# Estadística Descriptiva e Introducción a la Probabilidad

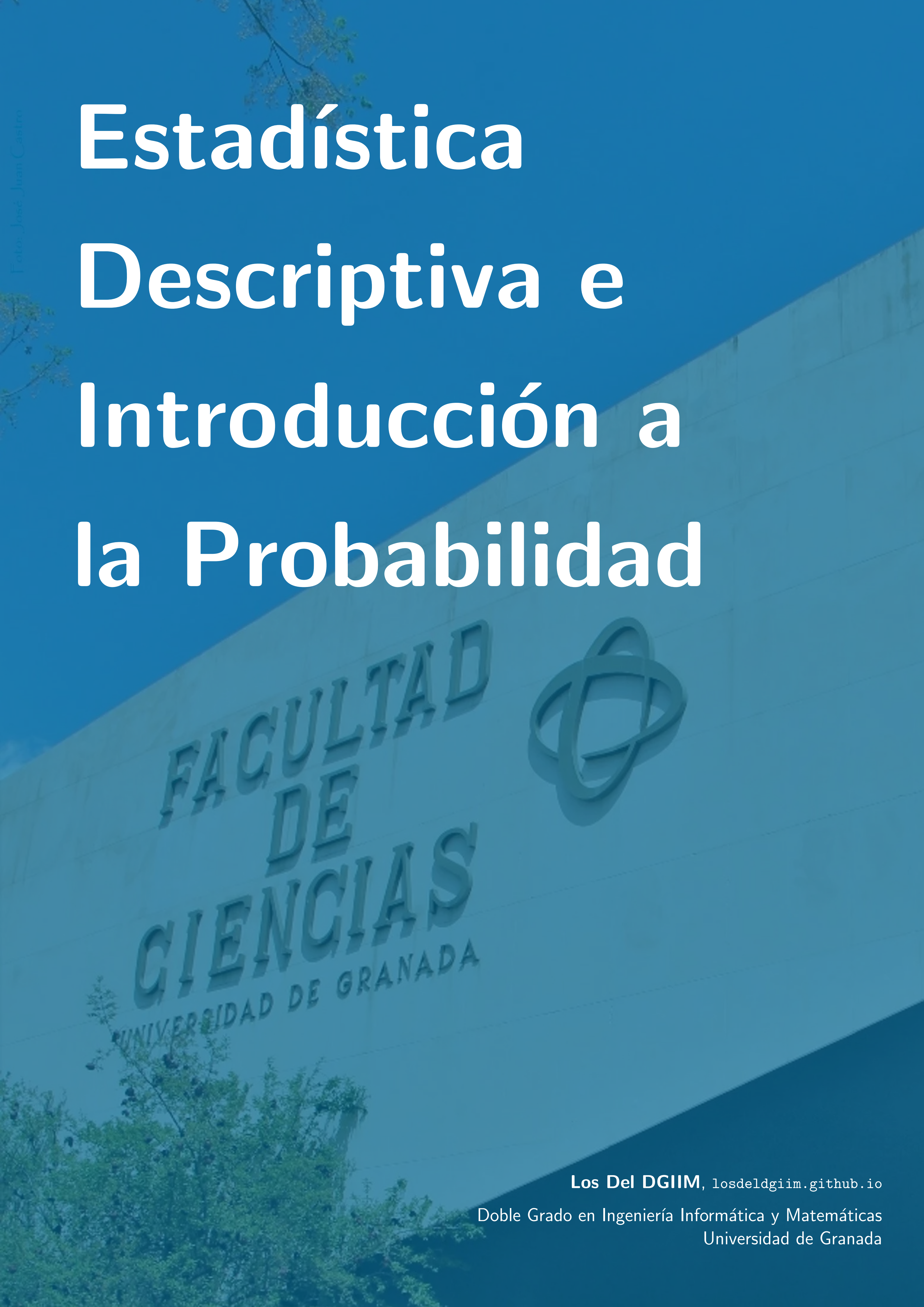


Foto: José Juan Castro

**Los Del DGIIM**, `losdeldgiim.github.io`

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada

# Estadística Descriptiva e Introducción a la Probabilidad

Los Del DGIIM, losdeldgiim.github.io

Arturo Olivares Martos
José Juan Urrutia Milán "JJ"

Granada, 2023

# Índice general

# 1. Estadística descriptiva unidimensional

## 1.1. Definiciones

- **Fenómenos determinísticos:** Los que dan lugar al mismo resultado si se hacen bajo condiciones idénticas.
- **Fenómenos aleatorios:** los resultados pueden variar incluso si el estudio se realiza con las mismas condiciones iniciales.
- **Población:** Conjunto de unidades con al menos una característica en común sobre la que se desea obtener cierta información.
- **Muestra:** Subconjunto de la población elegido en términos de representatividad.
- **Carácter:** Propiedad a ser estudiada, puede ser cualitativa o cuantitativa.
- **Modalidad:** Valores que se han presentado al medir una variable con una cierta escala.
  - Principio de incompatibilidad: Un individuo sólo puede tomar valor en una modalidad.
  - Principio de exhaustividad: Todos los individuos deben tomar al menos un valor.
- **Variable** Ente matemático capaz de captar las diferentes modalidades que puede tomar una característica a medir en una población. Se suelen relacionar con valores cuantitativos.
  - **Variables discretas:** Si el paso de un valor al siguiente representa un salto.
  - **Variables continuas:** Si se puede tomar cualquier valor entre dos valores dados.

**Notación.** Supongamos que tenemos una población de tamaño $n$ de la cual obtenemos una muestra para realizar un estudio. Queremos medir el carácter $X$ sobre dicha muestra. A cada modalidad le asignaremos un valor $x_i$. De esta forma, al número de individuos que nos responda que presentan la modalidad $x_i$ acerca del carácter $X$ lo notaremos $n_i$.

## 1.2. Escalas de medida

Cuando realizamos un estudio estadístico debemos identificar de forma precisa las modalidades y asignarles símbolos o números a dichas modalidades. Eso se denomina la medición del carácter.

- **Escala nominal:** Las modalidades sólo pueden ser iguales o diferentes y no se puede determinar un órden en estas. Un ejemplo son las modalidades de verdadero/falso o los colores de pelo.
- **Escala ordinal:** Las modalidades pueden ser iguales o diferentes y además se puede establecer un orden de mayor que. Esta escala es válida tanto para carácteres cualitativos y cuantitativos. Ejemplo: niveles de aceptación.
- **Escala de intervalo:** Existen diferencias, luego podemos introducir el término de resta pero no existe un cero absoluto. Un ejemplo son las tallas de ropa.
- **Escala de razón:** Aquella en la que podemos dividir y decir que una modalidad es $a$ veces otra modalidad. Existe el cero absoluto.

**Ejercicio.** Da ejemplos de cada una de las escalas de medida. Para cada ejemplo, indica el carácter, si es cuantitativo o cualitativo, y la población.

1. **Escala nominal**
   *a*) <u>Ser hijo único</u> Es cualitativo y la población son las personas. Las modalidades son ser hijo único o no.
   *b*) <u>Sexo al nacer</u> Es cualitativo y la población son las personas. Las modalidades son ser hombre o mujer.
   *c*) <u>Ser funcionario</u> Es cualitativo y la población es la población activa. Las modalidades son ser o no funcionario.
   *d*) <u>Consumir estupefacientes</u> Es cualitativo y la población son las personas. Las modalidades son consumirlas o no.
   *e*) <u>Provincia de origen de los estudiantes del DGIIM</u> Es cualitativo y la población son los estudiantes. Las modalidades son las distintas provincias.
   *f*) <u>Grupo Sanguíneo</u> Es cualitativo y la población son las personas. Las modalidades son A+, A-, B+, B-, AB+, AB-, 0+ y 0-.
2. **Escala ordinal**
   *a*) <u>Nivel de Estudios</u> Es cualitativo y la población son las personas. Las modalidades son analfabetismo, EP, ESO, FP, Universidad, etc.
   *b*) <u>Rango en el ejército</u> Es cualitativo y la población son los militares. Las modalidades son los distintos rangos.
3. **Escala de intervalo**
   *a*) <u>La temperatura en $^\circ C$</u> Es cuantitativo y la población son los distintos lugares del mundo. Es un carácter continuo.

b) Inteligencia Es cuantitativo y la población son las personas. Es un carácter continuo y se mide según el coeficiente intelectual ($CI$).

c) Tallaje de la ropa Es cuantitativo y la población son las prendas de ropa. Es un carácter discreto y las modalidades son las distintas tallas.

d) Grado de satisfacción de los políticos españoles[1] Es cualitativo y la población son los españoles. Las modalidades son los distintos gustos.

4. **Escala de razón**

a) Notas de 1° del DGIIM Es cuantitativo y la población son los estudiantes de dicha clase. Es un carácter continuo.

b) Número de hermanos Es cuantitativo y la población son las personas. Es un carácter discreto, y las modalidades son $\mathbb{N} \cup \{0\}$.

c) Número de segundos que tarda un coche en llegar a los $100km/h$ Es cuantitativo y la población son los coches. Es un carácter continuo.

d) Salario bruto anual Es cuantitativo y la población es la población activa. Es un carácter continuo.

e) El tamaño de la RAM de un PC Es cuantitativo y la población son los ordenadores. Es un carácter discreto.

f) Estatura Es cuantitativo y la población son las personas. Es un carácter continuo.

## 1.3. Distribución de frecuencias

En una población de $n$ individuos medimos la característica $X$, que puede adoptar las modalidades $x_1, x_2, ..., x_k$. Se denomina:

- **Frecuencia absoluta de $x_i$:** Número total de individuos en la población que representa dicho valor $x_i$. Se denota $n_i$. Además, se verifica que $\sum\limits_{i=1}^{k} n_i = n$.

- **Frecuencia relativa de $x_i$:** Es la proporción de individuos que presenta dicha modalidad: $f_i = \dfrac{n_i}{n}$. Además, se verifica que $\sum\limits_{i=1}^{k} f_i = 1$.

Si las modalidades se pueden ordenar, supuesto $x_1 < x_2 < \ldots < x_i < \ldots < x_k$:

- **Frecuencia absoluta acumulada:** Es el número de individuos que presentan una modalidad menor o igual que $x_i$: $N_i = \sum\limits_{j=1}^{i} n_j$.

- **Frecuencia relativa acumulada:** Es la proporción de individuos que presentan una modalidad menor o igual que $x_i$: $F_i = \dfrac{N_i}{n}$.

Llamaremos **distribución de frecuencias** de una variable al conjunto formado por cada uno de los valores modales junto con sus frecuencias:

[1] Ejemplo de escala de *Likert*

- Distribución de frecuencias absolutas: $\{x_i, n_i\}_{i=1,\ldots,k}$
- Distribución de frecuencias relativas: $\{x_i, f_i\}_{i=1,\ldots,k}$
- Distribución de frecuencias absolutas acumuladas: $\{x_i, N_i\}_{i=1,\ldots,k}$
- Distribución de frecuencias relativas acumuladas: $\{x_i, F_i\}_{i=1,\ldots,k}$

Si las modalidades son intervalos, llamaremos a estos límites intervalos de clases: $[e_0, e_1], ]e_1, e_2], ]e_2, e_3], \ldots, ]e_{k-1}, e_k]$ y definimos nuevos conceptos:

- **Marca de clase:** Es la media de los límites de cada intervalo: $x_i = \dfrac{e_{i-1} + e_i}{2}$.
- **Amplitud de clase:** $a_i = e_i - e_{i-1}$.
- **Densidad de frecuencia de modalidad i de variable X:** $h_i = d_i = \dfrac{n_i}{a_i}$.
  También se calcula a veces con la frecuencia relativa.

Cuando comencemos el estudio de una variable, indicaremos:
*Sea $X$ una variable estadística con población $n$ y modalidades $x_1, \ldots, x_k$ con* distribución de frecuencias*:*

$$\{x_i, n_i\}_{i=1,\ldots,k}$$

## 1.4. Representaciones gráficas de una variable estadística unidimensional

### 1.4.1. Variables cualitativas o atributos

- **Diagrama de sectores**.
  Es un círculo dividido en tantos sectores circulares como modalidades tenga el carácter, siendo el área de cada uno proporcional a la frecuencia absoluta o relativa de la modalidad.

- **Diagrama de rectángulos o barras**.
  Consiste en varios rectángulos (uno por modalidad) de base constante y alturas proporcionales a las frecuencias (absolutas o relativas) de cada modalidad.

- **Pictograma**.
  Se dibujan figuras, normalmente alusivas al carácter que se está estudiando, bien una para cada modalidad con tamaño proporcional a su frecuencia, o bien repitiendo la figura tantas veces como requieran las frecuencias.

### 1.4.2. Variables discretas

- **Diagramas de barras**.
  Similar al de atributos: en un sistema de ejes cartesianos se representa en el eje de abcisas los valores de la variable, y se trazan barras verticales con longitudes proporcionales a sus frecuencias (absolutas o relativas).

- **Función / Curva de distribución o acumulativa**:

$$F : \mathbb{R} \mapsto \mathbb{R} \mid F(x) = \begin{cases} 0 & \text{si } x < x_1 \\ F_i & \text{si } x_i \leqslant x < x_{i+1} \\ 1 & \text{si } x \geqslant x_k \end{cases}$$

Función definida en todo $\mathbb{R}$, no decreciente y continua por la derecha.

### 1.4.3. Variables continuas

**Histogramas:**
De los más utilizados. La base de los rectángulos son las diferentes clases o intervalos de la variable y la altura de estos son las densidades de frecuencia $h_i = \frac{f_i}{a_i}$ (o $h_i = \frac{n_i}{a_i}$).

Calculamos el área de los rectángulos de la siguiente forma en los histogramas:

$$Area = a_i \cdot \frac{f_i}{a_i} = f_i \quad Area = a_i \cdot \frac{n_i}{a_i} = n_i$$

**Poligonal de frecuencias:** Poligonal resultante de unir los puntos correspondientes a los techos de las marcas de clase de los intervalos en el histograma.

**Curva acumulativa o de distribución:** Es la proporción de individuos en la población cuyo valor de la variable es inferior o igual a cada modalidad. Esta función se conoce únicamente para los valores de $x$ que son cota superior de cada intervalo. Asumiendo que los datos son equidistantes en el intervalo, podemos unir dichos puntos.

$$F(e_i) = \sum_{j=1}^{i} f_j$$

Se trata de una función monótona no decreciente.

## 1.5. Características unidimensionales

Resúmenes cuantitativos de la información de los datos.
Propiedades deseables (Propiedades de Yule):

- Deben definirse de manera objetiva.
- Deben usar todas las observaciones.
- Deben tener un significado concreto, para ser de rápida y fácil interpretación.
- Deben ser sencillas de calcular.
- Deben prestarse fácilmente al cálculo algebraico.

- Deben ser poco sensibles a fluctuaciones muestrales, sin mucho cambio cuando alteramos las medidas de los extremos.

Pueden ser de varios tipos:

- **Medidas de posición:** permiten situar una distribucion en la recta real. Las más importantes son las de centralización o tendencia central, denominadas promedios (proporcionan un valor central representativo alrededor del cual se agrupan los datos) y cuantiles (que proporcionan valores representativos de parte de la distribución).
- **Medidas de dispersión:** Miden el grado de esparcimiento de los datos de una distribución.
- **Medidas de forma:** Caracterizan de manera precisa la forma de una distribución sin tener que llevar a cabo una representación gráfica.

### 1.5.1. Medidas de posición

*Observación.* Promediar es dar la media aritmética, mediana y moda.

#### Media aritmética

Nos sirve para tener un valor alrededor del cual se sitúan los valores, actuando como centro de gravedad de la distribución.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \sum_{i=1}^{k} f_i x_i$$

Si los datos están distribuidos en intervalos, se usan las marcas de clase.

**Proposición 1.1.** *La media de una variable unidimensional está acotada.*

*Demostración.* Es inmediata, ya que $x_1 \leqslant \overline{x} \leqslant x_k$. □

**Proposición 1.2.** *La media aritmética de las desviaciones de los datos respecto de la media aritmética es igual a* 0*:*

$$\sum_{i=1}^{k} f_i(x_i - \bar{x}) = 0$$

*Demostración.*

$$\sum_{i=1}^{k} f_i(x_i - \bar{x}) = \sum_{i=1}^{k}(f_i x - f_i\bar{x}) = \sum_{i=1}^{k} f_i x - \bar{x}\sum_{i=1}^{k} f_i = \bar{x} - 1\bar{x} = 0$$

□

**Proposición 1.3.** *Si se somete una variable $X$ a una transformación lineal afín, la media aritmética de la nueva variable es la imagen de la media de $X$ por la misma transformación:*

$$Y = aX + b \Longrightarrow \bar{Y} = a\bar{x} + b$$

*Demostración.*

$$\bar{Y} = \sum_{i=1}^{k} f_i y_i = \sum_{i=1}^{k} f_i(ax_i + b) = \sum_{i=1}^{k} [f_i(ax_i) + f_i b] =$$

$$= a\sum_{i=1}^{k} f_i x_i + b\sum_{i=1}^{k} f_i = a\bar{x} + 1b = a\bar{x} + b$$

□

**Proposición 1.4.** *La media aritmética de los cuadrados de las desviaciones respecto a la media aritmética es mínima:*

$$\sum_{i=1}^{k} f_i(x_i - \bar{x})^2 < \sum_{i=1}^{k} f_i(x_i - a)^2, \quad \forall a \neq \bar{x}$$

*Demostración.* Se trata de minimizar una parábola con coeficiente líder positivo, por lo que el valor que anule la primera derivada será el mínimo absoluto.

$$\frac{\partial}{\partial a}\sum_{i=1}^{k} f_i(x_i - a)^2 = -2\sum_{i=1}^{k}(f_i x_i - f_i a) = -2\bar{x} + 2\sum_{i=1}^{k} f_i a = 0 \Longleftrightarrow a = \bar{x}$$

□

Test de Yule:

- Es objetiva.
- Usa todas las observaciones.
- Representa el centro de gravedad de la distribución.
- Sencillo de calcular.
- Se presta a hacer cálculos para compararla.
- (Contra) Es sensible a fluctuaciones en los extremos.

### Media geométrica

Se usa cuando las variables sufren variaciones acumuladas (como en porcentajes), se usa cuando se desea promediar datos de una variable que tiene efectos multiplicativos acumulativos.

$$G = \sqrt[n]{x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}} = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}}$$

Si se trata de una variable continua, la calculamos con las marcas de clase.

**Proposición 1.5.** *El logaritmo de la media geométrica es la media aritmética de los logaritmos de los valores de la variable.*

*Demostración.*

$$\log G = \log \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}} = \frac{1}{n}\sum_{i=1}^{k} n_i \log x_i = \sum_{i=1}^{k} f_i \log x_i$$

□

*Observación.* Si algún valor de la variable es $0 \Rightarrow G = 0$, luego esta media no nos será de mucha utilidad.

Test de Yule:

- Es objetiva.
- Usa todas las observaciones.
- (Contra) Tiene significado concreto pero no es sencillo de interpretar.
- (Contra) Complicada de calcular.
- (Contra) Sufre también a fluctuaciones pero menos que la aritmética.

### Media armónica

Se usa para promediar datos de magnitudes que son cocientes de dos magnitudes; esto es, magnitudes relativas (su unidad de medida es referida a una unidad de otra variable). Por ejemplo, para promedir velocidades, rendimientos o productividades, etc.

Se define como la inversa de la media aritmética de los valores inversos de la variable.

$$H = \frac{n}{\frac{n_1}{x_1} + \frac{n_2}{x_2} + \ldots + \frac{n_k}{x_k}} = \frac{n}{\sum\limits_{i=1}^{k} \frac{n_i}{x_i}}$$

Tenemos el mismo problema con la geométrica, cuando algún $x_i$ vale 0 difícilmente vamos a poder realizar dicha media. Además, no se recomienda realizarla cuando sea cercano a 0, ya que nos dispararía la media, sensible para valores pequeños.

Test de Yule:

- Es objetiva.
- Usa todos los datos.
- (Contra) Tiene un significado, pero es difícil de manejar.
- (Contra) No es sencilla de calcular.

### Media cuadrática

Es la menos utilizada y, fundamentalmente, su uso se reduce al cálculo de promedios sobre superficies. Se define como la raíz cuadrada de la media aritmética de los cuadrados de los valores de la variable.

$$Q = \sqrt{\sum_{i=1}^{k} f_i x_i^2}$$

### Moda

La moda de una distribución es el valor de mayor frecuencia (absoluta o relativa), el que más se repite.

- Variables discretas: La moda es el valor $x_i \mid n_i \geqslant n_j$ o $f_i \geqslant f_j$ $\forall j \in \{1, \dots, k\}$.

  Si en variables discretas todas las variables $x_i$ se repiten $n_i$ veces con $n_i = n_j$ $\forall i, j \mid i \neq j$ entonces diríamos que no hay moda. Si la moda se repite varias veces, hablamos de una moda plurimodal.

- Variables continuas: La moda está en el denominado intervalo modal, el de mayor densidad de frecuencia $h_i = \frac{f_i}{a_i}$, es decir, el de mayor altura en el histograma.

  La moda en variables continuas se calcula haciendo interpolación y semejanza de triángulos:

  La distancia de la moda desde el intervalo que más se repite es inversamente proporcional a la frecuencia de los intervalos contiguos.

  Calculamos la moda sabiendo que:

$$\frac{h_i - h_{i-1}}{M_o - e_{i-1}} = \frac{h_i - h_{i+1}}{e_i - M_o} = \frac{(h_i - h_{i-1}) + (h_i - h_{i+1})}{a_i}$$

  Despejando:

$$M_o = \frac{a_i(h_i - h_{i-1})}{2h_i - h_{i-1} - h_{i+1}} + e_{i-1}$$

### Mediana

Como la media aritmética se presta a variaciones extremas, surge la mediana: Aquel valor que separa la distribución en dos efectivos de igual tamaño (supuestos ordenados por valor creciente de carácter).

Se puede calcular para variables discretas y para variables continuas.

La mediana también se puede calcular para características cualitativas pero han de estar en escala ordinal.

La mediana es la solución a la siguiente ecuación:

$$F(x) = \frac{1}{2} \qquad \left(\text{Equivalentemente, } N(x) = \frac{n}{2}\right)$$

- Variables discretas: Se busca $x_i \mid N_{i-1} < \frac{n}{2} \leqslant N_i$.
  - Si $N_i > \frac{n}{2} \Rightarrow Me = x_i$.
  - Si $N_i = \frac{n}{2} \Rightarrow Me = \dfrac{x_i + x_{i+1}}{2}$.

- Variables continuas: Se busca $I_i = (e_{i-1}, e_i] \mid N_{i-1} < \dfrac{n}{2} \leqslant N_i$ ó $F_{i-1} < \dfrac{1}{2} \leqslant F_i$.
  - Si $N_i = \dfrac{n}{2}$ ó $F_i = \dfrac{1}{2} \Rightarrow M_e = e_i$.
  - Si $N_i > \dfrac{n}{2}$ ó $F_i > \dfrac{1}{2} \Rightarrow$ La mediana está dentro de $I_i$, que se denomina el intervalo mediano y, para calcularla, hacemos una interpolación lineal:
  $$M_e = e_{i-1} + \frac{\frac{n}{2} - N_{i-1}}{n_i}(e_i - e_{i-1}) = e_{i-1} + \frac{\frac{1}{2} - F_{i-1}}{f_i}(e_i - e_{i-1})$$
  Dicha fórmula se deduce aplicando semejanza de triángulos a la gráfica de la curva de distribución.

**Proposición 1.6.** *La desviación absoluta media respecto a la mediana es mínima:*
$$\sum_{i=1}^{k} f_i|x_i - Me| < \sum_{i=1}^{k} f_i|x_i - a| \quad \forall a \neq Me$$

### Percentiles

El percentil de orden $r \in \{1, \dots, 100\}$ es un valor $P_r$ que divide al conjunto ordenado de datos en dos partes tales que el $r\,\%$ del total son inferiores o iguales a $P_r$.

- Variables discretas: Se busca:
$$x_i \mid N_{i-1} < \frac{nr}{100} \leqslant N_i \qquad \left(\text{Equivalentemente}, F_{i-1} < \frac{r}{100} \leqslant F_i\right)$$
  - Si $N_i > \frac{nr}{100} \Rightarrow P_r = x_i$.
  - Si $N_i = \frac{nr}{100}$, todo número del intervalo $[x_i, x_{i+1}[$ es percentil de oden $r$. Por lo que se suele tomar como $P_r$ el punto medio de dicho intervalo.

- Variables continuas: Se busca:
$$I_i = (e_{i-1}, e_i] \mid N_{i-1} < \frac{nr}{100} \leqslant N_i \qquad \left(\text{Equivalentemente}, F_{i-1} < \frac{r}{100} \leqslant F_i\right)$$
  - Si $N_i = \frac{nr}{100} \Rightarrow P_r = e_i$.
  - Si $N_i > \frac{nr}{100} \Rightarrow P_r = e_{i-1} + \dfrac{\frac{nr}{100} - N_{i-1}}{n_i}(e_i - e_{i-1})$

Los cuartiles $Q_1, Q_2, Q_3, Q_4$ coinciden con los percentiles $P_{25}, P_{50}, P_{75}, P_{100}$.
Además, el decil $D_r$ coincide con el percentil $P_{r0}$.
Observemos que la mediana es el percentil $P_{50}$.
El percentil $P_r$ nos indica que el r % de la población presenta al menos $P_r$ valor en la variable $X$.

### 1.5.2. Medidas de dispersión

Buscamos ver cuanto se desvían los datos respecto de alguna medida de posición central. Hay varios tipos de medidas de dispersión:

- Medidas de dispersión absolutas:
  Tienen un comportamiento global y dependen de las unidades de medida de la variable.
  - Recorrido o rango.
  - Recorrido intercuartílico.
  - Desviación absoluta respecto a la media aritmética.
  - Desviación absoluta respecto a la mediana.
  - Varianza.
  - Desviación típica.
- Medidas de dispersión relativas:
  Se definen de forma individualizada para cada distribución y son adimensionales, lo que nos permite comparar distribuciones.
  - Coeficiente de apertura.
  - Recorrido relativo.
  - Recorrido semi-intercuartílico.
  - Coeficiente de variación.
  - Índice de dispersión respecto a la mediana.

#### Recorrido o rango

Bajo el supuesto de que los valores de la variable estén ordenados en sentido creciente:

$$R = x_k - x_1$$

Cuanto mayor sea el rango, más dispersa será nuestra distribución.

#### Recorrido intercuartílico

$$R_I = Q_3 - Q_1$$

Indica la longitud del intervalo en el que está incluido el 50 % central de los datos.

#### Desviación absoluta media respecto a $\overline{x}$

$$D_{\overline{x}} = \frac{\sum\limits_{i=1}^{k} |x_i - \overline{x}| n_i}{n} = \sum_{i=1}^{k} f_i |x_i - \overline{x}|$$

Cuanto mayor sea esta, menos representativa será la media. Una desviación media elevada implica mucha variabilidad en los datos.

**Desviación absoluta media respecto a $M_e$**

$$D_{M_e} = \frac{\sum\limits_{i=1}^{k} |x_i - M_e| n_i}{n}$$

Cuanto mayor sea esta, menos representativa será la mediana.

**Varianza**

$$Var(X) = \sigma^2 = \frac{\sum\limits_{i=1}^{k} n_i (x_i - \overline{x})^2}{n} = \sum_{i=1}^{k} f_i (x_i - \overline{x})^2$$

Es la media del área de los cuadrados de lado $x_i - \overline{x}$, el cuadrado de la media cuadrática de la desviación de los datos a su media.

**Lema 1.7.** *La varianza nunca puede ser negativa ($\sigma^2 \geqslant 0$).*

*Demostración.* Se deduce directamente de su definición, ya que es la media aritmética de términos no negativos. $\square$

**Proposición 1.8.** *La varianza es la media cuadrática de dispersión óptima:*

$$\sum_{i=1}^{k} f_i (x_i - \overline{x})^2 < \sum_{i=1}^{k} f_i (x_i - a)^2, \qquad \forall\, a \neq \overline{x}$$

*Demostración.* Ver la Proposición 1.4. $\square$

**Proposición 1.9.** *La varianza está acotada superior e inferiormente en cada distribución de frecuencias:*

$$\text{mín}(x_i - \overline{x})^2 < \sigma^2 < \text{máx}(x_i - \overline{x})^2$$

**Teorema 1.10** (Teorema de König)**.** *Para cualquier $a \in \mathbb{R}$ y para cualquier variable estadística $X$, se verifica:*

$$\sum_{i=1}^{k} f_i (x_i - a)^2 = \sum_{i=1}^{k} f_i (x_i - \bar{x})^2 + (a - \bar{x})^2$$

*En concreto, para $a = 0$, tenemos una forma más cómoda de calcular la varianza:*

$$\sigma_x^2 = \sum_{i=1}^{k} f_i (x_i - \bar{x})^2 = \sum_{i=1}^{k} f_i x_i^2 - \bar{x}^2$$

*Demostración.*

$$\sum_{i=1}^{k} f_i(x_i-a)^2 = \sum_{i=1}^{k} f_i(x_i^2+a^2-2ax_i) = \sum_{i=1}^{k} f_i(x_i^2+a^2-2ax_i+\bar{x}^2-\bar{x}^2+2x_i\bar{x}-2x_i\bar{x}) =$$
$$= \sum_{i=1}^{k} f_i(x_i^2 - 2x_i\bar{x} + \bar{x}^2 + a^2 - 2ax_i - \bar{x}^2 + 2x_i\bar{x}) =$$
$$= \sum_{i=1}^{k} f_i(x_i + \bar{x})^2 + \sum_{i=1}^{k} f_i(a^2 - 2ax_i - \bar{x}^2 + 2x_i\bar{x}) =$$
$$= \sum_{i=1}^{k} f_i(x_i + \bar{x})^2 + a^2 - 2a\bar{x} - \bar{x}^2 + 2\bar{x}^2 = \sum_{i=1}^{k} f_i(x_i - \bar{x})^2 + (a - \bar{x})^2$$

□

**Proposición 1.11.** *Si se someten los datos a una transformación afín* $y_i = ax_i + b$ $i = 1, \ldots, k$, *la varianza de los datos transformados es* $\sigma_y^2 = a^2\sigma_x^2$.

*Es decir, no se ve afectada por cambios de origen pero sí se ve afectada por cambios de escala.*

*Demostración.*

$$y = ax+b \Longrightarrow \sigma_y^2 = \sum(y_i-\bar{y})^2 f_i = \sum((ax_i+\cancel{b})-(a\bar{x}+\cancel{b}))^2 f_i = \sum a^2(x_i-\bar{x})^2 f_i =$$
$$= a^2 \sum(x_i - \bar{x})^2 f_i = a^2\sigma_x^2$$

□

### Desviación típica

Se define como la raíz cuadrada positiva de la varianza:

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

Nos ayuda a tipificar distribuciones para compararlas. Representa la dispersión de los datos de la distribución respecto de la media. Esto quiere decir que la mayoría de los datos de la distribución se encuentran en el intervalo $[\overline{x} - \sigma, \overline{x} + \sigma]$

La gran mayoría de las propiedades se deducen directamente de las propiedades de la varianza.

**Lema 1.12.** *Es no negativa* ($\sigma \geqslant 0$).

**Proposición 1.13.** *Es una medida de dispersión óptima.*

$$\sigma < \sqrt{\sum_{i=1}^{k} f_i(x_i - b)^2} \qquad \forall\ b \neq \overline{x}$$

**Proposición 1.14.** *Está acotada superior e inferiormente en cada distribución.*

**Proposición 1.15.** *Si se someten los datos a una transformación afín* $y_i = ax_i + b$ $i = 1, \ldots, k$, *la desviación típica de los datos transformados es* $\sigma_y = |a|\sigma_x$.

*Es decir, no se ve afectada por cambios de origen pero sí se ve afectada por cambios de escala.*

*Demostración.*

$$y = ax + b \Longrightarrow \sigma_y^2 = a^2\sigma_x^2 \Longrightarrow \sigma_y = |a|\sigma_x$$

□

#### Coeficiente de apertura

Se define como el cociente entre los dos valores extremos de una distribución. Supuestos ordenamos crecientemente los valores:

$$C_A = \frac{x_k}{x_1}$$

Si trabajamos con variables continuas, es el extremo superior del último intervalo entre el extremo inferior del primero:

$$C_A = \frac{e_k}{e_0}$$

#### Recorrido relativo

Se define como el cociente entre el recorrido y la media aritmética:

$$R_R = \frac{R}{\overline{x}} = \frac{x_k - x_1}{\overline{x}}$$

#### Recorrido semi-intercuartílico

Se define como el cociente entre el recorrido intercuartílico y la suma del primer tercer cuartil:

$$R_{SI} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

#### Coeficiente de variación de Pearson

Se define como la relación por cociente entre la desviación típica y la media:

$$CV(X) = \frac{\sigma_x}{|\overline{x}|}$$

A mayor valor del coeficiente, mayor heterogeneidad de los valores de la variable. Se usa para comparar las dispersiones de las variables estadísticas.

#### Índice de dispersión respecto a la mediana

Se define como el cociente entre la desviación absoluta media respecto a la mediana y la mediana:

$$V_{M_e} = \frac{D_{M_e}}{M_e}$$

### 1.5.3. Momentos

Nos permiten simplificar los cálculos.

Sea $r \in \mathbb{N} \cup \{0\}$. Se llama momento de orden $r$ respecto al valor $a$ a la cantidad:

$$_a m_r = \sum_{i=1}^{k} f_i (x_i - a)^r = \frac{1}{n} \sum_{i=1}^{k} n_i (f_i - a)^r$$

Según los valores de $a$ distinguimos dos tipos de momentos:

- **Momentos no centrales** ($a = 0$):

$$m_r = \sum_{i=1}^{k} f_i x_i^r$$

Algunos momentos no centrales especiales son:

$$m_0 = 1 \qquad m_1 = \bar{x} \qquad m_2 = \sum_{i=1}^{k} f_i x_i^2$$

*Observación.* Notemos que el anterior $m_2$ es un término del Teorema de König.

- **Momentos centrales** ($a = \overline{x}$):

$$\mu_r = \sum_{i=1}^{k} f_i (x_i - \overline{x})^r$$

Algunos momentos centrales especiales son:

$$\mu_0 = 1 \qquad \mu_1 = 0 \qquad \mu_2 = \sigma^2$$

Para la simplificación del cálculo de momentos, podemos expresar momentos centrales en función de otros no centrales más simples y viceversa:

- Momentos centrales en función de los no centrales:

$$\mu_r = \sum_{i=1}^{k} f_i (x_i - m_1)^r = \sum_{i=1}^{k} f_i \sum_{t=0}^{r} (-1)^t \binom{r}{t} m_1^t x_i^{r-t} =$$
$$= \sum_{t=0}^{r} (-1)^t \binom{r}{t} m_1^r \sum_{i=1}^{k} f_i x_i^{r-t} = \sum_{t=0}^{r} (-1)^t \binom{r}{t} m_1^t m_{r-t}$$

- Momentos no centrales en función de los centrales y de $m_1$:

$$m_r = \sum_{i=1}^{k} f_i x_i^r = \sum_{i=1}^{k} f_i [(x_i - m_1) + m_1]^r = \sum_{i=1}^{k} f_i \sum_{t=0}^{r} \binom{r}{t} m_1^t (x_i - m_1)^{r-t} =$$
$$= \sum_{t=0}^{r} \binom{r}{t} m_1^t \sum_{i=1}^{k} f_i (x_i - m_1)^{r-t} = \sum_{t=0}^{r} \binom{r}{t} m_1^t \mu_{r-t}$$

De tal forma que:

$$\mu_2 = m_2 - m_1^2 \qquad\qquad m_2 = \mu_2 + m_1^2$$
$$\mu_3 = m_3 - 3m_2 m_1 + 2m_1^3 \qquad\qquad m_3 = \mu_3 + 3\mu_2 m_1 + m_1^3$$
$$\mu_4 = m_4 - 4m_3 m_1 + 6m_1^2 m_2 - 3m_1^4 \qquad\qquad m_4 = \mu_4 + 4\mu_3 m_1 + 6\mu_2 m_1^2 + m_1^4$$

### 1.5.4. Medidas de forma

#### Medidas de asimetría

Entendemos por asimetría a la falta de simetría respecto del eje vertical $x = \overline{x}$.

Diremos que una distribución es simétrica si la media divide a la distribución en dos partes iguales.

Diremos a su vez que una distribución es asimétrica por la izquierda (negativa) o por la derecha (positiva) si por la izquierda o por la derecha presenta más datos, respectivamente.

**Coeficiente de Fisher**

$$\gamma_1(X) = \frac{\mu_3}{\sigma_x^3} = \sum_{i=1}^{k} f_i \left( \frac{x_i - \overline{x}}{\sigma_x} \right)^3$$

- Si $\gamma_1(X) > 0$ la distribución es asimétrica por la derecha o positiva.
- Si $\gamma_1(X) < 0$ la distribución es asimétrica por la izquierda o negativa.
- Si la distribución es simétrica $\Rightarrow \gamma_1(X) = 0$.

  Sin embargo, se puede dar que $\gamma_1(X) = 0$ y que la distribución sea asimétrica. Por tanto, cuando esto suceda deberemos representar nuestra distribución.

**Coeficiente de Pearson**

Válida sólo para distribuciones campaniformes:

$$A_p = \frac{\overline{x} - Mo}{\sigma_x}$$

Se verifica (empíricamente) que $(\overline{x} - Mo) \approx 3(\overline{x} - Me)$, por lo que:

$$A_p^* = \frac{3(\overline{x} - Me)}{\sigma_x}$$

Este coeficiente tiene la misma interpretación que el de Fisher.

#### Medidas de apuntamiento o curtosis

Miden la concentración central de frecuencias de una distribución. Para ello, la comparamos con una distribución normal, con la misma media y desviación típica que nuestra distribución. Se presentan tres casos:

- Leptocúrtica: La distribución está más concentrada que la normal.
- Mesocúrtica: La distribución se asemeja a la normal.
- Platicúrtica: La distribución está menos concentrada que la normal.

**Coeficiente de curtosis de Fisher**

$$\gamma_2(X) = \frac{\mu_4}{\sigma^4} - 3$$

- Si $\gamma_2(X) > 0$ nuestra distribución es leptocúrtica.
- Si $\gamma_2(X) = 0$ nuestra distribución es mesocúrtica.
- Si $\gamma_2(X) < 0$ nuestra distribución es platicúrtica.

**Coeficiente de curtosis de Kelley**

$$K = \frac{1}{2}\frac{Q_3 - Q_1}{D_9 - D_1} - 0{,}263$$

Con la misma interpretación que el coeficiente de curtosis de Fisher.

**Teorema 1.16** (Desigualdad de Tchebychev)**.** *El porcentaje de datos en cualquier intervalo de la forma* $(\overline{x} - k\sigma_x, \overline{x} + k\sigma_x)$ *es de al menos* $100(1 - \frac{1}{k^2})$ *%.*

# 2. Estadística descriptiva bidimensional

## 2.1. Distribución conjunta de dos caracteres estadísticos

Sea una población formada por $n$ individuos en la que se desea estudiar simultáneamente dos caracteres, $X$ e $Y$. Dichos caracteres podrán ser ambos cualitativos, uno cualitativo y otro cuantitativo o ambos cuantitativos (los dos discretos, los dos continuos o uno discreto y otro continuo).

Si designamos por $x_1, x_2, \ldots, x_k$ las $k$ modalidades posibles del carácter $X$ y por $y_1, y_2, \ldots, y_p$ las $p$ modalidades posibles del carácter $Y$, las observaciones correspondientes a cada individuo serán de la forma $(x_i, y_j)$, par ordenado que representa las modalidades tomadas por dicho individuo en los caracteres $X$ e $Y$.

- $n_{ij}$: Número total de individuos en la población que presentan simultáneamente la modalidad $x_i$ del carácter $X$ y la modalidad $y_j$ del carácter $Y$. Le llamamos frecuencia absoluta del par $(x_i, y_j)$.
- $f_{ij}$: Proporción de individuos en la población que presentan simultáneamente la modalidad $x_i$ del carácter $X$ y la modalidad $y_j$ del carácter $Y$. Le llamamos frecuencia relativa del par $(x_i, y_j)$. Por la definición de proporción sobre el total, tenemos que:

$$f_{ij} = \frac{n_{ij}}{n} \qquad i \in \{1, 2, \ldots, k\} \; j \in \{1, 2, \ldots, p\}$$

Gracias al principio de incompatibilidad y exhaustividad de las modalidades, tenemos que:

$$\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij} = n \quad \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij} = 1$$

La distribución $\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\ldots,k \\ j=1,\ldots,p}}$ recibe el nombre de distribución conjunta de los caracteres $X$ e $Y$.

- $n_{i.}$: Número total de individuos que presentan la modalidad $x_i$ del carácter $X$ sin tener en cuenta las modalidades que puedan tomar para el carácter $Y$:

$$n_{i.} = \sum_{j=1}^{p} n_{ij} \qquad i \in \{1, \ldots, k\}$$

- $f_{i.}$: Proporción total de individuos que presentan la modalidad $x_i$ del carácter $X$ sin tener en cuenta el carácter $Y$:

$$f_{i.} = \sum_{j=1}^{p} f_{ij} = \frac{n_{i.}}{n} \qquad i \in \{1, \dots, k\}$$

- $n_{.j}$: Número total de individuos que presentan la modalidad $y_j$ del carácter $Y$ sin tener en cuenta las modalidades que puedan tomar para el carácter $X$:

$$n_{.j} = \sum_{i=1}^{k} n_{ij} \qquad j \in \{1, \dots, p\}$$

- $f_{.j}$: Proporción total de individuos que presentan la modalidad $y_j$ del carácter $Y$ sin tener en cuenta el carácter $X$:

$$f_{.j} = \sum_{i=1}^{k} f_{ij} = \frac{n_{.j}}{n} \qquad j \in \{1, \dots, p\}$$

Se tiene que:

$$\sum_{i=1}^{k} n_{i.} = \sum_{j=1}^{p} n_{.j} = n \qquad\qquad \sum_{i=1}^{k} f_{i.} = \sum_{j=1}^{p} f_{.j} = 1$$

## 2.2. Tablas estadísticas bidimensionales

Para agrupar nuestros datos estadísticos, usaremos una tabla de doble entrada como la siguiente:

| $X\backslash Y$ | $y_1$ | $y_2$ | $\dots$ | $y_j$ | $\dots$ | $y_p$ | $n_{i.}$ |
|---|---|---|---|---|---|---|---|
| $x_1$ | $n_{11}$ | $n_{12}$ | $\dots$ | $n_{1j}$ | $\dots$ | $n_{1p}$ | $n_{1.}$ |
| $x_2$ | $n_{21}$ | $n_{22}$ | $\dots$ | $n_{2j}$ | $\dots$ | $n_{2p}$ | $n_{2.}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\dots$ | $\vdots$ | $\dots$ | $\vdots$ | $\vdots$ |
| $x_i$ | $n_{i1}$ | $n_{i1}$ | $\dots$ | $n_{ij}$ | $\dots$ | $n_{ip}$ | $n_{i.}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\dots$ | $\vdots$ | $\dots$ | $\vdots$ | $\vdots$ |
| $x_k$ | $n_{k1}$ | $n_{k2}$ | $\dots$ | $n_{kj}$ | $\dots$ | $n_{kp}$ | $n_{k.}$ |
| $n_{.j}$ | $n_{.1}$ | $n_{.2}$ | $\dots$ | $n_{.j}$ | $\dots$ | $n_{.p}$ | $n$ |

En el caso de que uno de los carácteres sea cualitativo, esta tabla recibirá el nombre de tabla de contingencia.

## 2.3. Representaciones gráficas

### 2.3.1. Diagrama de dispersión o nube de puntos

Consiste en representar cada par de observaciones $(x_i, y_j)$ por un punto en un plano bidimensional. Para representar la frecuecia absoluta de cada punto, se suele incluir un número pequeño al lado de cada punto. En caso de representar variables cuantitativas continuas, usaremos las marcas de clase.

### 2.3.2. Estereogramas

Los estereogramas son gráficos tridimensionales formados por barras colocadas en cada punto $(x_i, y_j)$ con altura $n_{ij}$.

En el caso de las variables cuantitativas continuas, las barras pasarán a ser prismas, donde la base del prisma tiene dimensiones $e_i - e_{i-1}xe_j - e_{j-1}$. La altura de cada prisma vendrá dada por la fórmula:

$$h_{ij} = \frac{n_{ij}}{(L_i - L_{i-1})(L_j - L_{j-1})}$$

De tal forma que el volumen de cada prisma es igual a la frecuencia absoluta de cada pareja de intervalos de clase.

## 2.4. Distribuciones marginales

Las modalidades del carácter $X$ junto con las frecuencias $n_{i.}$ forman la distribución marginal del carácter $X$:

$$\{x_i, n_{i.}\}_{i=1,\ldots,k}$$

Mientras que las modalidades del carácter $Y$ junto con las frecuencias $n_{.j}$ forman la distribución marginal del carácter $Y$:

$$\{y_j, n_{.j}\}_{j=1,\ldots,p}$$

Ambas son distribuciones unidimensionales, a las que es posible dar el tratamiento visto en el tema anterior.

