en el tiempo y tomadas a intervalos regulares.

Lo que distingue una serie temporal de un conjunto de datos cualquiera es que **el orden importa**: las observaciones no son intercambiables ni independientes, y por eso las técnicas de los temas anteriores no bastan.

Frecuencia	Periodos por año
Anual	1
Trimestral	4
Mensual	12
Semanal	52
Diaria	365

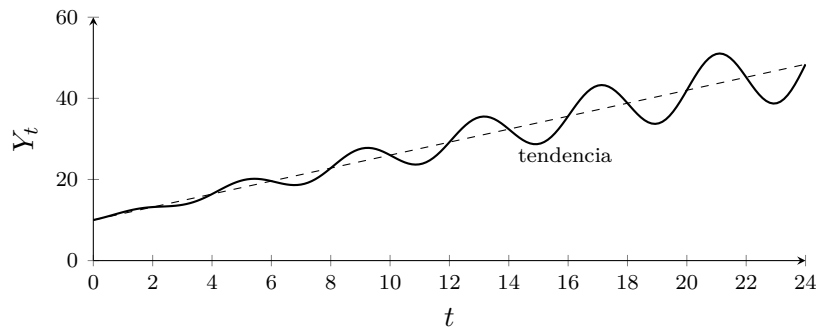
4.2 Componentes

Componente	Símbolo	Qué recoge
Tendencia	T_t	el movimiento de fondo a largo plazo
Estacionalidad	E_t	oscilaciones que se repiten cada año
Ciclo	C_t	oscilaciones de varios años, sin periodo fijo
Irregular	I_t	lo que queda: aleatorio e imprevisible

En la práctica el ciclo y la tendencia se estudian juntos, porque separarlos exige series muy largas.

4.2.1 Los dos modelos

Modelo	Expresión	Cuándo
Aditivo	$Y_t = T_t + E_t + I_t$	la oscilación estacional es de tamaño constante
Multiplicativo	$Y_t = T_t \cdot E_t \cdot I_t$	la oscilación crece con el nivel de la serie



Cómo se elige el modelo: se mira la gráfica. Si las oscilaciones alrededor de la tendencia mantienen su amplitud, aditivo; si se abren en abanico, multiplicativo. Y el multiplicativo se convierte en aditivo tomando logaritmos, que es lo que se hace cuando conviene trabajar con un solo esquema.

4.3 Tendencia secular

4.3.1 Ajuste por mínimos cuadrados

Se ajusta una función del tiempo con la técnica del tema 2, tomando t como variable explicativa:

Función	Cuándo
Lineal, $T_t = a + bt$	crecimiento constante en unidades
Parabólica, $T_t = a + bt + ct^2$	crecimiento acelerado o desacelerado
Exponencial, $T_t = ab^t$	crecimiento a tasa constante

La exponencial se ajusta tomando logaritmos: $\ln T_t = \ln a + t \ln b$, que es lineal.

Un truco de cálculo que ahorra mucho: codificar el tiempo de forma que $\sum t = 0$. Con un número impar de periodos se numera $\dots, -2, -1, 0, 1, 2, \dots$; con número par, $\dots, -3, -1, 1, 3, \dots$. Entonces $a = \bar{y}$ y $b = \sum ty / \sum t^2$, sin resolver ningún sistema.

4.3.2 Medias móviles

Suavizan la serie promediando cada observación con sus vecinas:

$$MM_k(t) = \frac{1}{k} \sum_{i=-(k-1)/2}^{(k-1)/2} Y_{t+i}$$

Elección de k	Efecto
Igual al número de periodos del año	elimina la estacionalidad por completo

Elección de k	Efecto
Pequeño	suaviza poco, conserva detalle
Grande	suaviza mucho, pierde los extremos de la serie

Con k **par** —trimestres o meses— la media móvil no queda centrada en un periodo, y hay que **centrarla** promediando dos consecutivas. Es el paso que más se olvida al desestacionalizar datos trimestrales.

Nota 4.1 . Toda media móvil **pierde datos en los extremos**: con $k = 4$ centrada se pierden dos observaciones al principio y dos al final. Y son justo las del final las que interesan para predecir, así que la media móvil sirve para describir el pasado y no para extrapolar.

4.4 Variación estacional

El procedimiento estándar, para un modelo multiplicativo con datos trimestrales:

1. Calcular la media móvil centrada de orden 4: es la tendencia-ciclo.
2. Dividir la serie original entre ella: queda estacionalidad más irregular.
3. Promediar esos cocientes por trimestre: se cancela lo irregular y quedan los **índices estacionales brutos**.
4. **Ajustarlos** para que sumen 400 —o 1200 con datos mensuales—, multiplicando por el factor que corresponda.

Ejemplo 4.1 . Los índices brutos de una serie trimestral son 92, 105, 118 y 81, que suman 396.

El factor de ajuste es $400/396 = 1,0101$, y los índices definitivos quedan 92,9, 106,1, 119,2 y 81,8, que ya suman 400.

Se leen así: el tercer trimestre está un 19,2 % por encima de lo que correspondería a la tendencia, y el cuarto un 18,2 % por debajo.

4.5 Desestacionalización

$$Y_t^{des} = \frac{Y_t}{E_t} \times 100 \quad (\text{multiplicativo}), \quad Y_t^{des} = Y_t - E_t \quad (\text{aditivo})$$

Es la operación que permite comparar periodos consecutivos: sin ella, un descenso del primer al segundo trimestre puede ser pura estacionalidad y no un empeoramiento.