Como ejemplo, la distribución marginal del carácter $X$ es la siguiente:

| $X$ | $n_{i.}$ | $f_{i.}$ |
|---|---|---|
| $x_1$ | $n_{1.}$ | $f_{1.}$ |
| $x_2$ | $n_{2.}$ | $f_{2.}$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $x_i$ | $n_{i.}$ | $f_{i.}$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $x_k$ | $n_{k.}$ | $f_{k.}$ |
| | $n$ | 1 |

## 2.5. Distribuciones condicionadas

En ocasiones es interesante el estudio de un carácter sólo sobre los individuos que presentan una modalidad (o varias) del otro carácter. por ejemplo, podría ser interesante el estudio del carácter $X$ en la subpoblación formada por los individuos que presentan la modalidad $y_j$ del carácter $Y$, una subpoblación de $n_{.j}$ individuos.

Decimos que la frecuencia relativa de la modalidad $x_i$ del carácter $X$ en aquellos individuos que presentan la modalidad $y_j$ del carácter $Y$ es:

$$f_{i/j} \equiv f_i^j = \frac{n_{ij}}{n_{.j}} \qquad i \in \{1, \ldots, k\}$$

Análogamente, podemos considerar la frecuencia relativa de la modalidad $y_j$ del carácter $Y$ en aquellos individuos que presentan la modalidad $x_i$ del carácter $X$:

$$f_{j/i} \equiv f_j^i = \frac{n_{ij}}{n_{i.}} \qquad j \in \{1, \ldots, p\}$$

De esta forma, existen $p$ distribuciones condicionadas para el carácter $X$ según una única modalidad del carácter $Y$ y $k$ distribuciones condicionadas para el carácter $Y$ según una única modalidad del carácter $X$.

Ejemplo de la distribución condicionada del carácter $X$ respecto a la modalidad $y_j$ del carácter $Y$, que denotaremos por $X/Y = y_j$:

| $X$ | $n_{ij}$ | $f_i^j$ |
|---|---|---|
| $x_1$ | $n_{1j}$ | $f_1^j$ |
| $x_2$ | $n_{2j}$ | $f_2^j$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $x_i$ | $n_{ij}$ | $f_i^j$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $x_k$ | $n_{kj}$ | $f_k^j$ |
| | $n_{.j}$ | 1 |

De las definiciones anteriores, tenemos que:

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

$$f_{ij} = f_{i.} f_j^i = f_{.j} f_i^j$$

## 2.6. Dependencia e Independencia estadística

Dos caracteres $X$ e $Y$ serán estadísticamente dependientes cuando la variación en uno de ellos influya en la distribución del otro.

Por otra parte, se dice que el carácter $X$ es estadísticamente independiente del carácter $Y$ si las distribuciones de $X$ condicionadas a cada valor $y_j$ de $Y$ $(X/Y = y_j)$ son idénticas para cualquier valor de $j$. En este caso, cada distribución condicionada es idéntica a la distribución marginal de $X$:

$$\frac{n_{i1}}{n_{.1}} = \frac{n_{i2}}{n_{.2}} = \ldots = \frac{n_{ij}}{n_{.j}} = \ldots = \frac{n_{ip}}{n_{.p}} \qquad \forall i = 1, \ldots, k$$

De donde tenemos que:

$$f_{i/j} \equiv f_i^j = \frac{n_{ij}}{n_{.j}} = \frac{n_{i1} + n_{i2} + \ldots + n_{ip}}{n_{,1} + n_{,2} + \ldots + n_{.p}} = \frac{n_{i.}}{n} = f_{i.}$$

Por lo que si $f_i^j = f_{i.}$ $\forall i \in \{1, \ldots, k\}$, entonces el carácter $X$ será independiente del carácter $Y$. (Análogamente, se define la independencia del carácter $Y$ con el carácter $X$).

**Proposición 2.1.** *Si $X$ es una variable independiente de $Y \Longrightarrow f_{i\cdot} = f_{i/j}$*

*Demostración.* Suponemos $X$ e $Y$ variables independientes.

$$f_{i\cdot} = \sum_{j=1}^{p} f_{ij} = \sum_{j=1}^{p} f_{i/j} f_{\cdot j} \overset{(*)}{=} f_{i/j} \sum_{j=1}^{p} f_{\cdot j} = f_{i/j} f_{\cdot\cdot} = f_{i/j}$$

donde en $(*)$ he usado que las variables son independientes, por lo que la frecuencia condicionada a $Y$ no depende de $j$. □

**Teorema 2.2** (Teorema de Caracterización de la Independencia)**.** *Sean $X$ e $Y$ dos variables estadísticas.*

$$X \text{ e } Y \text{ son independientes} \Longleftrightarrow f_{ij} = f_{i\cdot} f_{\cdot j} \Longleftrightarrow n_{ij} = \frac{n_{i\cdot} n_{\cdot j}}{n} \quad \forall i, j$$

*Demostración.* Demostramos mediante la doble implicación:

$\Longrightarrow$) Suponemos $X$ e $Y$ independientes, es decir, $f_{i/j} = f_{i\cdot}$

Por tanto,

$$f_{ij} = f_{i/j} f_{\cdot j} = f_{i\cdot} f_{\cdot j}$$

$\Longleftarrow$) Suponemos $f_{ij} = f_{i\cdot} f_{\cdot j}$.

Probemos que $X$ e $Y$ son linealmente independientes.

$$f_{\cdot j} = \frac{f_{ij}}{f_{i\cdot}} = f_{j/i}$$

□

**Proposición 2.3.** *Si el carácter $X$ es independiente del carácter $Y$, entonces $Y$ es independiente de $X$ (la independencia es una propiedad recíproca).*

*Demostración.* Supuesto $X$ independiente de $Y$, tenemos $f_{i\cdot} = f_{i/j} \quad \forall j = 1, \ldots, p$

$$f_{ij} = f_{i\cdot} f_{\cdot j} = f_{i\cdot} f_{j/i} \Longrightarrow f_{\cdot j} = f_{j/i}$$

demostrando así que $X$ es independiente de $Y$. □

Se dice que el carácter $X$ depende funcionalmente del carácter $Y$ si a cada modalidad de $y_j$ de $Y$ le corresponde una única modalidad posible de $X$ con frecuencia no nula. Es decir:

$$\forall j \in \{1, \ldots, p\}, n_{ij} = 0 \text{ excepto para un valor } i = \varphi(j) \mid n_{ij} = n_{\cdot j}$$

**Ejemplo.** Consideramos la siguiente tabla estadística bidimensional:

| $X \backslash Y$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
|---|---|---|---|---|---|---|
| $x_1$ | 3 | 0 | 6 | 0 | 0 | 9 |
| $x_2$ | 0 | 4 | 0 | 0 | 2 | 6 |
| $x_3$ | 0 | 0 | 0 | 5 | 0 | 5 |
| | 3 | 4 | 6 | 5 | 2 | 20 |

En dicha distribución conjunta, $X$ depende funcionalmente del carácter $Y$. Sin embargo, $Y$ no depende funcionalmente del carácter $X$.

Si sucede que la dependencia funcional es bidireccional, hablaremos de una dependencia funcional recíproca. Notemos que esta es de poco interés estadístico.

## 2.7. Momentos bidimensionales

Dada una variable estadística bidimensional $(X, Y)$ con una distribución conjunta $\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,k \\ j=1,\dots,p}}$, se definen los momentos conjunto central y no central de órdenes $r$ y $s$ $(r, s \in \mathbb{N} \cup \{0\})$ como:

$$\mu_{rs} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}(x_i - \overline{x})^r (y_j - \overline{y})^s$$

$$m_{rs} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij} x_i^r y_j^s$$

Los momentos centrales más utilizados son las varianzas marginales, $\mu_{20} = \sigma_x^2$ y $\mu_{02} = \sigma_y^2$ y el momento $\mu_{11}$, cuya importancia se describe a continuación:

### 2.7.1. Varianza

**Definición 2.1.** Dadas dos variables estadísticas unidimensionales, $X$ y $Y$, se define la covarianza de las variables $X$ e $Y$ como:

$$\sigma_{xy} = Cov(X, Y) = \mu_{11}$$

**Proposición 2.4.** *Dadas dos variables estadísticas unidimensionales, $X$ y $Y$, se tiene:*

$$\sigma_{xy} = \mu_{11} = m_{11} - m_{10}m_{01}$$

*Demostración.*

$$\sigma_{xy} = Cov(X, Y) = \mu_{11} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}(x_i - \bar{x})(y_j - \bar{y}) = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}(x_i y_j - x_i \bar{y} - \bar{x} y_j + \bar{x}\bar{y}) =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{p} x_i y_j f_{ij} - \sum_{i=1}^{k} \sum_{j=1}^{p} x_i \bar{y} f_{ij} - \sum_{i=1}^{k} \sum_{j=1}^{p} \bar{x} y_j f_{ij} + \sum_{i=1}^{k} \sum_{j=1}^{p} \bar{x}\bar{y} f_{ij} =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{p} x_i y_j f_{ij} - \bar{y} \sum_{i=1}^{k} x_i f_{i\cdot} - \bar{x} \sum_{j=1}^{p} y_j f_{\cdot j} + \bar{x}\bar{y} \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij} =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{p} x_i y_j f_{ij} - \bar{y}\bar{x} - \bar{x}\bar{y} + \bar{x}\bar{y} =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{p} x_i y_j f_{ij} - \bar{x}\bar{y} = m_{11} - m_{10}m_{01}$$

□

**Proposición 2.5.** *Si $X$ e $Y$ son independientes* $\Longrightarrow \begin{cases} m_{rs} = m_{r0}m_{0s} \\ \mu_{rs} = \mu_{r0}\mu_{0s} \end{cases}$

*Demostración.* Supongo $X$ e $Y$ independientes, por lo que $f_{ij} = f_{i.}f_{.j}$. Entonces:

$$m_{rs} = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}x_i^r y_j^s = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{i.}f_{.j}x_i^r y_j^s = \sum_{i=1}^{k} f_{i.}x_i^r \sum_{j=1}^{p} f_{.j}y_j^s = m_{r0}m_{0s}$$

$$\mu_{rs} = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(x_i - \bar{x})^r(y_j - \bar{y})^s = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{i.}f_{.j}(x_i - \bar{x})^r(y_j - \bar{y})^s =$$

$$= \sum_{i=1}^{k} f_{i.}(x_i - \bar{x})^r \sum_{j=1}^{p} f_{.j}(y_j - \bar{y})^s = \mu_{r0}\mu_{0s}$$

□

**Corolario 2.5.1.** *Si $X$ e $Y$ son independientes* $\Longrightarrow \sigma_{xy} = 0$

*Demostración.* Supongo $X$ e $Y$ independientes, por lo que $m_{rs} = m_{r0}m_{0s}$. Entonces:

$$\sigma_{xy} = m_{11} - m_{10}m_{01} = m_{10}m_{01} - m_{10}m_{01} = 0$$

□

**Proposición 2.6.** *Si se transforman los valores de $x_i$ e $y_j$ mediante transformaciones lineales dadas por:*

$$\begin{cases} x_i' = ax_i + b \\ y_j' = cy_j + d \end{cases}$$

*La covarianza queda como:*

$$\sigma_{x'y'} = ac\sigma_{xy}$$

*Demostración.*

$$\sigma_{x'y'} = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(x_i' - \bar{x}')(y_j' - \bar{y}') = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}[(ax_i + b) - (a\bar{x} + b)][(cy_j + d) - (c\bar{y} + d)] =$$

$$= ac\sum_{i=1}^{k}\sum_{j=1}^{p}(x_i - \bar{x})(y_j - \bar{y})f_{ij} = ac\sigma_{xy}$$

□

Si expresamos nuevas variables a partir de otras, podemos calcular su covarianza a partir de la otra:

$$x_i' = ax_i + b \quad y_j' = cy_j + d$$

$$\sigma_{X'Y'} = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(ax_i + b - \overline{x'})(cy_j + d - \overline{y'}) =$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(ax_i + b - (a\overline{x} + b))(cy_j + d - (c\overline{y} + d)) =$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(ax_i - a\overline{x})(cy_j - c\overline{y}) = ac\sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(x_i - \overline{x})(y_j - \overline{y}) = ac\sigma_{xy}$$

## 2.8. Regresión

Pretendemos buscar que un número de magnitudes $X_1, \ldots, X_n$ se relacionen con una variable $Y$ mediante la expresión:

$$Y = f(X_1, \ldots, X_n)$$

Podemos abordar el problema desde dos enfoques:

- Regresión: La determinación de la estructura de dependencia que mejor expresa la relación de la variable $Y$ con las demás.
- Correlación: El estudio del grado de dependencia que existe entre las variables.

Si dos variables presentan una dependencia estadística (es decir, una dependencia no funcional), no es posible encontrar una ecuación tal que los valores que puedan presentar dichas variables la satisfagan. Es decir, no es posible encontrar una función que pase por todos los puntos del diagrama de dispersión que representa esa distribución conjunta. Por tato, tendremos que ajustar lo mejor posible una función a una serie de valores observados, encontrando una curva que, aunque no pase por todos los puntos de la nube, más se aproxime a ellos. Dicho método recibe el nombre de ajuste por mínimos cuadrados.

### 2.8.1. Método de mínimos cuadrados

Sea $f(x_i, a_0, \ldots, a_n)$ la función que aproxima la variable $Y$ en función de los valores de $X$.

**Notación.** A los valores ajustados se les notará de la siguiente manera:

$$\hat{y_j} = f(x_i; a_0, \ldots, a_n)$$

**Definición 2.2** (Residuo)**.** Se define el residuo de la modalidad $y_j$ de la variable $Y$ como:

$$e_{ij} = y_j - \hat{y_j} = y_j - f(x_i, a_0, a_1, \ldots, a_n)$$

El método de mínimos cuadrados consiste en encontrar una función $f$ que minimice la media de los cuadrados de los residuos:

$$ECM(f(x_i, a_0, a_1, \ldots, a_n)) = \psi(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij} e_{ij}^2$$

La función $\psi$ se denomina el error cuadrático medio de la función $f$, denotada $ECM(f(x_i, a_0, a_1, \ldots, a_n))$. Como los parámetros $(x_i, a_0, a_1, \ldots, a_n)$ sólo están sometidos a sumas, productos y cuadrados dentro de $\psi$, dicha función es derivable respecto a cada $a_i$ $\forall i \in \{0, \ldots, n\}$. Además, se puede asegurar que el punto $(\hat{a_0}, \hat{a_1}, \ldots, \hat{a_n})$ donde se anulan las derivadas parciales primeras respecto de cada $a_i$ corresponde a un mínimo de la función $\psi$.

El cálculo de los parámetros de la función de ajuste óptima según el método de los mínimos cuadrados consiste en resolver el siguiente sistema, llamado sistema de ecuaciones normales:

$$\frac{\partial \psi}{\partial a_r} = 0 \Rightarrow \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij} e_{ij} \frac{\partial f}{\partial a_r} = 0 \qquad \forall r \in \{0, \dots n\}$$

Una de las funciones de regresión más utilizadas para expresar el comportamiento de una variable en función de la otra es un polinomio de grado $n$ (comenzaremos con $n = 1$).

### Ajuste lineal (recta de regresión)

$$Y = f(X; a, b) = a + bX$$

Supongamos que queremos ajustar por el método de mínimos cuadrados una recta que exprese $Y$ en función de $X$. La función sería $Y = f(X; a, b) = a + bX$, por lo que tendremos que calcular el mínimo en $a$ y $b$ de la función:

$$\psi(a, b) = ECM(a, b) = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij} [y_j - (a + bx_i)]^2$$

Obtenemos el sistema de ecuaciones normales:

$$\left\{ \begin{array}{l} \dfrac{\partial \psi}{\partial a} = 0 \Rightarrow \displaystyle\sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij} [y_j - (a + bx_i)] = 0 \\ \\ \dfrac{\partial \psi}{\partial b} = 0 \Rightarrow \displaystyle\sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij} [y_j - (a + bx_i)] x_i = 0 \end{array} \right\} \Longrightarrow \left\{ \begin{array}{l} m_{01} = a + b m_{10} \\ \\ m_{11} = a m_{10} + b m_{20} \end{array} \right.$$

La resolución de dicho sistema nos proporciona los coeficientes buscados:

$$\hat{a} = m_{01} - \frac{m_{11} - m_{10} m_{01}}{m_{02} - m_{01}^2} m_{10} = \overline{y} - \frac{\sigma_{xy}}{\sigma_x^2} \overline{x}$$

$$\hat{b} = \frac{m_{11} - m_{10} m_{01}}{m_{02} - m_{01}^2} = \frac{\sigma_{xy}}{\sigma_x^2}$$

Por tanto, **la recta de regresión de $Y$ sobre $X$** tiene por expresión:

$$Y = \frac{\sigma_{xy}}{\sigma_x^2} X + \overline{y} - \frac{\sigma_{xy}}{\sigma_x^2} \overline{x} \qquad \left( \text{Equivalentemente}, Y - \overline{y} = \frac{\sigma_{xy}}{\sigma_x^2} (X - \overline{x}) \right)$$

**Definición 2.3** (Coeficiente de regresión lineal)**.** Al coeficiente $\dfrac{\sigma_{xy}}{\sigma_x^2}$ se le denomina coeficiente de regresión lineal de $Y$ sobre $X$.

Análogamente, se define el coeficiente de regresión lineal de $X$ sobre $Y$.

Análogamente, la recta de regresión mínimo cuadrática de $X$ sobre $Y$ es la recta $X = h(Y; c, d) = c + dY$ que minimiza la función

$$\phi(c, d) = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}(x_i - \hat{x}_j)^2$$

donde $\hat{x}_j = c + dy_j$. Siguiente el procedimiento anterior, llegamos a que **la recta de regresión de $X$ sobre $Y$ es**:

$$X - \overline{x} = \frac{\sigma_{xy}}{\sigma_y^2}(Y - \overline{y})$$

Los coeficientes de regresión son las pendientes de las rectas de regresión. Los signos de dichos coeficientes son los mismos para ambas rectas e igual al signo de la covarianza. Cuando exista dependencia funcional lineal, las dos rectas de regresión coincidirán con la recta de dependencia.

Algunas propiedades de la recta de regresión son:

**Lema 2.7.** *Las rectas de regresión pasan por el punto $(\bar{x}, \bar{y})$.*

*Demostración.* La recta de regresión de $Y$ sobre $X$ tiene la forma de $y - \bar{y} = K(x - \bar{x})$. Para $x = \bar{x}$, vemos que $y = \bar{y}$.

Análogamente, la recta de regresión de $X$ sobre $Y$ tiene la forma de $x - \bar{x} = K(y - \bar{y})$. Para $y = \bar{y}$, vemos que $x = \bar{x}$. □

**Lema 2.8.** *La media de los valores ajustados coincide con la de los valores observados de la variable.*

*Demostración.*

$$\overline{\hat{y}} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}\hat{y}_j = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}(ax_i + b) = a\bar{x} + b = \bar{y}$$

□

**Corolario 2.8.1.** *La media de los residuos vale 0.*

*Demostración.*

$$\sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}e_{ij} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}(y_j - \hat{y}_j) = \bar{y} - \overline{\hat{y}} = 0$$

□

**Corolario 2.8.2.** *La media de los productos de los residuos por los valores de la variable explicativa vale cero.*

*Demostración.*

$$\sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}e_{ij}x_i = \bar{x} \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}e_{ij} = 0$$

□

**Corolario 2.8.3.** *La media de los productos de los residuos por los valores ajustados vale cero.*

*Demostración.*

$$\sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}e_{ij}\hat{y}_j = \bar{\hat{y}}\sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}e_{ij} = 0$$

donde se ha aplicado la Proposición 1.3. □

### Ajuste polinómico

$$Y = f(X; a_0, a_1, \ldots, a_n) = a_0 + a_1X + \cdots + a_nX^n$$

Si queremos aproximar mediante un polinomio de grado superior o igual a dos, el método de mínimos cuadrados nos conducirá al sistema de ecuaciones:

$$\begin{cases} m_{01} & = a_0 + a_1m_{10} + \ldots + a_nm_{n0} \\ m_{11} & = a_0m_{10} + a_1m_{20} + \ldots + a_nm_{n+1,0} \\ m_{21} & = a_0m_{20} + a_1m_{30} + \ldots + a_nm_{n+2,0} \\ & \vdots \\ m_{n1} & = a_0m_{n0} + a_1m_{n+1,0} + \ldots + a_nm_{2n,0} \end{cases}$$

Para ajustar a la nube otro tipo de función, intentaremos pasar a un ajuste polinómico. Ejemplos de esto son los siguientes ajustes:

### Ajuste hiperbólico

$$Y = f(X; a, b) = a + b\frac{1}{X}$$

Si queremos realizar un ajuste hiperbólico mediante una hipérbola equilátera realizamos la transformación $Z = \dfrac{1}{X}$, y realizamos el ajusta de mínimos cuadrados a la recta $Y = a + bZ$ sobre las variables $(Z, Y)$.

### Ajuste potencial

$$Y = f(X; a, b) = aX^b$$

De otra forma, si queremos aplicar el ajuste potencial hemos de aplicar el logaritmo y obtenemos la siguiente expresión:

$$\ln Y = \ln a + b\ln X$$

Llamando a las variables $V = \ln Y$, $U = \ln X$, $A = \ln a$, quedándonos la siguiente expresión a calcular el ajuste lineal:

$$V = A + bU$$

**Ajuste exponencial**

$$Y = f(X; a, b) = ab^x$$

De otra forma, si queremos aplicar el ajuste exponencial hemos de aplicar el logaritmo y obtenemos la siguiente expresión:

$$\ln Y = \ln a + X \ln b$$

Llamando a las variables $V = \ln Y$, $A = \ln a$ y $B = \ln b$, quedándonos la siguiente expresión a calcular el ajuste lineal:

$$V = A + BX$$

### 2.8.2. Regresión de tipo I

Podemos además realizar regresiones de una variable dependiente $Y$ dado el valor $x_i$ de una variable independiente asociada $X$. Es decir, predecir el comportamiento de la variable condicionada $Y/X = x_i$.

Teniendo en cuenta la representatividad de le media en lo que al comportamiento de una variable se refiere, se define la curva de regresión de tipo I de $Y/X$ como la curva que pasa por los puntos $(x_i, \overline{y_i})$ $\forall i \in \{1, \ldots, k\}$. Análogamente, se defien la curva de regresión de tipo I de $X/Y$ como la curva que pasa por los puntos $(\overline{x_j}, y_j)$ $\forall j \in \{1, \ldots, p\}$.

Estas curvas tienen la propiedad de ser entre todas las funciones las que mejor se ajustan a los datos observados según el método de mínimos cuadrados. Estas curvas no son de gran utilidad práctica, pues el hecho de conocerla solamente en puntos aislados hace que sea inútil para le predicción en los demás casos.

## 2.9. Correlación

El grado de asociación entre las variables nos indicará en qué medida la expresión encontrada mediante la regresión explica una variable en función de la otra. El estudio de la correlación también equivale al estudio de la bondad del ajuste de una curva a una nube de puntos.

Para ello, en primer lugar es importante diferenciar entre los ajustes lineales en los parámetros y los no lineales en los parámetros.

Los que sí son lineales en los parámetros son aquellos a los que no se les ha aplicado ninguna transformación a los parámetros. Ejemplo de estos son los ajustes mediante rectas, parábolas o hipérbolas equiláteras.

Los no lineales en los parámetros implican que a alguno de los parámetros se le ha aplicado alguna transformación. Ejemplos son el ajuste potencial o el exponencial.

### 2.9.1. Varianza residual. Coeficiente de determinación

El método de mínimos cuadrados toma como medida del error que se comente al ajustar una curva la siguiente medida:

**Definición 2.4** (Varianza Residual)**.** Se define la varianza residual del ajuste de $Y$ en función de $X$ como:

$$\sigma_{ry}^2 = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}e_{ij}^2 = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(y_j - \hat{y}_i)^2 = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}[y_j - f(x_i)]^2$$

Dicha cantidad se usa como medida de la bondad del ajuste. Por tanto, cuanto menor sea la varianza resiudal, mejor será el ajuste.

*Observación.* En funciones lineales de los parámetros la media de los residuos es cero (generalización del lema 2.8), por lo que la expresión anterior es precisamente la varianza de los residuos, o varianza residual.

En funciones no lineales en los parámetros, la media de los residuos no es nula, aunque se sigue denominando varianza residual. Por ello, es importante no confundir la varianza residual con la varianza de los residuos en los ajustes no lineales en los parámetros.

También se define la siguiente medida:

**Definición 2.5** (Varianza Explicada)**.** Se define la varianza residual de $Y$ como:

$$\sigma_{ey}^2 = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(\hat{y}_j - \overline{y})^2$$

Por norma general, se toma como medida del grado de ajuste el coeficiente de determinación.

**Definición 2.6** (Coeficiente de determinación)**.** El coeficiente de determinación, o razón de correlación, es la proporción de la varianza total de la variable $Y$ explicada por la regresión. Esto es, el cociente o razón entre la varianza explicada y la total.

$$\eta_{Y/X}^2 = \frac{\sigma_{ey}^2}{\sigma_y^2}$$

Por tanto, para comparar todo tipo de ajustes se puede emplear la **varianza residual** o el **coeficiente de determinación**, aunque se suele emplear la primera medida.

### Caso concreto de ajustes lineales en los parámetros

En este caso, como la varianza residual coincide con la varianza de los residuos, se tiene que es posible descomponer la varianza en una suma de la varianza residual y la varianza explicada por la regresión:

$$\sigma_y^2 = \sigma_{ey}^2 + \sigma_{ry}^2$$

En este caso, se tiene que:

$$\eta_{Y/X}^2 := \frac{\sigma_{ey}^2}{\sigma_y^2} = 1 - \frac{\sigma_{ry}^2}{\sigma_y^2}$$

**Interpretación del coeficiente de correlación**

De la misma expresión se deduce que: $0 \leqslant \eta^2_{Y/X} \leqslant 1$.

- $\eta^2_{Y/X} = 0 \Leftrightarrow \dfrac{\sigma^2_{ey}}{\sigma^2_y} = 0 \Leftrightarrow \sigma^2_{ey} = 0$. Es decir, el modelo no explica nada de $Y$ a partir de $X$. El ajuste es el peor posible que se puede hacer por mínimos cuadrados.
- $\eta^2_{Y/x} = 1 \Leftrightarrow \dfrac{\sigma^2_{ey}}{\sigma^2_y} = 1$. Es decir, todos los residuos son nulos y s explica la variable totalmente. El ajuste es perfecto.
- Para valores intermedios entre 0 y 1, según estén más próximos a un extremo o a otro nos indicarán un peor o mejor ajuste: Un ajute del 60 % explica que el 60 % de la variabilidad total de $Y$ la explica el modelo propuesto mediante la variable independiente.

### 2.9.2. Correlación en el caso lineal

En este caso, tenemos el siguiente resultado, muy útil para calcular la bondad de los ajustes lineales:

**Teorema 2.9.** *En el caso de un ajuste lineal, el ajuste de determinación viene dado por:*

$$\eta^2_{Y/X} = \eta^2_{X/Y} = r^2 = \frac{\sigma^2_{xy}}{\sigma^2_x \sigma^2_y}$$

*Demostración.* Demostramos para la recta de $Y$ sobre $X$, ya que en el otro caso sería análogo. La recta mencionada tiene por expresión:

$$Y - \overline{y} = \frac{\sigma_{xy}}{\sigma^2_x}(X - \overline{x}) \Longrightarrow Y = \overline{y} + \frac{\sigma_{xy}}{\sigma^2_x}(X - \overline{x})$$

Por tanto, la varianza residual en el caso de la recta de regresión es:

$$\begin{aligned}
\sigma^2_{ry} &= \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}[y_j - f(x_i)]^2 = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}\left[y_j - \left(\overline{y} + \frac{\sigma_{xy}}{\sigma^2_x}(x_i - \overline{x})\right)\right]^2 = \\
&= \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}\left[(y_j - \overline{y}) - \left(\frac{\sigma_{xy}}{\sigma^2_x}(x_i - \overline{x})\right)\right]^2 = \\
&= \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}\left[(y_j - \overline{y})^2 + \frac{\sigma^2_{xy}}{(\sigma^2_x)^2}(x_i - \overline{x})^2 - 2\frac{\sigma_{xy}}{\sigma^2_x}(y_j - \overline{y})(x_i - \overline{x})\right] = \\
&= \sigma^2_y + \frac{\sigma^2_{xy}}{\sigma_x} - 2\frac{\sigma^2_{xy}}{\sigma^2_x} = \sigma^2_y - \frac{\sigma^2_{xy}}{\sigma^2_x}
\end{aligned}$$

Por tanto, como en este caso estamos ante un ajuste lineal en los parámetros, tenemos que:

$$\eta^2_{Y/X} = 1 - \frac{\sigma^2_{ry}}{\sigma^2_y} = 1 - \frac{\sigma^2_y - \frac{\sigma^2_{xy}}{\sigma^2_x}}{\sigma^2_y} = 1 - 1 + \frac{\sigma^2_{xy}}{\sigma^2_x\sigma^2_y} = \frac{\sigma^2_{xy}}{\sigma^2_x\sigma^2_y} \qquad \square$$

Este resultado es de gran ayuda, ya que nos permite calcular $\eta^2_{Y/X}$ de una forma mucho más cómoda.

Es útil calcular la varianza residual en función de $r^2$, ya que el cálculo del segundo es mucho más sencillo. No obstante, es necesario a veces conocer la varianza residual para comparar con modelos no lineales en los parámetros. Por eso, se tiene que:

$$\eta^2_{Y/X} = 1 - \frac{\sigma^2_{ry}}{\sigma^2_y} = r^2 \Longrightarrow \sigma^2_{ry} = (1 - r^2)\sigma^2_y$$

Por último, se introduce un nuevo coeficiente. La raíz cuadrada del coeficiente de determinación lineal anterior (con el signo de la covarianza) recibe el nombre de coeficiente de correlación lineal:

$$r = \pm\sqrt{r^2} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Dicho coeficiente se usa para determinar el grado de dependencia lineal de la variable dependiente ante los valores de la variable independiente. Esta dependencia puede ser directa (o positiva) o indirecta (o negativa), según el signo de la covarianza. Adopta valroes entre $\pm 1$ y 0.

- Para la covarianza positiva:

  Si $r = 1$, existirá una dependencia lineal funcional, mientras que si $r = 0$ no existirá ninguna dependencia o asociación entre las variables de tipo lineal, aunque sí puede haberla de otra naturaleza, convirtiéndose las rectas de regresión paralelas a los ejes de coordenadas.

- Para la covarianza negativa: Si $r = -1$, existirá una correlación perfecta, con una dependencia funcional lineal, coincidiendo las dos rectas en una sola.

Para resumir, diremos que $-1 \leqslant r \leqslant 1$. Cuando varía de $-1$ a 0 estamos en una correlación negativa y la dependencia será mayor cuanto más se aproxime a $-1$ mientras que si varía de 0 a 1, la correlación es positiva y el grado de dependencia será mayor cuanto más se aproxime a 1.

## 2.10. Predicciones

Uno de los objetivos principales de la regresión y correlación es hacer predicciones de la variable dependiente en función de los valores que toma la variable independiente. Las predicciones se efectúan utilizando la función estimada por el método de mínimos cuadrados, $f$. Con la que obtenemos los valores teóricos que ajustan a los observados. La predicción será más fiable cuanto mayor sean los coeficientes de determinación correspondientes o razones de correlación, ya que menor será la varianza de los residuos, que nos indica la cuantía de la separación entre lo observado y estimado.

Hay que tener presente que la fiabilidad de las predicciones disminuye a medida que los valores de la variable independiente se aleja de su recorrido, pues puede que el modelo ajustado no sea válido para dicho valores en la medida dada por $\eta^2$

# 3. Combinatoria

A continuación, y antes de adentrarnos en la Probabilidad y como puente a ella, haremos una breve introducción a la combinatoria, aquella rama de las matemáticas que estudia las diversas formas de realizar agrupaciones de conjuntos a partir de uno.

Dependiendo de las agrupaciones que queramos realizar distinguimos distintos tipos de sucesos, dependiendo de si importa el orden o no.

$$\text{¿Importa el orden?} \begin{cases} \text{Sí} \begin{cases} \text{Variaciones} \\ \text{Permutaciones} \end{cases} \\ \\ \text{No} \left\{ \text{Conmutaciones} \right. \end{cases}$$

Donde cada una de ellas puede ser con y sin repeticiones. Es decir, hay que diferenciar si, tras escoger un elemento del conjunto, para la siguiente elección se vuelve a tener en cuenta o ya no.

## 3.1. Variaciones

Las variaciones consisten en, dado un conjunto de $m$ elementos, calcular de cuantas formas posibles somos capaces de formar grupos de $n$ elementos ($n, m \in \mathbb{N}$).

Se trata de una disciplina en la que importa el orden, por lo que diremos que dos variaciones de $n$ elementos (ambas obtenidas a partir del mismo conjunto de $m$ elementos) son distintas si y sólo si tienen distintos elementos o si estos están dispuestos en distinto orden.

Es decir, a partir de $\{0, 1, 2\}$:
Las variaciones $\{0, 1\}, \{0, 2\}, \{1, 0\}$ son distintas entre sí.

### 3.1.1. Variaciones sin repetición

Son aquellas en las que dado un conjunto padre de $m$ elementos distintos, sólo podremos obtener el elemento $a$ del conjunto padre una única vez. Es decir, cada conjunto hijo de tamaño $n$ contendrá 0 o 1 veces el elemento $a$ del conjunto padre. Por tanto, todas estas cumplen que $1 \leqslant n \leqslant m$.

Por tanto, dado el conjunto padre $\{0, 1, 2\}$, la variación $\{1, 1\}$ no es una variación sin repetición.

Para calcular el número de formas distintas en las que podemos generar variaciones sin repetición de tamaño $n$ a partir de un conjunto padre de tamaño $m$ (con $1 \leqslant n \leqslant m$) usaremos la expresion:

$$V_{m,n} = \frac{m!}{(m-n)!}$$

Que se deduce de la expresión:

$$V_{m,n} = m \cdot (m-1) \cdot (m-2) \cdot \ldots \cdot (m-n+1) = \frac{m!}{(m-n)!}$$

De todos los posible elementos $m$, cogemos 1 luego hemos tenido $m$ posibles opciones. A continuación, elegimos otro elemento de esos $m-1$ restantes, luego hemos vuelto a tener $m-1$ posibles opciones, repetimos el proceso hasta elegir $m-n+1$ (el último elemento que hay que coger para obtener un total de $n$) elementos y cogemos otro, luego hemos vuelto a tener $m-n+1$ opciones y ya tenemos los $n$ elementos deseados.

### 3.1.2. Variaciones con repetición

Esta disciplina se basa en, dado un conjunto padre de $m$ elementos distintos, obtenemos $n$ elementos de este, pudiendo repetir cualquier elemento $a$ del conjunto padre las veces que queramos. Estamos interesados en calcular el número de formas posibles en las que podemos formar conjuntos de $n$ elementos que contengan los elementos del conjunto padre de $m$ elementos. Notemos que, en este caso, $n$ puede ser mayor que $m$.

Un ejemplo de variación con repetición de un conjunto $\{0, 1\}$ pueden ser:
$\{0, 0, 0, 1, 1\}, \{1, 1, 1, 1, 1\}$.

Para calcular el número de formas distintas en las que podemos generar variaciones con repetición de tamaño $n$ a partir de un conjunto padre de $m$ elementos, usamos la expresión:

$$VR_{m,n} = m^n$$

Que se deduce del siguiente razonamiento: De los $m$ elementos del conjunto padre, para el primer elemento de nuestra variación podemos hacer $m$ elecciones. Para el segundo elemento de la variación seguimos pudiendo hacer $m$ elecciones, luego tenemos $m \cdot m$ posibilidades para los dos primeros elementos, ..., para el elemento $n$-ésimo seguimos teniendo $m$ posibilidades, luego hay

$$m \cdot \underbrace{\ldots}_{n \text{ veces}} \cdot m = m^n$$

## 3.2. Permutaciones

Las permutaciones consisten en, dado un conjunto de $m$ elementos, calcular de cuantas formas posibles somos capaces de reordenar dicho conjunto en conjuntos distintos.

Se trata de una disciplina en la que obviamente importa el orden, por lo que dos permutaciones de $m$ elementos serán distintas si sus elementos están dispuestos en distinto orden.

Por tanto, dado el conjunto $\{1, 2\}$ las permutaciones $\{1, 2\}, \{2, 1\}$ son distintas.

### 3.2.1. Permutaciones sin repetición

Esta disciplina cuenta el número de formas posibles en las que podemos reordenar un conjunto de $m$ elementos en el que cada elemento se repite una única vez. Es decir, no hay elementos repetidos en él. Por tanto, un conjunto de la forma $\{0, 1, 1\}$ sería un conjunto inválido para esta modalidad.

Notemos que las permutaciones de $m$ elementos sin repetición coinciden con las variaciones sin repetición de elegir $m$ elementos a partir de $m$ elementos del conjunto padre. Por tanto, igualando $n = m$ tenemos la expresión deseada para calcular el número de formas distintas en las que podemos reordenar un conjunto de $m$ elementos:

$$P_m = V_{m,m} = m!$$

Su razonamiento es análogo al de las variaciones sin repetición, esta vez cogiendo $m$ elementos.

### 3.2.2. Permutaciones con repetición

Esta disciplina trata de contar el número de formas posibles en las que podemos reordenar un conjunto de $m$ elementos en el que es posible que un elemento se repite varias veces. Es decir, el primer elemento se repite $n_1$ veces, el siguiente que es distinto al primero $n_2$ veces, ..., el último elemento distinto a los anteriores se repite $n_k$ veces. Notemos que $\sum\limits_{i=1}^{k} n_i = m$.

Para calcular el número de formas distintas en las que podemos reordenar un conjunto de $m$ elementos en el que hay $n_1$ elementos iguales, $n_2$ elementos iguales, ..., $n_k$ elementos iguales viene dado por la expresión:

$$PR_m^{n_1,n_2,\dots,n_k} = \frac{m!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}$$

Su razonamiento es, partiendo de una permutación sin repetición, tenemos $m!$ formas posibles de permutar el conjunto, lo que pasa es que tenemos elementos repetidos, el primer elemento se repite $n_1$ veces, el segundo distinto al primero $n_2$ veces, ... Por tanto, de todas las formas posibles en las que podemos ordenar el conjunto queremos quitarnos la posibilidad de intercambiar los elementos que son iguales entre sí, que lo conseguimos dividiendo entre $n_i!$ $\forall i \in \{1, \dots, k\}$.

## 3.3. Combinaciones

Las combinaciones de $n$ elementos a partir de un conjunto de $m$ elementos distintos son variaciones en las que no importa el orden, es decir, las combinaciones del conjunto padre $\{0, 1, 2\}$ dadas por $\{0, 1\}, \{1, 0\}$ son iguales. Por tanto diremos que dos combinacinoes son distintas si y sólo si tienen elementos distintos.

Notemos que las combinaciones (con y sin repetición) son variaciones (con y sin variación) en las que no importa el orden. Para adentrarnos en la diferencia entre las que son con repetición o sin repetición, es necesario definir el siguiente concepto:

**Definición 3.1** (Número combinatorio)**.** Dados $m, n \in \mathbb{N} \cup \{0\}$, con $m \geqslant n$, el número combinatorio "$m$ sobre $n$" es un número entero que resulta de evaluar la siguiente expresión:

$$\binom{m}{n} := \frac{m!}{(m-n)! \cdot n!}$$

### 3.3.1. Combinaciones sin repetición

Esta modalidad se basa en, dado un conjunto padre de $m$ elementos distintos, obtener $n$ elementos de dicho conjunto (cada elemento sólo se podrá obtener como máximo una vez), sin que importe el orden. Por tanto, debemos tener en cuenta que $1 \leqslant n \leqslant m$.

El número de formas distintas de las que podemos tomar $n$ elementos distintos de un conjunto padre con $m$ elementos distintos sin que importe el orden (con $1 \leqslant n \leqslant m$) viene dado por la expresión:

$$C_{m,n} = \frac{V_{m,n}}{P_n} = \frac{m!}{(m-n)! \cdot n!} =: \binom{m}{n}$$

Es decir, consideramos el número total de variaciones de $n$ elementos a partir de un conjunto de $m$ elementos distintos y le quitamos las distintas formas en las que podemos reordenar dicho conjunto de $n$ elementos, es decir, nos olvidamos de las permutaciones de $n$ elementos que podemos realizar ya que con ellas obtendríamos conjuntos iguales al no importar el orden.

### 3.3.2. Combinaciones con repetición

Esta modalidad es similar a la anterior, dada una variación con repetición de $n$ elementos a partir de un conjunto padre de $m$ elementos distintos, tratamos de olvidarnos de aquellos conjuntos que teniendo los mismos elementos están ordenados de formas distintas.

Para calcular el número de formas en las que podemos coger una variación de $n$ elementos a partir de un conjunto padre de $m$ elementos distintos sin que importe el orden usaremos la expresión:

$$CR_{m,n} = C_{m+n-1,n} = \binom{m+n-1}{n} = \frac{(m+n-1)!}{(m-1)! \cdot n!}$$

# 4. Axiomática probabilística

La Probabilidad es la ciencia que estudia cómo debe emplearse la información y cómo dar una guía de acción en situación prácticas que envuelven incertidumbre.

- **Fenómenos determinísticos:** Los que dan lugar al mismo resultado si se hacen bajo condiciones idénticas.
- **Fenómenos aleatorios:** los resultados pueden variar incluso si el estudio se realiza con las mismas condiciones iniciales.

Características de los fenómenos aleatorios:

- El experimento se puede repetir indefinidamente bajo idénticas condiciones.
- Cualquier modificación mínima en las condiciones iniciales de la repetición puede modificar completamente el resultado final del experimento.
- Se puede determinar el conjunto de posibles resultados del experimento, pero no se puede predecir previamente un resultado particular.
- Si el experimento se repite un número grande de veces, entonces aparece algún modelo de regularidad estadística en los resultados obtenidos.

## 4.1. Espacio Muestral

Si consideramos un experimento aleatorio arbitrario, cada uno de los posibles resultados que no puedan descomponerse en otros más simples recibirán el nombre de suceso elemental. En el ejemplo de tirar un dado, los posibles sucesos elementales son: 1, 2, 3, 4, 5 y 6.

El conjunto que contiene a todos los sucesos elementales asociados a un experimento aleatorio recibe el nombre de espacio muestral, $\Omega$. En el ejemplo anterior: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Los espacios muestrales pueden ser finitos, infinitos numerables o continuos.

Llamaremos suceso (o suceso aleatorio) a cualquier subconjunto del espacio muestral, es decir, un conjunto de sucesos elementales cuya aparición da lugar a un suceso. En el ejemplo del dado, podemos considerar el suceso A de sacar un número par: $A = \{2, 4, 6\} \subseteq \Omega$

### 4.1.1. Sucesos

Cabe destacar que existen diversos tipos de sucesos:

- **Suceso elemental:** Definido anteriormente, cada uno de los resultados posibles de nuestro experimento aleatorio, consta de un único elemento del espacio muestral.
- **Suceso compuesto:** Aquel que consta de dos o más sucesos.
- **Suceso seguro o universal:** Aquel que ocurre siempre. Consta con todos los sucesos elementales del espacio muestral y, por tanto, se identifica con él. En el ejemplo anterior, un suceso seguro es sacar en un dado un número del 1 al 6.
- **Suceso imposible:** Aquel que no puede ocurrir nunca. No contiene ningún elemento del espacio muestral, lo representamos con $\emptyset$. En el ejemplo del dado, un suceso imposible es sacar un 7 al tirar el dado.

### 4.1.2. Relaciones y Operaciones de sucesos

Dados dos sucesos $A$ y $B$ de un experimento aleatorio, diremos que el suceso $A$ está contenido en el suceso $B$, notado $A \subseteq B$ si siempre que ocurre el suceso $A$ ocurre el suceso $B$, dicho en lenguaje de conjuntos, si $\forall a \in A\ a \in B$.

Dados dos sucesos $A$ y $B$ de un experimento aleatorio, diremos que el suceso $A$ es igual al suceso $B$ si siempre que ocurre el suceso $A$ ocurre el suceso $B$ y siempre que ocurre el suceso $B$ ocurre el suceso $A$, es decir: $A \subseteq B\ \wedge\ B \subseteq A$.

Dados dos sucesos $A$ y $B$ de un experimento aleatorio, definimos la unión de $A$ y de $B$, notado $A \cup B$ como aquel suceso que ocurre siempre que ocurre $A$ o que ocurre $B$, es decir: $\forall a \in A \cup B \Rightarrow a \in A\ \vee\ a \in B$

Dados dos sucesos $A$ y $B$ de un experimento aleatorio, definimos la intersección de $A$ y de $B$, notado $A \cap B$ como aquel suceso que si ocurre implica que ocurre $A$ y que ocure $B$, es decir: $\forall a \in A \cap B \Rightarrow a \in A\ \wedge\ a \in B$

Dados dos sucesos $A$ y $B$ de un experimento aleatorio, definimos el complementario de $A$ en $B$, notado $B - A$ como aquel suceso que ocurre siempre que ocurre $B$ y no ocurre $A$, es decir: $\forall a \in B - A \Rightarrow a \in B\ \wedge\ a \notin A$.
Podemos hablar solamente del complementario de un suceso. En este caso, entenderems que el complementario de un suceso $A$ de un experimento aleatorio, notado $\overline{A}$ es igual al complementario de $A$ en el espacio muestral $\Omega$: $\overline{A} = \Omega - A$.