Por eso los organismos estadísticos publican las dos series. El dato bruto sirve para comparar con el mismo trimestre del año anterior, y el desestacionalizado para comparar con el trimestre inmediatamente anterior.

4.6 Predicción

Con las componentes estimadas, la predicción se construye recomponiéndolas:

$$\hat{Y}_{t+h} = \hat{T}_{t+h} \times \hat{E}_{t+h}$$

1. Extrapolar la tendencia al periodo futuro con la función ajustada.
2. Multiplicar por el índice estacional del trimestre o mes correspondiente.
3. El componente irregular no se predice: su esperanza es cero, o uno en el modelo multiplicativo.

Ejemplo 4.2 . La tendencia ajustada de una serie trimestral es $T_t = 200 + 5t$ con t en trimestres, y el índice estacional del tercer trimestre es 119,2. Para el trimestre $t = 25$:

$$\hat{Y}_{25} = (200 + 125) \times \frac{119,2}{100} = 325 \times 1,192 = 387,4$$

Nota 4.2 . Toda predicción por descomposición supone que **el patrón se mantiene**, y ese supuesto es tanto más frágil cuanto más lejos se extrapola. Una predicción a un trimestre es razonable; a cinco años es una extrapolación de la recta con adornos, y las crisis económicas son justamente los momentos en que el patrón se rompe.

4.7 Ejercicios

Ejercicio 4.7.1. Una serie trimestral tiene valores 100, 120, 140, 90 en el primer año. Calcular la media móvil centrada de orden 4 correspondiente al tercer trimestre, sabiendo que el primer trimestre del año siguiente vale 110.

Solución 4.7.1. La media móvil de orden 4 sin centrar de los cuatro primeros valores es $(100 + 120 + 140 + 90)/4 = 112,5$, situada entre el segundo y el tercer trimestre. La siguiente, con 120, 140, 90, 110, es 115, situada entre el tercero y el cuarto.

La media móvil **centrada** en el tercer trimestre es la media de esas dos: $(112,5 + 115)/2 = 113,75$.

Ejercicio 4.7.2. Los índices estacionales de una serie mensual suman 1188. Ajustarlos.

Solución 4.7.2. Con datos mensuales deben sumar 1200, así que el factor es $1200/1188 = 1,0101$. Cada índice se multiplica por él. Un índice bruto de 95 pasa a 95,96.

Ejercicio 4.7.3. El PIB trimestral desestacionalizado cae un 0,3 % respecto del trimestre anterior, y el dato bruto sube un 2 %. ¿Cómo se interpreta?

Solución 4.7.3. La subida del 2 % es lo que cabía esperar por la estacionalidad y algo más: comparado con lo que un trimestre así suele aportar, la economía se ha quedado por debajo. La señal relevante es la del dato desestacionalizado, que indica contracción. Comparar datos brutos de trimestres consecutivos mezcla la estacionalidad con la evolución real y es el error más habitual al leer estas cifras.

El análisis de series cronológicas está desarrollado en Martín Pliego, 2004 y Newbold et al., 2013, con problemas resueltos en Hermoso Gutiérrez y Hernández Bastida, 2000 y Amor Pulido, 2016.

Probabilidad

Tema 5 del programa. Definición y asignación de probabilidades, probabilidad condicionada, sucesos dependientes e independientes, y las fórmulas de la probabilidad total y de Bayes.

5.1 Experimentos y sucesos

Concepto	Definición
Experimento aleatorio	el resultado no se conoce de antemano
Espacio muestral Ω	conjunto de resultados posibles
Suceso	subconjunto de Ω
Suceso elemental	un solo resultado
Sucesos incompatibles	$A \cap B = \emptyset$
Sistema completo	sucesos incompatibles cuya unión es Ω

Las operaciones con sucesos son las de conjuntos, con las leyes de De Morgan incluidas:

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \quad \overline{A \cap B} = \bar{A} \cup \bar{B}$$

5.2 Definición y asignación

Definición 5.1 (Axiomas de Kolmogorov). Una probabilidad es una función P sobre los sucesos tal que

- $P(A) \geq 0$ para todo suceso A ;
- $P(\Omega) = 1$;
- si A y B son incompatibles, $P(A \cup B) = P(A) + P(B)$.

De los tres axiomas se deducen las propiedades de uso diario:

Propiedad	Expresión
Complementario	$P(\bar{A}) = 1 - P(A)$
Suceso imposible	$P(\emptyset) = 0$
Monotonía	si $A \subseteq B$, $P(A) \leq P(B)$
Acotación	$0 \leq P(A) \leq 1$
Unión general	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

La regla del complementario es la más rentable: calcular «al menos uno» casi siempre es más fácil por «ninguno».

5.2.1 Formas de asignar probabilidades

Método	Cómo	Cuándo vale
Laplace	casos favorables entre casos posibles	resultados equiprobables
Frecuentista	límite de la frecuencia relativa	el experimento se puede repetir
Subjetiva	grado de creencia, coherente con los axiomas	sucesos no repetibles

Nota 5.1 . La regla de Laplace exige equiprobabilidad, y aplicarla sin comprobarla es el error clásico. Al lanzar dos dados y sumar, los once resultados de 2 a 12 **no** son equiprobables: el 7 sale seis veces más que el 2, porque el espacio muestral equiprobable es el de los 36 pares ordenados, no el de las sumas.

5.3 Probabilidad condicionada

Definición 5.2 .

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$

Condicionar es **reducir el espacio muestral** a B y renormalizar. Es la operación que formaliza «sabiendo que ha ocurrido B ».

De la definición sale la **regla del producto**:

$$P(A \cap B) = P(B) P(A | B) = P(A) P(B | A)$$

y su generalización en cadena, que es la forma de calcular probabilidades de extracciones sucesivas sin reposición.

Ejemplo 5.1 . Una urna con 5 bolas blancas y 3 negras. La probabilidad de sacar dos blancas consecutivas sin reposición es

$$P(B_1 \cap B_2) = \frac{5}{8} \cdot \frac{4}{7} = \frac{20}{56} = \frac{5}{14}$$

La segunda fracción es condicionada: tras sacar una blanca quedan 4 blancas de 7 bolas.

5.4 Independencia

Definición 5.3 . A y B son independientes si

$$P(A \cap B) = P(A) P(B)$$

equivalentemente, si $P(A | B) = P(A)$ cuando $P(B) > 0$.

Confusión frecuente	Aclaración
Independientes = incompatibles	son cosas opuestas: si son incompatibles y tienen probabilidad no nula, saber que ocurre uno descarta el otro, luego dependen
Independientes dos a dos = independientes	la independencia mutua es más fuerte y exige comprobar también las intersecciones triples

Proposición 5.1 . Si A y B son independientes, también lo son $\bar{A}$ y B , A y $\bar{B}$, y $\bar{A}$ y $\bar{B}$.

La independencia se supone o se demuestra, nunca se intuye. En el lanzamiento repetido de una moneda es razonable suponerla; entre dos préstamos de la misma región económica no lo es, y suponerla fue una de las causas técnicas de la crisis financiera de 2008.

5.5 Probabilidad total y Bayes

Teorema 5.1 (Probabilidad total). Si $\{A_1, \dots, A_n\}$ es un sistema completo de sucesos con probabilidad no nula,

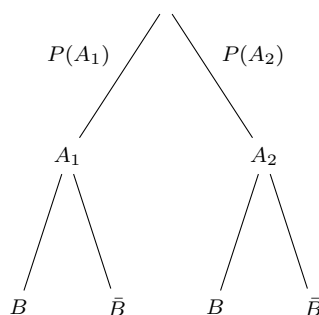
$$P(B) = \sum_{i=1}^n P(A_i) P(B | A_i)$$

Teorema 5.2 (Bayes). Con las mismas hipótesis y $P(B) > 0$,

$$P(A_j | B) = \frac{P(A_j) P(B | A_j)}{\sum_i P(A_i) P(B | A_i)}$$

Nombre	Papel
$P(A_i)$	probabilidades a priori : lo que se sabía antes
$P(B A_i)$	verosimilitudes: lo que cada causa predice
$P(A_j B)$	probabilidad a posteriori : lo que se cree después de observar

Bayes es la fórmula que **invierte el condicionamiento**: de la probabilidad del efecto dada la causa a la de la causa dado el efecto. Es la base del diagnóstico, del filtrado de correo y de la clasificación automática.



Ejemplo 5.2 (La trampa de la prueba diagnóstica). Una enfermedad afecta al 1 % de la población. Una prueba detecta al 99 % de los enfermos y da falso positivo en el 5 % de los sanos. Si alguien da positivo, ¿qué probabilidad tiene de estar enfermo?

$$P(E | +) = \frac{0,01 \cdot 0,99}{0,01 \cdot 0,99 + 0,99 \cdot 0,05} = \frac{0,0099}{0,0594} = 0,167$$

Solo un 16,7 %, pese a que la prueba acierta el 99 % de las veces en los enfermos. La razón es que los sanos son cien veces más numerosos, así que sus falsos positivos —cinco de cada cien— superan con mucho a los verdaderos positivos.

Ese resultado se llama **falacia de la tasa base**, y es el argumento decisivo contra el cribado masivo de enfermedades poco frecuentes: con prevalencia baja, la mayoría de los positivos son falsos por buena que sea la prueba.

5.6 Ejercicios

Ejercicio 5.6.1. Se lanzan tres monedas. Calcular la probabilidad de obtener al menos una cara.

Solución 5.6.1. Por el complementario: la probabilidad de ninguna cara es $(1/2)^3 = 1/8$, así que la pedida es $7/8$. Contar directamente los siete casos favorables entre ocho lleva al mismo sitio con más trabajo, y con cinco monedas la diferencia ya es notable.

Ejercicio 5.6.2. En una empresa, el 60 % de los empleados son mujeres y el 30 % de las mujeres y el 20 % de los hombres tienen titulación superior. Elegido un empleado con titulación superior, ¿cuál es la probabilidad de que sea mujer?

Solución 5.6.2. Por probabilidad total, $P(T) = 0,6 \cdot 0,3 + 0,4 \cdot 0,2 = 0,18 + 0,08 = 0,26$. Y por Bayes,

$$P(M | T) = \frac{0,18}{0,26} = 0,692$$

Un 69,2 %, superior al 60 % de partida porque la titulación es más frecuente entre las mujeres de esa empresa.

Ejercicio 5.6.3. Dos sucesos con $P(A) = 0,4$ y $P(B) = 0,5$ son incompatibles. ¿Son independientes?