Diremos que dos sucesos $A$ y $B$ son disjuntos o incompatibles si no pueden ocurrir simultáneamente, es decir, que su intersección sea vacía: $A \cap B = \emptyset$. En el ejemplo de tirar un dado, el suceso de que salga un número par es incompatible con el suceso de que salga impar.

Diremos que un conjunto de sucesos $\{A_1, A_2, \ldots, A_n\}$ es un sistema exhaustivo de sucesos si la unión de todos ellos es igual al espacio muestral: $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$

Diremos que un conjunto de sucesos $\{A_1, A_2, \ldots, A_n\}$ es un sistema completo de sucesos o una partición del espacio muestral si dicho conjunto constituye un sistema exhaustivo de sucesos y además son mutuamente excluyentes, es decir, son disjuntos dos a dos:

$$A_1 \dot{\cup} A_2 \dot{\cup} \ldots \dot{\cup} A_n = \Omega$$

## 4.2. Estructuras álgebra y $\sigma-$álgebra

Un conjunto de subconjuntos de $\Omega$ no trivial $\mathcal{A}$ se dice que tiene estructura de álgebra de sucesos o álgebra de Boole si verifica que:

- $\forall A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$
- $\forall A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cup A_2 \in \mathcal{A}$

Notemos que $\forall \mathcal{A} \Rightarrow \Omega \in \mathcal{A} \ \wedge \ \emptyset \in \mathcal{A}$.
Un ejemplo de álgebra de Boole es, para un cierto $\Omega$, $\mathcal{A} = P(\Omega)$.

Un conjunto de subconjuntos de $\Omega$ no trivial $\mathcal{A}$ se dice que tiene estructura de $\sigma-$álgebra si verifica que:

- $\forall A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$
- $\forall A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

Notemos, al igual que antes, que $\forall \mathcal{A} \Rightarrow \Omega \in \mathcal{A} \ \wedge \ \emptyset \in \mathcal{A}$. Además, cualquier $\sigma-$álgebra es un álgebra.

**Notación.** A lo largo de este documento, cada vez que aparezca $\mathcal{A}$ haremos referencia a una $\sigma-$álgebra.

## 4.3. Diferentes concepciones de Probabilidad

Definimos una probabilidad como una fución $P : \mathcal{A} \to \mathbb{R}$, que verifica unos ciertos axiomas. A continuación veremos algunas concepciones distintas de cómo calcular la imagen de cualquier elemento de la $\sigma-$álgebra.

### 4.3.1. Concepción clásica

Sea $A$ un suceso arbitrario asociado a un experimento aleatorio, se define la probabilidad del suceso $A$ mediante la **Regla de Laplace** como:

$$P(A) = \frac{|A|}{|\Omega|}$$

### 4.3.2. Concepción frecuentista

Si se realizan $N$ repeticiones de un experimento y un determinado suceso $A$ se ha presentado en $N_A$ ocasiones, se define la frecuencia relativa de $A$ en las $N$ pruebas como:

$$f_N(A) = \frac{N_A}{N}$$

Supongamos que el número de realizaciones del experimento crece indefinidamente y consideramos la sucesión de frecuencias relativas de $A$. Se define la probabilidad de $A$ como:

$$P(A) = \lim_{N\to\infty} f_N(A)$$

## 4.4. Definición axiomática de Kolmogorov

Sea $(\Omega, \mathcal{A})$ un espacio medible asociado a un experimento aleatorio, se define una probabilidad como una función

$$P : \mathcal{A} \to [0,1]$$

que verifica los siguiente axiomas:

- Axioma de la no negatividad: $P(A) \geqslant 0 \;\; \forall A \in \mathcal{A}$
- Axioma del suceso seguro: $P(\Omega) = 1$
- Axioma de $\sigma$-aditividad o aditividad numerable: Sea $A_1, A_2, \ldots \in \mathcal{A}$ una sucesión de sucesos incompatibles ($A_i \cap A_j = \emptyset \; \forall i \neq j$) entonces:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Algunas consecuencias de la definición axiomática de una probabilidad son:

**Proposición 4.1.** *La probabilidad del suceso imposible es nula:* $P(\emptyset) = 0$.

*Demostración.*

$$1 = P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset) = 1 + P(\emptyset) \Rightarrow P(\emptyset) = 0$$

□

**Proposición 4.2.** $\forall A \in \mathcal{A} \Rightarrow P(\overline{A}) = 1 - P(A)$.

*Demostración.*

$$P(A \cup \overline{A}) = P(\Omega) = 1 = P(A) + P(\overline{A}) \Rightarrow P(\overline{A}) = 1 - P(A)$$

□

**Proposición 4.3.** $P$ *es monótona no decreciente. Es decir,*

$$\forall A, B \in \mathcal{A} \mid A \subseteq B \Rightarrow P(A) \leqslant P(B)$$

*Además,* $P(B - A) = P(B) - P(A)$

*Demostración.*

$$P(B) = P(A \cup (B - A)) = P(A) + P(B - A) \Rightarrow P(B) \geqslant P(A)$$

Además, se tiene que:

$$P(B) = P(A) + P(B - A) \Rightarrow P(B - A) = P(B) - P(A)$$

□

**Proposición 4.4.** $\forall A \in \mathcal{A} \Rightarrow P(A) \leqslant 1$

*Demostración.*

$$\forall A \in \mathcal{A} \Rightarrow A \subseteq \Omega \Rightarrow 1 = P(\Omega) \geqslant P(A)$$

□

**Proposición 4.5.** $\forall A, B \in \mathcal{A} \Rightarrow P(B - A) = P(B) - P(A \cap B)$

*Demostración.*

$$P(B) = P((B \cap A) \cup (B - A)) = P(B \cap A) + P(B - A) \Rightarrow P(B - A) = P(B) - P(B \cap A)$$

□

**Proposición 4.6.** $\forall A, B \in \mathcal{A} \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$

*Demostración.*

$$\begin{aligned} P(B \cup A) &= P[(B \cup A) \cap \Omega] = P[(B \cup A) \cap (A \cup \overline{A})] = P[(B \cap \overline{A}) \cup A] = \\ &= P(B \cap \overline{A}) + P(A) = P(B - A) + P(A) = P(B) - P(A \cap B) + P(A) \end{aligned}$$

□

**Proposición 4.7** (Subaditividad finita)**.** $\forall A, B \in \mathcal{A}$

$$P(A \cup B) \leqslant P(A) + P(B)$$

*En general, dados $A_1, A_2, \ldots \in \mathcal{A}$ se verifica que:*

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} P(A_i)$$

*Demostración.*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow P(A \cup B) \leqslant P(A) + P(B)$$

Para demostrarlo de forma general, hacemos inducción sobre $n$.

- <u>Para $n = 1$</u>: Trivialmente es cierto.
- <u>Para $n = 2$</u>: Se acaba de probar previamente, en el caso particular.

- Supuesto cierto para $n-1$, comprobamos para $n$:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n-1} A_i \cup A_n\right) = P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n) - P\left(\bigcup_{i=1}^{n-1} A_i \cap A_n\right) \overset{(*)}{\leqslant}$$
$$\overset{(*)}{\leqslant} \sum_{i=1}^{n-1} P(A_i) + P(A_n) - P\left(\bigcup_{i=1}^{n-1} A_i \cap A_n\right) =$$
$$= \sum_{i=1}^{n} P(A_i) - P\left(\bigcup_{i=1}^{n-1} A_i \cap A_n\right) \leqslant \sum_{i=1}^{n} P(A_i)$$

donde en $(*)$ hemos aplicado la hipótesis de inducción.

$\square$

**Corolario 4.7.1** (Subaditividad numerable)**.** *Dada una colección de sucesos* $A_1, A_2, \ldots \in \mathcal{A}$*, se verifica:*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} P(A_i)$$

**Proposición 4.8** (Principio de inclusión-exclusión)**.** *Dada una colección de sucesos* $A_1, A_2, \ldots A_n \in \mathcal{A}$*, entonces:*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i \cap A_j) +$$
$$+ \sum_{i<j<k}^{n} P(A_i \cap A_j \cap A_k) + \ldots + (-1)^{n-1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

*Demostración.* Hacemos inducción sobre $n$:

- Para $n=2$: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$, cierto.

- Supuesto cierto para $n-1$, comprobamos para $n$:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n-1} A_i \cup A_n\right) = P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n) - P\left(\bigcup_{i=1}^{n-1} A_i \cap A_n\right) =$$
$$\overset{HI}{=} \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n-1} P(A_i \cap A_j) + \sum_{i<j<k}^{n-1} P(A_i \cap A_j \cap A_k) +$$
$$+ \ldots + (-1)^{n-2} P\left(\bigcap_{i=1}^{n-1} A_i\right) - P\left(\bigcup_{i=1}^{n-1} (A_i \cap A_n)\right) =$$
$$\overset{HI}{=} \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n-1} P(A_i \cap A_j) + \sum_{i<j<k}^{n-1} P(A_i \cap A_j \cap A_k) +$$
$$+ \ldots + (-1)^{n-2} P\left(\bigcap_{i=1}^{n-1} A_i\right) - \left[\sum_{i=1}^{n-1} P(A_i \cap A_n) - \sum_{i<j}^{n-1} P(A_i \cap A_j \cap A_n) +\right.$$
$$\left. + \sum_{i<j<k}^{n-1} P(A_i \cap A_j \cap A_k \cap A_n) + \ldots + (-1)^{n-2} P\left(\bigcap_{i=1}^{n-1} A_i \cap A_n\right)\right]$$

Lo que encontramos entre corchetes es la última iteración (la $n$) de los bucles anteriores, luego:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n-1} A_i \cup A_n\right) = P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n) - P\left(\bigcup_{i=1}^{n-1} A_i \cap A_n\right) =$$

$$= \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i \cap A_j) + \sum_{i<j<k}^{n} P(A_i \cap A_j \cap A_k) +$$

$$+ \ldots + (-1)^{n-1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

□

**Proposición 4.9** (Desigualdad de Bonferroni)**.** *Dada una colección de sucesos* $A_1, A_2, \ldots A_n \in \mathcal{A}$*, entonces:*

$$\sum_{i=1}^{n} P(A_i) \geqslant P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i \cap A_j)$$

*Demostración.* Hacemos inducción sobre $n$:

- Para $n = 2$: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$, siendo ciertas por tanto las dos desigualdades.
- Supuesto cierto para $n-1$, comprobamos para $n$:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n-1} A_i \cup A_n\right) = P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n) - P\left(\bigcup_{i=1}^{n-1} A_i \cap A_n\right) \geqslant$$

$$\overset{HI}{\geqslant} \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n-1} P(A_i \cap A_j) - P\left(\bigcup_{i=1}^{n-1} A_i \cap A_n\right) \geqslant$$

$$\overset{HI}{\geqslant} \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n-1} P(A_i \cap A_j) - \sum_{i=1}^{n-1} P(A_i \cap A_n) =$$

$$= \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i \cap A_j)$$

La primera desigualdad se demuestra con la primera línea, mientras que es necesario el resto para demostrar la segunda desigualdad.

□

**Proposición 4.10** (Desigualdad de Boole)**.** $\forall A, B \in \mathcal{A}$*, entonces:*

$$P(A \cap B) \geqslant 1 - P(\overline{A}) - P(\overline{B})$$

*Demostración.*

$$P(A \cap B) = P(\overline{\overline{A \cap B}}) = P(\overline{\overline{A} \cup \overline{B}}) = 1 - P(\overline{A} \cup \overline{B}) = 1 - P(\overline{A}) - P(\overline{B}) + P(\overline{A} \cap \overline{B})$$

Por tanto, tenemos que

$$P(A \cap B) \geqslant 1 - P(\overline{A}) - P(\overline{B})$$

□

# 5. Probablidad condicionada e Independencia

## 5.1. Definiciones

**Definición 5.1.** Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico arbitrario y $A \in \mathcal{A} \mid P(A) > 0$, definimos $\forall B \in \mathcal{A}$ la **probabilidad condicionada de** $B$ **sobre** $A$ como:

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

Es decir, acabamos de definir la función $P(\cdot \mid A) : \mathcal{A} \to [0,1]$ que nos lleva a cada $B \in \mathcal{A}$ en $P(B \mid A)$.

Veamos que esta nueva función cumple la axiomática de Kolmogorov y que, por tanto, puede definir una nueva probabilidad a partir de la ya definida.

**Proposición 5.1.** *Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico arbitrario y $A \in \mathcal{A} \mid P(A) > 0$. Para todo $B \in \mathcal{A}$, la probabilidad dondicionada de $B$ sobre $A$ cumple las tres condiciones de la Axiomática de Kolmogorov.*

*Demostración.* Demostramos cada una de las 3 condiciones:

1. Axioma de la no-negatividad

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} \geqslant 0 \quad \forall B \in \mathcal{A}$$

2. Axioma del suceso seguro

$$P(\Omega \mid A) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

3. Axioma de $\sigma$-aditividad o aditividad numerable

   Sea $A_1, A_2, \ldots \in \mathcal{A}$ una sucesión de sucesos incompatibles, entonces:

$$P\left(\bigcup_{i=1}^{\infty} B_i \mid A\right) = \frac{P\left[\left(\bigcup_{i=1}^{\infty} B_i\right) \cap A\right]}{P(A)} = \frac{P\left(\bigcup_{i=1}^{\infty} B_i \cap A\right)}{P(A)} =$$

$$= \frac{\sum_{i=1}^{\infty} P(B_i \cap A)}{P(A)} = \sum_{i=1}^{\infty} \frac{P(B_i \cap A)}{P(A)} = \sum_{i=1}^{\infty} P(B_i \mid A) \qquad \square$$

Por tanto, $P(\cdot \mid A)$ es también una probabilidad y por tanto, cumple todas las consecuencias vistas en la sección anterior.

## 5.2. Teoremas

Observemos que de la propia definición de la probabilidad condicionada se tiene que:

$$\begin{array}{lll} P(A\cap B) = P(A)P(B \mid A) & \text{Si} & P(A) > 0 \\ P(A\cap B) = P(B)P(A \mid B) & \text{Si} & P(B) > 0 \end{array}$$

Ambas ecuaciones se suman a la lista ya disponible de cálculo de probabilidades de intersecciones:

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

**Teorema 5.2** (de la probabilidad compuesta)**.** *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio de probabilidad y* $A_1, A_2, \dots, A_n \in \mathcal{A}$ *con* $P\left(\bigcap\limits_{i=1}^{n-1} A_i\right) > 0$*, entonces:*

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1\cap A_2)\cdot \ldots \cdot P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right)$$

*Demostración.* Realizamos inducción sobre $n$:

- <u>Para $n=2$</u>: Lo tenemos visto.
- <u>Supongámoslo cierto para $n-1$, veamos el caso $n$</u>:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(\bigcap_{i=1}^{n-1} A_i \cap A_n\right) = P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) P\left(\bigcap_{i=1}^{n-1} A_i\right) =$$

$$\stackrel{HI}{=} P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1\cap A_2)\cdot \ldots \cdot P\left(A_{n-1} \mid \bigcap_{i=1}^{n-2} A_i\right) P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right)$$

□

**Teorema 5.3** (de la Probabilidad Total)**.** *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio de probabilidad y sea* $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ *una partición de* $\Omega$ *con* $P(A_n) > 0 \ \ \forall n \in \mathbb{N}$*. Entonces,* $\forall B \in \mathcal{A}$*:*

$$P(B) = \sum_{i=1}^{\infty} P(A_i)P(B \mid A_i)$$

*Demostración.*

$$P(B) = P(B\cap\Omega) = P\left(B \cap \left(\bigcup_{i=1}^{\infty} A_i\right)\right) = P\left(\bigcup_{i=1}^{\infty} (B\cap A_i)\right) =$$

$$= \sum_{i=1}^{\infty} P(A_i \cap B) = \sum_{i=1}^{\infty} P(A_i)P(B \mid A_i)$$

□

**Ejemplo.** Tenemos dos cajas: A, que contiene 3 bolas rojas y 1 negra; y B, que contiene 5 bolas rojas y 8 negras.

La probabilidad de que cojamos una bola de la caja A es de 1/3 mientras que la probabilidad de coger una bola de la caja B es de 2/3. Calcular la probabilidad de que salga una bola negra.

$$P(A) = 1/3 \qquad P(B) = 2/3$$

$$P(N) = P(A)P(N \mid A) + P(B)P(N \mid B) = \frac{1}{3}\frac{1}{4} + \frac{2}{3}\frac{8}{13} = \frac{77}{156} \approx 0{,}49$$

**Teorema 5.4** (Regla de Bayes)**.** *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio de probabilidad y sea* $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ *una partición de* $\Omega$ *con* $P(A_n) > 0 \ \ \forall n \in \mathbb{N}$. *Entonces,* $\forall B \in \mathcal{A}$*:*

$$P(A_j \mid B) = \frac{P(B \mid A_j)P(A_j)}{\sum\limits_{i=1}^{\infty} P(B \mid A_i)P(A_i)} \qquad j \in \mathbb{N}$$

*Demostración.* Sea $j \in \mathbb{N}$:

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B \mid A_j)P(A_j)}{\sum\limits_{i=1}^{\infty} P(A_i)P(B \mid A_i)}$$

□

**Ejemplo.** Continuando con el ejemplo anteriormente propuesto, calculamos la probabilidad de que hayamos obtenido la bola en la caja A en el caso de que hayamos obtenido una bola negra:

$$P(A \mid N) = \frac{P(N \mid A)P(A)}{P(A)P(N \mid A) + P(B)P(N \mid B)} = \frac{\frac{1}{4}\frac{1}{3}}{\frac{1}{3}\frac{1}{4} + \frac{2}{3}\frac{8}{13}} = \frac{13}{77} \approx 0{,}1688$$

## 5.3. Independencia de sucesos

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y $A \in \mathcal{A}$ con $P(A) > 0$. La ocurrencia del suceso $A$ puede alterar la probabilidad de ocurrencia de cualquier otro suceso $B \in \mathcal{A}$. Al estudiar dichas probabilidades, puede darse que:

- $P(B \mid A) \neq P(B)$, es decir, la ocurrencia de $A$ modifica la probabilidad de $B$. Diremos entonces que $B$ **depende** de $A$.
  - Si $P(B \mid A) > P(B)$, se dice que el suceso $A$ favorece al $B$.
  - Si $P(B \mid A) < P(B)$, se dice que el suceso $A$ desfavorece al $B$.
- $P(B \mid A) = P(B)$, es decir, la ocurrencia de $A$ no tiene ningún efecto sobre el suceso $B$, se dice que el suceso $B$ es **independiente** del suceso $A$.

**Teorema 5.5** (Caracterización de la independencia)**.** *Sea* $A, B \in \mathcal{A} \mid P(B) > 0$. *Entonces:*

$$A \text{ es independiente a } B \Longleftrightarrow P(A \cap B) = P(A)P(B)$$

*Demostración.* Procedemos mediante doble implicación:

$\Longrightarrow$) Suponemos $A$ y $B$ independientes, es decir, $P(A|B) = P(A)$.

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

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

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

□

**Corolario 5.5.1** (Simetría de la independencia)**.** *Sean $A, B \in \mathcal{A}$. Entonces:*

$$A \textit{ es independiente a } B \Longleftrightarrow B \textit{ es independiente a } A$$

*Demostración.*

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

□

**Proposición 5.6.** *Si $A, B \in \mathcal{A}$ son independientes. Entonces, son equivalentes:*

*1. $A$ y $\overline{B}$ son independientes.*

*2. $\overline{A}$ y $B$ son independientes.*

*3. $\overline{A}$ y $\overline{B}$ son independientes.*

*Demostración.* Demostramos cada una de las partes, partiendo de que:

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

1. Probemos que $A$ y $\overline{B}$ son independientes:

$$\begin{aligned} P(A)P(\overline{B}) = P(A)(1 - P(B)) = P(A) - P(A)P(B) = \\ = P(A) - P(A \cap B) = P(A - B) = P(A \cap \overline{B}) \end{aligned}$$

2. $A$ y $B$ son independientes $\Leftrightarrow$ $B$ y $A$ son independientes $\Leftrightarrow$ $B$ y $\overline{A}$ son independientes $\Leftrightarrow$ $\overline{A}$ y $B$ son independientes.

3. $A$ y $B$ son independientes $\Leftrightarrow$ $\overline{A}$ y $B$ son independientes.

   Sea $C = \overline{A}$. Sabemos que $C$ y $B$ son independientes $\Leftrightarrow$ $C$ y $\overline{B}$ son independientes $\Leftrightarrow$ $\overline{A}$ y $\overline{B}$ son independientes.

□

**Proposición 5.7.** *Sean $A, B \in \mathcal{A}$ con $A, B \neq \emptyset$, entonces:*

$$\textit{Si } A \textit{ y } B \textit{ son independientes} \Rightarrow \textit{No son incompatibles } (A \cap B \neq \emptyset)$$

*Demostración.* $P(A) > 0 \;\wedge\; P(B) > 0$ Luego:

$$P(A \cap B) = P(A)P(B) > 0 \Rightarrow P(A \cap B) \neq 0 \Rightarrow A \cap B \neq \emptyset$$

$\square$

Por último, destacamos dos últimas definiciones en cuanto a sucesos independientes:

**Definición 5.2** (Independencia dos a dos)**.** Dado un espacio probabilístico $(\Omega, \mathcal{A}, P)$ y una clase de sucesos $\mathcal{U} \subset \mathcal{A}$ no vacía, diremos que sus sucesos son **independientes dos a dos** si

$$\forall A, B \in \mathcal{U} \mid A \neq B \Rightarrow A \text{ y } B \text{ son independientes}$$

**Definición 5.3** (Independencia mutua)**.** Dado un espacio probabilístico $(\Omega, \mathcal{A}, P)$ y una clase de sucesos $\mathcal{U} \subset \mathcal{A}$ no vacía, diremos que sus sucesos son **mutuamente independientes** si para cada subcolección finita $\{A_{i1}, A_{i2}, \ldots, A_{ik}\}$ de sucesos distintos de $\mathcal{U}$ se verifica:

$$P(A_{i1} \cap A_{i2} \cap \ldots \cap A_{ik}) = \prod_{j=1}^{k} P(A_{ij})$$

# 6. Variables aleatorias

## 6.1. Definiciones

Para dar la definición de variable aleatoria, es necesario pararnos a observar que hasta ahora hemos estado trabajando con el espacio probabilístico $(\Omega, \mathcal{A}, P)$. A continuación, nos será de especial importancia tener en cuenta el espacio probabilístico $(\mathbb{R}, \mathcal{B})$ donde $\mathbb{R}$ es el conjunto de números reales y $\mathcal{B}$ es una $\sigma$-álgebra sobre $\mathbb{R}$ que contiene a todos sus intervalos. Esta $\mathcal{B}$ recibe el nombre de $\sigma$-álgebra de Borel[1].

**Definición 6.1.** Una variable aleatoria $X$ es una función medible[2] sobre un espacio de probabilidad. Es decir, una función:

$$X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B})$$

tal que la imagen inversa de cualquier conjunto de Borel es medible (es decir, un suceso de la $\sigma$-álgebra $\mathcal{A}$):

$$X^{-1}(B) \in \mathcal{A} \qquad \forall B \in \mathcal{B}$$

Notemos que para comprobar que dicha variable sea aleatoria basta con que la imagen inversa de cualquier intervalo de la forma $]-\infty, x]$ sea medible (es decir, que pertenezca a $\mathcal{A}$).

Por tanto, $X$ será una variable aleatoria si es una función $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B})$ tal que

$$X^{-1}(]-\infty, x]) \in \mathcal{A} \qquad \forall x \in \mathbb{R}$$

**Ejemplo.** Consideramos el experimento aleatorio de lanzar una moneda:

$$\Omega = \{c, +\} \text{ (c cara, + cruz)}$$

$$\mathcal{A} = \{\emptyset, \{c\}, \{+\}, \Omega\}$$

Definimos $X$ como el número de caras al lanzar una moneda:

$$X(\omega) = \begin{cases} 0 & \text{si} \quad c \notin \omega \\ 1 & \text{si} \quad c \in \omega \end{cases} \qquad \omega \in \Omega$$

Comprobamos que se trata de una variable aleatoria. Trivialmente sabemos que se trata de una función del tipo:

$$X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B})$$

[1]Se estudiará a fondo en Análisis Matemático II.

[2]Este concepto se introducirá en la Teoría de Integración que se verá en Análisis Matemático II.

Comprobamos que $X^{-1}(B) \in \mathcal{A} \;\; \forall B \in \mathcal{B}$:

$$\begin{array}{ll} \text{Si } x < 0 & X^{-1}(]-\infty, x]) = \emptyset \in \mathcal{A} \\ \text{Si } 0 \leqslant x < 1 & X^{-1}(]-\infty, x]) = + \in \mathcal{A} \\ \text{Si } x \geqslant 1 & X^{-1}(]-\infty, x]) = \Omega \in \mathcal{A} \end{array}$$

## 6.2. Probabilidad inducida

**Definición 6.2** (Probabilidad inducida)**.** Sea $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B})$ una variable aleatoria, podemos definir una nueva función $P_X : (\mathbb{R}, \mathcal{B}) \to [0, 1]$ como sigue:

$$P_X(B) = P(X^{-1}(B)) \qquad \forall B \in \mathcal{B}$$

Dando lugar al espacio probabilístico $(\mathbb{R}, \mathcal{B}, P_X)$.

**Notación.** Podemos encontrar $P_X(B)$ notado como $P(X \in B)$.

**Proposición 6.1.** *La función $P_X$ es una función de probabilidad.*

*Demostración.* Demostramos cada una de las 3 condiciones, usando para ello que $P$ es una función de probabilidad:

- Axioma de no negatividad:

$$P_X(B) = P(X^{-1}(B)) \geqslant 0 \qquad \forall B \in \mathcal{B}$$

- Axioma del suceso seguro:

$$P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$$

- Axioma de $\sigma$-aditividad:

  Sean $B_1, B_2, \ldots \in \mathcal{B}$ disjuntos dos a dos, entonces:

$$P_X\left(\bigcup_{i=1}^{\infty} B_i\right) = P\left(X^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right)\right) = P\left(\bigcup_{i=1}^{\infty} X^{-1}(B_i)\right) =$$
$$= \sum_{i=1}^{\infty} P(X^{-1}(B_i)) = \sum_{i=1}^{\infty} P_X(B_i)$$

□

Notemos que era necesario imponer la medibilidad de nuestra función variable aleatoria $X$, ya que lo que pretendemos es poder asignar a cada valor del codominio de nuestra variable aleatoria una probabilidad inducida $P_X$ por la función probabilidad $P$ ya definida en nuestra $\sigma$-álgebra original, y para ello, naturalmente, necesitamos una correspondencia entre reales (codominio particular en este caso) y elementos de $\mathcal{A}$ (donde ya tenemos definida $P$).

Esta función es de gran relevancia pues es la pieza fundamental para la definición de la función de distribución, tomando $B = ]-\infty, x] \mid x \in \mathbb{R}$, y también para la definición de la función masa de probabilidad, tomando $B = [x, x] \mid x \in \mathbb{R} = \{x\} \in \mathbb{R}$.

Notemos que, tras haber demostrado que $P_X$ es una función de probabilidad, podemos aplicar los resultados ya conocidos de probabilidad a esta nueva función de probabilidad. Así, tenemos que:

- $P_X(]-\infty, x]) = P_X(]-\infty, x[) + P_X(x), \quad \forall x \in \mathbb{R}$

  Se ha usado la Proposición 4.6, donde hemos usado que

$$P_X(]-\infty, x[\, \cap \{x\}) = P_X(\emptyset) = 0$$

- $P_X(]-\infty, x]) = 1 - P_X(]x, +\infty[), \quad \forall x \in \mathbb{R}$

  Se ha usado la Proposición 4.2, donde hemos usado que $]x, +\infty[\, = \overline{]-\infty, x]}$.

- $P_X([a, b]) = P_X(]-\infty, b]) - P_X(]-\infty, a[), \quad \forall a, b \in \mathbb{R}, a \leqslant b$

  Se ha usado la Proposición 4.3, sabiendo que $[a, b] = ]-\infty, b] \setminus ]-\infty, a[$, y que $]-\infty, a] \subset ]-\infty, b]$.

## 6.3. Función de distribución

**Definición 6.3** (Función de distribución)**.** Se define la función de distribución de una variable aleatoria como:

$$\begin{array}{rccl} F_X : & \mathbb{R} & \longrightarrow & [0,1] \\ & x & \longmapsto & P_X\,(]-\infty, x]) \end{array}$$

**Notación.** Notemos que:

$$\begin{aligned} P_X(]-\infty, x]) = (P \circ X^{-1})(]-\infty, x]) = P(\{\omega \in \Omega \mid X(\omega) \in ]-\infty, x]\}) = \\ = P(\{\omega \in \Omega \mid X(\omega) \leqslant x\}) \end{aligned}$$

Por tanto, será más común encontrar $F_X(x)$ notado como $P(X \leqslant x)$ o como $P[X \leqslant x]$, siendo los paréntesis y los corchetes equivalentes.

**Definición 6.4** (Distribución de probabilidad)**.** Dada una variable aleatoria $X$, su distribución de probabilidad será el conjunto $\{(x_i, F_X(x_i))\}$.

### 6.3.1. Propiedades

- $F_X$ es creciente.

  Usamos la Proposición 4.3, y vemos que si $x \leqslant y$, entonces $]-\infty, x] \subseteq ]-\infty, y]$, y por tanto $P_X(]-\infty, x]) \leqslant P_X(]-\infty, y])$.

- $F_X$ es continua por la derecha. Es decir, para todo $x \in \mathbb{R}$, se cumple que:

$$\lim_{t \to x^+} F_X(t) = \lim_{t \to x^+} P_X(]-\infty, t]) = F_X(x)$$

- $\exists m, M \in \mathbb{R}$ tal que $\forall x, y \in \mathbb{R} \mid x \leqslant m \ \wedge \ y \geqslant M$ entonces $F_X(x) = 0$ y $F_X(y) = 1$.

  También notado como $F_X(-\infty) = 0$ y $F_X(+\infty) = 1$

- El conjunto de puntos de discontinuidad de una función de distribución es numerable.

- $\forall y \in \mathbb{R} \Rightarrow \lim\limits_{x \to y^-} F_X(x) = P_X(]-\infty, y[)$

  También notado como $F_X(x^-) = P(X < x)$.

- La función de distribución sólo puede presentar discontinuidades de salto. Además, la longitud del salto en un punto es la probabilidad que toma dicho punto:

$$P_X(x) = P_X(]-\infty, x]) - P_X(]-\infty, x[) = F_X(x) - \lim_{y \to x^-} F_X(x)$$

  También notado como $P_X(x) = F_X(x) - F_X(x^-)$.

- La función de distribución es continua en $x \in \mathbb{R} \iff P_X(x) = 0$.

## 6.4. Clasificaciones de variables aleatorias

Dependiendo de la forma de la distribución, podemos distinguir distintos tipos de variables aleatorias.

**Definición 6.5** (Recorrido de una variable aleatoria)**.** Dada una variable aleatoria $X$, definimos su recorrido, $Re_X$ como:

$$Re_X = \{x \in \mathbb{R} \mid \ \exists \omega \in \Omega \mid X(\omega) = x\} = Img(X)$$

### 6.4.1. Variables aleatorias discretas, función masa de probabilidad

**Definición 6.6** (Variable aleatoria discreta)**.** Una variable aleatoria se dice que es discreta si $Re_X = E \subset \mathbb{R}$ con $E$ numerable, es decir, que sólo toma una cantidad numerable de valores: $E = \{x_1, x_2, \dots\}$. Es decir, $P_X(E) = 1$.

**Definición 6.7** (Función masa de probabilidad)**.** Sea $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$ una variable aleatoria discreta con valores en $E$, definimos la función masa de probabilidad como:

$$\begin{array}{rccl} P: & E & \longrightarrow & [0,1] \\ & x_i & \longmapsto & P(x_i) = P_X(x_i) \end{array}$$

Esta verifica:

- $0 \leqslant p_i \leqslant 1 \qquad \forall x_i \in E.$

- $\sum\limits_{i=1}^{N} p_i = 1 \qquad$ con $N = |E|$

**Notación.** A continuación, a cada $P(x_i)$ lo notaremos por $p_i$.

La función de distribución de la variable $X$ a partir de la función masa de probabilidad queda como:

$$F_X(x_i) = P_X(]-\infty, x_i]) = \sum_{j=1}^{i} p_j \qquad \forall x_i \in E$$

También es posible expresar la función masa de probabilidad a partir de la función de distribución de la variable $X$:

$$p_i = F_X(x_i) - F_X(x_{i-1}) \qquad \forall x_i, x_{i-1} \in E$$

Notemos que la función de distribución de variables discretas es escalonada y creciente.

**Ejemplo.** Dado el experimento de lanzar una moneda, damos su espacio muestral y $\sigma$-álgebra asociados:

$$\Omega = \{c, +\} \qquad \mathcal{A} = \{\emptyset, \{c\}, \{+\}, \Omega\}$$

Con los que podemos definir la siguiente variable aleatoria, que nos da el número de caras, $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$ por:

$$X(\omega) = \begin{cases} 0 & c \notin \omega \\ 1 & c \in \omega \end{cases}$$

Esta definición de variable aleatoria, de recorrido $Re_X = \{0, 1\}$ nos induce una función masa de probabilidad $P$:

$$p_0 = P(0) = \frac{1}{2} \qquad p_1 = P(1) = \frac{1}{2}$$

### 6.4.2. Variables aleatorias continuas, función de densidad

**Definición 6.8.** Una variable aleatoria se dice que es continua si su función de distribución $F_X$ es absolutamente continua, es decir, si existe una función $f : \mathbb{R} \to \mathbb{R}$ tal que:

$$F_X(x) = \int_{-\infty}^{x} f(t)dt$$

Llamaremos a dicha función $f$ **función de densidad**, y verifica:

- $f(x) \geqslant 0 \qquad \forall x \in \mathbb{R}$.
- $f$ es Riemman integrable (ya que $F$ tiene a lo sumo un número finito de discontinuidades sobre cada intervalo finito de $\mathbb{R}$).
- $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1$

$$\int_{-\infty}^{+\infty} f(x)dx = \lim_{x} \to +\infty \int_{-\infty}^{x} f(t)dt = \lim_{x} \to +\infty F_X(x) = 1$$

donde hemos usado que $F_X(+\infty) = 1$.

Como consecuencia del Teorema Fundamental del Cálculo, podemos obtener la función de densidad a partir de la función de distribución:

$$f(x) = F'_X(x) \qquad \forall x \in \mathbb{R}$$

Esta nueva función nos permitirá también calcular probabilidades asociadas a un intervalo:

$$P_X([a,b]) = \int_a^b f(x)\ dx$$

Notemos, por tanto, que $P_X([a,b]) = P_X(]a,b[) = P_X([a,b[) = P_X(]a,b])$, puesto que no interfieren en los límites de integración.

Además, en el caso continuo, la probabilidad inducida en un punto será siempre 0:

$$P_X(x) = \int_x^x f(t)\ dt = 0 \qquad \forall x \in \mathbb{R}$$

En el caso de las variables continuas, el recorrido de la variable aleatoria será un intervalo $Re_X = [a,b] \subseteq \mathbb{R}$.

**Ejemplo.** Comprobar si la función $f : \mathbb{R} \to \mathbb{R}$ definida a continuación puede ser una función de densidad. En caso afirmativo, construir la función de distribución asociada a una variable aleatoria $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$.

$$f(x) = \begin{cases} 0 & x < 0 \\ 2x & 0 \leqslant x \leqslant 1 \\ 0 & x > 1 \end{cases}$$

Comprobamos que pueda tratarse de una función de densidad:

1. $f(x) \geqslant 0 \qquad \forall x \in \mathbb{R}$
2. Es Riemman Integrable, por ser la imagen acotada.
3. $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{0} 0\ dx + \int_0^1 2x\ dx + \int_1^{+\infty} 0\ dx = \int_0^1 2x\ dx = x^2]_0^1 = 1$

Por tanto, $f$ es una función de densidad, y podemos definir a partir de ella una función de distribución $F : \mathbb{R} \to \mathbb{R}$:

- Si $x < 0$:
$$F_X(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{-\infty}^{x} 0\ dt = 0$$

  Si $0 \leqslant x \leqslant 1$:
$$F_X(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{-\infty}^{0} 0\ dt + \int_0^x 2t\ dt = t^2]_0^x = x^2$$

  Si $x > 1$:
$$F_X(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{-\infty}^{0} 0\ dt + \int_0^1 2t\ dt + \int_1^{+\infty} 0\ dt = t^2]_0^1 = 1$$

Por tanto, la función de distribución es:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \leqslant x \leqslant 1 \\ 1 & x > 1 \end{cases}$$

### 6.4.3. Variables aleatorias mixtas

**Definición 6.9.** Una variable aleatoria se dice que es mixta si $Re_X = E \cup [a,b]$ con $E$ numerable y $[a,b] \subseteq \mathbb{R}$. Si se da que $E \subset [a,b] \Rightarrow X$ será una variable continua de recorrido $[a,b]$.

Dicha variable aleatoria tendrá asociadas una función masa de probabilidad y una función de densidad:

- Tendremos $P : E \to [0,1]$ dada por $p_i = P(x_i) = P_X(x_i) \qquad \forall x_i \in E$ tal que:
  - $0 \leqslant p_i \leqslant 1 \qquad \forall x_i \in E$.
  - $\sum\limits_{i=1}^{N} p_i = P \qquad$ con $P \in [0,1]$ y $N = |E|$.
- Y además existirá una función $f : \mathbb{R} \to \mathbb{R}$ Riemman integrable tal que:
  - $f(x) \geqslant 0 \qquad \forall x \in \mathbb{R}$.
  - $\displaystyle\int_{-\infty}^{+\infty} f(x)\ dx = 1 - P$

De esta forma, podemos calcular la probabilidad de cualquier intervalo $[c,d]$ definiendo el intervalo $[y,z] = [a,b] \cap [c,d]$ y mediante la siguiente expresión:

$$P_X([c,d]) = \sum_{\substack{x_i \in E \\ x_i \in [c,d]}} p_i + \int_y^z f(x)\ dx$$

También se puede calcular el valor de la función de distribución en un punto x:

$$F_X(x) = \sum_{\substack{x_i \in E \\ x_i \leqslant x}} p_i + \int_{-\infty}^{x} f(t)\ dt$$

## 6.5. Cambio de variable

En varias ocasiones, nos será de gran interés dar la distribución de probabilidad de una variable aleatoria $Y$, que depende de otra variable aleatoria $X$, en función de esta última.

Notemos que si conocemos la probabilidad inducida $P_X$ de una variable aleatoria $X$, entonces conocemos la función de distribución $F_X$ de dicha variable aleatoria y, por tanto, conocemos la distribución de probabilidad asociada a dicha variable.

**Teorema 6.2** (Teorema general de cambio de variable)**.** *Sea* $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$ *una variable aleatoria con una probabilidad inducida* $P_X$*. Sea* $Y = h(x)$ *otra variable aleatoria, con* $h : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ *medible, entonces la probabilidad inducida de* $Y$ *puede obtenerse a partir de* $P_X$ *como:*

$$P_Y(B) = P_X(h^{-1}(B)) \qquad \forall B \in \mathcal{B}$$

*Demostración.*

$$P_Y(B) = P(Y^{-1}(B)) = P(X^{-1}(h^{-1}(B))) = P_X(h^{-1}(B)) \qquad \forall B \in \mathcal{B}$$

□

Por este teorema, podemos obtener la función de distribución de $Y$ mediante:

$$F_Y(y) = P_X(h^{-1}(]-\infty, y])) \qquad \forall y \in \mathbb{R}$$

El cambio de variable en el caso de una variable discreta dará lugar a una variable discreta, mientras que si efectuamos un cambio de variable a una variable continua, podremos obtener una variable discreta, una variable continua o una variable mixta. Abordaremos dichos casos:

### 6.5.1. Cambio en variable aleatoria discreta

**Corolario 6.2.1.** *Sea* $X$ *una variable aleatoria discreta con valores en* $Re_X = \{x_1, x_2, \ldots\} \subset \mathbb{R}$ *y* $h : \mathbb{R} \to \mathbb{R}$ *medible tal que* $Y = h(x)$*, entonces* $Re_Y = g(Re_X)$ *y la función masa de probabilidad de* $Y$ *será:*

$$P[Y = y] = \sum_{x \in h^{-1}(y) \cap Re_X} P[X = x] \qquad \forall y \in Re_Y$$

*Demostración.* Como $h$ es medible y $Re_X$ es numerable, entonces $g(Re_X) = Re_Y$ es numerable, por lo que $Y$ es discreto. Por el Teorema de Cambio de Variable, tenemos que:

$$P[Y = y] = P[X \in h^{-1}(y)] = P_X(h^{-1}(y)) = \sum_{x \in h^{-1}(y) \cap Re_X} P[X = x], \qquad \forall y \in Re_Y$$

□

**Ejemplo.** Dada la variable aleatoria $X$ con $Re_X = \{-1, 0, 1\}$ y su función masa de probabilidad dada por:

$$P(x) = \begin{cases} 1/3 & x = -1 \\ 1/3 & x = 0 \\ 1/3 & x = 1 \end{cases}$$

Definimos la variable aleatoria $Y$ por $Y = h(X) = X^2$. Se pide dar la función masa de probabilidad de $Y$, $\hat{P}$.

El recorrido de $Y$ será: $Re_Y = \{h(-1), h(0), h(1)\} = \{0, 1\}$.

$$\hat{P}(y) = \begin{cases} P(0) = 1/3 & y = 0 \\ P(-1) + P(1) = 2/3 & y = 1 \end{cases}$$

### 6.5.2. Cambio de variable aleatoria continua a discreta

**Corolario 6.2.2.** *Sea $X$ una variable aleatoria continua con función de densidad $f_X$, y consideramos una función medible $h$ tal que $Y = h(x)$ es una variable aleatoria discreta. Entonces, $Re_Y = h(X)$ y la función masa de probabilidad de $Y$ será:*

$$P[Y = y] = \int_{h^{-1}(y)} f_X(x)\ dx \qquad \forall y \in Re_Y$$

*Demostración.* Esto sale de forma directa por el Teorema de Cambio de Variable, ya que:

$$P[Y = y] = P[X \in (h^{-1}(y))] = \int_{h^{-1}(y)} f_X(x)\ dx$$

□

El corolario anterior nos indica que, por ejemplo, en el caso en el que:

$$h([a, b] \cup [c, d]) = y \Rightarrow P(y) = \int_a^b f(x)\ dx + \int_c^d f(x)\ dx$$

**Ejemplo.** Dada la variable aleatoria $X$ con $Re_X = \mathbb{R}$ y su función de densidad $f : \mathbb{R} \to \mathbb{R}$ dada por:

$$f(x) = \begin{cases} 0 & x \leqslant -1 \\ 1/2 & -1 < x < 1 \\ 0 & x \geqslant 1 \end{cases}$$

Definimos la variable aleatoria $Y$ por $Y = h(x)$. Se pide dar la función masa de probabilidad de $Y$, sabiendo que:

$$h(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geqslant 0 \end{cases}$$

En primer lugar, damos el recorrido de $Y$: $Re_Y = \{h(\mathbb{R})\} = \{0, 1\}$.

$$P(y) = \begin{cases} \displaystyle\int_{-\infty}^0 f(x)\ dx = \int_{-\infty}^{-1} 0\ dx + \int_{-1}^0 1/2\ dx = \left.\frac{x}{2}\right]_{-1}^0 = 1/2 & y = 0 \\ \\ \displaystyle\int_0^{+\infty} f(x)\ dx = \int_0^1 1/2\ dx + \int_1^{+\infty} 0\ dx = \left.\frac{x}{2}\right]_0^1 = 1/2 & y = 1 \end{cases}$$

### 6.5.3. Cambio de variable aleatoria continua a continua

Sea $X$ una variable aleatoria continua con función de densidad $f$ y recorrido $Re_X = [a, b] \subseteq \mathbb{R}$. Si $h$ es una función medible estrictamente monótona y derivable en $[a, b]$, entonces $Y = h(x)$ es una variable aleatoria continua con función de densidad:

$$g(y) = \begin{cases} f(h^{-1}(y)) \cdot |(h^{-1})'(y)| & y \in h([a, b]) \\ \\ 0 & y \notin h([a, b]) \end{cases} \qquad \forall y \in \mathbb{R}$$

**Generalización.**

Una generalización del teorema puede hacerse en el caso de que $h$ no tenga una única inversa y cada valor de la variable $Y = h(x)$ proceda de un número finito o infinito de valores de $X$. En dicho caso, la función de densidad de $Y$ será:

$$g(y) = \sum_{k=1}^{\infty} f(h_k^{-1}(y)) \cdot |(h_k^{-1})'(y)|$$

Siendo $h_1^{-1}(y), h_2^{-1}(y), \dots$ las antiimágenes de $y \quad \forall y \in h([a, b])$.

**Ejemplo.** Dada la variable aleatoria $X$ con $Re_X = ]-1, 1[$ y su función de densidad $f : \mathbb{R} \to \mathbb{R}$ dada por:

$$f(x) = \begin{cases} 0 & x \leqslant -1 \\ \frac{1}{2} & -1 < x < 1 \\ 0 & x \geqslant 1 \end{cases}$$

Definimos la variable aleatoria $Y$ por $Y = h(X) = X^2$. Se pide dar la función de densidad de $Y$.

En primer lugar, calculamos el recorrido de $Y$: $Re_Y = \{h(]-1, 1[)\} = [0, 1[$. Además, al no ser $h$ inyectiva en el dominio de definición, tiene dos inversas: la raíz positiva y la negativa.

$$h^{-1}(y) = \pm\sqrt{y}$$
$$(h^{-1})'(y) = \pm\frac{1}{2\sqrt{y}}$$

Damos la función de densidad de $Y$:

$$g(y) = \sum_{k=1}^{\infty} f(h_k^{-1}(y)) \cdot |(h_k^{-1})'(y)| = \frac{1}{2}\frac{1}{2\sqrt{y}} + \left|-\frac{1}{2}\frac{1}{2\sqrt{y}}\right| = \frac{1}{2\sqrt{y}} \qquad \forall y \in ]0, 1[$$

### 6.5.4. Cambio de una variable aleatoria continua a mixta

Sea $X$ una variable aleatoria continua con función de densidad $f$ y recorrido $Re_X = [a, b] \subseteq \mathbb{R}$. Si $h$ es una función medible y definimos $Y = h(x)$ como una nueva variable aleatoria tal que $Y$ presenta una distribución mixta, entonces tendremos que su función de densidad viene dada por:

**Ejemplo.** Dada una variable aleatoria $X$ con $Re_X = ]-1, 1[$ y su función de densidad $f : \mathbb{R} \to \mathbb{R}$ dada por:

$$f(x) = \begin{cases} 0 & x \leqslant -1 \\ \frac{1}{2} & -1 < x < 1 \\ 0 & x \geqslant 1 \end{cases}$$

Definimos la variable aleatoria $Y$ por $Y = h(X)$. Se pide dar la función de distribución de $Y$.

$$h(x) = \begin{cases} x & x < 0 \\ 1 & x \geqslant 0 \end{cases}$$

Calculamos el recorrido de $Y$: $Re_Y = ]-1,0[\cup\{1\}$. Por tanto, aplicando el teorema del cambio de variable:

$$F_Y(y) = \begin{cases} 0 & y \leqslant -1 \\ \displaystyle\int_{-1}^{y} \frac{1}{2}\, dt = \left.\frac{t}{2}\right]_{-1}^{y} = \frac{y+1}{2} & -1 < y < 0 \\ \dfrac{1}{2} & 0 \leqslant y < 1 \\ 1 & y \geqslant 1 \end{cases}$$

## 6.6. Esperanza matemática

### 6.6.1. Variables aleatorias discretas

**Definición 6.10.** Sea $X$ una variable aleatoria discreta con valores en $Re_X = E = \{x_1, x_2, \ldots\} \subset \mathbb{R}$ y sea $P$ la función masa de probabilidad de $X$, se define la esperanza matemática, media o valor esperado de $X$, denotado $E[X]$ como:

$$E[X] = \sum_{x_i \in E} x_i P(x_i)$$

Esta existe siempre y cuando la serie previamente mencionada converja absolutamente.

*Observación.* A lo largo de este documento, cada vez que hagamos referencia a la esperanza de una variable aleatoria $X$, hemos supuesto previamente que existe. Es decir, siempre que encontremos que: Sea $X$ una variable aleatoria y aparezca $E[X]$, hemos supuesto implícitamente que su varianza existe. En caso contrario, no tendría sentido hablar de $E[X]$.

**Ejemplo.** Dada la variable aleatoria $X$ con $Re_X = \{-1, 0, 1\}$ y su función masa de probabilidad dada por:

$$P(x) = \begin{cases} 1/3 & x = -1 \\ 1/3 & x = 0 \\ 1/3 & x = 1 \end{cases}$$

Se pide calcular la esperanza matemática de $X$.

$$E[X] = -1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} - 1 \cdot \frac{1}{3} = 0$$

### 6.6.2. Variables aleatorias continuas

**Definición 6.11.** Sea $X$ una variable aleatoria continua con valores en $Re_X = [a, b]$ y $f$ la función de densidad de dicha variable, se define la esperanza matemática de $X$ como:

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\, dx$$

La esperanza de $X$ existe siempre y cuando la integral sea absolutamente convergente.

**Ejemplo.** Sea $X$ una variable aleatoria con valores en $Re_X = [0, \sqrt[3]{3}]$ y función de densidad $f : \mathbb{R} \to \mathbb{R}$ dada por:

$$f(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \leqslant x \leqslant \sqrt[3]{3} \\ 0 & x > \sqrt[3]{3} \end{cases}$$

Se pide calcular la esperanza matemática de $X$.

$$E[X] = \int_{-\infty}^{+\infty} xf(x)\ dx = \int_0^{\sqrt[3]{3}} xf(x)\ dx = \int_0^{\sqrt[3]{3}} x^3\ dx \approx 1{,}0817$$

### 6.6.3. Propiedades

**Proposición 6.3.** *La esperanza de una variable aleatoria constante es la constante. Es decir, dado $c \in \mathbb{R}$, si $X = c \quad \forall \omega \in \mathcal{A}$, entonces:*

$$E[X] = c$$

*Demostración.* En este caso no tiene sentido considerar el caso continuo, ya que $Re_X = \{c\}$. Para el caso discreto, como $\sum\limits_{i=1}^{|Re_X|} P(x_i) = 1$, tenemos que $P(c) = 1$. Por tanto,

$$E[X] = \sum_{x_i \in Re_X} x_i P(x_i) = cP(c) = c$$

□

**Proposición 6.4.** *Si una variable aleatoria está acotada; es decir, si $\exists M \in \mathbb{R} \mid |Im(X)| \leqslant M$, entonces $E[X] \leqslant M$.*

*Demostración.* Hemos de distinguir entre el caso discreto y el caso continuo:

- Caso discreto: Supongamos $Re_X = E$.

$$E[X] = \sum_{x_i \in E} x_i P(x_i) \leqslant \sum_{x_i \in E} M \cdot P(x_i) = M \sum_{x_i \in E} P(x_i) = M$$

- Caso continuo: Supongamos $Re_X = [a, b]$.

$$E[X] = \int_a^b xf(x)\ dx \leqslant \int_a^b Mf(x)\ dx = M \int_a^b f(x)\ dx = M$$

□

**Proposición 6.5.** *Sea $X$ una variable aleatoria. Si $X \geqslant 0$ y existe esperanza, entonces $E[X] \geqslant 0$.*

*Demostración.* Hemos de distinguir entre el caso discreto y el caso continuo:

- <u>Caso discreto</u>: Supongamos $Re_X = E$.

$$E[X] = \sum_{x_i \in E} x_i P(x_i) \geqslant 0$$

donde se ha empleado que $x_i \geqslant 0$ (hipótesis) y $P(x_i) \geqslant 0$ por definición.

- <u>Caso continuo</u>: Supongamos $Re_X = [a, b]$.

$$E[X] = \int_a^b x f(x)\, dx$$

donde se ha empleado que $x \geqslant 0$ (hipótesis) y $f(x) \geqslant 0$ por definición.

□

**Definición 6.12.** Sea $X$ una variable aleatoria. Se dice que $X$ es simétrica respecto del valor $c$ si $X - c$ y $c - X$ tienen la misma distribución. Es decir, en función del tipo:

- <u>Discreto</u>: $P[X = c + x] = P[X = c - x]$ para todo $x \in \mathbb{R}$.
- <u>Continuo</u>: $P[X \leqslant c - x] = P[X \geqslant c + x]$ para todo $x \in \mathbb{R}$.

**Proposición 6.6.** *Sea $X$ una variable aleatoria con distribución simétrica respecto del valor $c$. Entonces, si $\exists E[X] \Longrightarrow E[X] = c$.*

**Proposición 6.7.** *Sea $X$ una variable aleatoria. Consideramos $c \in \mathbb{R}$. Entonces, se tiene que $E[cX] = cE[X]$.*

*Demostración.* Distinguimos entre el caso disreto y el continuo:

- <u>Caso discreto</u>: $Re_X = E$.

$$E[cX] = \sum_{x_i \in E} cx_i P_Y(cx_i) = c \sum_{x_i \in E} x_i P_Y(cx_i) = c \sum_{x_i \in E} x_i P(x_i) = cE[X]$$

Donde $P_Y$ denota la función masa de probabilidad de la variable aleatoria discreta $Y = cX$. Consúltese el capítulo de cambio de variable discreta (capítulo 6.5.1) para ver que $P_Y(cx_i) = P(x_i)$ siendo $P$ la función masa de probabilidad de la variable aleatoria $X$.

- <u>Caso continuo</u>: $Re_X = [a, b]$.

$$E[cX] = \int_{-\infty}^{+\infty} cx_i f_Y(cx_i)\, dx = c \int_{-\infty}^{+\infty} x_i f_Y(cx_i)\, dx = c \int_{-\infty}^{+\infty} x_i f(x_i)\, dx = cE[X]$$

Donde $f_Y$ denota la función de densidad de la variable aleatoria continua $Y = cX$. Consúltese el capítulo de cambio de variable continua (capítulo 6.5.3) para ver que $f_Y(cx_i) = f(x_i)$ siendo $f$ la función de densidad de la variable aleatoria $X$.

□

**Proposición 6.8.** *Sean $X$ e $Y$ variables aleatorias con recorrido $Re_X, Re_Y$, entonces:*

$$E[X+Y] = E[X] + E[Y]$$

*Demostración.* Para el caso discreto, consideramos $Re_X = E, Re_Y = F$ ambos numerables.

Sea $P_X$ la función masa de probabilidad de $X$ y $P_Y$ la de $Y$.

Sea $P_Z$ la función masa de probabilidad de la variable aleatoria $Z = X+Y$, $P_Z : Re_X \times Re_Y \to [0,1]$.

$$\begin{aligned}
E[X+Y] &= \sum_{x_i \in E} \sum_{y_j \in F} (x_i + y_j) P_Z(x_i, y_j) = \\
&= \sum_{x_i \in E} \sum_{y_j \in F} x_i P_Z(x_i, y_j) + \sum_{x_i \in E} \sum_{y_j \in F} y_j P_Z(x_i, y_j) = \\
&= \sum_{x_i \in E} x_i \sum_{y_j \in F} P_Z(x_i, y_j) + \sum_{x_i \in E} y_j \sum_{y_j \in F} P_Z(x_i, y_j) = \\
&= \sum_{x_i \in E} x_i P_X(x_i) + \sum_{y_j \in F} y_j P_Y(y_j) = E[X] + E[y]
\end{aligned}$$

Donde hemos usado que $\sum\limits_{y_j \in F} P_Z(x_i, y_j) = P_X(x_i)$, que excede el conocimiento de este curso.

La demostración en el caso continuo es análoga. □

**Corolario 6.8.1** (Linealidad)**.** *Sea $X_i$ una variable aleatoria $\quad \forall i = 1, \dots, n$. Si $\exists E[X_i]$, entonces:*

$$\exists E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E[X_i]$$

*Demostración.* Se deduce directamente de las dos proposiciones anteriores.

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} E[a_i X_i] = \sum_{i=1}^{n} a_i E[X_i]$$

□

**Proposición 6.9.** *Dada $X$ una variable aleatoria y se consideran dos funciones de $X$, $g(X), h(X)$, variables aleatorias. Si $g(X) \leqslant h(X)$, entonces:*

$$E[g(X)] \leqslant E[h(X)]$$

*Demostración.* Realizamos distinción entre caso discreto y continuo:

- <u>Caso discreto</u>. Suponemos $Re_X = E$.

  Sabemos que $f(x_i) \geqslant 0 \ \forall x_i \in E$. Entonces, no cambia el sentido de la desigualdad, por lo que:

  $$g(x_i)f(x_i) \leqslant h(x_i)f(x_i) \qquad \forall x_i \in E$$

Por tanto, tenemos que:

$$\sum_{x_i \in E} g(x_i)f(x_i) \leqslant \sum_{x_i \in E} h(x_i)f(x_i)$$

ya que se cumple para todos los sumandos. Por tanto, usando las definiciones, tenemos que:

$$E[g(X)] \leqslant E[h(X)]$$

- Caso continuo. Suponemos $Re_X = [a, b]$.

  Sabemos que $f(x) \geqslant 0 \ \forall x \in [a, b]$. Entonces, no cambia el sentido de la desigualdad, por lo que:

  $$g(x)f(x) \leqslant h(x)f(x) \qquad \forall x_i \in [a, b]$$

  Por tanto, ya que el operador integral mantiene las relaciones de orden, tenemos que:

  $$\int_a^b g(x)f(x)\ dx \leqslant \int_a^b h(x)f(x)\ dx$$

  Por tanto, usando las definiciones, tenemos que:

  $$E[g(X)] \leqslant E[h(X)]$$

□

**Corolario 6.9.1.** *Sea $X$ una variable aleatoria y $g, h$ funciones reales tales que $\exists E[g(x)], E[h(x)]$, entonces:*

$$\exists E[\alpha g(X) + \beta h(X)] = \alpha E[g(X)] + \beta E[h(X)] \qquad \forall \alpha, \beta \in \mathbb{R}$$

**Proposición 6.10.** *La esperanza matemática minimiza el error cuadrático medio:*

$$\min_{a \in R} E[(X-a)^2] = E[(X - E[X])^2]$$

*Demostración.* Sea $a \in \mathbb{R}$, y definimos el error cuadrático medio respecto al valor $a$:

$$\mu(a) = E[(X-a)^2] = E[X^2 - 2aX + a^2] = E[X^2] - 2aE[X] + a^2$$

Para minimizar dicha función, hemos de calcular los puntos que anulan la primera derivada. Además, al tratarse de una parábola con coeficiente líder positivo, dicho punto crítico será un mínimo.

$$\mu'(a) = 2E[X] - 2a = 0 \Longleftrightarrow a = E[X]$$

Por tanto, $a = E[X]$ es un mínimo de la expresión $\mu(a)$ y, al ser el único, concluimos que es mínimo absoluto. Luego, acabamos de ver que la esperanza matemática minimiza el error cuadrático medio. □

**Definición 6.13** (Independencia de Variables Aleatorias)**.** Dadas $n$ variables aleatorias $X_1, X_2, \dots, X_n$, se dice que son independientes si y solo si:

$$P[X_1 \leqslant x_1, X_2 \leqslant x_2, \dots, X_n \leqslant x_n] = P[X_1 \leqslant x_1]P[X_2 \leqslant x_2] \cdots P[X_n \leqslant x_n] \qquad \forall x_1, x_2, \dots, x_n \in \mathbb{R}$$

**Ejemplo.** Sea $X = c$ una variable aleatoria degenerada, y sea $Y$ otra variable aleatoria cualquiera. Entonces, $X$ y $Y$ son independientes, ya que:

$$P[X \leqslant x] = \begin{cases} 0 & \text{si } x < c \\ 1 & \text{si } x \geqslant c \end{cases} \qquad P[X \leqslant x, Y \leqslant y] = \begin{cases} 0 & \text{si } x < c \\ P[Y \leqslant y] & \text{si } x \geqslant c \end{cases}$$

Por tanto, $P[X \leqslant x, Y \leqslant y] = P[X \leqslant x]P[Y \leqslant y]$.

**Proposición 6.11.** *Sean $X$ e $Y$ variables aleatorias. Entonces, se tiene que:*

$$X \text{ e } Y \text{ son independientes} \iff E[XY] = E[X]E[Y]$$

**Teorema 6.12** (de Markov: desigualdad básica)**.** *Sea $X$ una variable aleatoria. Si $X \geqslant 0$ y $\exists E[X]$, entonces:*

$$P[X \geqslant \varepsilon] \leqslant \frac{E[X]}{\varepsilon} \qquad \forall \varepsilon > 0$$

*Demostración.* Realizamos la distinción entre el caso discreto y el continuo:

- Caso discreto: Supongamos que $Re_X = E$. Entonces:

$$E[X] = \sum_{x_i \in E} x_i P(x_i) = \sum_{x_i \in E | x_i < \varepsilon} x_i P(x_i) + \sum_{x_i \in E | x_i \geqslant \varepsilon} x_i P(x_i)$$

  Por la Proposición 6.5, tenemos que $\displaystyle\sum_{x_i \in E | x_i < \varepsilon} x_i P(x_i) \geqslant 0$. Por tanto,

$$\begin{aligned} E[X] = \sum_{x_i \in E} x_i P(x_i) &= \sum_{x_i \in E | x_i < \varepsilon} x_i P(x_i) + \sum_{x_i \in E | x_i \geqslant \varepsilon} x_i P(x_i) \geqslant \sum_{x_i \in E | x_i \geqslant \varepsilon} x_i P(x_i) \geqslant \\ &\geqslant \sum_{x_i \in E | x_i \geqslant \varepsilon} \varepsilon P(x_i) = \varepsilon \sum_{x_i \in E | x_i \geqslant \varepsilon} P(x_i) = \varepsilon P[X \geqslant \varepsilon] \end{aligned}$$

  Por tanto, despejando $P[X \geqslant \varepsilon]$, tenemos que:

$$P[X \geqslant \varepsilon] \leqslant \frac{E[X]}{\varepsilon}$$

- Caso continuo: Supongamos $Re_X = [a, b]$.

$$E[X] = \int_a^b xP(x)\,dx = \int_a^\varepsilon xP(x)\,dx + \int_\varepsilon^b xP(x)\,dx \overset{(*)}{\geqslant} \varepsilon \int_\varepsilon^b P(x)\,dx = \varepsilon P[X \geqslant \varepsilon]$$

  donde en $(*)$ hemos empleado la Proposición 6.5.

  Por tanto, despejando $P[X \geqslant \varepsilon]$, tenemos que:

$$P[X \geqslant \varepsilon] \leqslant \frac{E[X]}{\varepsilon}$$

□

**Corolario 6.12.1** (Desigualdad de Markov)**.** *Sea $X$ es una variable aleatoria y sea $\alpha \in \mathbb{R}^+$. Si $\exists E[|X|^\alpha]$, se tiene:*

$$P[|X| \geqslant k] \leqslant \frac{E[|X|^\alpha]}{k^\alpha}, \qquad \forall k > 0$$

*Demostración.* En primer lugar, es importante señalar que:

$$P[|X| \geqslant k] = P[|X|^\alpha \geqslant k^\alpha]$$

Esto se debe al Teorema de Cambio de Variable y a que la trasformación $h(|X|) = |X|^\alpha$ es inyectiva, ya que la base y el exponente son positivos.

Además, aplicando la desigualdad básica a la variable $|X|^\alpha$, con $\varepsilon = k^\alpha$, se tiene que:

$$P[|X^\alpha \geqslant k^\alpha] \leqslant \frac{E[|X|^\alpha]}{k^\alpha}$$

Uniendo ambas ecuaciones, se tiene lo pedido. □

**Corolario 6.12.2** (Desigualdad de Chebychev)**.** *Si $X$ es una variable aleatoria tal que $\exists E[X^2]$, se tiene*

$$P[|X - E[X]| \geqslant k] \leqslant \frac{\mathrm{Var}[X]}{k^2}, \qquad \forall k > 0$$

*Demostración.* Se basa en aplicar la desigualdad básica a la variable $(X - E[X])^2$. Para $\varepsilon = k^2$, se tiene que:

$$P[|X - E[X]| \geqslant k] = P[(X - E[X])^2 \geqslant k^2] \leqslant \frac{E[(X - E[X])^2]}{k^2} = \frac{\mathrm{Var}[X]}{k^2}$$

□

**Expresiones alternativas de la desigualdad**:

1. De forma directa, se deduce que:

$$P[|X - E[X]| < k] \geqslant 1 - \frac{\mathrm{Var}[X]}{k^2}, \qquad \forall k > 0$$

   Esta expresión proporciona una cota inferior para la probabilidad de que una variable tome valores en cualquier intervalo real centrado en la media de la variable.

2. Empleando ahora $\varepsilon = \mathrm{Var}[X]k^2$, de forma idéntica se deduce que:

$$P[|X - E[X]| \geqslant \sqrt{\mathrm{Var}[X]}k] \leqslant \frac{1}{k^2}, \qquad \forall k > 0$$

3. De forma directa, se deduce que:

$$P[|X - E[X]| < \sqrt{\mathrm{Var}[X]}k] \geqslant 1 - \frac{1}{k^2}, \qquad \forall k > 0$$

   Esta expresión proporciona una cota inferior para la probabilidad de que una variable tome valores en cualquier intervalo real centrado en la media de la variable con una amplitud de $k\sqrt{\mathrm{Var}[X]}$.

### 6.6.4. Cambio de variable aleatoria discreta

Sea $X$ una variable aleatoria discreta con $Re_X = E$ e $Y = h(X)$ otra variable aleatoria discreta, entonces, si existe $E[Y]$ (existe si la serie converje absolutamente), entonces:

$$E[Y] = \sum_{x_i \in E} h(x_i)P(x_i)$$

### 6.6.5. Cambio de variable aleatoria continua

Sea $X$ una variable aleatoria continua con $Re_X = [a, b]$ e $Y = h(X)$ otra variable aleatoria continua, entonces, si existe $E[Y]$ (existe si la integral converje absolutamente), entonces:

$$E[Y] = \int_{-\infty}^{+\infty} h(x)f(x)\, dx$$

## 6.7. Moda

### 6.7.1. Variables aleatorias discretas

**Definición 6.14.** Sea $X$ una variable aleatoria discreta con recorrido $Re_X = E$, se define la moda de $X$ como aquel valor del recorrido de $X$ cuya imagen por su función masa de probabilidad sea mayor, es decir:

$$Mo_X = \text{máx}\{x \in E \mid P(x) \geqslant P(y) \quad \forall y \in E\}$$

### 6.7.2. Variables aleatorias continuas

**Definición 6.15.** Sea $X$ una variable aleatoria continua con recorrido $Re_X = [a, b]$, se define la moda de $X$ como aquel valor del recorrido de $X$ cuya imagen por su función de densidad sea mayor, es decir:

$$Mo_X = \text{máx}\{x \in [a, b] \mid f(x) \geqslant f(y) \quad \forall y \in [a, b]\}$$

Dicho de otra forma, la moda de $X$ es la abscisa en la que se alcanza el máximo absoluto de $f$.

**Ejemplo.** Sea $X$ una variable aleatoria con función de densidad:

$$f(x) = \begin{cases} 0 & x < 1 \\ \dfrac{8}{7x^2} & 1 \leqslant x \leqslant 8 \\ 0 & x > 8 \end{cases}$$

Se pide calcular la moda de $X$.

En este caso, hemos de de maximizar $f(X)$.

$$f'(x) = -2\frac{8}{7x^3} = -\frac{16}{7x^3} < 0$$

Por tanto, como $f(x)$ es estrictamente decreciente y positiva en $[1, 8]$, tenemos que la moda es el mínimo de dicho intervalo, es decir:

$$Mo_X = 1$$

## 6.8. Percentiles

### 6.8.1. Variables aleatorias discretas

**Definición 6.16.** Sea $X$ una variable aleatoria discreta con $Re_X = E$. Definimos el percentil $q \in [0,1]$ de la variable aleatoria $X$ como el valor $X_q \in E$ tal que:

$$\begin{gathered} P_X(]-\infty, X_q]) \geqslant q \\ P_X([X_q, +\infty[) \geqslant 1-q \end{gathered}$$

En el caso de que el valor buscado esté entre dos elementos del recorrido de $X$, $x_i, x_j \in E$, entonces el percentil $X_q$ buscado es la media de ambos elementos: $X_q = \dfrac{x_i + x_j}{2}$.

**Notación.** Para simplificar la notación, el percentil $X_{80}$ será el valor $X_{80} \in E$ tal que:

$$\begin{gathered} P_X(]-\infty, X_q]) \geqslant 0{,}8 \\ P_X([X_q, +\infty[) \geqslant 0{,}2 \end{gathered}$$

### 6.8.2. Variables aleatorias continuas

**Definición 6.17.** Sea $X$ una variable aleatoria continua con $Re_X = [a,b]$. Definimos el percentil $q \in [0,1]$ de la variable aleatoria $X$ como el valor $X_q \in [a,b]$ tal que:

$$\begin{gathered} P_X(]-\infty, X_q]) = q \\ P_X([X_q, +\infty[) = 1-q \end{gathered}$$

**Notación.** Para simplificar la notación, el percentil $X_{80}$ será el valor $X_{80} \in [a,b]$ tal que:

$$\begin{gathered} P_X(]-\infty, X_q]) = 0{,}8 \\ P_X([X_q, +\infty[) = 0{,}2 \end{gathered}$$

### 6.8.3. Mediana

**Definición 6.18.** Dada una variable aleatoria $X$ con recorrido $Re_X$, se define la media de la variable aleatoria, notada $Me$ como:

$$Me = X_{50}$$

**Ejemplo.** Sea $X$ una variable aleatoria discreta con $Re_X = [0,3] \cap \mathbb{Z}$ con función masa de densidad:

$$P(x) = \begin{cases} 3/16 & x = 0 \\ 7/16 & x = 1 \\ 5/16 & x = 2 \\ 1/16 & x = 3 \end{cases}$$

Se pide calcular la mediana de la variable aleatoria.

Buscamos $Me \in [0,3] \cap \mathbb{Z}$ tal que:

$$\begin{gathered} P_X([0, Me] \cap \mathbb{Z}) = F_X(Me) \geqslant 1/2 \\ P_X([Me, 3] \cap \mathbb{Z}) \geqslant 1/2 \end{gathered}$$

Supongamos que $Me = 1$, veamos que es cierto:

$$P_X([0,1] \cap \mathbb{Z}) = F_X(1) = 10/16 \geqslant 1/2$$
$$P_X([1,3] \cap \mathbb{Z}) = 13/16 \geqslant 1/2$$

Luego:

$$Me = 1$$

## 6.9. Momentos de una variable aleatoria

Sea $X$ una variable aleatoria. Definimos el momento de orden $k \in \mathbb{N}$ centrado en el punto $a \in \mathbb{R}$ como:

$$_am_k = E[(X-a)^k] \qquad k \in \mathbb{N} \cup \{0\} \ \wedge \ a \in \mathbb{R}$$

### 6.9.1. Momentos no centrales

**Definición 6.19** (Momentos no centrales)**.** Sea $X$ una variable aleatoria con recorrido $Re_X$, definimos el momento no centrado de orden $k \in \mathbb{N} \cup \{0\}$, notado $m_k$ como la cantidad:

$$m_k = E[X^k]$$

Algunos momentos no centrales relevantes son:

$$m_0 = E[X^0] = E[1] = 1$$
$$m_1 = E[X^1] = E[X]$$
$$m_2 = E[X^2]$$

### 6.9.2. Momentos centrales

**Definición 6.20** (Momentos centrales)**.** Sea $X$ una variable aleatoria, definimos el momento centrado de orden $k \in \mathbb{N} \cup \{0\}$, notado $\mu_k$ como la cantidad:

$$\mu_k = E[(X - E[X])^k]$$

Algunos momentos centrales relevantes son:

$$\mu_0 = E[(X - E[X])^0] = E[1] = 1$$
$$\mu_1 = E[X - E[X]] = E[X] - E[E[X]] = E[X] - E[X] = 0$$
$$\mu_2 = E[(X - E[X])^2]$$

Tenemos los siguientes resultados demostrados para la estadística unidimensional. La demostración en este caso es análoga:

$$m_k = \sum_{i=0}^{k} \binom{k}{i} \mu_{k-i} m_1^i \qquad\qquad \mu_k = \sum_{i=0}^{k} \binom{k}{i} (-1)^i m_{k-i} m_1^i$$

### 6.9.3. Varianza

**Definición 6.21.** Sea $X$ una variable aleatoria, definimos la varianza de $X$, notada $\sigma_X^2$ o $\mathrm{Var}[X]$ por:

$$\sigma_X^2 = \mathrm{Var}[X] := \mu_2(X) = E[(X - E[X])^2]$$

**Proposición 6.13.** *Dada una variable aleatoria $X$, tenemos que:*

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2$$

*Demostración.* Teniendo en cuenta que $E[X]$ es una constante, es decir, $E[XE[X]] = E[X]^2$, tenemos que:

$$\mathrm{Var}[X] = E[(X-E[X])^2] = E[X^2+(E[X])^2-2XE[X]] = E[X^2]+E[E[X]^2]-E[2XE[X]] = \\ = E[X^2] + (E[X])^2 - 2(E[X])^2 = E[X^2] - (E[X])^2 = m_2 - m_1^2$$

□

**Proposición 6.14.** *Sea $X$ una variable aleatoria cuya varianza existe, y consideramos $Y = aX + b, \quad a, b \in \mathbb{R}$. Entonces:*

$$\exists\, \mathrm{Var}[Y] := \mathrm{Var}[aX + b] = a^2\,\mathrm{Var}[X]$$

*Demostración.*

$$\mathrm{Var}[aX+b] = E[(aX+b)^2] - E[aX+b]^2 = E[a^2X^2+b^2+2abX] - (aE[X]+b)^2 = \\ = a^2E[X^2]+b^2+2abE[X]-a^2E[X]^2-b^2-2abE[X] = a^2(E[X^2]-E[X]^2) = a^2\,\mathrm{Var}[X]$$

□

**Proposición 6.15.** *Sean $X$ e $Y$ variables aleatorias independientes. Entonces:*

$$\exists\, \mathrm{Var}[X \pm Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

*Demostración.* Tenemos que:

$$\begin{aligned}\mathrm{Var}[X \pm Y] &= E[(X \pm Y)^2] - E[X \pm Y]^2 = \\ &= E[X^2 + Y^2 \pm 2XY] - E[X]^2 - E[Y]^2 \mp 2E[X]E[Y] = \\ &= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 \pm 2E[XY] \mp 2E[X]E[Y] \overset{(*)}{=} \\ &\overset{(*)}{=} \mathrm{Var}[X] + \mathrm{Var}[Y]\end{aligned}$$

donde en $(*)$ hemos usado que $E[XY] = E[X]E[Y]$ por la Proposición 6.11. □

**Corolario 6.15.1.** *Sean $X_1, \dots, X_n$ variables aleatorias independientes. Entonces:*

$$\exists Var\left[\sum_{i=1}^{n} a_iX_i\right] = \sum_{i=1}^{n} a_i^2\,\mathrm{Var}[X_i]$$

*Demostración.*

$$Var\left[\sum_{i=1}^{n} a_i X_i\right] \overset{(*)}{=} \sum_{i=1}^{n} \text{Var}[a_i X_i] = \sum_{i=1}^{n} a_i^2 \,\text{Var}[X_i]$$

donde en $(*)$ hemos usado el resultado anterior. □

**Proposición 6.16.** *Sea $X$ una variable aleatoria. Entonces:*

$$\text{Var}[X] = 0 \Longleftrightarrow X = c \in \mathbb{R}$$

*Es decir, la varianza de una variable aleatoria es nula si y solo si dicha variable aleatoria es una constante.*

*Demostración.* Procedemos mediante doble implicación:

$\Longleftarrow$) Suponemos $X = c \in \mathbb{R}$. Entonces,

$$\text{Var}[X] = E[X^2] - E[X]^2 = c^2 - c^2 = 0$$

$\Longrightarrow$) Suponemos $\text{Var}[X] = 0$. Entonces,

- <u>Caso discreto.</u> Consideramos $Re_X = E$:

  $$\text{Var}[X] = 0 = E[(X - E[X])^2] = \sum_{x_i \in E} (x_i - E[X])^2 f(x_i)$$

  Como tenemos que $f(x_i) > 0 \ \forall x_i$, entonces tenemos que todos los términos son no-negativos. Por tanto, para que se igualen a 0 todos ellos han de ser nulos. Por tanto,

  $$x_i = E[X] \qquad \forall x_i \in E$$

  Por tanto, llamando $E[X] = c \in \mathbb{R}$, tenemos $X = c$.

- <u>Caso continuo.</u> Consideramos $Re_X = [a, b]$:

  $$\text{Var}[X] = 0 = E[(X - E[X])^2] = \int_a^b (x - E[X])^2 f(x)\, dx$$

  Como tenemos que $f(x) > 0 \ \forall x$, entonces tenemos que el integrando es no-negativo. Por tanto, para que se iguale la integral a 0 el integrando ha de ser nulo. Por tanto,

  $$x = E[X] \qquad \forall x \in [a, b]$$

  Por tanto, llamando $E[X] = c \in \mathbb{R}$, tenemos $X = c$.

□

## 6.10. Función generatriz de momentos

Sea $X$ una variable aleatoria. Si $\exists t_0 \in \mathbb{R} \mid \forall t \in ]-t_0, t_0[$ existe la esperanza $E[e^{tX}]$, se dice entonces que existe la **función generatriz de momentos de** $X$, notada $M_X$ y definida como:

$$M_X(t) = E[e^{tX}] = \begin{cases} \displaystyle\sum_{x_i \in E} e^{tx_i} P(x_i) & \text{caso discreto} \\ \\ \displaystyle\int_{-\infty}^{+\infty} e^{tx} f(x)\, dx & \text{caso continuo} \end{cases} \qquad \forall t \in ]-t_0, t_0[$$

La función generatriz de momentos de una variable aleatoria no tiene por qué existir (por ejemplo, para la distribución de Cauchy). Pero si la variable aleatoria está acotada, entonces siempre existe.

**Teorema 6.17** (de unicidad)**.** *La función generatriz de momentos de una variable aleatoria, si existe, es única y determina de forma única la distribución de la variable.*

Es decir, una variable aleatoria no puede tener dos funciones generatrices de momentos y dos variable aleatorias con distinta distribución tampoco pueden tener la misma función generatriz de momentos.

**Teorema 6.18** (Relación con los momentos)**.** *Si existe la función generatriz de momentos de una variable aleatoria, entonces:*

1. *Existen todos los momentos de la variable aleatoria.*
2. $\forall t \in ]-t_0, t_0[$:
$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n E[X^n]}{n!}$$
3. *Existe la derivada de todos los órdenes de $M_X(t)$ en un entorno de cero y:*
$$M_X^{n)}(t)\Big|_{t=0} = E[X^n] \qquad n = 1, 2, \dots$$

*Demostración.* Demostramos cada resultado por separado:

1. Se deja como ejercicio al lector.
2. Sabemos[3] que $e^x = \sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$. Por tanto,

$$M_X(t) = E[e^{tX}] = E\left[\sum_{n=0}^{\infty} \frac{(tX)^n}{n!}\right] = E\left[\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right] =$$
$$= \sum_{n=0}^{\infty} E\left[\frac{t^n X^n}{n!}\right] = \sum_{n=0}^{\infty} \frac{t^n E\left[X^n\right]}{n!}$$

[3] Se demuestra mediante el desarrollo en serie de Taylor de la exponencial, conceptos estudiados en Cálculo II.

3. Tenemos que:

$$M_X^{n)}(t)\Big|_{t=0} = \frac{d^n}{dt^n}E[e^{tX}]\Big|_{t=0} = E\left[\frac{d^n}{dt^n}e^{tX}\right]\Big|_{t=0} =$$
$$= E\left[X^n e^{tX}\right]\Big|_{t=0} = E[X^n]$$

□

**Ejemplo.** Sea $X$ una variable aleatoria continua con recorrido $Re_X = \mathbb{R}^+$ y función de densidad:

$$f(x) = \begin{cases} 0 & x \leqslant 0 \\ e^{-x} & x > 0 \end{cases}$$

Se pide calcular la función generatriz de momentos de $X$, $M_X(t)$.

$$M_X(t) = E[e^{tX}] = \int_0^{+\infty} e^{tx}e^{-x}\ dx = \int_0^{+\infty} e^{x(t-1)}\ dx =$$
$$= \frac{1}{t-1}\int_0^{+\infty}(t-1)e^{x(t-1)}\ dx = \left.\frac{e^{x(t-1)}}{t-1}\right]_0^{+\infty} = \frac{1}{1-t}$$

Con $t < 1$ para que la integral converja y exista la esperanza.

Comprobamos que $M_X(t) = \dfrac{1}{1-t}$, $\quad t < 1$ es nuestra función generatriz de momentos (al menos para $m_0, m_1, m_2$):

$$M_X(0) = 1 = \int_0^{+\infty} f(x)\ dx = \int_0^{+\infty} e^{-x}\ dx = m_0$$

$$M_X'(t) = \frac{1}{(1-t)^2} \Longrightarrow M_X'(0) = 1 = \int_0^{+\infty} xf(x)\ dx = \int_0^{+\infty} xe^{-x}\ dx = m_1$$

$$M_X''(t) = \frac{2}{(1-t)^3} \Longrightarrow M_X''(0) = 2 = \int_0^{+\infty} x^2f(x)\ dx = \int_0^{+\infty} x^2e^{-x}\ dx = m_2$$

Algunas de sus propiedades son:

**Lema 6.19.** *Sea $X$ una variable aleatoria y sea $M_X(t)$ su función generatriz de momentos. Entonces:*

$$M_X(0) = 1$$

*Demostración.*

$$M_X(0) = E[e^{0X}] = E[1] = 1$$

□

**Proposición 6.20.** *Sea $X$ una variable aleatoria con función generatriz de momentos $M_X(t)$ $\forall t \in ]-t_0, t_0[$, sea $Y = aX + b$ $a, b \in \mathbb{R}$, entonces la función generatriz de momentos de $Y$, para $t$ tal que $at \in ]-t_0, t_0[$ verifica:*

$$M_Y(t) = e^{bt}M_X(at)$$

*Demostración.* Realizamos distinción entre caso continuo y caso discreto:

- <u>Caso discreto</u>. Suponemos $Re_X = E$. Entonces:

$$M_Y(t) = E[e^{ty}] = E[e^{t(ax+b)}] = \sum_{x_i \in E} e^{t(ax_i+b)} f(x_i) = e^{tb} \sum_{x_i \in E} e^{t(ax_i)} f(x_i) = \\ = e^{tb} E[e^{atx}] = e^{bt} M_X(at)$$

- <u>Caso continuo</u>. Suponemos $Re_X = [a, b]$. Entonces:

$$M_Y(t) = E[e^{ty}] = E[e^{t(ax+b)}] = \int_a^b e^{t(ax+b)} f(x)\, dx = e^{tb} \int_a^b e^{t(ax)} f(x)\, dx = \\ = e^{tb} E[e^{atx}] = e^{bt} M_X(at)$$

□

**Proposición 6.21.** *Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con función generatriz de momentos $M_{X_i}(t)$ para todo $i = 1, \ldots, n$. Entonces, la función generatriz de momentos de la variable aleatoria $X = \sum\limits_{i=1}^{n} X_i$ es:*

$$M_X(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

*Demostración.*

$$M_X(t) = E[e^{tX}] = E\left[e^{t\sum\limits_{i=1}^{n} X_i}\right] = E\left[\prod_{i=1}^{n} e^{tX_i}\right] \overset{(*)}{=} \prod_{i=1}^{n} E[e^{tX_i}] = \prod_{i=1}^{n} M_{X_i}(t)$$

donde en $(*)$ hemos empleado la Proposición 6.11. □

# 7. Modelos de distribuciones discretas

## 7.1. Distribución degenerada

**Definición 7.1.** Sea $X$ una variable aleatoria. Decimos que es degenerada si $X = c$, es decir, toma tan solo un valor constante.

Su función masa de probabilidad es:

$$f(x) = P[X = x] = \begin{cases} 1 & x = c \\ 0 & x \neq c \end{cases}$$

Su función de distribución es:

$$F_X(x) = P[X \leqslant x] \begin{cases} 0 & x < c \\ 1 & c \leqslant x \end{cases}$$

Sus momentos no centrados son:

$$m_k = E[X^k] = c^k P[X = c] = c^k \qquad \forall k \in \mathbb{N}$$

Sus momentos centrados son:

$$\mu_k = E[(X - E[X])^k] = E[(X - c)^k] = (c - c)^k P[X = c] = 0 \qquad \forall k \in \mathbb{N}$$

Su función generatriz de momentos es:

$$M_X(t) = E[e^{tX}] = e^{tc} P[X = c] = e^{tc} \qquad \forall t \in \mathbb{R}$$

Como se vio en la Proposición 6.16, tenemos que $\mathrm{Var}[X] = 0$ caracteriza a las variables degeneradas.

## 7.2. Distribución uniforme discreta

**Definición 7.2.** Sea $X$ una variable aleatoria. Se dice que la variable aleatoria $X$ se distribuye uniformemente alrededor de los puntos $x_1, \ldots, x_n$ si dichos valores son equiprobables.

Se notará como sigue:

$$X \rightsquigarrow U(x_1, \ldots, x_n)$$

Su función masa de probabilidad es:

$$f(x) = P[X = x] = \begin{cases} \frac{1}{n} & x = 1, \dots, n \\ 0 & \text{en caso contrario} \end{cases}$$

Su función de distribución es:

$$F_X(x) = P[X \leqslant x] \begin{cases} 0 & x < x_1 \\ \frac{k}{n} & x \in [x_k, , x_{k+1}[ \qquad \forall k = 1, \dots, n-1 \\ 1 & x_n \leqslant x \end{cases}$$

Sus momentos no centrados son:

$$m_k = E[X^k] = \sum_{x_i \in Re_X} x_i^k f(x_i) = \frac{1}{n} \sum_{x_i \in Re_X} x_i^k \qquad \forall k \in \mathbb{N}$$

En particular, $E[X] = m_1 = \frac{1}{n} \sum_{x_i \in Re_X} x_i = \bar{x}$.

Sus momentos centrados son:

$$\mu_k = E[(X - E[X])^k] = E[(X - \bar{x})^k] = \frac{1}{n} \sum_{x_i \in Re_X} (x_i - \bar{x})^k \qquad \forall k \in \mathbb{N}$$

En particular, $\text{Var}[X] = \mu_2 = \frac{1}{n} \sum_{x_i \in Re_X} (x_i - \overline{x})^2 = \sigma_x^2$.

Su función generatriz de momentos es:

$$M_X(t) = E[e^{tX}] = \frac{1}{n} \sum_{x_i \in Re_X} e^{tx_i} \qquad \forall t \in \mathbb{R}$$

Cuando $Re_X = \{1, 2, \dots, n\}$, veamos algunos casos particulares de series[1]:

$$\begin{aligned} E[X] &= \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2} \\ E[X^2] &= \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6} \\ Var(X) &= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12} \end{aligned}$$

## 7.3. Distribución de Bernouilli

**Definición 7.3** (Experimentos de Bernouilli)**.** Un experimento de Bernouilli es un experimento aleatorio que da lugar a dos posibles resultados mutuamente excluyentes y exhaustivos. Estos son denominados como éxito ($E$) y fracaso ($F = \bar{E}$).

$$\Omega = \{E, F\}$$

[1]Se demuestran fácilmente mediante inducción, pero se deja como ejercicio por no ser materia de la asignatura de EDIP, sino de Cálculo I.

**Definición 7.4** (Distribución de Bernouilli)**.** Se define la variable aleatoria con distribución de Bernouilli como:

$$X = \begin{cases} 1 & \text{ocurre el suceso } E \\ 0 & \text{ocurre el suceso } F \end{cases}$$

Denotando $p$ como $P(E) \in [0,1]$, tenemos que la variable aleatoria en cuestión tiene distribución de Bernouilli, y se notará como sigue:

$$X \rightsquigarrow B(1,p)$$

Su función masa de probabilidad es:

$$f(x) = P[X = x] = \begin{cases} p & x = 1 \\ 1-p & x = 0 \\ 0 & x \neq 0,1 \end{cases}$$

Análogamente, es común escribirlo como:

$$f(x) = p^x(1-p)^{1-x} \qquad x = 0,1$$

Su función de distribución es:

$$F_X(x) = P[X \leqslant x] \begin{cases} 0 & x < 0 \\ 1-p & 0 \leqslant x < 1 \\ 1 & 1 \leqslant x \end{cases}$$

Sus momentos no centrados son:

$$m_k = E[X^k] = \sum_{x_i \in Re_X} x_i^k f(x_i) = 0^k \cdot (1-p) + 1^k \cdot p = p \qquad \forall k \in \mathbb{N}$$

Sus momentos centrados son:

$$\mu_k = E[(X-E[X])^k] = E[(X-p)^k] = \sum_{x_i \in Re_X} (x_i-p)^k f(x_i) = (1-p)^k \cdot p + (0-p)^k(1-p) =$$
$$= p(1-p)^k + (-p)^k(1-p) \qquad \forall k \in \mathbb{N}$$

En particular,

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = p - p^2 = p(1-p)$$

Su función generatriz de momentos es:

$$M_X(t) = E[e^{tX}] = \sum_{x_i \in Re_X} e^{tx_i} f(x_i) = e^t \cdot p + e^0 \cdot (1-p) = 1-p+e^t p = 1+p(e^t-1) \qquad \forall t \in \mathbb{R}$$

## 7.4. Distribución binomial

**Definición 7.5** (Distribución binomial)**.** Se dice que una variable aleatoria $X$ sigue una distribución binomial de parámetros $n$ y $p$, $n \in \mathbb{N}$, $p \in ]0,1[$ si modela el número de éxitos en $n$ repeticiones independientes de un ensayo de Bernouilli con probabilidad $p$ de éxito, manteniéndose esta constante en las $n$ repeticiones. Se notará como

$$X \rightsquigarrow B(n,p)$$

*Observación.* Es fácil ver que la distribución de Bernouilli es un caso particular de una distribución binomial, considerando $n = 1$.