Solución 5.6.3. No. Al ser incompatibles, $P(A \cap B) = 0$, mientras que $P(A)P(B) = 0,20 \neq 0$. De hecho son fuertemente dependientes: saber que ha ocurrido A hace imposible B . Dos sucesos de probabilidad no nula no pueden ser incompatibles e independientes a la vez.

La teoría de la probabilidad está desarrollada en Canavos, 1989 y Castillo Manrique y Guijarro Garvi, 2006, con problemas resueltos en Hermoso Gutiérrez y Hernández Bastida, 2000 y Amor Pulido, 2016.

Variables aleatorias y distribuciones de probabilidad

Tema 6 del programa. Concepto de variable aleatoria, función de distribución, casos discreto y continuo, valor esperado, momentos, otras características y variables bidimensionales.

6.1 Concepto

Definición 6.1 (Variable aleatoria). Función que asigna un número real a cada resultado del experimento, $X : \Omega \rightarrow \mathbb{R}$.

Es lo que permite pasar de sucesos a números y aplicar todo el aparato de los temas 1 y 2. La diferencia con una variable estadística está en el punto de partida: la estadística describe datos ya observados, y la aleatoria describe el **modelo** que los genera.

Tipo	Valores	Se describe con
Discreta	conjunto finito o numerable	función de masa $P(X = x_i)$
Continua	cualquier valor de un intervalo	función de densidad $f(x)$

6.2 Función de distribución

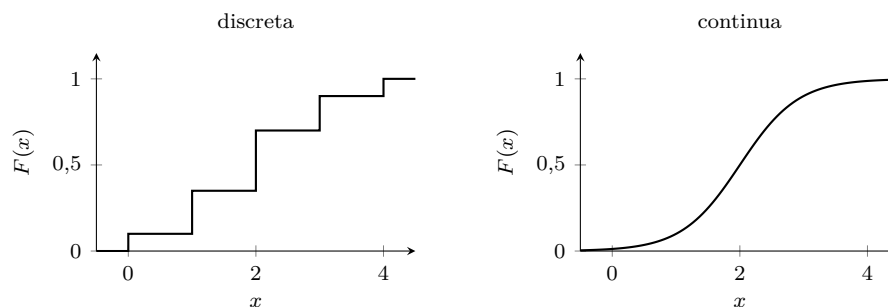
Definición 6.2 .

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

Proposición 6.1 (Propiedades). F es no decreciente, continua por la derecha, y cumple $\lim_{x \rightarrow -\infty} F(x) = 0$ y $\lim_{x \rightarrow +\infty} F(x) = 1$.

Existe siempre, sea la variable discreta o continua, y por eso es la herramienta que unifica los dos casos.

	Discreta	Continua
$F(x)$	escalonada	continua
Se obtiene	$F(x) = \sum_{x_i \leq x} P(X = x_i)$	$F(x) = \int_{-\infty}^x f(t) dt$
Y al revés	el salto en x_i es $P(X = x_i)$	$f(x) = F'(x)$
$P(X = a)$	el salto, puede ser positivo	siempre 0
$P(a < X \leq b)$	$F(b) - F(a)$	$F(b) - F(a)$



Que $P(X = a) = 0$ en el caso continuo tiene una consecuencia práctica: **los signos de desigualdad son indiferentes**, $P(X < a) = P(X \leq a)$. En el discreto no, y confundirlos cambia el resultado.

6.2.1 La densidad

$$f(x) \geq 0, \quad \int_{-\infty}^{\infty} f(x) dx = 1, \quad P(a \leq X \leq b) = \int_a^b f(x) dx$$

Nota 6.1 . La densidad **no es una probabilidad** y puede superar 1: lo que vale 1 es el área total, no la altura. En una uniforme sobre $[0, 0.5]$ la densidad vale 2 en todo el intervalo. Lo que tiene significado probabilístico es siempre el área bajo la curva.

6.3 Valor esperado

$$E[X] = \sum_i x_i P(X = x_i) \quad \text{o} \quad E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Es el análogo de la media del tema 1, y se interpreta como el valor medio a largo plazo si el experimento se repite muchas veces.

Propiedad	Expresión
Linealidad	$E[aX + b] = aE[X] + b$
Suma	$E[X + Y] = E[X] + E[Y]$, siempre
Producto	$E[XY] = E[X]E[Y]$ solo si son independientes
De una función	$E[g(X)] = \sum g(x_i)P(x_i)$ o la integral

La linealidad de la suma no exige independencia y la del producto sí: es la asimetría que más se olvida.

Nota 6.2 . $E[g(X)] \neq g(E[X])$ en general. Con g convexa, la desigualdad de Jensen da $E[g(X)] \geq g(E[X])$. De ahí que la rentabilidad media de una inversión volátil sea menor que la que sugiere aplicar la fórmula al rendimiento medio, y que en finanzas la volatilidad tenga coste aunque la media no cambie.

6.4 Momentos y otras características

$$\alpha_r = E[X^r], \quad \mu_r = E[(X - \mu)^r]$$

Medida	Definición
Media	$\mu = E[X] = \alpha_1$
Varianza	$\sigma^2 = \text{Var}(X) = E[(X - \mu)^2] = E[X^2] - \mu^2$
Desviación típica	σ
Coeficiente de variación	σ/μ
Asimetría	μ_3/σ^3
Curtosis	$\mu_4/\sigma^4 - 3$
Moda	donde la masa o la densidad es máxima
Mediana	el valor con $F(x) = 0,5$
Cuantil p	el valor con $F(x) = p$

Propiedad de la varianza	Expresión
Traslación	$\text{Var}(X + b) = \text{Var}(X)$
Escala	$\text{Var}(aX) = a^2 \text{Var}(X)$
Suma	$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y)$
Suma con independencia	$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

La última fila es el fundamento de la **diversificación**: la varianza de una cartera de activos independientes crece con el número de activos, pero su media también, y el riesgo relativo baja como $1/\sqrt{n}$.