Razonemos la función masa de probabilidad. Como la variable $X$ modela el número de éxitos, tenemos que en total se producen $x$ éxitos y $n - x$ fracasos. Como la probabilidad se mantiene constante, tenemos que la probabilidad para cierta ordenación de los éxitos es $p^x(1-p)^{n-x}$. No obstante, las distintas formas de reorganizar los sucesos son combinaciones de $x$ éxitos entre un total de $n$ sucesos; es decir, combinaciones sin repetición. Por tanto, tenemos que en total hay $\binom{n}{x}$ formas distintas. Por tanto, tenemos el siguiente resultado:

**Proposición 7.1.** *Sea $X \rightsquigarrow B(n,p)$. Entonces, la función masa de probabilidad de la distribución binomial es:*

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x \in \mathbb{N} \cup \{0\}$$

*Demostración.* Comprobemos que es una función masa de probabilidad. En primer lugar, es fácil ver que $f(x) \geqslant 0 \ \forall x \in \mathbb{N} \cup \{0\}$. Veamos ahora que $\sum\limits_{x=0}^{n} f(x) = 1$:

$$\sum_{x=0}^{n} f(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = [p + (1-p)]^n = 1^n = 1$$

donde he aplicado que $f(x)$ es el término $x-$ésimo del binomio de Newton. □

**Proposición 7.2.** *Sea $X \rightsquigarrow B(n,p)$. Entonces, tenemos que:*

$$E[X] = np$$

*Demostración.* Tenemos que:

$$E[X] = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

Buscamos emplear el binomio de Newton, que dice que:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \qquad \forall a, b \in \mathbb{R}, \ \forall n \in \mathbb{N}$$

Para poder “deshacernos” de la $x$ que acompaña al número combinatorio, aplico que:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{n}{x} \cdot \frac{(n-1)!}{(x-1)!(n-1-x+1)!} = \frac{n}{x}\binom{n-1}{x-1}$$

Por tanto,

$$E[X] = n\sum_{x=1}^{n}\binom{n-1}{x-1}p^x(1-p)^{n-x}$$

Realizo el cambio de variable, por notación, $x-1=h$. Entonces:

$$E[X] = n\sum_{h=0}^{n-1}\binom{n-1}{h}p^{h+1}(1-p)^{n-h-1} = np(p+1-p)^{n-1} = np$$

□

**Lema 7.3.** *Sea $X \rightsquigarrow B(n,p)$. Entonces, tenemos que:*

$$E[X^2] = n(n-1)p^2 + np$$

*Demostración.* Tenemos que:

$$E[X(X-1)] = \sum_{x=0}^{n} x(x-1)\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=2}^{n} x(x-1)\binom{n}{x}p^x(1-p)^{n-x}$$

Buscamos emplear el binomio de Newton, que dice que:

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \qquad \forall a,b\in\mathbb{R},\ \forall n\in\mathbb{N}$$

Para poder "deshacernos" de la $x$ que acompaña al número combinatorio, aplico que:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{n(n-1)}{x(x-1)}\cdot\frac{(n-2)!}{(x-2)!(n-2-x+2)!} = \frac{n(n-1)}{x(x-1)}\binom{n-2}{x-2}$$

Por tanto,

$$E[X(X-1)] = n(n-1)\sum_{x=2}^{n}\binom{n-2}{x-2}p^x(1-p)^{n-x}$$

Realizo el cambio de variable, por notación, $x-2=h$. Entonces:

$$E[X(X-1)] = n(n-1)\sum_{h=0}^{n-2}\binom{n-2}{h}p^{h+2}(1-p)^{n-h-2} = n(n-1)p^2(p+1-p)^{n-2} = n(n-1)p^2$$

Por tanto, tenemos que $E[X^2-X] = E[X^2]-E[X] = n(n-1)p^2$. Por tanto, usando que $E[X]=np$,

$$E[X^2] = n(n-1)p^2 + np$$

□

**Corolario 7.3.1.** *Sea $X \rightsquigarrow B(n,p)$. Entonces, tenemos que:*

$$\mathrm{Var}[X] = np(1-p)$$

*Demostración.* Tenemos que:

$$\begin{aligned}\mathrm{Var}[X] = E[X^2] - E[X]^2 = n(n-1)p^2 + np - n^2p^2 = np[p(n-1) + 1 - np] = \\ = np[pn - p + 1 - np] = np(1-p)\end{aligned}$$

□

**Proposición 7.4.** *Sea $X \rightsquigarrow B(n,p)$. La función generatriz de momentos es:*

$$M_X(t) = [1 + p(e^t - 1)]^n$$

*Demostración.* Buscamos emplear el binomio de Newton, que dice que:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \qquad \forall a, b \in \mathbb{R},\ \forall n \in \mathbb{N}$$

Tenemos que:

$$\begin{aligned}E[e^{tX}] = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (e^t p)^x (1-p)^{n-x} = \\ = (e^t p + 1 - p)^n = [1 + p(e^t - 1)]^n\end{aligned}$$

□

**Proposición 7.5** (Simetría)**.** *Si $X \rightsquigarrow B(n,p)$, entonces la variable que contabiliza el número de fracasos, $Y = n - X \rightsquigarrow B(n, 1-p)$ y, además,*

$$P[X = x] = P[Y = n - x]$$

*Demostración.* Sabemos que se trata de nuevo de una distribución binomial con las mismas repeticiones. No obstante, por tener solo dos sucesos elementales tenemos que $P(f) = 1 - p$, siendo $f$ el fallo.

La segunda expresión es también trivialmente cierta, ya que el número de fallos ha de ser $n$ menos el número de éxitos. □

**Ejemplo.** Tenemos una muestra de tornillos defectuosos, y sabemos que la probabilidad de que un tornillo sea defectuoso es $p = 0{,}05$. Calcular la probabilidad de que en una muestra de 30 tornillos haya exactamente 5 defectuosos.

Tenemos que el espacio muestral es $\Omega = \{T.Op, T.Def\}$

Sea la variable aleatoria siguiente:

$X$="Número de tornillos defectuosos en una muestra de 30"

Tenemos que $X$ sigue una distribución binomial de la forma:

$$X \rightsquigarrow B(30; 0{,}05)$$

Por tanto,

$$P[X = 5] = \binom{30}{5} \cdot 0{,}05^5 \cdot 0{,}95^{25}$$

## 7.5. Distribución Geométrica

**Definición 7.6** (Distribución geométrica)**.** Se dice que una variable aleatoria $X$ sigue una distribución geométrica de parámetro $p$, $p \in ]0,1[$ si modela el número de fracasos antes de llegar al primer éxito en un ensayo de Bernouilli con probabilidad $p$ de éxito, manteniéndose esta constante en todas las repeticiones. Se notará como

$$X \rightsquigarrow \mathcal{G}(p)$$

Veamos cuál es su función masa de probabilidad. Como la variable $X$ modela el número de fracasos antes de llegar al primer éxito, tenemos que en total se producen $x$ fracasos y un éxito. Como la probabilidad se mantiene constante, tenemos que la probabilidad para cierta ordenación de los sucesos es $(1-p)^x p$. Por tanto:

$$f(x) = (1-p)^x p \qquad x \in \mathbb{N} \cup \{0\}$$

Veamos que, efectivamente, se trata de una función masa de probabilidad.

**Proposición 7.6.** *Sea $X \rightsquigarrow \mathcal{G}(p)$. Entonces, la función masa de probabilidad de la distribución geométrica es:*

$$f(x) = (1-p)^x p \qquad x \in \mathbb{N} \cup \{0\}$$

*Demostración.* Comprobemos que es una función masa de probabilidad. En primer lugar, es fácil ver que $f(x) \geqslant 0 \ \forall x \in \mathbb{N} \cup \{0\}$. Veamos ahora que $\sum\limits_{x=0}^{\infty} f(x) = 1$:

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} (1-p)^x p = p \sum_{x=0}^{\infty} (1-p)^x \overset{(*)}{=} p \cdot \frac{1}{1-(1-p)} = 1$$

donde en $(*)$ he aplicado la fórmula de la suma de una serie geométrica de razón $(1-p)$. □

**Proposición 7.7.** *Sea $X \rightsquigarrow \mathcal{G}(p)$. Entonces, su función de distribución es:*

$$F_X(x) = P[X \leqslant x] = \begin{cases} 1-(1-p)^{x+1} & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

*Demostración.* Tenemos que:

$$F_X(x) = P[X \leqslant x] = \sum_{k=0}^{x} f(k) = \sum_{k=0}^{x} (1-p)^k p = p \sum_{k=0}^{x} (1-p)^k$$

Por tanto, y usando la fórmula de la suma parcial de una serie geométrica de razón $(1-p)$, tenemos que:

$$F_X(x) = p \cdot \frac{1-(1-p)^{x+1}}{1-(1-p)} = 1-(1-p)^{x+1}$$

□

**Proposición 7.8.** *Sea $X \rightsquigarrow \mathcal{G}(p)$. Entonces, su función generatriz de momentos es:*

$$M_X(t) = \frac{p}{1-(1-p)e^t} \qquad t < -\ln(1-p)$$

*Demostración.* Tenemos que:

$$M_X(t) = E[e^{tX}] = \sum_{x=0}^{\infty} e^{tx} f(x) = \sum_{x=0}^{\infty} e^{tx}(1-p)^x p = p\sum_{x=0}^{\infty}[e^t(1-p)]^x$$

Veamos ahora qué ha de cumplir $t$ para que la serie converga. Como sabemos que $e^t(1-p) > 0$, tan solo comprobamos que $e^t(1-p) < 1$:

$$e^t(1-p) < 1 \iff e^t < \frac{1}{1-p} \iff t < \ln\left(\frac{1}{1-p}\right) = -\ln(1-p)$$

Por tanto, para $t < -\ln(1-p)$, la serie converge y, por tanto, la función generatriz de momentos es:

$$M_X(t) = \frac{p}{1-(1-p)e^t}$$

□

**Corolario 7.8.1.** *Sea $X \rightsquigarrow \mathcal{G}(p)$. Entonces, tenemos que:*

$$E[X] = \frac{1-p}{p}$$

*Demostración.* Tenemos que:

$$E[X] = M_X'(t)\Big|_{t=0} = \frac{p}{(1-(1-p)e^t)^2} \cdot (1-p)e^t\Big|_{t=0} = \frac{p(1-p)}{p^2} = \frac{1-p}{p}$$

□

**Corolario 7.8.2.** *Sea $X \rightsquigarrow \mathcal{G}(p)$. Entonces, tenemos que:*

$$\mathrm{Var}[X] = \frac{1-p}{p^2}$$

*Demostración.* Tenemos que:

$$E[X^2] = M_X''(t)\Big|_{t=0} = p(1-p)\cdot\frac{e^t[1-(1-p)e^t]^{\cancel{2}} + e^t\cdot 2(1-p)e^t\,\cancel{(1-(1-p)e^t)}}{(1-(1-p)e^t)^{\cancel{4}\,3}}\Bigg|_{t=0} =$$

$$= p(1-p)e^t\cdot\frac{1-(1-p)e^t+2(1-p)e^t}{(1-(1-p)e^t)^3}\Bigg|_{t=0} = p(1-p)e^t\cdot\frac{1+(1-p)e^t}{(1-(1-p)e^t)^3}\Bigg|_{t=0} =$$

$$= p(1-p)\cdot\frac{1+(1-p)}{p^3} = (1-p)\cdot\frac{2-p}{p^2}$$

Por tanto, tenemos que:

$$\mathrm{Var}[X] = E[X^2]-E[X]^2 = \frac{(1-p)(2-p)}{p^2} - \frac{(1-p)^2}{p^2} = \frac{(1-p)(2-p+1-p)}{p^2} = \frac{1-p}{p^2}$$

□

**Proposición 7.9** (Falta de memoria)**.** *Sea* $X \rightsquigarrow \mathcal{G}(p)$*. Entonces, se cumple que:*

$$P(X \geqslant h+k \mid X \geqslant h) = P(X \geqslant k) \qquad \forall h,k \in \mathbb{N} \cup \{0\}$$

*Demostración.* Tenemos que:

$$P(X \geqslant h+k \mid X \geqslant h) = \frac{P(X \geqslant h+k, X \geqslant h)}{P(X \geqslant h)} = \frac{P(X \geqslant h+k)}{P(X \geqslant h)} \overset{(*)}{=}$$
$$\overset{(*)}{=} \frac{(1-p)^{h+k}}{(1-p)^h} = (1-p)^k \overset{(*)}{=} P(X \geqslant k)$$

donde en $(*)$ hemos usado que:

$$P(X \geqslant x) = 1 - P(X < x) = 1 - P(X \leqslant x-1) = 1 - (1-(1-p)^x) = (1-p)^x$$

□

## 7.6. Distribución binomial negativa

**Definición 7.7** (Distribución binomial negativa)**.** Se dice que una variable aleatoria $X$ sigue una distribución binomial negativa de parámetros $r$ y $p$, $r \in \mathbb{N}$, $p \in ]0,1[$ si modela el número de fracasos antes de llegar al $r-$ésimo éxito en un ensayo de Bernouilli con probabilidad $p$ de éxito, manteniéndose esta constante en todas las repeticiones. Se notará como

$$X \rightsquigarrow BN(r,p)$$

*Observación.* Es fácil ver que la Distribución Geométrica es un caso particular de una distribución binomial negativa, considerando $r = 1$.

Razonemos la función masa de probabilidad. Como la variable $X$ modela el número de fracasos antes de llegar al $r-$ésimo éxito, tenemos que en total se producen $x$ fracasos y $r$ fracasos. Como la probabilidad se mantiene constante, tenemos que la probabilidad para cierta ordenación de los sucesos es $p^r(1-p)^x$. No obstante, las distintas formas de reorganizar los sucesos son combinaciones de $x$ fracasos y $r-1$ éxitos[2]; es decir, combinaciones sin repetición. Por tanto, tenemos que en total hay $\binom{x+r-1}{x}$ formas distintas. Por tanto, a priori deducimos que la función masa de probabilidad de la distribución binomial negativa es:

$$f(x) = \binom{x+r-1}{x}(1-p)^x p^r \qquad x \in \mathbb{N} \cup \{0\}$$

No obstante, hay una expresión alternativa de la función masa de probabilidad muy útil para las demostraciones. Para ello, se introduce la siguiente definición:

**Definición 7.8.** Dado $r \in \mathbb{R}, x \in \mathbb{N}$, se define[3] el siguiente número combinatorio:

$$\binom{-r}{x} = \frac{(-r)(-r-1)\cdots(-r-x+1)}{x!}; \ \binom{-\alpha}{0} = 1, \qquad \forall r \in \mathbb{R}, x \in \mathbb{N}$$

[2] El éxito $r-$ésimo no se añade porque su posición está fijada, ya que debe ser el último suceso

[3] El número combinatorio está ya definido, pero con más restricciones. Esta definición concuerda con la anterior en los supuestos anteriores.

Habiendo definido dicho número combinatorio, podemos introducir la expresión alternativa de la función masa de probabilidad:

**Proposición 7.10** (Expresión alternativa)**.** *Se cumple la siguiente igualdad:*

$$f(x) = \binom{x+r-1}{x}(1-p)^x p^r = \binom{-r}{x}(p-1)^x p^r \qquad \forall x \in \mathbb{N} \cup \{0\}$$

*Demostración.* En primer lugar, veamos el siguiente resultado:

$$\binom{x+r-1}{x} = \frac{(x+r-1)!}{x!(r-1)!} = \frac{(x+r-1)(x+r-2)\cdots(\cancel{x}+r-\cancel{x})\cancel{(r-1)!}}{x!\cancel{(r-1)!}} =$$
$$= \frac{(x+r-1)(x+r-2)\cdots(r)}{x!} \overset{(*)}{=} (-1)^x \cdot \frac{(-r)(-r-1)\cdots(-r-x+1)}{x!} = (-1)^x \binom{-r}{x}$$

donde en $(*)$ he incluido en el numerador $x$ signos negativos y, por tanto, se incluye también el factor $(-1)^x$.

Por tanto, se tiene que:

$$\binom{x+r-1}{x}(1-p)^x p^r = (-1)^x \binom{-r}{x}(1-p)^x p^r = \binom{-r}{x}(p-1)^x p^r$$

□

Recordemos que tenemos lo que hemos deducido que es la función masa de probabilidad y su expresión alternativa. No obstante, no hemos demostrado que, efectivamente, se trata de una función masa de probabilidad.

**Proposición 7.11.** *Sea $X \rightsquigarrow BN(r,p)$. Entonces, la función masa de probabilidad de la distribución binomial negativa es:*

$$f(x) = \binom{x+r-1}{x}(1-p)^x p^r = \binom{-r}{x}(p-1)^x p^r \qquad x \in \mathbb{N} \cup \{0\}$$

*Demostración.* Comprobemos que es una función masa de probabilidad. En primer lugar, es fácil ver que $f(x) \geqslant 0 \ \forall x \in \mathbb{N} \cup \{0\}$ a partir de la primera expresión.

Veamos ahora que $\sum\limits_{x=0}^{\infty} f(x) = 1$. Para ello, usaremos el desarrollo en serie de potencias de la función $(1+t)^\alpha$:

$$(1+t)^\alpha = \sum_{x=0}^{\infty} \binom{\alpha}{x} t^x, \qquad |t| < 1, \alpha \in \mathbb{R} \tag{7.1}$$

Usando la expresión alternativa de la función masa de probabilidad obtenida anteriormente y aplicando este desarrollo tenemos:

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \binom{-r}{x}(p-1)^x p^r = p^r \sum_{x=0}^{\infty} \binom{-r}{x}(p-1)^x \overset{Ec.\ 7{,}1}{=} p^r[1+(p-1)]^{-r} = p^r p^{-r} = 1$$

donde he aplicado que $f(x)$ es el término $x-$ésimo del binomio de Newton. □

**Proposición 7.12.** *Sea $X \rightsquigarrow BN(r,p)$. Entonces, tenemos que:*

$$E[X] = \frac{r(1-p)}{p}$$

*Demostración.* Demostramos haciendo uso de la expresión alternativa y de la Ecuación 7.1:

$$\begin{aligned}
E[X] &= \sum_{x=0}^{\infty} xf(x) \overset{(*)}{=} \sum_{x=1}^{\infty} xf(x) = \sum_{x=1}^{\infty} x\binom{-r}{x}(p-1)^x p^r = p^r \sum_{x=1}^{\infty} x\binom{-r}{x}(p-1)^x = \\
&= p^r \sum_{x=1}^{\infty} x\frac{(-r)(-r-1)\cdots(-r-x+1)}{x!}(p-1)^x = \\
&= (-r)(p-1)p^r \sum_{x=1}^{\infty} \frac{(-r-1)\cdots(-r-x+1)}{(x-1)!}(p-1)^{x-1} = \\
&= r(1-p)p^r \sum_{x=1}^{\infty} \binom{-r-1}{x-1}(p-1)^{x-1} \overset{[y=x-1]}{=} r(1-p)p^r \sum_{y=0}^{\infty} \binom{-r-1}{y}(p-1)^y = \\
&\overset{Ec.\ 7,1}{=} r(1-p)p^r[1+(p-1)]^{-r-1} = r(1-p)p^r p^{-r-1} = r(1-p)p^{-1} = \frac{r(1-p)}{p}
\end{aligned}$$

donde en $(*)$ he aplicado que la primera iteración es nula por ser $x=0$. $\square$

**Lema 7.13.** *Sea $X \rightsquigarrow BN(r,p)$. Entonces, tenemos que:*

$$E[X^2] = \frac{r(r+1)(1-p)^2}{p^2} + \frac{r(1-p)}{p}$$

*Demostración.* Calculamos en primer lugar lo siguiente:

$$\begin{aligned}
E[X(X-1)] &= \sum_{x=0}^{\infty} x(x-1)f(x) = \sum_{x=2}^{\infty} x(x-1)\binom{-r}{x}(p-1)^x p^r \\
&= p^r \sum_{x=2}^{\infty} x(x-1)\frac{(-r)(-r-1)\cdots(-r-x+1)}{x!}(p-1)^x \\
&= (-r)(-r-1)p^r(p-1)^2 \sum_{x=2}^{\infty} \frac{(-r-2)\cdots(-r-x+1)}{(x-2)!}(p-1)^{x-2} \\
&= r(r+1)p^r(p-1)^2 \sum_{x=2}^{\infty} \binom{-r-2}{x-2}(p-1)^{x-2} = \\
&\overset{y=x-2}{=} r(r+1)p^r(p-1)^2 \sum_{y=0}^{\infty} \binom{-r-2}{y}(p-1)^y \\
&\overset{Ec.\ 7,1}{=} r(r+1)p^r(p-1)^2[1+(p-1)]^{-r-2} = r(r+1)p^r(p-1)^2 p^{-r-2} \\
&= r(r+1)p^{-2}(p-1)^2 = \frac{r(r+1)(1-p)^2}{p^2}
\end{aligned}$$

Por tanto, tenemos que:

$$E[X(X-1)] = E[X^2 - X] = E[X^2] - E[X] \Longrightarrow$$
$$\Longrightarrow E[X^2] = E[X(X-1)] + E[X] = \frac{r(r+1)(1-p)^2}{p^2} + \frac{r(1-p)}{p}$$

□

**Corolario 7.13.1.** *Sea $X \rightsquigarrow BN(r,p)$. Entonces, tenemos que:*

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

*Demostración.* Por la linealidad de la esperanza, tenemos que:

$$\begin{aligned}
\text{Var}[X] &= E[X^2] - E[X]^2 = \frac{r(r+1)(1-p)^2}{p^2} + \frac{r(1-p)}{p} - \frac{r^2(1-p)^2}{p^2} = \\
&= \frac{\cancel{r^2(1-p)^2} + r(1-p)^2 - \cancel{r^2(1-p)^2}}{p^2} + \frac{r(1-p)}{p} = \frac{r(1-p)^2 + pr(1-p)}{p^2} = \\
&= \frac{[r(1-p)][(1-p)+p]}{p^2} = \frac{r(1-p)}{p^2}
\end{aligned}$$

□

**Proposición 7.14.** *Sea $X \rightsquigarrow BN(r,p)$. Entonces, tenemos que:*

$$M_X(t) = \frac{p^r}{[1-(1-p)e^t]^{-r}} \qquad \forall t < -\ln(1-p)$$

*Demostración.* Usando la expresión alternativa de la función masa de probabilidad:

$$\begin{aligned}
M_X(t) = E[e^{tX}] &= \sum_{x=0}^{\infty} e^t x \binom{-r}{x} (p-1)^x p^r = p^r \sum_{x=0}^{\infty} \binom{-r}{x} [e^t(p-1)]^x = \\
&\overset{Ec.\ 7{,}1}{=} p^r[1 + [e^t(p-1)]]^{-r} = \frac{p^r}{[1-(1-p)e^t]^{-r}}
\end{aligned}$$

donde hemos empleado la Ecuación 7.1. No obstante, esto solo es válido si:

$$|e^t(p-1)| < 1 \Longleftrightarrow e^t(1-p) < 1 \Longleftrightarrow e^t < \frac{1}{1-p} \Longleftrightarrow t < -\ln(1-p)$$

□

**Ejemplo.** Sabemos que la probabilidad de que te multen al conducir es $P(M) = 0{,}01$. Además, a las 3 te quitan el carnet. Calcular la probabilidad de que te quiten el carnet la décima vez que te subes al coche.

Sea $X =$"Número de veces que te subes al coche antes de la tercera multa."

Tenemos que $X \rightsquigarrow BN(3, 0{,}01)$. El número de fracasos antes del tercer éxito son 7. Por tanto,

$$P(X=7) = \binom{9}{7} 0{,}01^3 0{,}99^7$$

## 7.7. Distribución Hipergeométrica

**Definición 7.9** (Distribución hipergeométrica)**.** Supongamos una población de $N$ individuos en dos categorías de $N_1$ y $N_2 = N - N_1$ individuos cada una. Se elige una muestra de $n$ individuos de la población (sin reemplazamiento o simultáneamente). La variable aleatoria $X$ que contabiliza el número de individuos de la primera categoría en la muestra se dice sigue una distribución hipergeométrica de parámetros $N$, $N_1$ y $n$ y se denota:

$$X \rightsquigarrow H(N, N_1, n) \qquad N, N_1, n \in \mathbb{N}, \quad N_1, n \leqslant N$$

Tenemos que su función masa de probabilidad viene determinada por la Ley de Laplace:

$$f(x) = \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}} \qquad x \in [\text{máx}\{0, n-(N-N_1)\}, \text{mín}\{n, N_1\}]$$

El denominador son las combinaciones totales de $n$ individuos de un total de $N$. En el denominador, tenemos que son las combinaciones de los $x$ individuos de la población $N_1$ por las combinaciones de los $n-x$ de la población $N_2$.

Comprobemos que se trata de una función masa de probabilidad:

**Proposición 7.15.** *Sea $X \rightsquigarrow H(N, N_1, n)$. Entonces, su función masa de probabilidad es:*

$$f(x) = \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}} \qquad x \in [\text{máx}\{0, n-(N-N_1)\}, \text{mín}\{n, N_1\}]$$

*Demostración.* Como todos los términos son positivos, es fácil ver que $f(x) \geqslant 0 \quad \forall x$. Veamos ahora que su suma es igual a 1. Para ello, empleamos el siguiente resultado, que no se demuestra al no entrar dentro del temario de EDIP sino de Cálculo.

$$\sum_{x=0}^{n} \binom{a}{x}\binom{b}{n-x} = \binom{a+b}{n}, \qquad \forall a, b \in \mathbb{R},\ n \in \mathbb{N}$$

Notemos que en la suma del primer miembro algunos sumandos pueden ser nulos. En efecto, si $x > a$, entonces $\binom{a}{x} = 0$ (puesto que habrá un término nulo en $a(a-1)\cdots(a-x+1)$). De igual forma, si $n - x > b$ se verifica que $\binom{b}{n-x} = 0$. En consecuencia, los únicos sumandos no nulos son aquellos comprendidos entre $\text{máx}\{0, n-b\}$ y $\text{mín}\{n, a\}$ y, por tanto:

$$\sum_{x=0}^{n} \binom{a}{x}\binom{b}{n-x} = \sum_{x=\text{máx}\{0,n-b\}}^{\text{mín}\{n,a\}} \binom{a}{x}\binom{b}{n-x} = \binom{a+b}{n}, \qquad \forall a, b \in \mathbb{R},\ n \in \mathbb{N}$$

Por tanto,

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}} = \sum_{\text{máx}\{0,n-(N-N_1)\}}^{\text{mín}\{n,N_1\}} \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}} = \frac{\binom{N_1+N-N_1}{x+n-x}}{\binom{N}{n}} = \frac{\binom{N}{n}}{\binom{N}{n}} = 1$$

□

**Proposición 7.16.** *Sea* $X \rightsquigarrow H(N, N_1, n)$*. Entonces, tenemos que:*

$$E[X] = n\frac{N_1}{N}$$

**Proposición 7.17.** *Sea* $X \rightsquigarrow H(N, N_1, n)$*. Entonces, tenemos que:*

$$\text{Var}[X] = \frac{n(N-n)N_1(N-N_1)}{N^2(N-1)}$$

**Proposición 7.18** (Aproximación de Hipergeométrica a Binomial)**.** *Sea* $X \rightsquigarrow H(N, N_1, n)$*. Entonces, tenemos que, definiendo* $p = \frac{N_1}{N}$*,* $X$ *se aproxima a una distribución binomial* $B(n, p)$*.*

*De forma empírica, se ha demostrado que esta aproximación es adecuada si* $N_1 \leqslant 0{,}1N$ *o, equivalentemente,* $p \leqslant 0{,}1$*.*

*Demostración.* Tomando $N_1 = Np$, consideremos la función masa de probabilidad de la variable $X$ y la desarrollamos de la siguiente forma:

$$\begin{aligned}
P(X=x) &= \frac{\binom{Np}{x}\binom{N(1-p)}{n-x}}{\binom{N}{n}} = \frac{(Np)!}{x!(Np-x)!}\frac{(N(1-p))!}{(n-x)!(N(1-p)-n+x)!}\frac{n!(N-n)!}{N!} = \\
&= \frac{n!}{x!(n-x)!}\frac{(Np)!}{(Np-x)!}\frac{(N(1-p))!}{(N(1-p)-n+x)!}\frac{(N-n)!}{N!} = \\
&= \binom{n}{x}\cdot\frac{Np(Np-1)\cdots(Np-x+1)}{N(N-1)\cdots(N-n+1)}\cdot \\
&\quad\cdot\frac{N(1-p)(N(1-p)-1)\cdots(N(1-p)-n+x+1)}{N(N-1)\cdots(N-n+1)}
\end{aligned}$$

Si observamos la última expresión, tanto el numerador como el denominador son polinomios en $N$ de grado $n$, cuyos términos de mayor grado son, respectivamente, $(Np)^x(N(1-p))^{n-x}$ y $N^n$. Por tanto, tomando límites cuando $N$ tiende a infinito, se tiene:

$$\begin{aligned}
\lim_{N\to\infty} f(x) &= \lim_{N\to\infty}\binom{n}{x}\cdot\frac{Np(Np-1)\cdots(Np-x+1)}{N(N-1)\cdots(N-n+1)}\cdot \\
&\quad\cdot\frac{N(1-p)(N(1-p)-1)\cdots(N(1-p)-n+x+1)}{N(N-1)\cdots(N-n+1)} = \\
&= \binom{n}{x}\cdot\lim_{N\to\infty}\frac{N^x(Np)^x(N(1-p))^{n-x}}{N^n} = \binom{n}{x}\lim_{N\to\infty}\frac{N^np^x(1-p)^{n-x}}{N^n} = \\
&= \binom{n}{x}p^x(1-p)^{n-x}
\end{aligned}$$

□

**Ejemplo.** Tenemos un grupo de 20 empleados en el que hay 15 hombres y 5 mujeres y elegimos 6 personas para formar un grupo de forma completamente aleatoria. Calcular la probabilidad de que en el grupo haya 2 mujeres.

Definimos nuestra variable aleatoria $X$ como el número de mujeres en el grupo que cogemos. Notemos que:

$$H \rightsquigarrow (20, 5, 6)$$

Calculamos la probabilidad de que haya dos mujeres en el grupo:

$$P[X = 2] = \frac{\binom{5}{2}\binom{15}{4}}{\binom{20}{6}} = 0{,}3521$$

## 7.8. Distribución de Poisson

**Definición 7.10** (Distribución de Poisson)**.** Sea $X$ una variable aleatoria. Se dice que $X$ sigue una distribución de Poisson si representa el número de de ocurrencias de un determinado suceso durante un periodo de tiempo fijo o una región fija del espacio.

Ha de cumplir las siguientes condiciones:

1. El número de ocurrencias en un intervalo o región específicas ha de ser independiente de las ocurrencias en otras zonas o intervalos de tiempo.

2. Si se considera un intervalo de tiempo muy pequeño (o una región muy pequeña):

   - La probabilidad de una ocurrencia es proporcional a la longitud del intervalo (el volumen de la región).
   - La probabilidad de dos o más ocurrencias es despreciable.

Dicho de otra manera, si de media se tiene que en determinado intervalo se producen $\lambda$ ocurrencias, tenemos que sigue una distribución de Poisson de parámetro $\lambda$, notado por:

$$X \rightsquigarrow \mathcal{P}(\lambda)$$

El razonamiento de su función masa de probabilidad no se explica en el presente documento, ya que se basa en la teoría de ecuaciones diferenciales. No obstante, demostramos que, efectivamente, es una función masa de probabilidad:

**Proposición 7.19.** *Sea $X \rightsquigarrow \mathcal{P}(\lambda)$. Entonces, su función masa de probabilidad es:*

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!}$$

*Demostración.* Como todos los términos son positivos, tenemos que $f(x) \geqslant 0$ para todo $x \in \mathbb{N} \cup \{0\}$. Comprobemos ahora que la suma es la unidad:

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} \overset{(*)}{=} e^{-\lambda}e^{\lambda} = 1$$

donde en $(*)$ he aplicado el desarrollo en serie de Taylor de la exponencial. □

**Proposición 7.20.** *Sea $X \rightsquigarrow \mathcal{P}(\lambda)$. Entonces, tenemos que:*

$$M_X(t) = e^{\lambda(e^t-1)}$$

*Demostración.* Tenemos que la función generatriz de momentos es:

$$M_X(t) = E[e^{tX}] = \sum_{x=0}^{\infty} e^{tx} f_X(x) = \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \lambda)^x}{x!} = e^{-\lambda} e^{e^t \lambda} = e^{\lambda(e^t - 1)}$$

□

**Proposición 7.21.** *Sea $X \rightsquigarrow \mathcal{P}(\lambda)$. Entonces, tenemos que:*

$$E[X] = \lambda$$

*Demostración.* Tenemos dos opciones:

**Opción 1** Partiendo de la definición de esperanza matemática:

$$E[X] = \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \cdot \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = e^{-\lambda} \cdot \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \overset{(*)}{=} e^{-\lambda} \lambda e^{\lambda} = \lambda$$

donde en $(*)$ he aplicado el desarrollo en serie de Taylor de la exponencial.

**Opción 2** Usando la función generatriz de momentos:

$$E[X] = M_X'(0) = \left. \frac{d}{dt} e^{\lambda(e^t - 1)} \right|_{t=0} = \left. \lambda e^t \cdot e^{\lambda(e^t - 1)} \right|_{t=0} = \lambda$$

□

**Lema 7.22.** *Sea $X \rightsquigarrow \mathcal{P}(\lambda)$. Entonces, tenemos que:*

$$E[X^2] = \lambda^2 + \lambda$$

*Demostración.* Igualmente, tenemos dos opciones:

**Opción 1** Partiendo de la definición de esperanza matemática:

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \cdot \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = e^{-\lambda} \cdot \lambda^2 \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \lambda^2 e^{\lambda} = \lambda^2$$

donde he usado la expresión del polinomio de Taylor de la exponencial. Por tanto,

$$E[X(X-1)] = E[X^2] - E[X] = \lambda^2 \Longrightarrow E[X^2] = \lambda^2 + \lambda$$

**Opción 2** Usando la función generatriz de momentos:

$$\begin{aligned} E[X^2] = M_X''(0) &= \left. \frac{d^2}{dt^2} e^{\lambda(e^t - 1)} \right|_{t=0} = \left. \frac{d}{dt} \lambda \cdot e^{\lambda(e^t - 1) + t} \right|_{t=0} = \\ &= \left. \lambda e^{\lambda(e^t - 1) + t} (\lambda e^t + 1) \right|_{t=0} = \lambda^2 + \lambda \end{aligned}$$

□

**Corolario 7.22.1.** *Sea $X \rightsquigarrow \mathcal{P}(\lambda)$. Entonces, tenemos que:*

$$\mathrm{Var}[X] = \lambda$$

*Demostración.* Tenemos que la varianza es:

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

□

**Proposición 7.23** (Aproximación de Poisson a la Binomial)**.** *Sea $X \rightsquigarrow B(n,p)$. Entonces, tenemos que, si $n$ es muy grande y $p \approx 0$, podemos aproximarla mediante una distribución de Poisson.*

*Definiendo $\lambda = np$, $X$ se aproxima a una distribución de Poisson $P(\lambda)$.*

*De forma empírica, se ha demostrado que esta aproximación es adecuada si $np \leqslant 5$.*

*Demostración.* Reescribimos la función de masa de probabilidad de la binomial como sigue:

$$\begin{aligned}P[X=x] &= \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!}\frac{(np)^x}{n^x}(1-p)^{n-x} = \\ &= \frac{(np)^x}{x!}\frac{n(n-1)\cdots(n-x+1)}{n^x}(1-p)^{n-x}\end{aligned}$$

Consideramos ahora $n \to \infty$ y $np \to \lambda$. Entonces, tenemos que:

- $\displaystyle\lim_{n\to\infty} \frac{n(n-1)\cdots(n-x+1)}{n^x} = 1$
- $\displaystyle\lim_{np\to\lambda} \frac{(np)^x}{x!} = \frac{\lambda^x}{x!}$
- Como $n \to \infty$ y $np \to \lambda$, tenemos que $p \to 0$, de donde:

  $$\lim_{\substack{n\to\infty\\ np\to\lambda}} (1-p)^{n-x} \overset{(*)}{=} e^{\lim\limits_{\substack{n\to\infty\\ np\to\lambda}} (n-x)(-p)} = e^{-\lambda}$$

  donde en $(*)$ hemos empleado la Fórmula de Moivre[4].

Por tanto, tenemos que:

$$\lim_{\substack{n\to\infty\\ np\to\lambda}} P[X=x] = \frac{\lambda^x}{x!}e^{-\lambda}$$

Por tanto, hemos demostrado que, si $n$ es muy grande y $p \approx 0$, la distribución binomial se aproxima a una distribución de Poisson. □

*Observación* (Cambio de Intervalo de Tiempo). Podremos cambiar el intervalo de tiempo al que está relacionada nuestra variable aleatoria de distribución de Poisson, alterando el valor de $\lambda$ y pasando a un nuevo valor $\lambda t$, donde $t$ es la cantidad de veces que se repite el nuevo intervalo en el antiguo.

[4]Concepto visto en Cálculo I para las indeterminaciones del tipo $1^\infty$.

**Ejemplo.** Un dispositivo falla 1,2 veces en 6 meses. Se pide calcular la probabilidad de que dicho dispositivo falle 2 veces en 6 meses y calcular la probabilidad de que dicho dispositivo falle menos de 5 veces en 2 años y medio.

Sea $X$ la variable aleatoria que determina el número de veces que falla en 6 meses. Tenemos que $X \rightsquigarrow \mathcal{P}(1{,}2)$.

Entonces, tenemos que la probabilidad de que dicho dispositivo falle 2 veces en 6 meses es:

$$P[X = 2] = 0{,}2169$$

Por tanto, si $X'$ es una variable aleatoria que determina el número de veces que un dispositivo falla en 2,5 años, tenemos que $X \rightsquigarrow \mathcal{P}(5 \cdot 1{,}2) = \mathcal{P}(6)$. Por tanto,

$$P[X' < 5] = P[X' \leqslant 4] = \sum_{x_i=0}^{4} f(x_i) = 0{,}0025+0{,}0149+0{,}0446+0{,}0892+0{,}1339 = 0{,}2851$$

## 7.9. Distribuciones Reproductivas

**Definición 7.11** (Reproductividad de Variables Aleatorias)**.** Una familia $X_1, X_2, \dots, X_n$ de variables aleatorias independientes con el mismo tipo de distribución se dice que es reproductiva si la variable aleatoria $X = \sum\limits_{i=1}^{n}$ sigue una distribución del mismo tipo.

**Proposición 7.24** (Distribución Reproductiva - Binomial)**.** *Sean $X_1, \dots, X_n$ variables aleatorias independientes y $X_i \rightsquigarrow B(k_i, p)$ para $i = 1, \dots, n$. Entonces:*

$$\sum_{i=1}^{n} X_i \rightsquigarrow B\left(\sum_{i=1}^{n} k_i, p\right)$$

*Demostración.* Sabemos que la función generatriz de momentos de cada una de las variables aleatorias $X_i$ es:

$$M_{X_i}(t) = (1 - p + pe^t)^{k_i} \qquad i = 1, \dots, n$$

Usando el Teorema 6.21, tenemos que:

$$M_{\sum\limits_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} (1 - p + pe^t)^{k_i} = (1 - p + pe^t)^{\sum\limits_{i=1}^{n} k_i}$$

Esta es la función generatriz de momentos de una variable aleatoria que sigue una distribución binomial con parámetros $\left(\sum\limits_{i=1}^{n} k_i, p\right)$. Por tanto, hemos demostrado que:

$$\sum_{i=1}^{n} X_i \rightsquigarrow B\left(\sum_{i=1}^{n} k_i, p\right)$$

$\square$

**Proposición 7.25** (Distribución Reproductiva - Poisson)**.** *Sean $X_1, \dots, X_n$ variables aleatorias independientes y $X_i \rightsquigarrow \mathcal{P}(\lambda_i)$ para $i = 1, \dots, n$. Entonces:*

$$\sum_{i=1}^{n} X_i \rightsquigarrow \mathcal{P}\left(\sum_{i=1}^{n} \lambda_i\right)$$

*Demostración.* Sabemos que la función generatriz de momentos de cada una de las variables aleatorias $X_i$ es:

$$M_{X_i}(t) = e^{\lambda_i(e^t-1)} \qquad i = 1, \dots, n$$

Usando el Teorema 6.21, tenemos que:

$$M_{\sum\limits_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} e^{\lambda_i(e^t-1)} = e^{\left(\sum\limits_{i=1}^{n} \lambda_i\right)(e^t-1)}$$

Esta es la función generatriz de momentos de una variable aleatoria que sigue una distribución de Poisson con parámetro $\left(\sum\limits_{i=1}^{n} \lambda_i\right)$. Por tanto, hemos demostrado que:

$$\sum_{i=1}^{n} X_i \rightsquigarrow \mathcal{P}\left(\sum_{i=1}^{n} \lambda_i\right)$$

□

**Proposición 7.26** (Distribución Reproductiva - Binomial Negativa)**.** *Sean $X_1, \dots, X_n$ variables aleatorias independientes y $X_i \rightsquigarrow BN(k_i, p)$ para $i = 1, \dots, n$. Entonces:*

$$\sum_{i=1}^{n} X_i \rightsquigarrow BN\left(\sum_{i=1}^{n} k_i, p\right)$$

*Demostración.* Sabemos que la función generatriz de momentos de cada una de las variables aleatorias $X_i$ es:

$$M_{X_i}(t) = \left(\frac{p}{1-(1-p)e^t}\right)^{k_i} \qquad i = 1, \dots, n$$

Usando el Teorema 6.21, tenemos que:

$$M_{\sum\limits_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} \left(\frac{p}{1-(1-p)e^t}\right)^{k_i} = \left(\frac{p}{1-(1-p)e^t}\right)^{\sum\limits_{i=1}^{n} k_i}$$

Esta es la función generatriz de momentos de una variable aleatoria que sigue una distribución binomial negativa con parámetros $\left(\sum\limits_{i=1}^{n} k_i, p\right)$. Por tanto, hemos demostrado que:

$$\sum_{i=1}^{n} X_i \rightsquigarrow BN\left(\sum_{i=1}^{n} k_i, p\right)$$

□

# 8. Relaciones de ejercicios

## 8.1. Variables Estadísticas Unidimensionales

**Ejercicio 8.1.1.** El número de hijos de las familias de una determinada barriada de una ciudad es una variable estadística de la que se conocen los siguientes datos:

| $x_i$ | $n_i$ | $N_i$ | $f_i$ |
|---|---|---|---|
| 0 | 80 | | 0,16 |
| 1 | 110 | | |
| 2 | | 320 | |
| 3 | | | 0,18 |
| 4 | 40 | | |
| 5 | | | |
| 6 | 20 | | |

Sea el número de hijos de una familia, $X$, una variable estadística con población $n = 500$ y modalidades $x_1, \ldots, x_6$ con *distribución de frecuencias*:

$$\{x_i, n_i\}_{i=1,\ldots,7}$$

1. Completar la tabla de frecuencias.

| $x_i$ | $n_i$ | $N_i$ | $f_i$ |
|---|---|---|---|
| 0 | 80 | 80 | 0,16 |
| 1 | 110 | 190 | 0,22 |
| 2 | 130 | 320 | 0,26 |
| 3 | 90 | 410 | 0,18 |
| 4 | 40 | 450 | 0,08 |
| 5 | 30 | 480 | 0,06 |
| 6 | 20 | 500 | 0,04 |

2. Representar la distribución mediante un diagrama de barras y la curva de distribución.

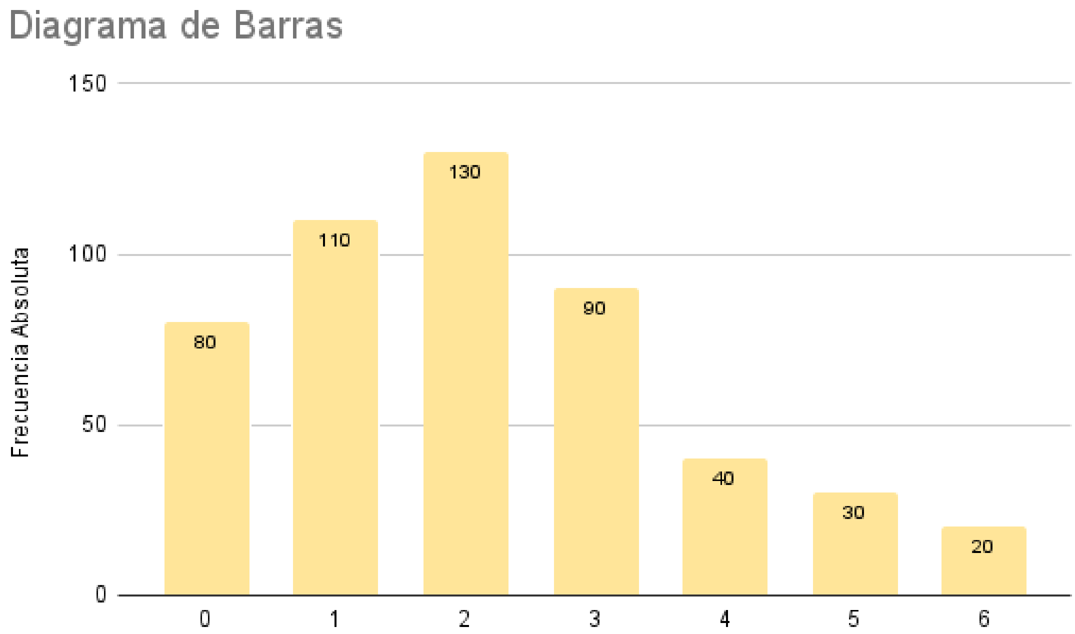


Figura 8.1: Diagrama de Barras

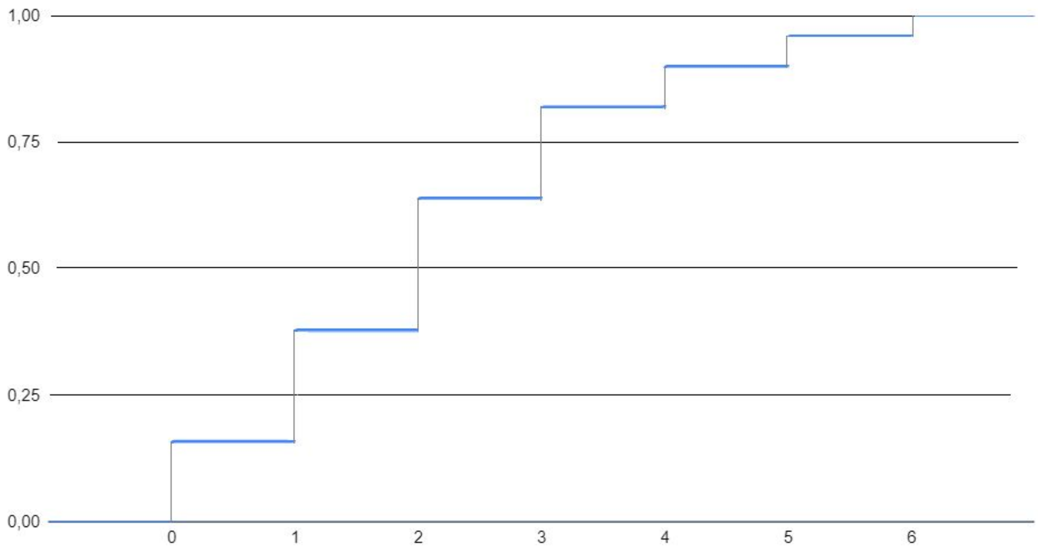


Figura 8.2: Curva de distribución

3. Promediar los valores de la variable mediante diferentes medidas. Interpretarlas.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{7} n_i x_i = \frac{1070}{500} = 2{,}14 \qquad Me = 2 \qquad Mo = 2$$

La media aritmética nos informa de que en la población estudiada la media de hijos es aproximadamente 2.

La moda nos informa de que el valor más usual, es decir, el número de hijos más repetidos, es 2.

La mediana nos indica que más de la mitad de los encuestados tienen 2 o menos hijos.

**Ejercicio 8.1.2.** La puntuación obtenida por 50 personas que se presentaron a una prueba de selección, sumadas las puntuaciones de los distintos tests, fueron:

174, 185, 166, 176, 145, 166, 191, 175, 158, 156, 156, 187, 162, 172, 197, 181, 151, 161, 183, 172, 162, 147, 178, 176, 141, 170, 171, 158, 184, 173, 169, 162, 172, 181, 187, 177, 164, 171, 193, 183, 173, 179, 188, 179, 167, 178, 180, 168, 148, 173.

1. Agrupar los datos en intervalos de amplitud 5 desde 140 a 200 y dar la tabla de frecuencias.
   Sea la puntuación obtenida en una prueba de selección, $X$, una variable estadística continua con población $n = 50$ e intervalos $I_1, \ldots, I_{12}$ con *distribución de frecuencias*:

$$\{I_i, n_i\}_{i=1,\ldots,12}$$

| $I_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $a_i$ | $h_i$ | $c_i$ |
|---|---|---|---|---|---|---|---|
| $[140, 145]$ | 2 | 2 | 0,04 | 0,04 | 5 | 0,4 | 142,5 |
| $(145, 150]$ | 2 | 4 | 0,04 | 0,08 | 5 | 0,4 | 147,5 |
| $(150, 155]$ | 1 | 5 | 0,02 | 0,1 | 5 | 0,2 | 152,5 |
| $(155, 160]$ | 4 | 9 | 0,08 | 0,18 | 5 | 0,8 | 157,5 |
| $(160, 165]$ | 5 | 14 | 0,1 | 0,28 | 5 | 1 | 162,5 |
| $(165, 170]$ | 6 | 20 | 0,12 | 0,4 | 5 | 1,2 | 167,5 |
| $(170, 175]$ | 10 | 30 | 0,2 | 0,6 | 5 | 2 | 172,5 |
| $(175, 180]$ | 8 | 38 | 0,16 | 0,76 | 5 | 1,6 | 177,5 |
| $(180, 185]$ | 6 | 44 | 0,12 | 0,88 | 5 | 1,2 | 182,5 |
| $(185, 190]$ | 3 | 47 | 0,06 | 0,94 | 5 | 0,6 | 187,5 |
| $(190, 195]$ | 2 | 49 | 0,04 | 0,98 | 5 | 0,4 | 192,5 |
| $(195, 200]$ | 1 | 50 | 0,02 | 1 | 5 | 0,2 | 197,5 |

2. Representar la distribución mediante un histograma, poligonal de frecuencias y curva de distribución.

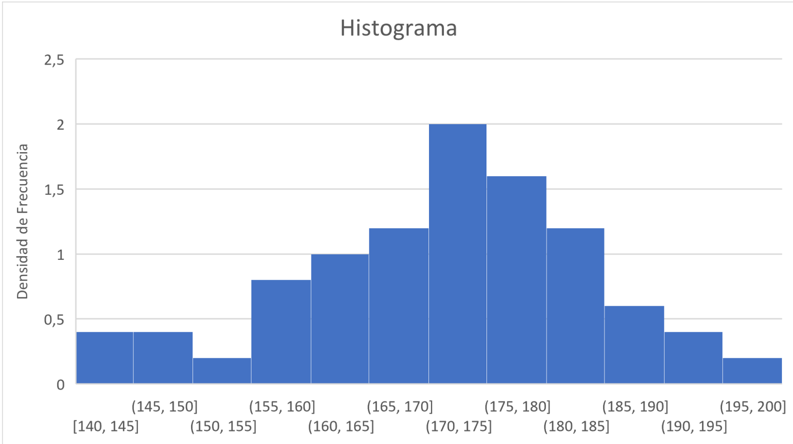


Figura 8.3: Histograma

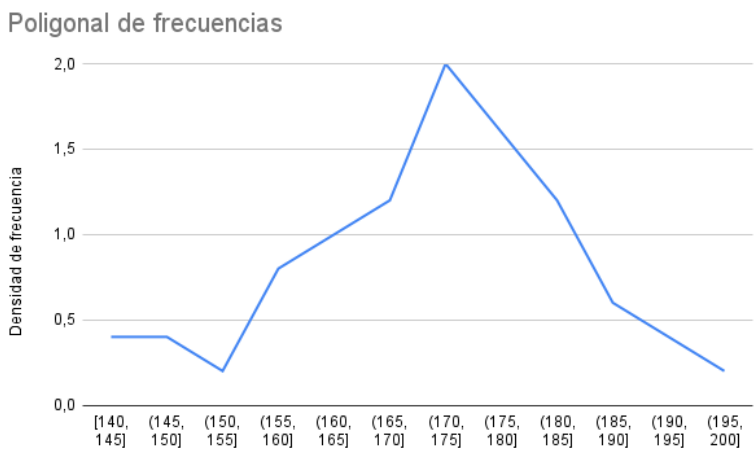


Figura 8.4: Poligonal de frecuencias

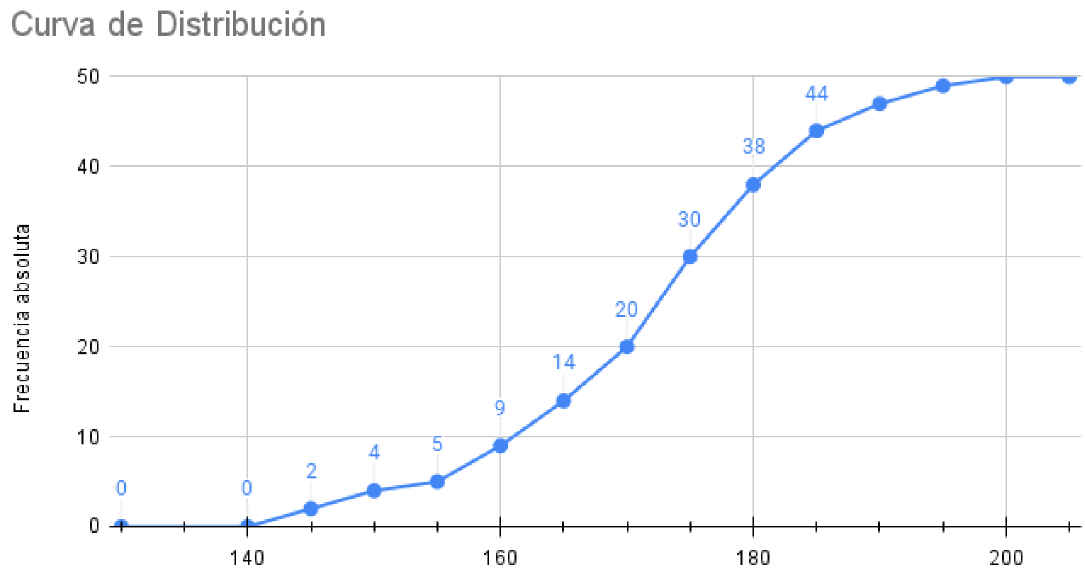


Figura 8.5: Curva de distribución

**Ejercicio 8.1.3.** La distribución de la renta familiar en el año 2003 por comunidades autónomas se recoge en la siguiente tabla:

| $I_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $c_i$ | $a_i$ | $h_i$ |
|---|---|---|---|---|---|---|---|
| (8300, 9300] | 2 | | | | | | |
| , 10200] | | 5 | | | | | |
| | | | | 10/18 | | | |
| | | | 2/18 | | 12000 | | |
| | 4 | | | | | | 0,005/18 |
| | | 18 | | | | | 0,002/18 |

1. Completar la tabla. Sea la renta familiar en el año 2003, $X$, una variable estadística continua con población $n = 18$ e intervalos $I_1, \dots, I_6$ con *distribución de frecuencias*:

$$\{I_i, n_i\}_{i=1,\dots,6}$$

| $I_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $c_i$ | $a_i$ | $h_i$ |
|---|---|---|---|---|---|---|---|
| (8300, 9300] | 2 | 2 | 2/18 | 2/18 | 8800 | 1000 | 0,002/18 |
| (9300, 10200] | 3 | 5 | 3/18 | 5/18 | 9750 | 900 | $0{,}00\overline{3}/18$ |
| (10200, 11300] | 5 | 10 | 5/18 | 10/18 | 10750 | 1100 | $0{,}00\overline{45}/18$ |
| (11300, 12700] | 2 | 12 | 2/18 | 12/18 | 12000 | 1400 | 0,001428/18 |
| (12700, 13500] | 4 | 16 | 4/18 | 16/18 | 13100 | 800 | 0,005/18 |
| (13500, 14500] | 2 | 18 | 2/18 | 1 | 14000 | 1000 | 0,002/18 |

2. Representar la distribución mediante un histograma, poligonal de frecuencias y curva de distribución.

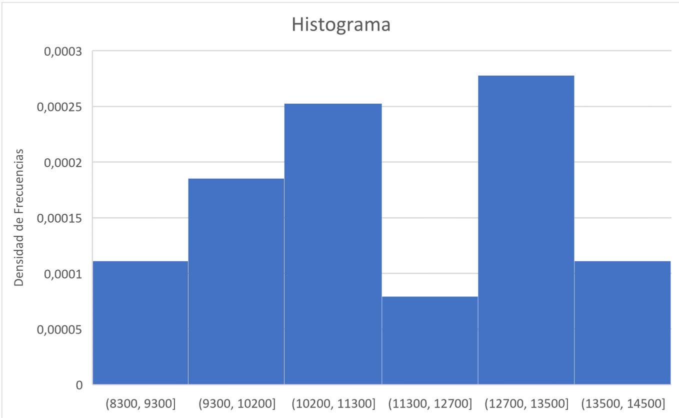


Figura 8.6: Histograma

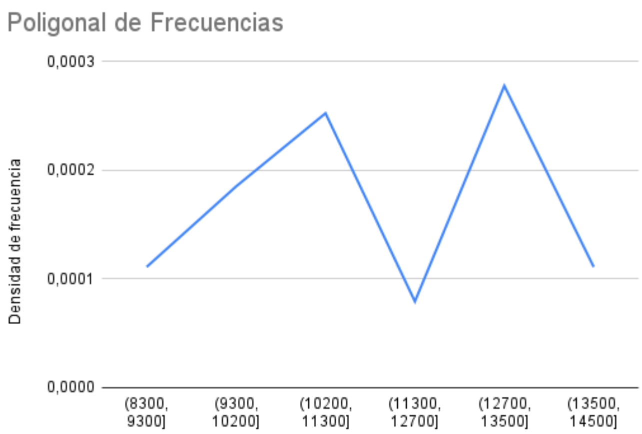


Figura 8.7: Poligonal de Frecuencias

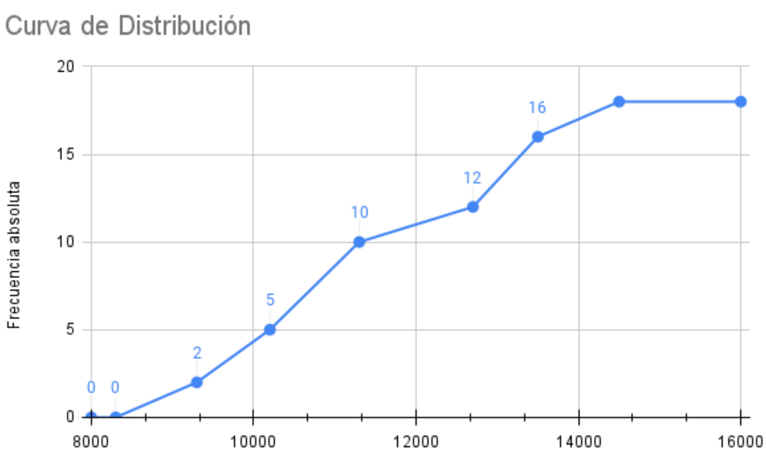


Figura 8.8: Curva de Distribución

3. ¿Cuántas comunidades presentan una renta menor o igual que 12700 euros? ¿Y cuántas superior a 11300 euros?

   Observando la tabla, es fácil ver que hay 12 comunidades autónomas cuya renta familiar es menor que 12700 euros. Esto se debe a que $N_4 = 12$, y $e_4 = 12700$.

   Además, se puede ver que hay 8 comunidades con renta mayor que 11300 euros. Esto se debe a que $e_3 = 11300$ y que $n - N_3 = 18 - 10 = 8$.

**Ejercicio 8.1.4.** En una determinada empresa se realiza un estudio sobre la calidad de su producción. La distribución siguiente informa sobre el número de piezas defectuosas encontradas en 100 cajas examinadas con 50 unidades cada una de ellas:

| $x_i$ | N$^o$ piezas defectuosas | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_i$ | N$^o$ de cajas | 6 | 9 | 10 | 11 | 14 | 16 | 16 | 9 | 4 | 3 | 2 |

1. Calcular el número medio de piezas defectuosas por caja.
   Sea el número de piezas defectuosas por caja, $X$, una variable estadística con población $n = 100$ y modalidades $x_1, \ldots, x_11$ con *distribución de frecuencias*:

$$\{x_i, n_i\}_{i=1,\ldots,11}$$

   La media aritmética es:

$$\bar{x} = \frac{\sum_{i=1}^{11} n_i x_i}{\sum_{i=1}^{11} n_i} = \frac{436}{100} = 4{,}36 \text{ piezas defectuosas por caja.}$$

2. ¿Cuántas piezas defectuosas se encuentran más frecuentemente en las cajas examinadas?
   Hay dos valores modales: $Mo = \{5, 6\}$, con una frecuencia absoluta $f_i = 16$.

3. ¿Cuál es el número mediano de piezas defectuosas por caja?

| $x_i$ | N$^o$ piezas defectuosas | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_i$ | N$^o$ de cajas | 6 | 9 | 10 | 11 | 14 | 16 | 16 | 9 | 4 | 3 | 2 |
| $N_i$ | | 6 | 15 | 25 | 36 | 50 | 66 | 82 | 91 | 95 | 98 | 100 |

   El valor mediano es $\frac{n}{2} = \frac{100}{2} = 50$.

   La mediana, como $N_5 = 50$, es:

$$Me = \frac{x_5 + x_6}{2} = \frac{4+5}{2} = 4{,}5$$

4. Calcular los cuartiles de la distribución. Interpretarlos.

   - $\underline{Q_1 = 2{,}5}$: Esto significa que el 25 % de la población, es decir, el 25 % de las cajas de la muestra, tienen menos de 2,5 piezas defectuosas.
   - $\underline{Q_2 = Me = 4{,}5}$: Esto significa que el 50 % de la población, es decir, la mitad de las cajas de la muestra, tienen menos de 4,5 piezas defectuosas.

- $\underline{Q_3 = 6}$: Esto significa que el 75 % de la población, es decir, el 75 % de las cajas de la muestra, tienen menos de 6 piezas defectuosas.

5. Calcular los deciles de orden 3 y 7. Interpretarlos.

   - $\underline{D_3 = P_{30} = 3}$: Esto significa que el 30 % de la población, es decir, el 30 % de las cajas de la muestra, tienen menos de 3 piezas defectuosas.
   - $\underline{D_7 = P_{70} = 6}$: Esto significa que el 70 % de la población, es decir, el 70 % de las cajas de la muestra, tienen menos de 6 piezas defectuosas.

6. Cuantificar la dispersión de la distribución utilizando diferentes medidas, interpretando los resultados y señalando las ventajas e inconvenientes de cada una.

   - Medidas de dispersión absolutas:
     - $R = x_{11} - x_1 = 10$
       El rango representa la diferencia entre el máximo de las modalidades y entre el último.
     - $R_I = Q_3 - Q_1 = 3{,}5$
       El rango intercuartílico es la longitud del intervalo en el que están el 50 % central de los datos.
     - $D_{\bar{x}} = \frac{\sum_{i=1}^{11} |x_i - \bar{x}| n_i}{n} = 2$ Es la desviación absoluta media respecto a la media aritmética.
     - $D_{Me} = \frac{\sum_{i=1}^{11} |x_i - Me| n_i}{n} = 2$
       Es la desviación absoluta media respecto a la mediana.
     - $\sigma^2 = \frac{\sum_{i=1}^{11} n_i (x_i - \bar{x})^2}{n} = 5{,}8704$
       La varianza es la media cuadrática dispersión óptima. Como no tenemos datos muy por encima ni muy por debajo es representativa.
     - $\sigma = +\sqrt{\sigma^2} = 2{,}42289$
       La desviación típica es el promedio de las desviaciones individuales con respecto a la media.
   - Medidas de dispersión relativas:
     - $C_A = \frac{x_k}{x_1} = \frac{10}{0} = \infty \Longrightarrow$ El coeficiente de apertura es una medida no representativa en este caso.
     - $R_R = \frac{R}{\bar{x}} = \frac{10}{4{,}36} = 2{,}2936$
       El recorrido relativo es el número de veces que el recorrido contiene a la media aritmética.
     - $R_{SI} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{3{,}5}{8{,}5} = 0{,}4118$
       El recorrido semi-intercuartílico es el 50 % de los valores centrales dividido entre los restantes.
     - $C.V.(X) = \frac{\sigma_X}{|\bar{x}|} = 0{,}5557$
       El coeficiente de variación mide la variación de los datos respecto a la media. Se usa para saber si una medida e homogénea.
     - $V_{Me} = \frac{D_{Me}}{Me} = \frac{2}{4{,}5} = \frac{4}{9} = 0.\bar{4}$
       Mide el número de veces que la mediana está contenida en la desviación respecto a la mediana. Hay una buena representatividad de la mediana.

**Ejercicio 8.1.5.** Dadas las siguientes distribuciones:

| $I_i^{(1)}$ | $(0,1]$ | $(1,2]$ | $(2,3]$ | $(3,4]$ | $(4,5]$ |
|---|---|---|---|---|---|
| $n_i^{(1)}$ | 12 | 13 | 11 | 8 | 6 |

| $I_i^{(2)}$ | $(0,1]$ | $(1,3]$ | $(3,6]$ | $(6,10]$ | $(10,12]$ |
|---|---|---|---|---|---|
| $n_i^{(2)}$ | 1 | 6 | 7 | 12 | 2 |

Calcular para cada una de ellas:

1. Medias aritmética, armónica y geométrica.
   Sea $X_1$ una variable estadística continua con población $n = 50$ e intervalos $I_1^1, \dots, I_5^1$ con *distribución de frecuencias*:
   $$\{I_i^1, n_i^1\}_{i=1,\dots,5}$$
   Y sea $X_2$ una variable estadística continua con población $n = 28$ e intervalos $I_1^2, \dots, I_5^2$ con *distribución de frecuencias*:
   $$\{I_i^2, n_i^2\}_{i=1,\dots,5}$$

   - <u>Media aritmética:</u>
   $$\bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{5} c_i^1 n_i^1 = \frac{1}{50}\cdot 108 = 2{,}16$$
   $$\bar{x}_2 = \frac{1}{n_2}\sum_{i=1}^{5} c_i^2 n_i^2 = \frac{1}{28}\cdot 162 = 5{,}7857$$
   - <u>Media armónica:</u>
   $$H_1 = \frac{n_1}{\sum_{i=1}^{6}\frac{n_i^1}{c_i^1}} = \frac{50}{40{,}6857} = 1{,}2289$$
   $$H_2 = \frac{n_2}{\sum_{i=1}^{6}\frac{n_i^2}{c_i^2}} = \frac{28}{8{,}237} = 3{,}3991$$
   - <u>Media geométrica:</u>
   $$G_1 = \sqrt[n_1]{\prod_{i=1}^{6}(c_i^1)^{n_i^1}} = \sqrt[50]{2{,}118\cdot 10^{11}} = 1{,}6847$$
   $$G_2 = \sqrt[n_2]{\prod_{i=1}^{6}(c_i^2)^{n_i^2}} = \sqrt[28]{9{,}94\cdot 10^{18}} = 4{,}7696$$

2. El valor más frecuente.

| $I_i^{(1)}$ | $(0,1]$ | $(1,2]$ | $(2,3]$ | $(3,4]$ | $(4,5]$ |
|---|---|---|---|---|---|
| $n_i^{(1)}$ | 12 | 13 | 11 | 8 | 6 |
| $a_i^{(1)}$ | 1 | 1 | 1 | 1 | 1 |
| $h_i^{(1)}$ | 12 | 13 | 11 | 8 | 6 |

| $I_i^{(2)}$ | $(0,1]$ | $(1,3]$ | $(3,6]$ | $(6,10]$ | $(10,12]$ |
|---|---|---|---|---|---|
| $n_i^{(2)}$ | 1 | 6 | 7 | 12 | 2 |
| $a_i^{(2)}$ | 1 | 2 | 3 | 4 | 2 |
| $h_i^{(2)}$ | 1 | 3 | $\frac{7}{3}$ | 3 | 1 |

$$Mo^1 = \frac{h_i(e_i + e_{i-1}) - e_i h_{i-1} - e_{i-1}h_{i+1}}{2h_i - h_{i+1} - h_{i-1}} = \frac{4}{3} = 1.\bar{3}$$

En el caso de $X_2$, hay dos intervalos modales. Por tanto, habrá dos modas.

$$Mo_1^2 = \frac{h_i(e_i + e_{i-1}) - e_i h_{i-1} - e_{i-1}h_{i+1}}{2h_i - h_{i+1} - h_{i-1}} = 2{,}5$$

$$Mo_2^2 = \frac{h_i(e_i + e_{i-1}) - e_i h_{i-1} - e_{i-1}h_{i+1}}{2h_i - h_{i+1} - h_{i-1}} = 7$$

3. El valor superado por el 50 % de las observaciones.

| $I_i^{(1)}$ | $(0,1]$ | $(1,2]$ | $(2,3]$ | $(3,4]$ | $(4,5]$ |
|---|---|---|---|---|---|
| $n_i^{(1)}$ | 12 | 13 | 11 | 8 | 6 |
| $N_i^{(1)}$ | 12 | 25 | 36 | 44 | 50 |

| $I_i^{(2)}$ | $(0,1]$ | $(1,3]$ | $(3,6]$ | $(6,10]$ | $(10,12]$ |
|---|---|---|---|---|---|
| $n_i^{(2)}$ | 1 | 6 | 7 | 12 | 2 |
| $N_i^{(2)}$ | 1 | 7 | 14 | 26 | 28 |

Para la variable $X_1$, $N_2 = 25 = \frac{n_1}{2} \Longrightarrow Me_1 = e_2 = 2$.

Para la variable $X_2$, $N_3 = 14 = \frac{n_2}{2} \Longrightarrow Me_2 = e_3 = 6$.

4. Recorrido, recorrido intercuartílico y desviación típica. Interpretarlos. ¿Qué distribución es más homogénea?

- Recorrido

  $$R_1 = 5 - 0 = 5 \qquad\qquad R_2 = 12 - 0 = 12$$

  Es la diferencia entre el máximo de las modalidades y entre el último. Como podemos ver, en la primera distribución la totalidad de los datos se encuentra en un intervalo menor.

- Rango Intercuartílico

  Es la longitud del intervalo en el que están el 50 % central de los datos. En primer lugar, tengo que hallar $Q_1$ y $Q_3$ para $X_1$.

  $$Q_1^1 = e_{i-1} + \frac{50 \cdot \frac{1}{4} - N_{i-1}}{n_i} \cdot a_i = 1 + \frac{12{,}5 - 12}{13} \cdot 1 = \frac{27}{26} = 1{,}038$$

  $$Q_3^1 = e_{i-1} + \frac{50 \cdot \frac{3}{4} - N_{i-1}}{n_i} \cdot a_i = 3 + \frac{37{,}5 - 36}{8} \cdot 1 = \frac{51}{16} = 3{,}1875$$

  Por tanto, $R_I^1 = Q_3^1 - Q_1^1 = 2{,}1495$.

En segundo lugar, tengo que hallar $Q_1$ y $Q_3$ para $X_2$.

$$Q_1^2 = e_2 = 3, \text{ ya que } \frac{n}{4} = 7 = N_2$$

$$Q_3^1 = e_{i-1} + \frac{28 \cdot \frac{3}{4} - N_{i-1}}{n_i} \cdot a_i = 6 + \frac{21 - 14}{12} \cdot 4 = \frac{25}{3} = 8.\bar{3}$$

Por tanto, $R_I^2 = Q_3^2 - Q_1^2 = 5.\bar{3}$.

- Desviación Típica:
  Promedio de las desviaciones individuales con respecto a la media.

$$\sigma_{x_1} = +\sqrt{\sigma_{x_1}^2} = +\sqrt{\frac{1}{50}\sum_{i=1}^{5} n_i c_i^2 - \bar{x}_1^2} = +\sqrt{\frac{1}{50}\sum_{i=1}^{5} n_i c_i^2 - (2{,}16)^2} = 1{,}3208$$

$$\sigma_{x_2} = +\sqrt{\sigma_{x_2}^2} = +\sqrt{\frac{1}{28}\sum_{i=1}^{5} n_i c_i^2 - \bar{x}_2^2} = +\sqrt{\frac{1}{28}\sum_{i=1}^{5} n_i c_i^2 - (5{,}7857)^2} = 2{,}92$$

- Coeficiente de Variación de Pearson:

$$C.V.(X_1) = \frac{\sigma_{x_1}}{|\bar{x}_1|} = 0{,}61148$$

$$C.V.(X_2) = \frac{\sigma_{x_2}}{|\bar{x}_2|} = 0{,}5047$$

  Por tanto, como $C.V.(X_2) < C.V.(X_1)$, la distribución asociada a la variable $X_2$ es más homogénea.

**Ejercicio 8.1.6.** Un móvil efectúa un recorrido de 100 km en dos sentidos. En uno va a una velocidad constante de $V_1 = 60$ km/h y en el otro va a una velocidad constante de $V_2 = 70$ km/h. Calcular la velocidad media del recorrido.

Al ser un cociente entre magnitudes, ya que la velocidad es el cociente entre el espacio y el tiempo, se opta por la media armónica.

La media armónica es es:

$$H = \frac{2}{\frac{1}{60} + \frac{1}{70}} = 64{,}62 \ km/h$$

**Ejercicio 8.1.7.** Las acciones de una empresa han producido los siguientes rendimientos netos anuales:

| Año | Rentabilidad |
|---|---|
| 1994 | 12 % |
| 1995 | 10 % |
| 1996 | 7 % |
| 1997 | 6 % |
| 1998 | 5 % |

Obtener el rendimiento neto medio en esos cinco años.

Como la rentabilidad tiene efectos multiplicativos acumulativos en la evolución del valor de la empresa a partir de una cantidad inicial fija, que es el valor inicial en 1994, se opta por la media geométrica.

La media geométrica es:

$$G = \sqrt[5]{1{,}12 \cdot 1{,}10 \cdot 1{,}07 \cdot 1{,}06 \cdot 5} = 1{,}7968$$

Es decir, se obtiene un rendimiento neto de un 7.968 %.

**Ejercicio 8.1.8.** Un profesor califica a sus alumnos según el criterio siguiente: 40 % de suspensos, 30 % de aprobados, 15 % notables, 10 % sobresalientes y 5 % de matrículas. Las notas obtenidas son las siguientes:

| $(0,1]$ | $(1,2]$ | $(2,3]$ | $(3,4]$ | $(4,5]$ | $(5,6]$ | $(6,7]$ | $(7,8]$ | $(8,9]$ | $(9,10]$ |
|---|---|---|---|---|---|---|---|---|---|
| 34 | 74 | 56 | 81 | 94 | 70 | 41 | 28 | 16 | 4 |

Calcular las notas máximas para obtener cada una de las calificaciones.

Sean las notas de los alumnos de ese profesor, $X$, una variable estadística continua con población $n = 498$ e intervalos $I_1, \ldots, I_10$ con *distribución de frecuencias*:

$$\{I_i, n_i\}_{i=1,\ldots,10}$$

| $I_i$ | $(0,1]$ | $(1,2]$ | $(2,3]$ | $(3,4]$ | $(4,5]$ | $(5,6]$ | $(6,7]$ | $(7,8]$ | $(8,9]$ | $(9,10]$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $n_i$ | 34 | 74 | 56 | 81 | 94 | 70 | 41 | 28 | 16 | 4 |
| $N_i$ | 34 | 108 | 164 | 245 | 339 | 409 | 450 | 478 | 494 | 498 |

Se están pidiendo los percentiles $P_{40}$, $P_{70}$, $P_{85}$, y $P_{95}$.

$$P_{40} = e_{i-1} + \frac{\frac{40n}{100} - N_{i-1}}{n_i} \cdot a_i = 3 + \frac{199{,}2 - 164}{81} \cdot 1 = 3{,}4346$$

$$P_{70} = e_{i-1} + \frac{\frac{70n}{100} - N_{i-1}}{n_i} \cdot a_i = 5 + \frac{348{,}6 - 339}{70} \cdot 1 = 5{,}1371$$

$$P_{85} = e_{i-1} + \frac{\frac{85n}{100} - N_{i-1}}{n_i} \cdot a_i = 6 + \frac{423{,}3 - 409}{41} \cdot 1 = 6{,}3488$$

$$P_{95} = e_{i-1} + \frac{\frac{95n}{100} - N_{i-1}}{n_i} \cdot a_i = 7 + \frac{473{,}1 - 450}{28} \cdot 1 = 7{,}825$$

Por tanto, las notas máximas para cada calificación son:

- Suspensos: $P_{40} = 3{,}4346$.
- Aprobados: $P_{70} = 5{,}1371$.
- Notables: $P_{85} = 6{,}3488$.
- Sobresalientes: $P_{90} = 7{,}825$.

**Ejercicio 8.1.9.** Se ha medido la altura de 110 jóvenes, obteniendo:

| Altura | (1,55, 1,60] | (1,60, 1,70] | (1,70, 1,80] | (1,80, 1,90] | (1,90, 2,00] |
|---|---|---|---|---|---|
| N$^o$ jóvenes | 18 | 31 | 24 | 20 | 17 |

1. Si se consideran bajos el 3 % de los individuos de menor altura, ¿cuál es la altura máxima que pueden alcanzar?

   Sea la altura de los jóvenes, $X$, una variable estadística continua con población $n = 110$ e intervalos $I_1, \dots, I_5$ con *distribución de frecuencias*:

   $$\{I_i, n_i\}_{i=1,\dots,5}$$

   | $I_i$ | (1,55, 1,60] | (1,60, 1,70] | (1,70, 1,80] | (1,80, 1,90] | (1,90, 2,00] |
   |---|---|---|---|---|---|
   | $n_i$ | 18 | 31 | 24 | 20 | 17 |
   | $N_i$ | 18 | 49 | 73 | 93 | 110 |

   Se pide calcular $P_3$.

   $$P_3 = e_{i-1} + \frac{0{,}03n - N_{i-1}}{n_i} \cdot a_i = 1{,}55 + \frac{3{,}3 - 0}{18} \cdot 0{,}05 = 1{,}5592$$

   Por tanto, la altura máxima de los bajos sería 1,5592 $m$.

2. Si se consideran altos el 18 % de los individuos de mayor altura, ¿cuál es su altura mínima?
   $100 - 18 = 82\,\% \Longrightarrow$ Se pide calcular el $P_{82}$:

   $$P_{82} = e_{i-1} + \frac{0{,}80n - N_{i-1}}{n_i} \cdot a_i = 1{,}8 + \frac{90{,}2 - 73}{20} \cdot 0{,}1 = 1{,}886$$

   Por tanto, la altura mínima de los altos sería 1,886 $m$.

3. ¿Qué altura es superada sólo por 1/4 de los jóvenes?
   Se pide calcular el $Q_3$:

   $$Q_3 = e_{i-1} + \frac{0{,}75n - N_{i-1}}{n_i} \cdot a_i = 1{,}8 + \frac{82{,}5 - 73}{20} \cdot 0{,}1 = 1{,}84$$

   Por tanto, la altura superada sólo por 1/4 de los jóvenes sería 1,84 $m$.

4. Calcular el número de jóvenes cuya altura es superior a 1,75.

   $$P_\alpha = 1{,}70 + \frac{\frac{\alpha 110}{100} - 49}{24} \cdot 0{,}1 = 1{,}75 \Longrightarrow \frac{\frac{\alpha 110}{100} - 49}{24} = 0{,}5 \Longrightarrow \alpha = 55.\overline{45}\,\%$$

   Por tanto, el número de jóvenes con altura superior a $P_\alpha = 1{,}75\ m$ es:

   $$n - \frac{\alpha n}{100} = 49 \text{ personas}$$

5. Calcular la altura máxima de los 11 jóvenes más bajos.
$\frac{\alpha n}{100} = 11 \Longrightarrow 10\,\%$. Por tanto, se pide $P_{10}$:

$$P_{10} = 1{,}55 + \frac{11-0}{18} \cdot 0{,}05 = 1{,}5806\ m$$

Por tanto, la altura máxima de los 11 jóvenes más bajos es 1,5806 $m$.

6. Calcular la altura mínima de los 11 jóvenes más altos.
$\frac{\alpha n}{100} = 110 - 11 = 99 \Longrightarrow 10\,\%$. Por tanto, se pide $P_{90}$:

$$P_{90} = 1{,}90 + \frac{99-93}{17} \cdot 0{,}1 = 1{,}9353\ m$$

Por tanto, la altura mínima de los 11 jóvenes más altos es 1,9353 $m$.

**Ejercicio 8.1.10.** Realizando una prueba para el estudio del cáncer a 150 personas se obtuvo la siguiente tabla según la edad de los enfermos:

| Edad | $(10,30]$ | $(30,40]$ | $(40,50]$ | $(50,60]$ | $(60,90]$ |
|---|---|---|---|---|---|
| N° enfermos | 15 | 22 | 48 | 40 | 25 |

1. Calcular la edad más común de los individuos estudiados.

Sea el número de personas con cáncer, $X$, una variable estadística continua con población $n = 150$ e intervalos $I_1, \dots, I_5$ con *distribución de frecuencias*:

$$\{I_i, n_i\}_{i=1,\dots,5}$$

| $I_i$ | $(10,30]$ | $(30,40]$ | $(40,50]$ | $(50,60]$ | $(60,90]$ |
|---|---|---|---|---|---|
| $n_i$ | 15 | 22 | 48 | 40 | 25 |
| $N_i$ | 15 | 37 | 85 | 125 | 150 |
| $a_i$ | 20 | 10 | 10 | 10 | 30 |
| $h_i$ | 0,75 | 2,2 | 4,8 | 4 | $0{,}8\overline{3}$ |

Por tanto, el intervalo modal es $I_3$.

$$Mo = \frac{h_i(e_i + e_{i-1}) - e_i h_{i-1} - e_{i-1} h_{i+1}}{2h_i - h_{i+1} - h_{i-1}} = \frac{4{,}8(90) - 50 \cdot 2{,}2 - 40 \cdot 4}{2 \cdot 4{,}8 - 4 - 2{,}2} = \frac{810}{17} = 47{,}647$$

Por tanto, la edad más común es $Mo = 47{,}647$ años.

2. Calcular la edad mínima y máxima del 30 % central de los individuos.
Se pide hallar $P_{35}$ y $P_{65}$:

$$P_{35} = 40 + \frac{52{,}5 - 37}{48} \cdot 10 = 43{,}2292$$

$$P_{65} = 50 + \frac{97{,}5 - 85}{40} \cdot 10 = 53{,}125$$

Por tanto, las edades mínimas y máximas del 30 % central de los individuos son 43,2292 y 53,125 años respectivamente.

3. Calcular el recorrido intercuartílico y la desviación típica.
   Para calcular $R_I$, calculo primero $Q_1$ y $Q_3$:

$$Q_1 = 40 + \frac{37{,}5 - 37}{48} \cdot 10 = 40{,}1$$

$$Q_3 = 50 + \frac{112{,}5 - 85}{40} \cdot 10 = 56{,}875$$

   Por tanto, $R_I = Q_3 - Q_1 = 16{,}775$. Para calcular $\sigma$, calculo antes $\bar{x}$:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{5} c_i n_i = 48{,}7$$

$$\sigma_x = +\sqrt{\sigma_x^2} = +\sqrt{\frac{1}{n}\sum_{i=1}^{5} n_i c_i^2 - \bar{x}^2} = +\sqrt{\frac{1}{150}\sum_{i=1}^{5} n_i c_i^2 - (48{,}7)^2} = 15{,}4965$$

4. Calcular e interpretar los valores de los coeficientes de asimetría y curtosis.

   - <u>Coeficientes de Asimetría</u>:

$$\gamma_1(X) = \frac{1}{n}\sum_{i=1}^{5} n_i \left(\frac{x_i - \bar{x}}{\sigma}\right)^3 = \frac{13{,}755}{150} = 0{,}0917 > 0$$

$$A_p = \frac{\bar{x} - Mo}{\sigma} = 0{,}068 > 0$$

     Como los coeficientes de asimetría son positivos, eso implica que la distribución es asimétrica por la derecha o positiva, es decir, la media aritmética es mayor que la moda. Además, como son cercanos a 0, la distribución no es muy asimétrica. Esto se aprecia bien en la figura 8.9.

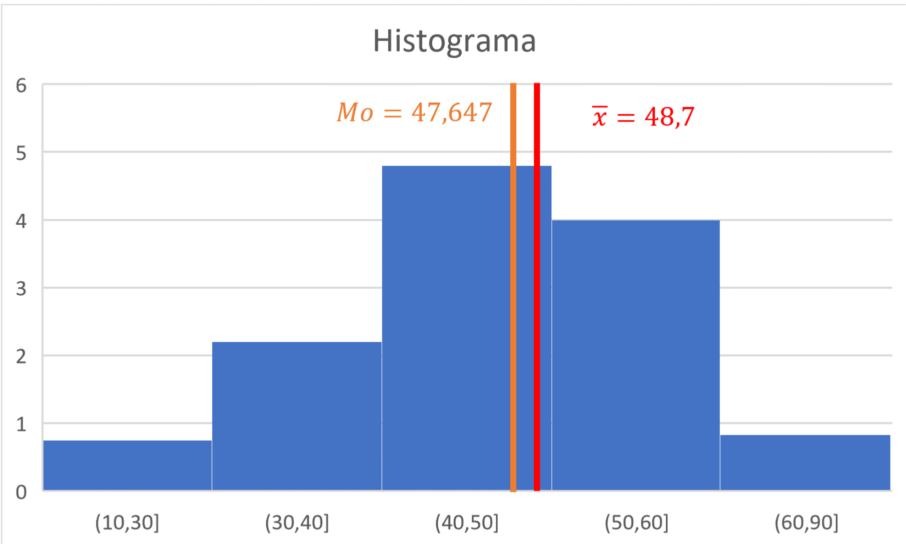


Figura 8.9: Histograma mostrando la media aritmética y la moda.

- Coeficientes de Curtosis:
  Para el coeficiente de curtosis de Fischer, es necesario calcular $\mu_4$:

$$\mu_4 = \frac{1}{n}\sum_{i=1}^{5} n_i(x_i - \bar{x})^4 = 153232{,}455$$

$$\gamma_2(X) = \frac{\mu_4}{\sigma^4} - 3 = 2{,}6571 - 3 = -0{,}3428 < 0$$

Para el coeficiente de curtosis de Kelley, es necesario calcular $D_9$ y $D_1$:

$$D_9 = 60 + \frac{135-125}{25}\cdot 30 = 72 \qquad D_1 = 30 = e_1, \text{ ya que } n_1 = \frac{n}{9} = 15$$

$$K = \frac{1}{2}\frac{Q_3 - Q_1}{D_9 - D_1} - 0{,}263 = \frac{1}{2}\frac{16{,}775}{42} - 0{,}263 = -0{,}0632 < 0$$

Como los coeficientes de curtosis son negativos, eso implica que la distribución es platicúrtica. Es decir, que los intervalos más centrales no sobresalen demasiado de la curva normal con la misma media aritmética y desviación típica. Además, las colas (intervalos $I_1$ e $I_5$) se encuentran por encima de esta distribución normal, como se aprecia bien en la figura 8.10.

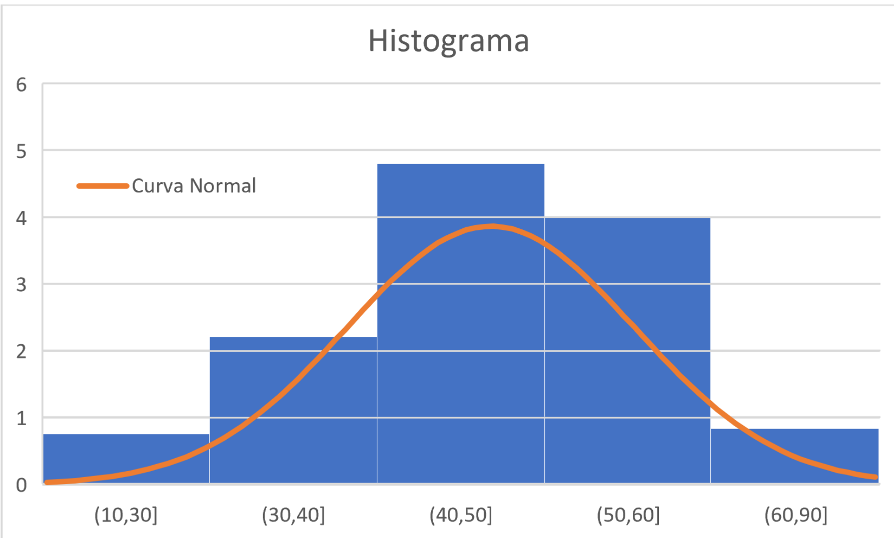


Figura 8.10: Histograma frente a la distribución normal.