Teorema 6.1 (Desigualdad de Chebyshev). Para cualquier variable con media μ y varianza σ^2 finitas y todo $k > 0$,

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Es una cota **universal**: vale sin conocer la distribución. Con $k = 2$ dice que como mucho el 25 % de la masa está a más de dos desviaciones típicas de la media, sea la distribución la que sea. A cambio de esa generalidad es muy poco fina: para una normal el valor real es un 4,6 %.

6.5 Variables aleatorias bidimensionales

Un par (X, Y) se describe con su distribución conjunta, y de ella salen las marginales y las condicionadas igual que en el tema 2.

Concepto	Discreto
Conjunta	$p_{ij} = P(X = x_i, Y = y_j)$
Marginal	$p_{i\cdot} = \sum_j p_{ij}$
Condicionada	$P(X = x_i Y = y_j) = p_{ij}/p_{\cdot j}$

Definición 6.3 (Independencia de variables aleatorias). X e Y son independientes si

$$F(x, y) = F_X(x) F_Y(y) \quad \text{para todo } (x, y)$$

equivalentemente, si la conjunta factoriza como producto de las marginales en todos los puntos.

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y], \quad \rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Independencia implica covarianza nula, y no al revés, exactamente igual que en el tema 2. La única familia donde las dos condiciones coinciden es la normal multivariante, y por eso allí «incorrelado» sí significa «independiente».

Ejemplo 6.1 . Un inversor reparte su capital entre dos activos con rentabilidades X e Y , con $E[X] = E[Y] = 8$ y $\sigma_X = \sigma_Y = 10$. La cartera es $C = 0,5X + 0,5Y$.

Su rentabilidad esperada es 8 en cualquier caso. Su varianza,

$$\text{Var}(C) = 0,25\sigma_X^2 + 0,25\sigma_Y^2 + 0,5 \text{Cov}(X, Y) = 50 + 50\rho$$

Con $\rho = 1$ da 100, la misma que un solo activo. Con $\rho = 0$, 50, y la desviación típica baja de 10 a 7,07. Con $\rho = -1$, cero: el riesgo desaparece. **Diversificar solo sirve si los activos no están perfectamente correlados**, y ese es todo el argumento de la teoría de carteras.

6.6 Ejercicios

Ejercicio 6.6.1. Una variable continua tiene densidad $f(x) = kx$ en $[0, 2]$ y 0 fuera. Hallar k , la función de distribución, la media y la varianza.

Solución 6.6.1. $\int_0^2 kx \, dx = 2k = 1$, así que $k = 1/2$.

$F(x) = \int_0^x t/2 \, dt = x^2/4$ en $[0, 2]$, con 0 antes y 1 después.

$E[X] = \int_0^2 x^2/2 \, dx = 4/3$. Y $E[X^2] = \int_0^2 x^3/2 \, dx = 2$, luego $\text{Var}(X) = 2 - 16/9 = 2/9$.

Ejercicio 6.6.2. Un juego paga 10 euros con probabilidad 0,1 y cobra 2 euros con probabilidad 0,9. ¿Conviene jugar?

Solución 6.6.2. $E[G] = 10 \cdot 0,1 - 2 \cdot 0,9 = 1 - 1,8 = -0,8$ euros por partida. El juego es desfavorable: a largo plazo se pierden 80 céntimos por jugada. Que exista una ganancia posible de 10 euros no lo cambia; lo que decide es la esperanza.

Ejercicio 6.6.3. Una variable tiene media 100 y desviación típica 15. Acotar la probabilidad de que tome un valor entre 70 y 130.

Solución 6.6.3. El intervalo es $\mu \pm 2\sigma$. Por Chebyshev, $P(|X - 100| \geq 30) \leq 1/4$, luego $P(70 < X < 130) \geq 0,75$. Es una cota válida para cualquier distribución; si además se supiera que es normal, el valor exacto sería 0,954.

Las variables aleatorias están desarrolladas en Canavos, 1989 y Castillo Manrique y Guijarro Garvi, 2006, con problemas resueltos en Hermoso Gutiérrez y Hernández Bastida, 2000 y Newbold et al., 2013.

Distribuciones discretas de probabilidad

Tema 7 del programa. Las distribuciones uniforme discreta, binomial, de Poisson, hipergeométrica y geométrica.

7.1 Uniforme discreta

Todos los valores $1, 2, \dots, n$ tienen la misma probabilidad.

$$P(X = k) = \frac{1}{n}, \quad E[X] = \frac{n+1}{2}, \quad \text{Var}(X) = \frac{n^2-1}{12}$$

Modela el dado equilibrado, el sorteo puro y la generación de números aleatorios. Es la distribución de referencia cuando no hay ninguna razón para preferir un resultado.

7.2 Binomial

Definición 7.1 . $X \sim B(n, p)$ cuenta los éxitos en n pruebas de Bernoulli independientes con probabilidad de éxito p constante:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

$$E[X] = np, \quad \text{Var}(X) = np(1-p)$$

Las tres hipótesis son las que hay que comprobar antes de usarla:

Hipótesis	Qué exige
Número fijo de pruebas	n se decide de antemano
Independencia	el resultado de una no afecta a las demás
Probabilidad constante	p no cambia entre pruebas

Muestrear **sin reposición** rompe las dos últimas, y entonces la distribución correcta es la hipergeométrica.

Proposición 7.1 (Propiedades). La suma de binomiales independientes con la misma p es binomial: $B(n_1, p) + B(n_2, p) = B(n_1 + n_2, p)$. La distribución es simétrica si $p = 0,5$, asimétrica a

la derecha si $p < 0,5$ y a la izquierda si $p > 0,5$.

Ejemplo 7.1 . Una máquina produce un 3 % de piezas defectuosas. En un lote de 20 piezas, la probabilidad de que haya como mucho una defectuosa es

$$P(X \leq 1) = 0,97^{20} + 20 \cdot 0,03 \cdot 0,97^{19} = 0,5438 + 0,3364 = 0,8802$$

Es decir, un 88 % de los lotes pasan un control que rechaza a partir de dos defectuosas.

7.3 Poisson

Definición 7.2 . $X \sim P(\lambda)$ cuenta sucesos raros en un intervalo fijo de tiempo o espacio:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

$$E[X] = \text{Var}(X) = \lambda$$

Que la media y la varianza coincidan es su rasgo distintivo, y sirve de comprobación empírica: si en unos datos de recuento la varianza supera con mucho a la media, el modelo de Poisson no es adecuado.

Dónde aparece: llamadas a una centralita por hora, clientes que entran en una tienda, averías de una máquina, erratas por página, siniestros de una cartera de seguros.

Teorema 7.1 (Aproximación de la binomial). Si n es grande y p pequeño con $np = \lambda$ moderado, entonces $B(n, p) \approx P(\lambda)$. En la práctica se aplica con $n \geq 30$ y $p \leq 0,1$.

La aproximación es útil porque el factorial de la binomial se vuelve inmanejable con n grande, mientras que la fórmula de Poisson no depende de n .

Ejemplo 7.2 . Una centralita recibe 3 llamadas por minuto de media. La probabilidad de recibir 5 en un minuto es

$$P(X = 5) = \frac{e^{-3} 3^5}{5!} = \frac{0,0498 \cdot 243}{120} = 0,1008$$

Y la de no recibir ninguna, $e^{-3} = 0,0498$.

Para un intervalo de dos minutos hay que **reescalar** el parámetro: $\lambda = 6$, no $\lambda = 3$. Es el error más frecuente con esta distribución.

7.4 Hipergeométrica

Definición 7.3 . Extracciones **sin reposición**: de una población de N elementos con K éxitos, se toman n y se cuenta cuántos son éxitos.

$$P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

$$E[X] = n \frac{K}{N}, \quad \text{Var}(X) = n \frac{K}{N} \left(1 - \frac{K}{N}\right) \frac{N-n}{N-1}$$

La media es la misma que la de una binomial con $p = K/N$; lo que cambia es la varianza, multiplicada por el **factor de corrección para poblaciones finitas** $(N-n)/(N-1)$, que es menor que 1: muestrear sin reposición reduce la variabilidad porque cada extracción aporta información sobre las siguientes.

Proposición 7.2 . Si n es pequeño frente a N —en la práctica, $n \leq 0,05N$ — la hipergeométrica se aproxima bien por la binomial con $p = K/N$.

Es la razón de que en una encuesta a mil personas de un país entero se use la binomial sin más: la población es tan grande que reponer o no reponer es indiferente.

Ejemplo 7.3 . Un lote de 50 piezas contiene 8 defectuosas. Se toman 5 al azar sin reposición. La probabilidad de que ninguna sea defectuosa es

$$P(X = 0) = \frac{\binom{8}{0} \binom{42}{5}}{\binom{50}{5}} = \frac{850\,668}{2\,118\,760} = 0,4015$$

La binomial con $p = 0,16$ daría $0,84^5 = 0,4182$. La diferencia es apreciable porque 5 frente a 50 es el 10 % de la población, por encima del umbral del 5 %.

7.5 Geométrica

Definición 7.4 . Número de pruebas hasta el primer éxito, en pruebas de Bernoulli independientes:

$$P(X = k) = (1 - p)^{k-1} p, \quad k = 1, 2, 3, \dots$$

$$E[X] = \frac{1}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2}$$

La media $1/p$ es muy intuitiva: si la probabilidad de éxito es $1/6$, se esperan 6 intentos.

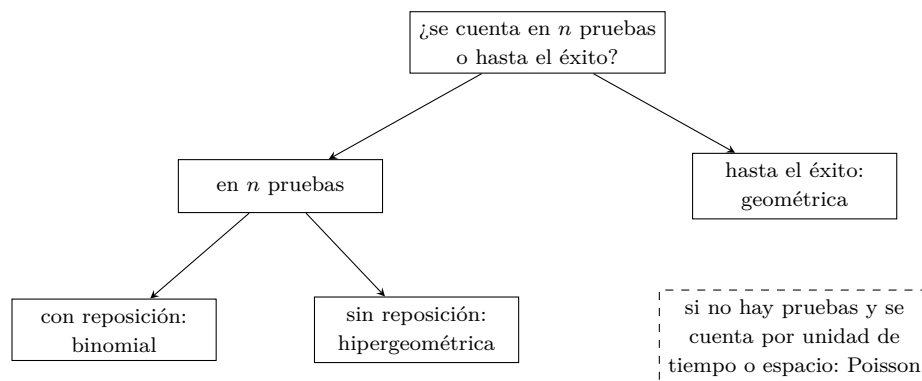
Proposición 7.3 (Falta de memoria).

$$P(X > m + n \mid X > m) = P(X > n)$$

Es la única distribución discreta con esa propiedad, y su lectura es tajante: **haber fallado muchas veces no acerca el éxito**. La creencia contraria es la falacia del jugador, y esta fórmula es su refutación exacta.