## 8.2. Variables Estadísticas Bidimensionales

**Ejercicio 8.2.1.** Se han lanzado dos dados varias veces, obteniendo los resultados que se presentan en la siguiente tabla, donde $X$ designa el resultado del primer dado e $Y$ el resultado del segundo:

| $X$ | 1 | 2 | 2 | 3 | 5 | 4 | 1 | 3 | 3 | 4 | 1 | 2 | 5 | 4 | 3 | 4 | 4 | 5 | 3 | 1 | 6 | 5 | 4 | 6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Y$ | 2 | 3 | 1 | 4 | 3 | 2 | 6 | 4 | 1 | 6 | 6 | 5 | 1 | 2 | 5 | 1 | 1 | 2 | 6 | 6 | 2 | 1 | 2 | 5 |

1. Construir la tabla de frecuencias.
   Sean los valores obtenidos en un dado, $X$; y los valores obtenidos en el segundo, $Y$; dos variables estadísticas con población $n = 24$ y modalidades $x_1, \ldots, x_6$ e $y_1, \ldots, y_6$ respectivamente, con *distribución de frecuencias*:

$$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\ldots,6 \\ j=1,\ldots,6}}$$

| $X \backslash Y$ | 1 | 2 | 3 | 4 | 5 | 6 | $n_{i\cdot}$ |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 | 0 | 3 | 4 |
| 2 | 1 | 0 | 1 | 0 | 1 | 0 | 3 |
| 3 | 1 | 0 | 0 | 2 | 1 | 1 | 5 |
| 4 | 2 | 3 | 0 | 0 | 0 | 1 | 6 |
| 5 | 2 | 1 | 1 | 0 | 0 | 0 | 4 |
| 6 | 0 | 1 | 0 | 0 | 1 | 0 | 2 |
| $n_{\cdot j}$ | 6 | 6 | 2 | 2 | 3 | 5 | 24 |

2. Calcular las puntuaciones obtenidas medias con cada dado y ver cuáles son más homogéneas.

   En primer lugar, calculo las puntuaciones medias mediante las medias aritméticas.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{6} n_{i\cdot}x_i = \frac{81}{24} = 3{,}375 \qquad \overline{y} = \frac{1}{n}\sum_{j=1}^{6} n_{\cdot j}y_j = \frac{77}{24} = 3{,}208\overline{3}$$

   Para estudiar la homogeneidad, necesito saber el coeficiente de variación de Pearson. Para calcular este, necesito en primer lugar calcular las desviaciones típicas.

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{6} n_{i\cdot}(x_i - \overline{x})^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{6} n_{i\cdot}x_i^2 - \overline{x}^2} = \sqrt{\frac{329}{24} - \overline{x}^2} = \sqrt{2{,}3177} = 1{,}5224$$

$$\sigma_y = \sqrt{\frac{1}{n}\sum_{j=1}^{6} n_{\cdot j}(y_j - \overline{y})^2} = \sqrt{\frac{1}{n}\sum_{j=1}^{6} n_{\cdot j}y_j^2 - \overline{y}^2} = \sqrt{\frac{335}{24} - \overline{x}^2} = \sqrt{3{,}6649} = 1{,}9144$$

   Teniendo la media y las desviaciones típicas, podemos calcular el coeficiente de variación de Pearson:

$$C.V.(X) = \frac{\sigma_x}{|\overline{x}|} = \frac{1{,}5224}{3{,}375} = 0{,}4511 \qquad C.V.(Y) = \frac{\sigma_y}{\overline{y}} = \frac{1{,}9144}{3{,}2083} = 0{,}5967$$

   Puesto que $C.V.(X) < C.V.(X) \Longrightarrow$, $X$ es más homogénea que $Y$.

3. ¿Qué resultado del segundo dado es más frecuente cuando en el primero se obtiene un 3?

   Se pide la moda de la distribución condicionada del carácter $Y$ a aquellos que presentan la modalidad $x_3$. La tabla de la distribución condicionada a $x_3$ es:

| $y_j$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $n_{3j}$ | 0 | 0 | 0 | 2 | 1 | 1 |

   Por tanto, $Mo(Y)^{i=3} = 4$.

4. Calcular la puntuación máxima del 50 % de las puntuaciones más bajas obtenidas con el primer dado si con el segundo se ha obtenido un 2 o un 5.

| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $n_i^{j=2,5} = n_{i2} + n_{i5}$ | 1 | 1 | 1 | 3 | 1 | 2 |
| $N_i^{j=2,5}$ | 1 | 2 | 3 | 6 | 7 | 9 |

   Nos piden $Me(X)^{j=2,5}$.

   Como $\dfrac{n^{j=2,5}}{2} = \dfrac{9}{2} = 4{,}5$, la mediana es $Me(X)^{j=2,5} = 4$.

**Ejercicio 8.2.2.** Medidos los pesos, $X$ (en $Kg$), y las alturas, $Y$ (en $cm$), a un grupo de individuos, se han obtenido los siguientes resultados:

| $X\backslash Y$ | 160 | 162 | 164 | 166 | 168 | 170 | $n_{i.}$ |
|---|---|---|---|---|---|---|---|
| 48 | 3 | 2 | 2 | 1 | 0 | 0 | 8 |
| 51 | 2 | 3 | 4 | 2 | 2 | 1 | 14 |
| 54 | 1 | 3 | 6 | 8 | 5 | 1 | 24 |
| 57 | 0 | 0 | 1 | 2 | 8 | 3 | 14 |
| 60 | 0 | 0 | 0 | 2 | 4 | 4 | 10 |
| $n_{.j}$ | 6 | 8 | 13 | 15 | 19 | 9 | 70 |

1. Calcular el peso medio y la altura media y decir cuál es más representativo.

   Sean el peso en Kg, $X$; y las alturas en cm, $Y$; dos variables estadísticas con población $n = 70$ y modalidades $x_1, \dots, x_6$ e $y_1, \dots, y_6$ respectivamente, con *distribución de frecuencias*:

$$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,6 \\ j=1,\dots,6}}$$

   Las medias aritméticas son:

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{6} x_i n_{i.} = \frac{3792}{70} = 54{,}1714\ Kg \quad \bar{y} = \frac{1}{n}\sum_{j=0}^{6} y_j n_{.j} = \frac{11600}{70} = 165{,}7143\ cm$$

   La más representativa será la distribución más homogénea. Para ello, compararemos el Coeficiente de Variación de Pearson, para lo cual calculamos la desviación típica.

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=0}^{6} n_{i.}x_i^2 - \bar{x}^2} = \sqrt{\frac{206316}{70} - \bar{x}^2} = 3{,}582$$

$$\sigma_y = \sqrt{\frac{1}{n}\sum_{j=0}^{6} j_{.j} y_j^2 - \bar{y}^2} = \sqrt{\frac{1922896}{70} - \bar{y}^2} = 2{,}9519$$

Por tanto, los coeficientes de variación son:

$$C.V.(X) = \frac{\sigma_x}{|\bar{x}|} = 0{,}0661 \qquad C.V.(Y) = \frac{\sigma_x}{|\bar{x}|} = 0{,}0178$$

Por tanto, como $C.V.(Y) < C.V.(X)$, entonces la distribución $Y$ es más homogénea y, por tanto, su media es más representativa.

2. Calcular el porcentaje de individuos que pesan menos de 55 $Kg$ y miden más de 165 $cm$.

   Como se pide $x_i < x_4 = 57$, solo se tiene en cuenta $i = 1, 2, 3$.

   Además, como $y_j > y_3 = 164$, solo se toman $j = 4, 5, 6$. Por tanto, el porcentaje es:

$$\frac{\displaystyle\sum_{\substack{i=1,2,3\\ j=4,5,6}} n_{ij}}{n} \cdot 100\,\% = \frac{20}{70} \cdot 100\,\% = 28{,}5714\,\%$$

3. Entre los que miden más de 165 $cm$, ¿cuál es el porcentaje de los que pesan más de 52 $Kg$?

   Como se pide que midan más de 165 cm, tomamos la distribución condicionada a $j = 4, 5, 6$.

| $X\backslash Y = 4,5,6$ | 166 | 168 | 170 | $n_{i.}$ |
|---|---|---|---|---|
| 48 | 1 | 0 | 0 | 1 |
| 51 | 2 | 2 | 1 | 5 |
| 54 | 8 | 5 | 1 | 14 |
| 57 | 2 | 8 | 3 | 13 |
| 60 | 2 | 4 | 4 | 10 |
| $n_{.j}$ | 15 | 19 | 9 | 43 |

   De esta distribución condicionada, calculemos el porcentaje de personas que pesan menos de 52 Kg. Como $\nexists x_i = 52$, sabemos que $P_\alpha = 51 = x_2$.

$$P_\alpha = 51 = x_2 \Longrightarrow N_{2.} = 6 = \frac{n^{j=4,5,6}}{100}\alpha \Longrightarrow \alpha = 6 \cdot \frac{100}{43} = 13{,}9535\,\%$$

   Por tanto, el porcentaje de personas que pesan más de 52 Kg entre los que miden más de 165 cm son:

$$100 - P_\alpha = 86{,}0465\,\%$$

4. ¿Cuál es la altura más frecuente entre los individuos cuyo peso oscila entre 51 y 57 $Kg$?

Se pide la distribución condicionada a $i = 2, 3, 4$. Por tanto, la distribución de frecuencias es:

| $X\backslash Y$ | 160 | 162 | 164 | 166 | 168 | 170 | $n_{i.}^{i=2,3,4}$ |
|---|---|---|---|---|---|---|---|
| 51 | 2 | 3 | 4 | 2 | 2 | 1 | 14 |
| 54 | 1 | 3 | 6 | 8 | 5 | 1 | 24 |
| 57 | 0 | 0 | 1 | 2 | 8 | 3 | 14 |
| $n_{.j}^{i=2,3,4}$ | 3 | 6 | 11 | 12 | 15 | 5 | 52 |

Por tanto, podemos ver que $Mo(Y)^{i=2,3,4} = 168\ cm$.

5. ¿Qué peso medio es más representativo, el de los individuos que miden 164 $cm$ o el de los que miden 168 $cm$?

   La distribución condicionada a $j = 3$ ($y_3 = 164\ cm$) queda como la tabla de la izquierda; mientras que la distribución condicionada a $j = 5$ ($y_5 = 168\ cm$) queda como la tabla de la derecha:

| $X\backslash Y$ | 164 |
|---|---|
| 48 | 2 |
| 51 | 4 |
| 54 | 6 |
| 57 | 1 |
| 60 | 0 |
| $n_{.j}^{j=3}$ | 13 |

| $X\backslash Y$ | 168 |
|---|---|
| 48 | 0 |
| 51 | 2 |
| 54 | 5 |
| 57 | 8 |
| 60 | 4 |
| $n_{.j}^{j=5}$ | 19 |

   Las medias aritméticas:

$$\bar{x}_{j=3} = \frac{1}{n^{j=3}} \sum_{i=0}^{6} x_i n_{i.}^{j=3} = 52{,}3846 \qquad \bar{x}_{j=5} = \frac{1}{n^{j=5}} \sum_{i=0}^{6} x_i n_{i.}^{j=5} = 56{,}2105$$

   Las desviaciones típicas:

$$\sigma_x^{j=3} = \sqrt{\frac{1}{n^{j=3}} \sum_{i=0}^{6} n_{i.}^{j=3} x_i^2 - \bar{x}_{j=3}^2} = 2{,}5283$$

$$\sigma_x^{j=5} = \sqrt{\frac{1}{n^{j=5}} \sum_{i=0}^{6} n_{i.}^{j=5} x_i^2 - \bar{x}_{j=5}^2} = 2{,}7662$$

   El coeficiente de variación de Pearson:

$$C.V.(X)^{j=3} = \frac{\sigma_x^{j=3}}{|\bar{x}_{j=3}|} = 0{,}04826 \qquad C.V.(X)^{j=5} = \frac{\sigma_x^{j=5}}{|\bar{x}_{j=5}|} = 0{,}04921$$

   Por tanto, como los coeficientes de Pearson son similares, el peso medio tiene una representatividad similar. No obstante, para $j = 3$, es decir, 164 $cm$, es más homogénea la distribución condicionada.

**Ejercicio 8.2.3.** En una encuesta de familias sobre el número de individuos que la componen ($X$) y el número de personas activas en ellas ($Y$) se han obtenido los siguientes resultados:

| $X\backslash Y$ | 1 | 2 | 3 | 4 | $n_{i.}$ |
|---|---|---|---|---|---|
| 1 | 7 | 0 | 0 | 0 | 7 |
| 2 | 10 | 2 | 0 | 0 | 12 |
| 3 | 11 | 5 | 1 | 0 | 17 |
| 4 | 10 | 6 | 6 | 0 | 22 |
| 5 | 8 | 6 | 4 | 2 | 20 |
| 6 | 1 | 2 | 3 | 1 | 7 |
| 7 | 1 | 0 | 0 | 1 | 2 |
| 8 | 0 | 0 | 1 | 1 | 2 |
| $n_{.j}$ | 48 | 21 | 15 | 5 | 89 |

1. Calcular la recta de regresión de $Y$ sobre $X$.
   Sean el número de individuos que componen la familia, $X$; y el número de personas activas en ellas, $Y$; dos variables estadísticas con población $n = 89$ y modalidades $x_1, \dots, x_8$ e $y_1, \dots, y_4$ respectivamente, con *distribución de frecuencias*:
   $$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,8\\ j=1,\dots,4}}$$
   Obtengo, en primer lugar, las medias aritméticas marginales:
   $$\bar{x} = \frac{1}{n}\sum_{i=0}^{8} x_i n_{i.} = \frac{342}{89} = 3{,}8427 \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{4} y_j n_{.j} = \frac{155}{89} = 1{,}7416$$
   Obtengo ahora la covarianza:
   $$\sigma_{xy} = \frac{1}{n}\sum_{i=0}^{8}\sum_{j=0}^{4} x_i y_j n_{ij} - \bar{x}\bar{y} = \frac{666}{89} - \bar{x}\bar{y} = 0{,}7907$$
   Obtengo ahora la varianza de la variable estadística $X$:
   $$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{8} n_{i.} x_i^2 - \bar{x}^2 = \frac{1538}{89} - \bar{x}^2 = 2{,}51455$$
   Por tanto, la recta de regresión de $Y$ sobre $X$ es:
   $$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Longrightarrow y = 0{,}31445x + 0{,}533$$

2. ¿Es adecuado suponer una relación lineal para explicar el comportamiento de $Y$ a partir de $X$?

   Para ver si nuestra recta de regresión es correcta o no, se emplea el coeficiente de determinación.
   $$\eta^2_{Y/X} = \frac{\sigma^2_{ey}}{\sigma^2_y}$$

Además, en el caso de que se trabaje en el caso lineal, tenemos que:

$$\eta^2_{Y/X} = \frac{\sigma^2_{ey}}{\sigma^2_y} = r^2 = \frac{\sigma^2_{xy}}{\sigma^2_x \sigma^2_y}$$

Calculamos por tanto $\sigma^2_y$:

$$\sigma^2_y = \frac{1}{n}\sum_{j=0}^{4} n_{.j} y_j^2 - \bar{y}^2 = \frac{347}{89} - \bar{x}^2 = 0{,}8657$$

Por tanto:

$$\eta^2_{Y/X} = \frac{\sigma^2_{ey}}{\sigma^2_y} = r^2 = \frac{\sigma^2_{xy}}{\sigma^2_x \sigma^2_y} = 0{,}287$$

Como $\eta^2_{Y/X} = 0{,}28719$ dista bastante del 1, entonces podemos afirmar que no es adecuado suponer una relación lineal, puesto que el modelo explica menos del 29 % de los casos.

Podemos ver la representación en forma de nube de putos, junto con la recta de regresión de Y sobre X, y junto con el coeficiente $r^2$, en la figura 8.11.

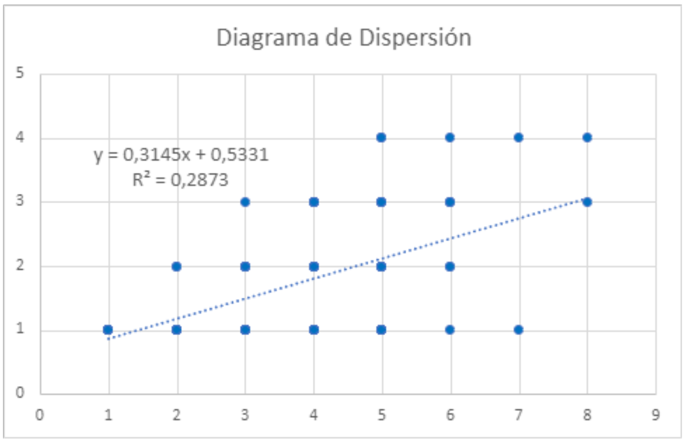


Figura 8.11: Diagrama de dispersión del ejercicio 8.2.3

**Ejercicio 8.2.4.** Se realiza un estudio sobre la tensión de vapor de agua ($Y$ , en *ml. de Hg.*) a distintas temperaturas ($X$, en $°C$). Efectuadas 21 medidas, los resultados son:

| $X\backslash Y$ | $c_i^x$ | $(0{,}5, 1{,}5]$ | $(1{,}5, 2{,}5]$ | $(2{,}5, 5{,}5]$ | $n_{i.}$ |
|---|---|---|---|---|---|
| $c_j^y$ | | 1 | 2 | 4 | |
| $(1, 15]$ | 8 | 4 | 2 | 0 | 6 |
| $(15, 25]$ | 20 | 1 | 4 | 2 | 7 |
| $(25, 30]$ | 27,5 | 0 | 3 | 5 | 8 |
| $n_{.j}$ | | 5 | 9 | 7 | 21 |

Explicar el comportamiento de la tensión de vapor en términos de la temperatura mediante una función lineal. ¿Es adecuado asumir este tipo de relación?

Sea la temperatura, $X$; y la tensión de vapor de agua, $Y$; dos variables estadísticas con población $n = 21$ y modalidades $I_1^x, \dots, I_3^x$ e $I_1^y, \dots, I_3^y$ respectivamente, con *distribución de frecuencias*:

$$\left\{(I_i^x, I_j^x), n_{ij}\right\}_{\substack{i=1,\dots,3\\ j=1,\dots,3}}$$

En este caso, se pide la recta de regresión de $Y$ sobre $X$. Para ello, calculamos previamente varios datos.

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{3} c_i^x n_{i.} = \frac{408}{21} = 19{,}4286 \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{3} c_j^y n_{.j} = \frac{51}{21} = 2{,}4286$$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{3}(c_i^x)^2 n_{i.} - \bar{x}^2 = \frac{9234}{21} - \bar{x}^2 = 62{,}2437$$

$$\sigma_y^2 = \frac{1}{n}\sum_{j=0}^{3}(c_j^y)^2 n_{.j} - \bar{y}^2 = \frac{153}{21} - \bar{y}^2 = 1{,}3876$$

$$\sigma_{xy} = \frac{1}{n}\sum_{i,j=0}^{3} c_i^x c_j^y n_{ij} - \bar{x}\bar{y} = \frac{1119}{21} - \bar{x}\bar{y} = 6{,}1014$$

La recta de regresión de $Y$ sobre $X$ es:

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Longrightarrow y = 0{,}098x + 0{,}5239$$

Para ver si asumir este tipo de relación es correcto, vemos el valor del coeficiente de determinación. Haciendo uso de que el ajuste empleado ha sido lineal:

$$\eta_{Y/X}^2 = \frac{\sigma_{ey}^2}{\sigma_y^2} = r^2 = \frac{\sigma_{x,y}^2}{\sigma_x^2\sigma_y^2} = 0{,}431$$

Como $r^2 = 0{,}431$, tenemos que explica solo el 43,1 % de los resultados, menos del 50 %. Por tanto, como dista mucho del 1, entonces podemos afirmar que en este caso no se ajusta correctamente mediante el ajuste lineal.

Estos resultados los podemos ver en el diagrama de dispersión de la figura 8.12.

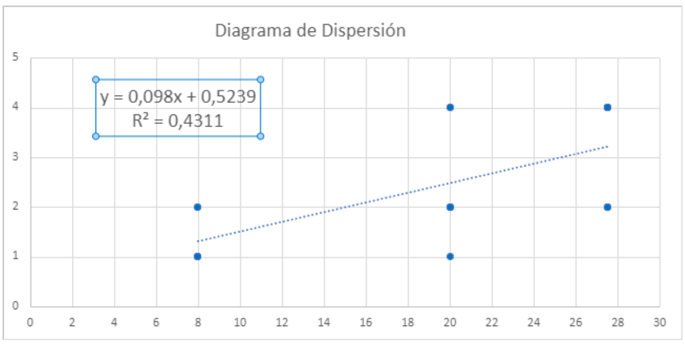


Figura 8.12: Diagrama de dispersión del ejercicio 8.2.4

**Ejercicio 8.2.5.** Estudiar la dependencia o independencia de las variables en cada una de las siguientes distribuciones. Dar, en cada caso, las curvas de regresión y la covarianza de las dos variables.

Distribución $A$

| $X\backslash Y$ | 1 | 2 | 3 | 4 | 5 | $n_{i.}$ |
|---|---|---|---|---|---|---|
| 10 | 2 | 4 | 6 | 10 | 8 | 30 |
| 20 | 1 | 2 | 3 | 5 | 4 | 15 |
| 30 | 3 | 6 | 9 | 15 | 12 | 45 |
| 40 | 4 | 8 | 12 | 20 | 16 | 60 |
| $n_{.j}$ | 10 | 20 | 30 | 50 | 40 | 150 |

Distribución $B$

| $X\backslash Y$ | 1 | 2 | 3 | $n_{i.}$ |
|---|---|---|---|---|
| $-1$ | 0 | 1 | 0 | 1 |
| 0 | 1 | 0 | 1 | 2 |
| 1 | 0 | 1 | 0 | 1 |
| $n_{.j}$ | 1 | 2 | 1 | 4 |

Sean $X^A$ y $Y^A$ dos variables estadísticas con población $n = 150$ y modalidades $x_1^A, \dots, x_4^A$ e $y_1^A, \dots, y_5^A$ respectivamente, con *distribución de frecuencias*:

$$\left\{(x_i^A, y_j^A), n_{ij}\right\}_{\substack{i=1,\dots,4 \\ j=1,\dots,5}}$$

Por el teorema de la caracterización de la independencia estadística, como $n_{ij} = \frac{n_{i.}n_{.j}}{n}$ $\forall i, j$, tenemos que las variables $X$ e $Y$ son estadísticamente independientes en la distribución $A$. Además, esto también se puede ver, ya que:

$$f_1^j = \frac{1}{5} = f_{1.} \qquad f_2^j = \frac{1}{10} = f_{2.} \qquad f_3^j = \frac{3}{10} = f_{3.} \qquad f_4^j = \frac{2}{5} = f_{4.}$$

Como vemos que $f_i^j$ no depende de $j$, tenemos que $X^A$ e $Y^A$ son estadísticamente independientes.

Sean ahora $X^B$ y $Y^B$ dos variables estadísticas con población $n = 4$ y modalidades $x_1^B, \dots, x_3^B$ e $y_1^B, \dots, y_3^B$ respectivamente, con *distribución de frecuencias*:

$$\left\{(x_i^B, y_j^B), n_{ij}\right\}_{\substack{i=1,\dots,3 \\ j=1,\dots,3}}$$

Como tenemos que $n_{11} = 0 \neq \frac{1}{4} = \frac{n_{1.}n_{.1}}{n}$, entonces las variables $X, Y$ no son estadísticamente independientes. Tampoco depende $X$ de $Y$ (a $y_2$ le corresponden $x_1$ y $x_3$) ni viceversa (a $x_2$ le corresponden $y_1$ e $y_3$).

Calculamos ahora las curvas de regresión.

1. Distribución A

   La covarianza es:

   $$\sigma_{xy} = 0, \text{ ya que } X^A \text{ e } Y^A \text{ son estadísticamente independientes.}$$

   Al ser las variables independientes, no tiene sentido calcular la curva de regresión, ya que ninguna de las variables explica la otra.

   No obstante, las calculamos. Como son independientes, tenemos que $\bar{x}_j = \bar{x}$, y $\bar{y}_i = \bar{y}$.

   Calculamos las medias aritméticas:

   $$\bar{x} = \frac{1}{n}\sum_{i=0}^{4} x_i n_{i.} = \frac{4350}{150} = 29 \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{5} y_j n_{.j} = \frac{540}{150} = 3{,}6$$

La curva de regresión $X/Y$ pasa por los puntos $(\bar{x}_j, y_j)$. Por tanto, los puntos son:

$$(29, 1) \qquad (29, 2) \qquad (29, 3) \qquad (29, 4) \qquad (29, 5)$$

La curva de regresión de $X/Y$ la podemos ver en la figura 8.13.

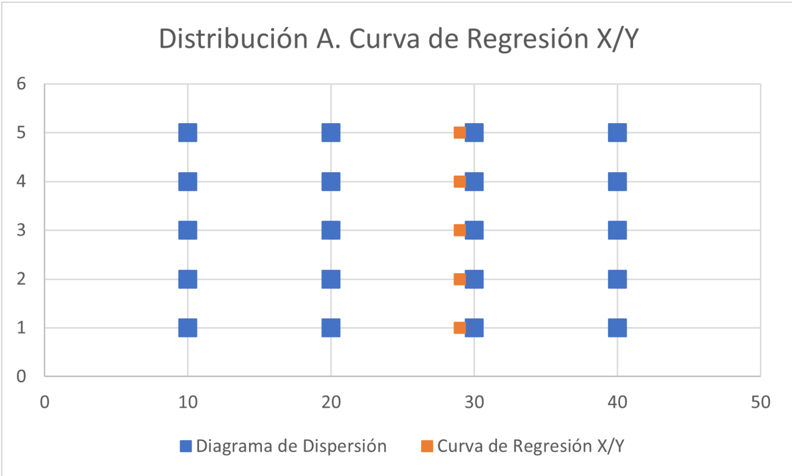


Figura 8.13: Curva de regresión de tipo I de $X/Y$

La curva de regresión $Y/X$ pasa por los puntos $(x_i, \bar{y}_i)$. Por tanto, los puntos son:

$$(10, 3{,}6) \qquad (20, 3{,}6) \qquad (30, 3{,}6) \qquad (40, 3{,}6)$$

La curva de regresión de $Y/X$ la podemos ver en la figura 8.14.

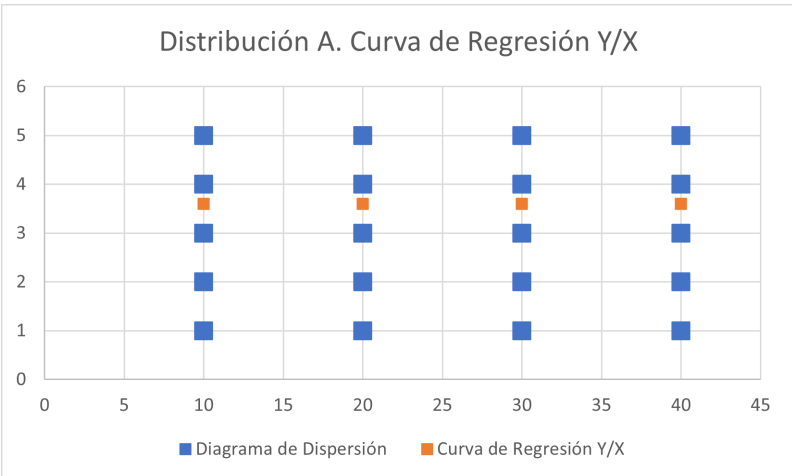


Figura 8.14: Curva de regresión de tipo I de $Y/X$

Como podemos ver, al ser las variables independientes, la curva de regresión no aporta información relevante. Los puntos se sitúan paralelos a los ejes a la altura de la media que se quiere explicar.

2. Distribución B

La covarianza es:

$$\sigma_{xy} = \frac{1}{n}\sum_{i,j=1}^{3} x_i y_j n_{ij} - \bar{x}\bar{y} = \frac{0}{4} - 0 = 0$$

Por tanto, aquí podemos ver que si las variables son independientes, entonces su covarianza es nula, pero que el recíproco no es cierto. Ejemplo de lo segundo es esto.

Calculamos ahora la curva de regresión de tipo I de $Y/X$, que pasa por los puntos $(x_i, \bar{y}_i)$:

$$\bar{y}_1 = \frac{1}{n_{1.}}\sum_{j=1}^{3} y_j n_{1j} = \frac{2}{1} = 2 \qquad \bar{y}_2 = \frac{1}{n_{2.}}\sum_{j=1}^{3} y_j n_{2j} = \frac{4}{2} = 2$$

$$\bar{y}_3 = \frac{1}{n_{3.}}\sum_{j=1}^{3} y_j n_{3j} = \frac{2}{1} = 2$$

Por tanto, los puntos son:

$$(-1, 2) \qquad (0, 2) \qquad (1, 2)$$

La curva de regresión de $Y/X$ la podemos ver en la figura 8.15.

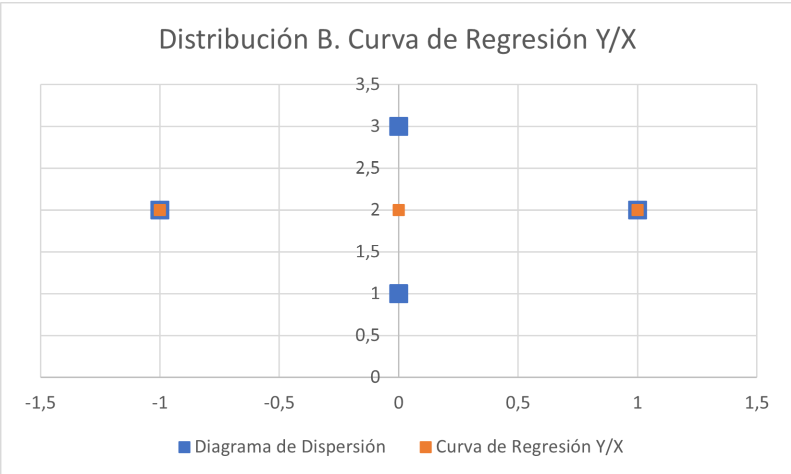


Figura 8.15: Curva de regresión de tipo I de $Y/X$

Calculamos ahora la curva de regresión de tipo I de $X/Y$, que pasa por los puntos $(\bar{x}_j, y_j)$:

$$\bar{x}_1 = \frac{1}{n_{,1}}\sum_{i=1}^{3} x_i n_{i1} = \frac{0}{1} = 0 \qquad \bar{x}_2 = \frac{1}{n_{,2}}\sum_{i=1}^{3} x_i n_{i2} = \frac{0}{2} = 0$$

$$\bar{x}_3 = \frac{1}{n_{,3}}\sum_{i=1}^{3} x_i n_{i3} = \frac{0}{1} = 0$$

Por tanto, los puntos son:

$$(0,1) \qquad (0,2) \qquad (0,3)$$

La curva de regresión de $X/Y$ la podemos ver en la figura 8.16.

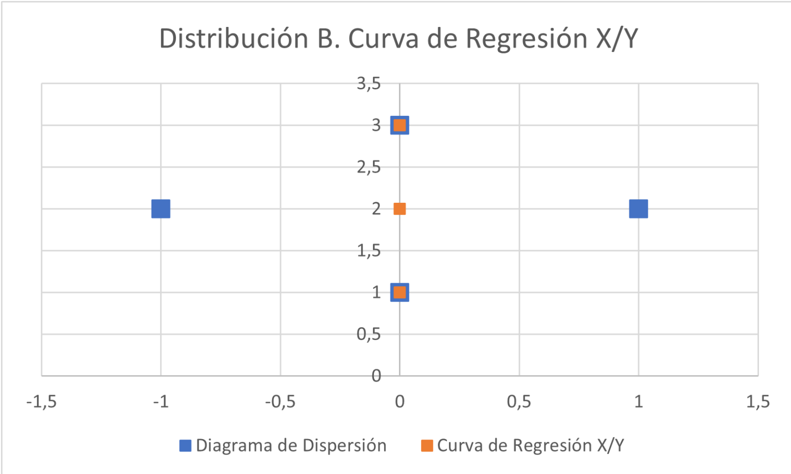


Figura 8.16: Curva de regresión de tipo I de $X/Y$

**Ejercicio 8.2.6.** Dada la siguiente distribución bidimensional:

| $X\backslash Y$ | 1 | 2 | 3 | 4 | $n_{i\cdot}$ |
|---|---|---|---|---|---|
| 10 | 1 | 3 | 0 | 0 | 4 |
| 12 | 0 | 1 | 4 | 3 | 8 |
| 14 | 2 | 0 | 0 | 2 | 4 |
| 16 | 4 | 0 | 0 | 0 | 4 |
| $n_{\cdot j}$ | 7 | 4 | 4 | 5 | 20 |

1. ¿Son estadísticamente independientes $X$ e $Y$?

   Sean $X$ y $Y$ dos variables estadísticas con población $n = 20$ y modalidades $x_1, \dots, x_4$ e $y_1, \dots, y_4$ respectivamente, con *distribución de frecuencias*:

   $$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,4 \\ j=1,\dots,4}}$$

   Veamos ahora si son independientes:

   $$f_1^1 = \frac{n_{11}}{n_{,1}} = \frac{1}{7} \qquad f_{1\cdot} = \frac{n_{1\cdot}}{n} = \frac{4}{20}$$

   Como $f_1^1 \neq f_{1\cdot} \Longrightarrow X$ y $Y$ por lo que no son independientes.

2. Calcular y representar las curvas de regresión de $X/Y$ e $Y/X$.

   Calculo en primer lugar la curva de regresión de $X/Y$. Esto son los puntos $(\bar{x}_j, y_j)\ \forall j = 1, \dots, 4$.

   $$\bar{x_1} = \frac{1}{n_{\cdot 1}} \sum_{i=0}^{n} x_i n_{i1} = \frac{102}{7} = 14{,}5715 \qquad \bar{x_2} = \frac{1}{n_{\cdot 2}} \sum_{i=0}^{n} x_i n_{i2} = \frac{42}{4} = 10{,}5$$

$$\bar{x_3} = \frac{1}{n_{\cdot 3}} \sum_{i=0}^{n} x_i n_{i3} = \frac{48}{4} = 12 \qquad \bar{x_4} = \frac{1}{n_{\cdot 4}} \sum_{i=0}^{n} x_i n_{i4} = \frac{64}{5} = 12{,}8$$

Por tanto, la curva de regresión de tipo I de $X/Y$ es la que pasa por los puntos:

$$(14{,}5715,\ 1) \qquad (10{,}5,\ 2) \qquad (12,\ 3) \qquad (12{,}8,\ 4)$$

Su representación se encuentra en la figura 8.17.

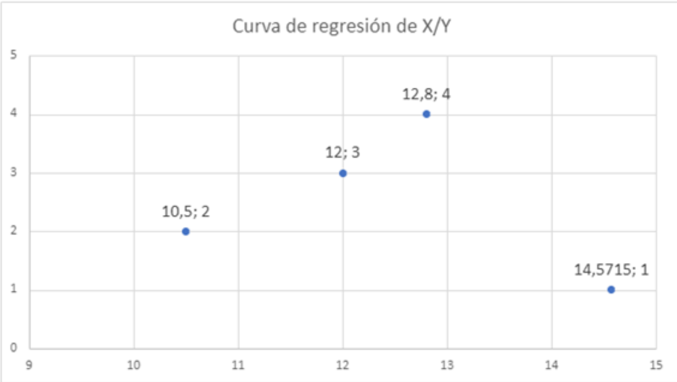


Figura 8.17: Curva de regresión de tipo I de $X/Y$

Calculo ahora la curva de regresión de $Y/X$. Esto son los puntos $(x_i, \bar{y_i})$ $\forall i = 1, \ldots, 4$.

$$\bar{y_1} = \frac{1}{n_{1\cdot}} \sum_{j=0}^{n} y_j n_{1j} = \frac{7}{4} = 1{,}75 \qquad \bar{y_2} = \frac{1}{n_{2\cdot}} \sum_{j=0}^{n} y_j n_{2j} = \frac{26}{8} = 3{,}25$$

$$\bar{y_3} = \frac{1}{n_{3\cdot}} \sum_{j=0}^{n} y_j n_{3j} = \frac{10}{4} = 2{,}5 \qquad \bar{y_4} = \frac{1}{n_{4\cdot}} \sum_{j=0}^{n} y_j n_{4j} = \frac{4}{4} = 1$$

Por tanto, la curva de regresión de tipo I de $Y/X$ es la que pasa por los puntos:

$$(10,\ 1{,}75) \qquad (12,\ 3{,}25) \qquad (14,\ 2{,}5) \qquad (16,\ 1)$$

Su representación se encuentra en la figura 8.18.

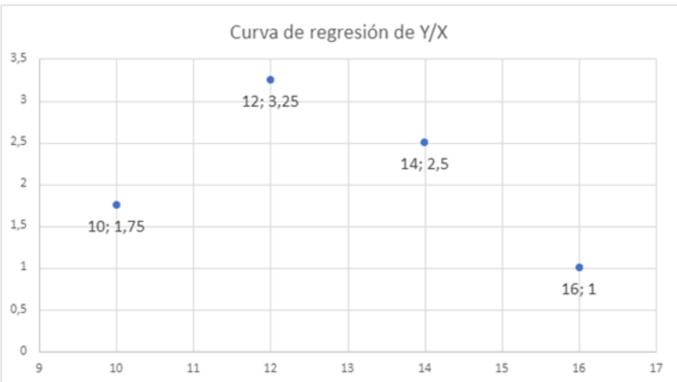


Figura 8.18: Curva de regresión de tipo I de $Y/X$

3. Cuantificar el grado en que cada variable es explicada por la otra mediante la correspondiente curva de regresión.

   En primer lugar, calculo la varianza explicada por cada curva de regresión. Para ello, necesito previamente las medias aritméticas.

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{4} x_i n_{i.} = \frac{256}{20} = 12{,}8 \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{4} y_j n_{.j} = \frac{47}{20} = 2{,}35$$

$$\sigma_{ey}^2 = \frac{1}{n}\sum_{i,j=1}^{4} n_{ij}(\hat{y}_j - \bar{y})^2 = \frac{1}{n}\sum_{i,j=1}^{4} n_{ij}(\bar{y}_i - \bar{y})^2 = \frac{1}{n}\sum_{i=1}^{4} n_{i.}(\bar{y}_i - \bar{y})^2 = \frac{15{,}3}{20} = 0{,}765$$

$$\sigma_{ex}^2 = \frac{1}{n}\sum_{i,j=1}^{4} n_{ij}(\hat{x}_i - \bar{x})^2 = \frac{1}{n}\sum_{i,j=1}^{4} n_{ij}(\bar{x}_j - \bar{x})^2 = \frac{1}{n}\sum_{j=1}^{4} n_{.j}(\bar{x}_j - \bar{x})^2 = \frac{45{,}6875}{20} = 2{,}2845$$

   Para poder comparar, necesito el coeficiente de determinación. Para ello, calculo también las varianzas de cada variable.

$$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{4} n_{i.}x_i^2 - \bar{x}^2 = \frac{3360}{20} - \bar{x}^2 = 4{,}16$$

$$\sigma_y^2 = \frac{1}{n}\sum_{j=0}^{4} n_{.j}y_j^2 - \bar{y}^2 = \frac{139}{20} - \bar{y}^2 = 1{,}4275$$

   Por tanto, los coeficientes de determinación son:

$$\eta_{X/Y}^2 = \frac{\sigma_{ex}^2}{\sigma_x^2} = 0{,}5492 \qquad \eta_{Y/X}^2 = \frac{\sigma_{ey}^2}{\sigma_y^2} = 0{,}5359$$

   Como podemos ver, la curva de regresión de $X/Y$ tiene un coeficiente de determinación menor, por lo que la variable $X$ es mejor explicada por la variable $Y$ que al revés.

4. ¿Están $X$ e $Y$ correlacionadas linealmente? Dar las expresiones de las rectas de regresión.

   Para calcular el coeficiente de determinación, necesito previamente la covarianza.

$$\sigma_{xy} = \frac{1}{n}\sum_{i,j=0}^{4} n_{ij}x_i y_j - \bar{x}\bar{y} = \frac{586}{20} - \bar{x}\bar{y} = -0{,}78$$

   Por tanto, la recta de regresión de $Y$ sobre $X$ es:

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Longrightarrow y = -0{,}188x + 4{,}75$$

   La recta de regresión de $X$ sobre $Y$ es:

$$x - \bar{x} = \frac{\sigma_{xy}}{\sigma_y^2}(y - \bar{y}) \Longrightarrow x = -0{,}546y + 14{,}083$$

Para ver si están correlacionadas linealmente, calculo el coeficiente de correlación lineal:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = -0{,}32008$$

Como el valor de $r$ no se acerca ni a $+1$ ni a $-1$, concluimos que el nivel de correlación entre ambas variables es bajo.

**Ejercicio 8.2.7.** Para cada una de las distribuciones:

Distribución $A$

| $X\backslash Y$ | 10 | 15 | 20 |
|---|---|---|---|
| 1 | 0 | 2 | 0 |
| 2 | 1 | 0 | 0 |
| 3 | 0 | 0 | 3 |
| 4 | 0 | 1 | 0 |

Distribución $B$

| $X\backslash Y$ | 10 | 15 | 20 |
|---|---|---|---|
| 1 | 0 | 2 | 0 |
| 2 | 1 | 0 | 0 |
| 3 | 0 | 0 | 3 |

Distribución $C$

| $X\backslash Y$ | 10 | 15 | 20 | 25 |
|---|---|---|---|---|
| 1 | 0 | 3 | 0 | 1 |
| 2 | 0 | 0 | 1 | 0 |
| 3 | 2 | 0 | 0 | 0 |

1. ¿Dependen funcionalmente $X$ de $Y$ o $Y$ de $X$?

Sean $X^A$ y $Y^A$ dos variables estadísticas con población $n = 7$ y modalidades $x_1^A, \ldots, x_4^A$ e $y_1^A, \ldots, y_3^A$ respectivamente, con *distribución de frecuencias*:

$$\left\{(x_i^A, y_j^A), n_{ij}\right\}_{\substack{i=1,\ldots,4 \\ j=1,\ldots,3}}$$

En el caso de la distribución $A$, la variable $Y^A$ depende funcionalmente de la variable $X^A$, ya que a cada modalidad de $X^A$ le corresponde una única modalidad de $Y^A$ con frecuencia no nula.

No obstante, la variable $X^A$ no depende funcionalmente de $Y^A$, ya que a la modalidad $y_2^A$ le corresponden dos modalidades con frecuencias no nulas, $x_1^A$ y $x_4^A$.

Sean ahora $X^B$ y $Y^B$ dos variables estadísticas con población $n = 6$ y modalidades $x_1^B, \ldots, x_3^B$ e $y_1^B, \ldots, y_3^B$ respectivamente, con *distribución de frecuencias*:

$$\left\{(x_i^B, y_j^B), n_{ij}\right\}_{\substack{i=1,\ldots,3 \\ j=1,\ldots,3}}$$

En el caso de la distribución $B$, la dependencia lineal es recíproca, ya que a cada modalidad de cada una de las variables le corresponde una única modalidad de la otra variable con frecuencia no nula.

Sean por último $X^C$ y $Y^C$ dos variables estadísticas con población $n = 7$ y modalidades $x_1^C, \ldots, x_3^C$ e $y_1^C, \ldots, y_4^C$ respectivamente, con *distribución de frecuencias*:

$$\left\{(x_i^C, y_j^C), n_{ij}\right\}_{\substack{i=1,\ldots,3 \\ j=1,\ldots,4}}$$

En el caso de la distribución $C$, la variable $X^C$ depende funcionalmente de la variable $Y^C$, ya que a cada modalidad de $Y^C$ le corresponde una única modalidad de $X^C$ con frecuencia no nula.

No obstante, la variable $Y^C$ no depende funcionalmente de $X^C$, ya que a la modalidad $x_1^C$ le corresponden dos modalidades con frecuencias no nulas, $y_2^C$ y $y_4^C$.

2. Calcular las curvas de regresión y comentar los resultados.

   a) Distribución A
   La curva de regresión de tipo I de $X/Y$ (ver figura 8.19) pasa por los puntos $(\bar{x}_j, y_j)$:

$$\bar{x}_1 = \frac{1}{n_{,1}} \sum_{i=0}^{4} x_i^A n_{i1} = \frac{2}{1} = 2 \qquad \bar{x}_2 = \frac{1}{n_{,2}} \sum_{i=0}^{4} x_i^A n_{i2} = \frac{6}{3} = 2$$

$$\bar{x}_3 = \frac{1}{n_{,3}} \sum_{i=0}^{4} x_i^A n_{i3} = \frac{9}{3} = 3$$

   Los puntos son:

$$(2, 10) \qquad (2, 15) \qquad (3, 20)$$

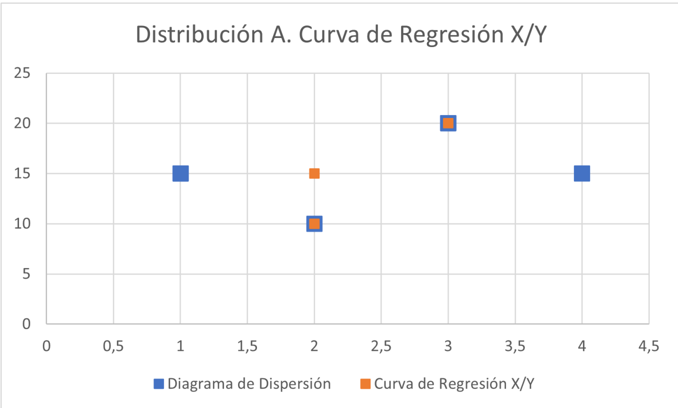


Figura 8.19: Curva de Regresión $X/Y$ de la distribución A.