7.6 Cómo se elige

Situación	Distribución
Resultados equiprobables	uniforme discreta
Éxitos en n pruebas, con reposición	binomial
Éxitos en n pruebas, sin reposición	hipergeométrica
Sucesos raros por unidad de tiempo o espacio	Poisson
Pruebas hasta el primer éxito	geométrica



7.7 Ejercicios

Ejercicio 7.7.1. Un examen tipo test tiene 10 preguntas con 4 opciones cada una. Respondiendo al azar, ¿cuál es la probabilidad de acertar al menos 5?

Solución 7.7.1. $X \sim B(10, 0,25)$. Se calcula por el complementario:

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0,9219 = 0,0781$$

Un 7,8 %. La media es $np = 2,5$ aciertos, así que acertar 5 es ya el doble de lo esperado.

Ejercicio 7.7.2. Una web recibe 120 visitas por hora. ¿Cuál es la probabilidad de que en un minuto concreto no reciba ninguna?

Solución 7.7.2. Hay que reescalar: 120 por hora son $\lambda = 2$ por minuto. Entonces $P(X = 0) = e^{-2} = 0,1353$. Usar $\lambda = 120$ daría un número absurdo, y es el error que la reescala evita.

Ejercicio 7.7.3. Un comercial cierra una venta en el 20 % de sus visitas. ¿Cuántas visitas necesita de media hasta la primera venta? ¿Cuál es la probabilidad de necesitar más de 5?

Solución 7.7.3. Es geométrica con $p = 0,2$, así que $E[X] = 1/0,2 = 5$ visitas.

$P(X > 5) = (1 - p)^5 = 0,8^5 = 0,3277$: casi un tercio de las veces hará falta más de cinco. Y por la falta de memoria, si ya lleva cinco fracasos, la probabilidad de necesitar más de cinco adicionales sigue siendo 0,3277.

Las distribuciones discretas están desarrolladas en Canavos, 1989 y Castillo Manrique y Guijarro Garvi, 2006, con problemas resueltos en Hermoso Gutiérrez y Hernández Bastida, 2000, Newbold et al., 2013 y Amor Pulido, 2016.

Temario práctico

La guía indica que el temario práctico coincide con el teórico: se resuelven ejercicios a mano y con software estadístico, y el **objetivo es hacer un análisis descriptivo completo de datos reales** obtenidos de bases de datos públicas.

8.1 De dónde salen los datos

Fuente	Qué ofrece
INE	población, IPC, EPA, contabilidad nacional
Eurostat	lo mismo comparable entre países de la UE
Banco de España	tipos, crédito, balanza de pagos
Datos abiertos de ayuntamientos y comunidades	padrón, tráfico, presupuestos
Banco Mundial, FMI, OCDE	series internacionales largas

Lo que se pide antes de calcular nada:

1. **Identificar la unidad de observación:** qué es una fila.
2. Determinar el tipo de cada variable, que decide qué medidas valen.
3. Anotar el periodo y la fuente exacta, para poder rehacer el trabajo.
4. Detectar **datos ausentes** y decidir qué hacer con ellos, diciéndolo.

Nota 8.1 . Lo que más falsea un análisis no es un cálculo mal hecho, es un dato mal entendido. Una serie de «paro registrado» y otra de «tasa de paro EPA» miden cosas distintas y no se pueden comparar ni encadenar. Leer la definición metodológica de la fuente es parte del trabajo, no un trámite.

8.2 Práctica 1. Análisis descriptivo unidimensional

Actividad	Qué se obtiene
Tabla de frecuencias	absolutas, relativas y acumuladas
Histograma o diagrama de barras	la forma de la distribución
Medidas de posición	media, mediana, moda, cuartiles
Medidas de dispersión	recorrido, varianza, desviación típica, <i>CV</i>
Medidas de forma	asimetría y curtosis
Diagrama de caja	mediana, cuartiles y valores atípicos

Lo que se pide interpretar, y es donde está la nota:

- **Si media y mediana difieren mucho**, la distribución es asimétrica y la media no representa al conjunto. Con datos de renta o de facturación ocurre siempre.
- **Si el coeficiente de variación pasa de 1**, la dispersión es del orden de la propia media y hablar de «un valor típico» es engañoso.
- **Qué hacer con los atípicos**: comprobarlos primero —un error de captura se corrige, un valor extremo real no se borra— y decir explícitamente qué se ha hecho.

8.3 Práctica 2. Análisis descriptivo bidimensional

Actividad	Qué se obtiene
Diagrama de dispersión	la forma de la relación
Covarianza y correlación	signo e intensidad de la relación lineal
Recta de regresión	el modelo lineal
R^2	la bondad del ajuste
Análisis de residuos	si el modelo lineal es adecuado

El **diagrama de dispersión** va siempre antes que el coeficiente, no después. Es la única forma de detectar que la relación es curva, que hay dos subpoblaciones mezcladas o que un solo punto extremo está sosteniendo toda la correlación.

El **análisis de residuos** cierra la práctica: se representan los residuos frente a los valores predichos, y si el modelo es adecuado no debe verse ningún patrón. Una forma de embudo indica que la dispersión crece con el nivel, y una curva indica que faltaba un término no lineal.

8.4 Práctica 3. Números índices y deflación

Actividad	Fuente típica
Construir un índice simple de una serie propia	cualquier serie del INE
Calcular tasas de variación interanual y media	ídem
Deflactar una serie nominal con el IPC	salarios, ventas, PIB
Enlazar dos series con distinta base	cambios de base del IPC
Descomponer la variación por grupos	IPC por divisiones

El ejercicio que más enseña: **tomar una serie de salarios medios de los últimos veinte años, deflactarla y comparar las dos curvas**. La nominal sube siempre; la real suele tener tramos planos y caídas, y esa diferencia es todo el tema 3 en un gráfico.

8.5 Práctica 4. Series cronológicas

Actividad	Qué se obtiene
Representar la serie	tendencia y estacionalidad a la vista
Media móvil centrada	la tendencia-ciclo

Actividad	Qué se obtiene
Índices estacionales	ajustados a 400 o 1200
Serie desestacionalizada	comparable entre periodos consecutivos
Predicción	tendencia extrapolada por índice estacional

Series con estacionalidad clara y fáciles de conseguir: pernoctaciones hoteleras, matriculaciones de vehículos, consumo eléctrico, afiliación a la Seguridad Social.

Se pide comparar la predicción con lo que realmente ocurrió, reservando los últimos periodos de la serie y no usándolos para ajustar. **Predecir sobre datos que se han usado para ajustar no mide nada.**

8.6 Práctica 5. Probabilidad y distribuciones

Actividad	Órdenes típicas
Simular experimentos	generar muestras aleatorias
Comprobar la frecuentista	frecuencia relativa frente a probabilidad al crecer n
Calcular probabilidades binomiales	función de masa y acumulada
Ídem Poisson, hipergeométrica y geométrica	ídem
Comprobar aproximaciones	binomial por Poisson, hipergeométrica por binomial

Los tres experimentos que se piden:

- **Ley de los grandes números:** simular lanzamientos de una moneda y representar la frecuencia relativa acumulada frente al número de lanzamientos. Se ve converger a 0,5 y, sobre todo, se ve cuánto tarda.
- **Comparar binomial y Poisson** para n creciente con np fijo, y medir a partir de qué n la diferencia es despreciable.
- **Comprobar el factor de corrección** de la hipergeométrica: calcular la misma probabilidad con las dos distribuciones para n/N creciente y ver cuándo deja de valer la aproximación.

Nota 8.2. Una simulación con pocas repeticiones no demuestra nada: la variabilidad del propio experimento domina. Antes de concluir que un resultado simulado contradice la teoría hay que repetirlo con órdenes de magnitud más de muestras y comprobar si la diferencia persiste.

8.7 Sobre la memoria

Lo que se entrega:

1. Origen de los datos, con la referencia exacta y la fecha de descarga.
2. Descripción de las variables y su tipo.
3. Tablas y gráficos, **numerados y con título**.
4. Cálculos, indicando las fórmulas usadas.
5. **Interpretación**, que es lo que se evalúa.
6. Limitaciones del análisis.

El punto 5 no es repetir los números en palabras. Es responder a qué significa que la asimetría sea positiva, por qué la media no es representativa, o qué implica un R^2 de 0,3 para la predicción que se pretende hacer.

Y el punto 6 distingue una memoria buena: decir que la muestra no es representativa, que la correlación no implica causalidad, o que la predicción supone que el patrón se mantiene, es parte del análisis y no una excusa.

El material de las prácticas sigue Amor Pulido, 2016 y Hermoso Gutiérrez, 2016, con los conjuntos de problemas de Hermoso Gutiérrez y Hernández Bastida, 2000 y Arnaldos García et al., 2003.

Bibliografía

- Amor Pulido, R. (2016). *Técnicas Cuantitativas I*. Técnica Avicam Fleming.
- Arnaldos García, F., Díaz Delfa, M. T., & Faura Martínez, Ú. (2003). *Estadística descriptiva para Economía y Administración de Empresas*. Thomson.
- Canavos, G. C. (1989). *Probabilidad y estadística: aplicaciones y métodos*. McGraw-Hill.
- Castillo Manrique, I., & Guijarro Garvi, M. (2006). *Estadística descriptiva y cálculo de probabilidades*. Pearson Prentice Hall.
- Hermoso Gutiérrez, J. A. (2016). *Apuntes de Técnicas Cuantitativas I*. Técnica Avicam Fleming.
- Hermoso Gutiérrez, J. A., & Hernández Bastida, A. (2000). *Curso básico de estadística descriptiva y probabilidad. Teoría y problemas*. Némesis.
- Hernández Bastida, A. (2007). *Curso elemental de estadística descriptiva*. Pirámide.
- Martín Pliego, F. J. (2004). *Introducción a la estadística económica y empresarial* (3.^a ed.). Thomson.
- Newbold, P., Carlson, W. L., & Thorne, B. (2013). *Estadística para administración y economía* (8.^a ed.). Pearson.