   La curva de regresión de tipo I de $Y/X$ (ver figura 8.20) pasa por los puntos $(x_i, \bar{y}_i)$:

$$\bar{y}_1 = \frac{1}{n_{1.}} \sum_{j=0}^{3} y_j^A n_{1j} = \frac{30}{2} = 15 \qquad \bar{y}_2 = \frac{1}{n_{2.}} \sum_{j=0}^{3} y_j^A n_{2j} = \frac{10}{1} = 10$$

$$\bar{y}_3 = \frac{1}{n_{3.}} \sum_{j=0}^{3} y_j^A n_{3j} = \frac{60}{3} = 20 \qquad \bar{y}_4 = \frac{1}{n_{4.}} \sum_{j=0}^{3} y_j^A n_{4j} = \frac{15}{1} = 15$$

   Los puntos son:

$$(1, 15) \qquad (2, 10) \qquad (3, 20) \qquad (4, 15)$$

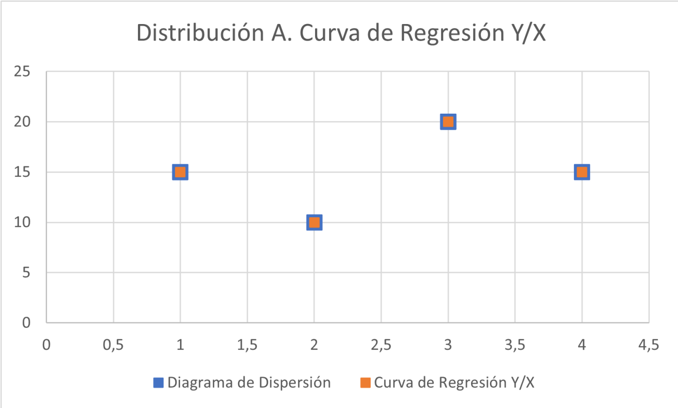


Figura 8.20: Curva de Regresión $Y/X$ de la distribución A.

*b*) Distribución B

La curva de regresión de tipo I de $X/Y$ (ver figura 8.21) pasa por los puntos $(\bar{x}_j, y_j)$:

$$\bar{x}_1 = \frac{1}{n_{,1}} \sum_{i=0}^{3} x_i^B n_{i1} = \frac{2}{1} = 2 \qquad \bar{x}_2 = \frac{1}{n_{,2}} \sum_{i=0}^{3} x_i^B n_{i2} = \frac{2}{2} = 1$$

$$\bar{x}_3 = \frac{1}{n_{,3}} \sum_{i=0}^{3} x_i^B n_{i3} = \frac{9}{3} = 3$$

Los puntos son:

$$(2, 10) \qquad (1, 15) \qquad (3, 20)$$

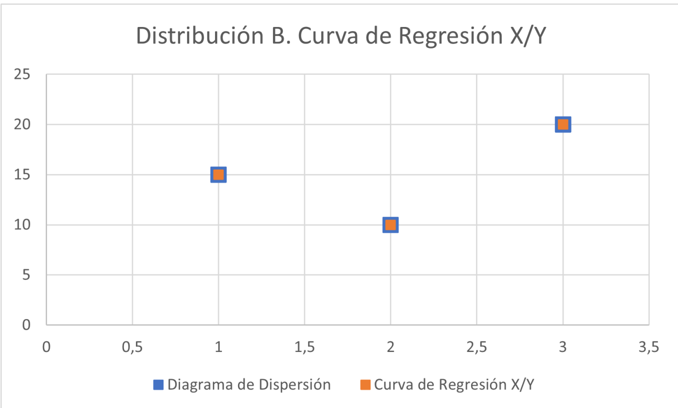


Figura 8.21: Curva de Regresión $X/Y$ de la distribución B.

La curva de regresión de tipo I de $Y/X$ (ver figura 8.22) pasa por los puntos $(x_i, \bar{y}_i)$:

$$\bar{y}_1 = \frac{1}{n_{1.}} \sum_{j=0}^{3} y_j^B n_{1j} = \frac{30}{2} = 15 \qquad \bar{y}_2 = \frac{1}{n_{2.}} \sum_{j=0}^{3} y_j^B n_{2j} = \frac{10}{1} = 10$$

$$\bar{y}_3 = \frac{1}{n_{3.}} \sum_{j=0}^{3} y_j^B n_{3j} = \frac{60}{3} = 20$$

Los puntos son:

$$(1, 15) \qquad (2, 10) \qquad (3, 20)$$

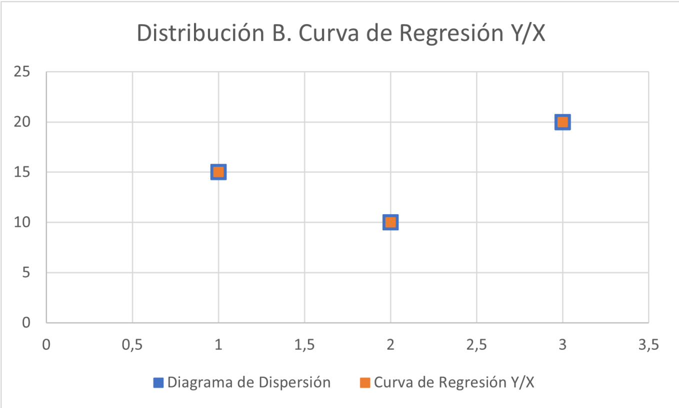


Figura 8.22: Curva de Regresión $Y/X$ de la distribución B.

c) Distribución C
La curva de regresión de tipo I de $X/Y$ (ver figura 8.23) pasa por los puntos $(\bar{x}_j, y_j)$:

$$\bar{x}_1 = \frac{1}{n_{,1}} \sum_{i=0}^{3} x_i^C n_{i1} = \frac{6}{2} = 3 \qquad \bar{x}_2 = \frac{1}{n_{,2}} \sum_{i=0}^{3} x_i^C n_{i2} = \frac{3}{3} = 1$$

$$\bar{x}_3 = \frac{1}{n_{,3}} \sum_{i=0}^{3} x_i^C n_{i3} = \frac{2}{1} = 2 \qquad \bar{x}_4 = \frac{1}{n_{,4}} \sum_{i=0}^{3} x_i^C n_{i4} = \frac{1}{1} = 1$$

Los puntos son:

$$(3, 10) \qquad (1, 15) \qquad (2, 20) \qquad (1, 25)$$

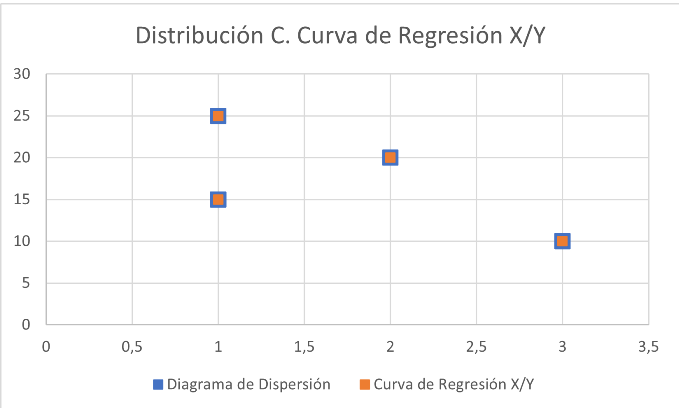


Figura 8.23: Curva de Regresión $X/Y$ de la distribución C.

La curva de regresión de tipo I de $Y/X$ (ver figura 8.24) pasa por los puntos $(x_i, \bar{y}_i)$:

$$\bar{y}_1 = \frac{1}{n_{1.}}\sum_{j=0}^{4} y_j^C n_{1j} = \frac{55}{4} = 17{,}5 \qquad \bar{y}_2 = \frac{1}{n_{2.}}\sum_{j=0}^{4} y_j^C n_{2j} = \frac{20}{1} = 20$$

$$\bar{y}_3 = \frac{1}{n_{3.}}\sum_{j=0}^{4} y_j^C n_{3j} = \frac{20}{2} = 10$$

Los puntos son:

$$(1,\ 17{,}5) \qquad (2, 20) \qquad (3, 10)$$

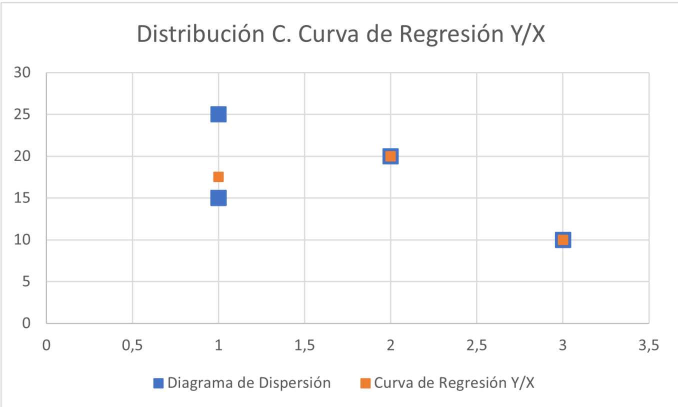


Figura 8.24: Curva de Regresión $Y/X$ de la distribución C.

Podemos ver que, en los casos en los que hay dependencia funcional, la nube de puntos se superpone con la curva de regresión. Además, en el caso de la dependencia funcional recíproca, la curva de regresión de $Y/X$ coincide con la curva de regresión de $X/Y$.

**Ejercicio 8.2.8.** De una muestra de 24 puestos de venta en un mercado de abastos se ha recogido información sobre el número de balanzas ($X$) y el número de dependientes ($Y$). Los resultados aparecen en la siguiente tabla:

| $X\backslash Y$ | 1 | 2 | 3 | 4 | $n_{i.}$ |
|---|---|---|---|---|---|
| 1 | 1 | 2 | 0 | 0 | 3 |
| 2 | 1 | 2 | 3 | 1 | 7 |
| 3 | 0 | 1 | 2 | 6 | 9 |
| 4 | 0 | 0 | 2 | 3 | 5 |
| $n_{.j}$ | 2 | 5 | 7 | 10 | 24 |

1. Determinar las rectas de regresión.

   Sea el número de balanzas en un puesto de venta de un mercado, $X$; y el número de dependientes en dicho puesto de venta, $Y$ dos variables estadísticas con población $n = 24$ y modalidades $x_1, \dots, x_4$ e $y_1, \dots, y_4$ respectivamente, con *distribución de frecuencias*:

$$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,4 \\ j=1,\dots,4}}$$

Para calcular el coeficiente de regresión lineal, antes son necesarios algunos cálculos:

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{4} x_i n_{i.} = \frac{64}{24} = 2.\bar{6} \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{4} y_j n_{.j} = \frac{73}{24} = 3{,}041\bar{6}$$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{4} n_{i.} x_i^2 - \bar{x}^2 = \frac{192}{24} - \bar{x}^2 = 0.\bar{8}$$

$$\sigma_y^2 = \frac{1}{n}\sum_{j=0}^{4} n_{.j} y_j^2 - \bar{y}^2 = \frac{245}{24} - \bar{y}^2 = 0{,}9566$$

$$\sigma_{xy} = \frac{1}{n}\sum_{i,j=0}^{n} n_{ij} x_i y_j - \bar{x}\bar{y} = \frac{209}{24} - \bar{x}\bar{y} = 0{,}597\bar{2}$$

Por tanto, la recta de regresión de $Y$ sobre $X$ es:

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Longrightarrow y = 0{,}6719x + 1{,}25$$

La recta de regresión de $X$ sobre $Y$ es:

$$x - \bar{x} = \frac{\sigma_{xy}}{\sigma_y^2}(y - \bar{y}) \Longrightarrow x = 0{,}6243y + 0{,}7677$$

2. ¿Es apropiado suponer que existe una relación lineal entre las variables?

Para ver cómo de bueno es el ajuste, usamos el coeficiente de determinación.

$$r^2 = \frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2} = 0{,}4195$$

Además, vemos el coeficiente de correlación lineal:

$$r = +\sqrt{r} = 0{,}64766$$

Como $r^2$ dista de 1, no es un buen ajuste. En concreto, se explican el 41,92 % de los casos, menos de la mitad. Además, como $r$ también dista de 1, podemos ver que no hay una correlación lineal alta entre ambas variables.

3. Predecir, a partir de los resultados, el número de balanzas que puede esperarse en un puesto con seis dependientes. ¿Es fiable esta predicción?

Como se da una nueva modalidad de la variable $Y$, usamos la recta de regresión de $X$ sobre $Y$:

$$x = 0{,}6243y + 0{,}7677 = 0{,}6243(6) + 0{,}7677 = 4{,}5135$$

Por tanto, se predice que habrá entre 4 y 5 balanzas. No obstante, como se ha razonado en el apartado anterior, esta predicción no es muy fiable.

**Ejercicio 8.2.9.** Se eligen 50 matrimonios al azar y se les pregunta la edad de ambos al contraer matrimonio. Los resultados se recogen en la siguiente tabla, en la que $X$ denota la edad del hombre e $Y$ la de la mujer:

| $X\backslash Y$ | $c_i^x$ | $(10,20]$ | $(20,25]$ | $(25,30]$ | $(30,35]$ | $(35,40]$ | $n_{i.}$ |
|---|---|---|---|---|---|---|---|
| $c_j^y$ | | 15 | 22,5 | 27,5 | 32,5 | 37,5 | |
| $(15,18]$ | 16,5 | 3 | 2 | 3 | 0 | 0 | 8 |
| $(18,21]$ | 19,5 | 0 | 4 | 2 | 2 | 0 | 8 |
| $(21,24]$ | 22,5 | 0 | 7 | 10 | 6 | 1 | 24 |
| $(24,27]$ | 25,5 | 0 | 0 | 2 | 5 | 3 | 10 |
| $n_{.j}$ | | 3 | 13 | 17 | 13 | 4 | 50 |

Estudiar la interdependencia lineal entre ambas variables.

Sea la edad a la que el hombre contrajo matrimonio, $X$; y la edad a la que la mujer contrajo matrimonio, $Y$; dos variables estadísticas con población $n = 50$ y modalidades $I_1^x, \dots, I_3^x$ e $I_1^y, \dots, I_3^y$ respectivamente, con *distribución de frecuencias*:

$$\left\{(I_i^x, I_j^x), n_{ij}\right\}_{\substack{i=1,\dots,4 \\ j=1,\dots,5}}$$

Para estudiar la interdependencia lineal, es necesario calcular el coeficiente de correlación lineal:

$$r = \pm\sqrt{r^2} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Calculamos por tanto los valores necesarios:

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{4} c_i^x n_{i.} = \frac{1083}{50} = 21{,}66 \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{5} c_j^y n_{.j} = \frac{1377{,}5}{50} = 27{,}55$$

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=0}^{4} n_{i.}(c_i^x)^2 - \bar{x}^2} = \sqrt{\frac{23872{,}5}{50} - \bar{x}^2} = \sqrt{8{,}2944} = 2{,}88$$

$$\sigma_y = \sqrt{\frac{1}{n}\sum_{j=0}^{5} n_{.j}(c_j^y)^2 - \bar{y}^2} = \sqrt{\frac{39468{,}75}{50} - \bar{y}^2} = \sqrt{30{,}3725} = 5{,}5111$$

$$\sigma_{xy} = \frac{1}{n}\sum_{i=0}^{4}\sum_{j=0}^{5} n_{ij} c_i^x c_j^y - \bar{x}\bar{y} = \frac{30318{,}75}{50} - \bar{x}\bar{y} = 9{,}642$$

Por tanto, tenemos que

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = 0{,}6075$$

De este resultado, obtenemos que la correlación es positiva. Además, la interdependencia lineal no es muy alta, ya que el valor de $r$ no es muy cercano al 1. Por tanto, el nivel de correlación no termina de ser muy alto, sino que es más bien intermedio.

Además, el ajuste lineal no es adecuado, ya que $r^2 = 0{,}3690$, y este es más cercano al 0 que al 1. De hecho, solo se predicen el 36,9 % de los valores.

**Ejercicio 8.2.10.** Calcular el coeficiente de correlación lineal de dos variables cuyas rectas de regresión son:

1. $x + 4y = 1$
2. $x + 5y = 2$

Sean $X$ e $Y$ dos variables estadísticas con población $n$ y modalidades $x_1, \ldots, x_k$ e $y_1, \ldots, y_p$ respectivamente, con distribución de frecuencias absolutas

$$\{(x_i, y_j); n_{ij}\}_{\substack{i=1,\ldots,k \\ j=1,\ldots,p}}$$

Supongamos que la primera recta es la recta de regresión de $Y/X$, mientras que la segunda es la recta de regresión de $X/Y$. Posteriormente comprobaremos si esta suposición es correcta.

Como la primera recta es la recta de regresión de $Y/X$:

$$y = -\frac{1}{4}x + \frac{1}{4} = \frac{\sigma_{xy}}{\sigma_x^2}x - \frac{\sigma_{xy}}{\sigma_x^2}\bar{x} + \bar{y}$$

Por tanto,

$$-\frac{1}{4} = \frac{\sigma_{xy}}{\sigma_x^2} \qquad \frac{1}{4} = -\frac{\sigma_{xy}}{\sigma_x^2}\bar{x} + \bar{y} = \frac{1}{4}\bar{x} + \bar{y}$$

Como la segunda recta es la recta de regresión de $X/Y$:

$$x = -5y + 2 = \frac{\sigma_{xy}}{\sigma_y^2}x - \frac{\sigma_{xy}}{\sigma_y^2}\bar{y} + \bar{x}$$

Por tanto,

$$-5 = \frac{\sigma_{xy}}{\sigma_y^2} \qquad 2 = -\frac{\sigma_{xy}}{\sigma_y^2}\bar{y} + \bar{x} = 5\bar{y} + \bar{x}$$

El coeficiente de determinación lineal es, por tanto:

$$r^2 = \frac{\sigma_{xy}}{\sigma_x^2} \cdot \frac{\sigma_{xy}}{\sigma_y^2} = -\frac{1}{4} \cdot (-5) = \frac{5}{4}$$

Como $0 \leqslant r^2 \leqslant 1$, deducimos que esta suposición **no** es correcta. Por tanto, sabemos que la primera recta es la recta de regresión de $X/Y$, mientras que la segunda es la recta de regresión de $Y/X$.

Como la primera recta es la recta de regresión de $X/Y$:

$$x = -4y + 1 = \frac{\sigma_{xy}}{\sigma_y^2}y - \frac{\sigma_{xy}}{\sigma_y^2}\bar{y} + \bar{x}$$

Por tanto,

$$-4 = \frac{\sigma_{xy}}{\sigma_y^2} \qquad 1 = -\frac{\sigma_{xy}}{\sigma_y^2}\bar{y} + \bar{x} = 4\bar{y} + \bar{x}$$

Como la segunda recta es la recta de regresión de $Y/X$:

$$y = -\frac{1}{5}x + \frac{2}{5} = \frac{\sigma_{xy}}{\sigma_x^2}x - \frac{\sigma_{xy}}{\sigma_x^2}\bar{x} + \bar{y}$$

Por tanto,

$$-\frac{1}{5} = \frac{\sigma_{xy}}{\sigma_x^2} \qquad \frac{2}{5} = -\frac{\sigma_{xy}}{\sigma_x^2}\bar{x} + \bar{y} = \frac{1}{5}\bar{x} + \bar{y}$$

El coeficiente de determinación lineal es, por tanto:

$$r^2 = \frac{\sigma_{xy}}{\sigma_x^2} \cdot \frac{\sigma_{xy}}{\sigma_y^2} = -\frac{1}{5} \cdot (-4) = \frac{4}{5}$$

Como $0 \leqslant r^2 \leqslant 1$, deducimos que esta suposición **sí** es la correcta.

Por tanto, como se pide el coeficiente de regresión lineal, tenemos que:

$$r = -\sqrt{r^2} = -0{,}8944$$

donde es necesario que hemos tomado la raíz con signo negativo, ya que la covarianza, que determina el signo del coeficiente de regresión lineal, tiene signo negativo.

**Ejercicio 8.2.11.** Consideremos una distribución bidimensional en la que la recta de regresión de $Y$ sobre $X$ es $y = 5x - 20$, y $\sum y_j^2 n_{\cdot j} = 3240$. Supongamos, además, que la distribución marginal de $X$ es:

| $x_i$ | 3 | 5 | 8 | 9 |
|---|---|---|---|---|
| $n_{i\cdot}$ | 5 | 1 | 2 | 1 |

Determinar la recta de regresión de $X$ sobre $Y$, y la bondad de los ajustes lineales.

Sean $X$ y $Y$ dos variables estadísticas con población $n = 9$ y modalidades $x_1, \dots, x_4$ e $y_1, \dots, y_p$ respectivamente, con *distribución de frecuencias*:

$$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,4 \\ j=1,\dots,p}}$$

De la distribución marginal de $X$, obtengo directamente los siguientes dos resultados:

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{4} n_{i\cdot} x_i = \frac{45}{9} = 5$$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{4} n_{i\cdot} x_i^2 - \bar{x}^2 = \frac{279}{9} - \bar{x}^2 = 6$$

La ecuación de la recta de regresión de $Y$ sobre $X$ es:

$$y = ax + b = \frac{\sigma_{xy}}{\sigma_x^2}x - \frac{\sigma_{xy}}{\sigma_x^2}\bar{x} + \bar{y} = 5x - 20$$

Igualando términos:

$$a = \frac{\sigma_{xy}}{\sigma_x^2} = 5 \Longrightarrow \sigma_{xy} = 5 \cdot \sigma_x^2 = 30$$

$$b = -\frac{\sigma_{xy}}{\sigma_x^2}\bar{x} + \bar{y} = -20 \Longrightarrow \bar{y} = -20 + \frac{\sigma_{xy}}{\sigma_x^2}\bar{x} = -20 + 5\bar{x} = 5$$

Teniendo $\bar{y}$ y usando la información dada por el enunciado, calculamos la varianza de la variable $Y$:

$$\sigma_y^2 = \frac{1}{n}\sum_{j=0}^{p} y_j^2 n_{\cdot j} - \bar{y}^2 = \frac{3240}{9} - \bar{y}^2 = 335$$

Por tanto, calculamos la recta de regresión de $X$ sobre $Y$:

$$x - \bar{x} = \frac{\sigma_{xy}}{\sigma_y^2}(y - \bar{y}) \Longrightarrow x = 0{,}08955y + 4{,}5522$$

Para hallar la bondad de los ajustes lineales, calculo el coeficiente de determinación lineal.

$$\eta^2_{X/Y} = \eta^2_{Y/X} = r^2 = \frac{\sigma_{xy}^2}{\sigma_x \sigma_y} = 0{,}4478$$

Como $r^2$ es más cercano a 0 que al 1, entonces explica menos de la mitad de los datos. Por tanto, suponer una relación lineal no es lo ideal.

Además, el coeficiente de correlación lineal es $r = +\sqrt{r^2} = 0{,}6692$. Como no es muy cercano a 1, la interdependencia lineal entre ambas variables no es muy alta.

**Ejercicio 8.2.12.** De las estadísticas de “Tiempos de vuelo y consumos de combustible” de una compañía aérea, se han obtenido datos relativos a 24 trayectos distintos realizados por el avión DC-9. A partir de estos datos se han obtenido las siguientes medidas:

$$\sum y_i = 219{,}719 \qquad \sum y_i^2 = 2396{,}504 \qquad \sum x_i y_i = 349{,}486$$

$$\sum x_i = 31{,}470 \qquad \sum x_i^2 = 51{,}075 \qquad \sum x_i^2 y_i = 633{,}993$$

$$\sum x_i^4 = 182{,}977 \qquad \sum x_i^3 = 93{,}6$$

La variable $Y$ expresa el consumo total de combustible, en miles de libras, correspondiente a un vuelo de duración $X$ (el tiempo se expresa en horas, y se utilizan como unidades de orden inferior fracciones decimales de la hora).

1. Ajustar un modelo del tipo $Y = aX + b$. ¿Qué consumo total se estimará para un programa de vuelos compuesto de 100 vuelos de media hora, 200 de una hora y 100 de dos horas? ¿Es fiable esta estimación?

   Sea la duración del vuelo, $X$, en horas; y el consumo de combustible, $Y$, en miles de libras; dos variables estadísticas con población $n = 24$ y modalidades $x_1, \dots, x_{24}$ e $y_1, \dots, y_{24}$ respectivamente, con *distribución de frecuencias*:

   $$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,24 \\ j=1,\dots,24}}$$

   Como son vuelos distintos, supongo las siguientes frecuencias:

   $$n_{ij} = \delta_{ij} = \begin{cases} 1 & si \quad i = j \\ 0 & si \quad i \neq j \end{cases}$$

Para calcular la recta de regresión de $Y$ sobre $X$, calculo previamente algunos datos.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{24} n_{i.}x_i = \frac{1}{24}\sum x_i = 1{,}31125 \qquad \bar{y} = \frac{1}{n}\sum_{j=1}^{24} n_{.j}y_j = \frac{1}{24}\sum y_i = 9{,}15496$$

$$\sigma_{xy} = \frac{1}{n}\sum_{i=0}^{24} x_i y_i n_{ii} - \bar{x}\bar{y} = \frac{1}{24}\sum x_i y_i - \bar{x}\bar{y} = 2{,}557475$$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{24} n_{i.}x_i^2 - \bar{x}^2 = \frac{1}{24}\sum x_i^2 - \bar{x}^2 = 0{,}40875$$

$$\sigma_y^2 = \frac{1}{n}\sum_{j=1}^{24} n_{.j}y_j^2 - \bar{y}^2 = \frac{1}{24}\sum y_i^2 - \bar{y}^2 = 16{,}04104$$

Por tanto, la recta de regresión de $Y$ sobre $X$ es:

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Longrightarrow y = 6{,}25682x + 0{,}9507$$

Para ver la bondad del ajuste lineal, estudio $r^2$:

$$r^2 = \frac{\sigma_{xy}^2}{\sigma_x^2\sigma_y^2} = 0{,}997545$$

Estimamos ahora para el programa de vuelos pedido.

- Para un vuelo de media hora, tenemos un consumo de:

  $$y = 6{,}25682x + 0{,}9507 = 6{,}25682 \cdot 0{,}5 + 0{,}9507 = 4{,}07911 \text{ miles de libras.}$$

- Para un vuelo de una hora, tenemos consumo de:

  $$y = 6{,}25682x + 0{,}9507 = 6{,}25682 \cdot 1 + 0{,}9507 = 7{,}20752 \text{ miles de libras.}$$

- Para un vuelo de dos horas, tenemos un consumo de:

  $$y = 6{,}25682x + 0{,}9507 = 6{,}25682 \cdot 2 + 0{,}9507 = 13{,}46434 \text{ miles de libras.}$$

Por tanto, para el programa total de vuelos pedido, el consumo será de:

$$y_T = 100 \cdot 4{,}07911 + 200 \cdot 7{,}20752 + 100 \cdot 13{,}46434 = 3195{,}849 \text{ miles de libras.}$$

Como $r^2 = 0{,}997545 \approx 1$, vemos que es un ajuste de buena calidad y, por tanto, es una estimación muy fiable.

2. Ajustar un modelo del tipo $Y = a+bX+cX^2$. ¿Qué consumo total se estimará para el mismo programa de vuelos del apartado $a$)?

   Para encontrar los parámetros $a, b, c$ mediante el ajuste de mínimos cuadrados, hemos de resolver el siguiente sistema de ecuaciones ortonormales:

$$\left\{\begin{array}{rcrcrcr} m_{01} & = & a & + & bm_{10} & + & cm_{20} \\ m_{11} & = & am_{10} & + & bm_{20} & + & cm_{30} \\ m_{21} & = & am_{20} & + & bm_{30} & + & cm_{40} \end{array}\right.$$

   equivalentemente, y usando la definición de los momentos conjuntos respecto al origen, tenemos que el sistema de ecuaciones ortonormales es:

$$\left\{\begin{array}{rcrcrcr} \bar{y} & = & a & + & b\bar{x} & + & cm_{20} \\ m_{11} & = & a\bar{x} & + & bm_{20} & + & cm_{30} \\ m_{21} & = & am_{20} & + & bm_{30} & + & cm_{40} \end{array}\right.$$

   Calculo por tanto los momentos que faltan:

$$m_{30} = \frac{1}{n}\sum_{i,j=1}^{24} n_{ij}x_i^3y_j^0 = \frac{1}{24}\sum x_i^3 = 3{,}9 \qquad m_{40} = \frac{1}{n}\sum_{i,j=1}^{24} n_{ij}x_i^4y_j^0 = \frac{1}{24}\sum x_i^4 = 7{,}62404$$

$$m_{20} = \frac{1}{n}\sum_{i,j=1}^{24} n_{ij}x_i^2y_j^0 = \frac{1}{24}\sum x_i^2 = \frac{51{,}075}{0{,}24} = 2{,}1281$$

$$m_{11} = \frac{1}{n}\sum_{i,j=1}^{24} n_{ij}x_i^1y_j^1 = \frac{1}{24}\sum x_iy_i = \frac{349{,}486}{24} = 14{,}5619$$

$$m_{21} = \frac{1}{n}\sum_{i,j=1}^{24} n_{ij}x_i^2y_j^1 = \frac{1}{24}\sum x_i^2y_i = 26{,}416375$$

   Por tanto, el sistema queda:

$$\left\{\begin{array}{rcrcrcr} 9{,}15496 & = & a & + & 1{,}31125b & + & 2{,}1281c \\ 14{,}5619 & = & 1{,}31125a & + & 2{,}1281b & + & 3{,}9c \\ 26{,}416375 & = & 2{,}1281a & + & 3{,}9b & + & 7{,}62404c \end{array}\right.$$

   Resolviendo, obtenemos que

$$a = 0{,}79855 \qquad b = 6{,}54482 \qquad c = -0{,}10596$$

   Por tanto, tenemos que la parábola de regresión de $Y$ sobre $X$ es:

$$y = 0{,}79855 + 6{,}54482x - 0{,}10596x^2$$

   Estimamos ahora para el programa de vuelos pedido.

   - Para un vuelo de media hora, tenemos un consumo de:

$$y = 0{,}79855 + 6{,}54482 \cdot (0{,}5) - 0{,}10596 \cdot (0{,}5)^2 = 4{,}04447 \text{ miles de libras.}$$

- Para un vuelo de una hora, tenemos consumo de:

$$y = 6{,}25682x + 0{,}9507 = 6{,}25682 \cdot 1 + 0{,}9507 = 7{,}23741 \text{ miles de libras.}$$

- Para un vuelo de dos horas, tenemos un consumo de:

$$y = 6{,}25682x + 0{,}9507 = 6{,}25682 \cdot 2 + 0{,}9507 = 13{,}46435 \text{ miles de libras.}$$

Por tanto, para el programa total de vuelos pedido, el consumo será de:

$$y_T = 100 \cdot 4{,}04447 + 200 \cdot 7{,}23741 + 100 \cdot 13{,}46435 = 3198{,}364 \text{ miles de libras.}$$

3. ¿Cuál de los dos modelos se ajusta mejor? Razonar la respuesta.
   Ambos ajustes tienen el mismo coeficiente de determinación, por lo que se podría pensar en un primer momento que ambos son iguales. No obstante, al haber una tercera ecuación (ser de grado 2), incluye un tercer parámetro y entonces añade información. Por tanto, el ajuste parabólico se ajusta mejor[1].

   No obstante, podemos ver que las diferencias son insignificantes, ya que las predicciones están muy próximas debido al alto valor de $r^2$.

**Ejercicio 8.2.13.** La curva de Engel, que expresa el gasto en un determinado bien en función de la renta, adopta en ocasiones la forma de una hipérbola equilátera. Ajustar dicha curva a los siguientes datos, en los que $X$ denota la renta en miles de euros e $Y$ el gasto en euros. Cuantificar la bondad del ajuste:

| $X$ | 10 | 12,5 | 20 | 25 |
|---|---|---|---|---|
| $Y$ | 50 | 90 | 160 | 180 |
| $Z$ | 0,1 | 0,08 | 0,05 | 0,04 |

Sea la renta, $X$, en miles de euros; y el gasto en determinado bien, $Y$, en euros; dos variables estadísticas con población $n = 4$ y modalidades $x_1, \dots, x_4$ e $y_1, \dots, y_4$ respectivamente, con *distribución de frecuencias*:

$$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,4 \\ j=1,\dots,4}}$$

Como el ajuste es de la forma $y = \frac{a}{x} + b$, realizo el cambio de variable $z = \frac{1}{x}$. Por tanto, calculo la recta de regresión de $Y$ sobre $Z$.

Para calcular el coeficiente de regresión de $Y$ sobre $Z$, necesito calcular antes varios datos:

$$\bar{z} = \frac{1}{n}\sum_{i=0}^{4} n_{i.}z_i = \frac{0{,}27}{4} = 0{,}0675 \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{4} n_{.j}y_j = \frac{480}{4} = 120$$

$$\sigma_{zy} = \frac{1}{n}\sum_{i,j=0}^{4} n_{ij}z_iy_j - \bar{z}\bar{y} = \frac{1}{n}\sum_{i,j=0}^{4} n_{ij}\frac{y_j}{x_i} - \bar{z}\bar{y} = \frac{27{,}4}{4} - \bar{z}\bar{y} = -1{,}25$$

[1] En el caso de tener las distintas modalidades, podríamos calcular $\sigma^2_{ey}$ para cada ajuste y veríamos que, en el caso de la parábola, es mayor.

$$\sigma_z^2 = \frac{1}{n}\sum_{i=0}^{4} n_{i.}z_i^2 - \bar{z}^2 = \frac{1}{n}\sum_{i=0}^{4} \frac{1}{x_i^2} - \bar{z}^2 = \frac{0{,}0205}{4} - \bar{z}^2 = 5{,}6875 \cdot 10^{-4}$$

$$\sigma_y^2 = \frac{1}{n}\sum_{j=0}^{4} n_{.j}y_j^2 - \bar{y}^2 = \frac{68600}{4} - \bar{y}^2 = 2750$$

Por tanto, la recta de regresión de $Y$ sobre $Z$ es:

$$y - \bar{y} = \frac{\sigma_{zy}}{\sigma_z^2}(z - \bar{z}) \Longrightarrow y = -2197{,}8022z + 268{,}3516$$

Deshaciendo el cambio de variable, obtengo la curva de regresión hiperbólica de $Y$ sobre $X$:

$$y = -\frac{2197{,}8022}{x} + 268{,}3516$$

Para estudiar la bondad del ajuste hiperbólico, necesitamos interpretar el coeficiente de determinación.

$$r^2 = \frac{\sigma_{zy}^2}{\sigma_z^2\sigma_y^2} = 0{,}999$$

Por tanto, este ajusta el 99,9 % de los casos, por lo que el ajuste es prácticamente ideal, ya que $r^2 \approx 1$.

**Ejercicio 8.2.14.** Se dispone de la siguiente información referente al gasto en espectáculos ($Y$, en euros) y la renta disponible mensual ($X$, en cientos de euros) de 6 familias:

| $Y$ | 30 | 50 | 70 | 80 | 120 | 140 |
|---|---|---|---|---|---|---|
| $X$ | 9 | 10 | 12 | 15 | 22 | 32 |

Explicar el comportamiento de $Y$ por $X$ mediante:

1. Relación lineal.

   Sea la renta disponible mensual, $X$, en cientos de euros; y el gasto en espectáculos, $Y$, en euros; dos variables estadísticas con población $n = 6$ y modalidades $x_1, \dots, x_6$ e $y_1, \dots, y_6$ respectivamente, con *distribución de frecuencias*:

   $$\{(x_i, y_j), n_{ij}\}_{\substack{i=1,\dots,6\\ j=1,\dots,6}}$$

   Para calcular la recta de regresión, hemos de hallar en primer lugar otros datos.

   $$\bar{x} = \frac{1}{n}\sum_{i=0}^{6} x_i n_{i.} = \frac{100}{6} = 16.\bar{6} \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{6} y_j n_{.j} = \frac{490}{6} = 81.\bar{6}$$

   $$\sigma_{xy} = \frac{1}{n}\sum_{i,j=0}^{6} x_i y_j n_{ij} - \bar{x}\bar{y} = \frac{9930}{6} - \bar{x}\bar{y} = 293.\bar{8}$$

   $$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{6} n_{ij}x_i^2 - \bar{x}^2 = \frac{2058}{6} - \bar{x}^2 = 65.\bar{2}$$

$$\sigma_y^2 = \frac{1}{n}\sum_{j=0}^{6} n_{ij}y_j^2 - \bar{y}^2 = \frac{48700}{6} - \bar{y}^2 = 1447.\bar{2}$$

Por tanto, recta de regresión de $Y$ sobre $X$ es:

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Longrightarrow y = 4{,}5059x + 6{,}56729$$

Vara ver si es correcto el ajuste o no, calculamos el coeficiente de determinación:

$$\eta_{Y/X}^2 = \frac{\sigma_{ey}^2}{\sigma_y^2} = r^2 = \frac{\sigma_{xy}^2}{\sigma_x^2\sigma_y^2} = 0{,}91503$$

La varianza residual, al ser el ajuste lineal en los parámetros, queda:

$$\sigma_{ry}^2 = (1 - r^2)\sigma_y^2 = 122{,}97$$

Estos resultados los podemos ver en el diagrama de dispersión de la figura 8.25.

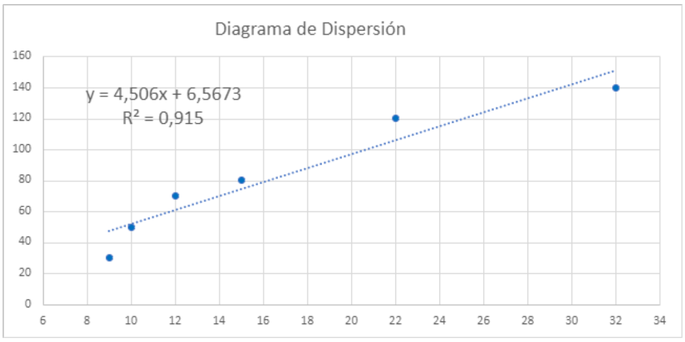


Figura 8.25: Diagrama de dispersión del ejercicio 8.2.14.1

2. Hipérbola equilátera.

   Como la recta será de la forma $y = \frac{a}{x} + b$, es necesario realizar en primer lugar el cambio de variable $z = \frac{1}{x}$. Queda por tanto de la siguiente manera:

| $Y$ | 30 | 50 | 70 | 80 | 120 | 140 |
|---|---|---|---|---|---|---|
| $Z$ | $0.\bar{1}$ | 0,1 | $0{,}08\bar{3}$ | $0{,}0\bar{6}$ | $0{,}0\overline{45}$ | 0,03125 |
| $X$ | 9 | 10 | 12 | 15 | 22 | 32 |

   Por tanto, calculamos la recta de regresión de $Y$ sobre $Z$:

$$\bar{z} = \frac{1}{n}\sum_{i=0}^{6} z_i n_{i.} = \frac{0{,}4378}{6} = 0{,}073 \qquad \bar{y} = \frac{1}{n}\sum_{j=0}^{6} y_j n_{.j} = \frac{490}{6} = 81.\bar{6}$$

$$\sigma_{zy} = \frac{1}{n}\sum_{i,j=0}^{6} z_i y_j n_{ij} - \bar{z}\bar{y} = \frac{29{,}3295}{6} - \bar{x}\bar{y} = -1{,}0734$$

$$\sigma_z^2 = \frac{1}{n}\sum_{i=0}^{6} n_{ij} z_i^2 - \bar{z}^2 = \frac{0{,}036777}{6} - \bar{z}^2 = 8{,}00541 \cdot 10^{-4}$$

$$\sigma_y^2 = \frac{1}{n}\sum_{j=0}^{6} n_{ij} y_j^2 - \bar{y}^2 = \frac{48700}{6} - \bar{y}^2 = 1447.\bar{2}$$

Por tanto, recta de regresión de $Y$ sobre $Z$ es:

$$y - \bar{y} = \frac{\sigma_{zy}}{\sigma_z^2}(z - \bar{z}) \Longrightarrow y = -1340{,}843z + 179{,}55$$

Deshaciendo el cambio de variable:

$$y = -\frac{1340{,}843}{x} + 179{,}55$$

Para ver si es correcto el ajuste o no, calculamos el coeficiente de determinación:

$$r^2 = \frac{\sigma_{zy}^2}{\sigma_z^2 \sigma_y^2} = 0{,}98102$$

La varianza residual, al ser el ajuste lineal en los parámetros, queda:

$$\sigma_{ry}^2 = (1 - r^2)\sigma_y^2 = 27{,}46$$

3. Curva potencial.

   Como la curva será de la forma $y = bx^a$, es necesario realizar en primer lugar un cambio de variable. Para ello, aplico el ln y establecemos $y' = \ln y$, $b' = \ln b$, $x' = \ln x$. Queda por tanto de la siguiente manera:

| $Y$ | 30 | 50 | 70 | 80 | 120 | 140 |
|---|---|---|---|---|---|---|
| $X$ | 9 | 10 | 12 | 15 | 22 | 32 |
| $Y'$ | 3,4012 | 3,912 | 4,2485 | 4,38203 | 4,7875 | 4,9416 |
| $X'$ | 2,197 | 2,3026 | 2,4849 | 2,7086 | 3,091 | 2,4849 |

   Por tanto, calculamos la recta de regresión de $Y'$ sobre $X'$:

$$\bar{x'} = \frac{1}{n}\sum_{i=0}^{6} x_i' n_{i.} = \frac{1}{n}\sum_{i=0}^{6} \ln(x_i) n_{i.} = \frac{1}{n}\ln\left(\prod_{i=0}^{6} x_i\right) = \frac{\ln 11404800}{6} = 2{,}70826$$

$$\bar{y'} = \frac{1}{n}\sum_{j=0}^{6} y_j' n_{.j} = \frac{1}{n}\sum_{j=0}^{6} \ln(y_j) n_{.j} = \frac{1}{n}\ln\left(\prod_{j=0}^{6} y_j\right) = \frac{\ln 14112 \cdot 10^7}{6} = 4{,}2788$$

$$\sigma_{x'y'} = \frac{1}{n}\sum_{i,j=0}^{6} x_i' y_j' n_{ij} - \bar{x'}\bar{y'} = \frac{70{,}8296}{6} - \bar{x'}\bar{y'} = 0{,}2168$$

$$\sigma_{x'}^2 = \frac{1}{n}\sum_{i=0}^{6} n_{ij}(x_i')^2 - \bar{x'}^2 = \frac{45{,}20386}{6} - \bar{x'}^2 = 0{,}1993$$

$$\sigma_{y'}^2 = \frac{1}{n}\sum_{j=0}^{6} n_{ij}(y'_j)^2 - \bar{y'}^2 = \frac{111{,}4638}{6} - \bar{y'}^2 = 0{,}26918$$

Por tanto, recta de regresión de $Y'$ sobre $X'$ es:

$$y' - \bar{y'} = \frac{\sigma_{x'y'}}{\sigma_{x'}^2}(x' - \bar{x'}) \Longrightarrow y' = 1{,}0878x' + 1{,}3327$$

Deshaciendo el cambio de variable:

$$\ln y = 1{,}0878 \ln x + 1{,}3327 \Longrightarrow y = e^{1{,}0878 \ln x + 1{,}3327} = e^{1{,}3327}x^{1{,}0878} = 3{,}7914x^{1{,}0878}$$

$$y = 3{,}7914x^{1{,}0878}$$

Para estudiar la bondad del ajuste, calculamos la varianza residual teniendo en cuenta que el ajuste **no** es lineal en los parámetros:

$$\begin{aligned}\sigma_{ry}^2 = \frac{1}{n}\sum_{i,j=1}^{6} n_{ij}(y_j - f(x_i))^2 &= \\ &= \frac{(30-41{,}3833)^2 + (50-46{,}4087)^2 + (70-56{,}5891)^2}{6} + \\ &+ \frac{(80-72{,}1359)^2 + (120-109{,}4176)^2 + (140-164{,}4757)^2}{6} = \\ &= \frac{1095{,}2203}{6} = 182{,}5367\end{aligned}$$

Estos resultados los podemos ver en el diagrama de dispersión de la figura 8.26.

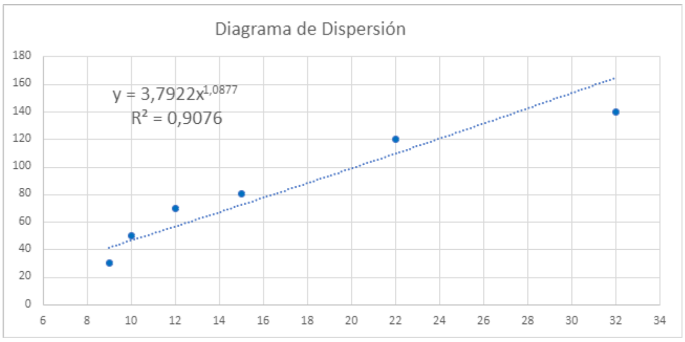


Figura 8.26: Diagrama de dispersión del ejercicio 8.2.14.3

4. Curva exponencial.

   Como la curva será de la forma $y = ba^x$, es necesario realizar en primer lugar un cambio de variable. Para ello, aplico el ln y establecemos $y' = \ln y$, $b' = \ln b$. El cambio de variable, por tanto, es:

$$\ln y = \ln(ba^x) = \ln b + \ln(a)x \Longrightarrow y' = b' + a'x$$

Queda por tanto de la siguiente manera:

| $Y$ | 30 | 50 | 70 | 80 | 120 | 140 |
|---|---|---|---|---|---|---|
| $X$ | 9 | 10 | 12 | 15 | 22 | 32 |
| $Y'$ | 3,4012 | 3,912 | 4,2485 | 4,38203 | 4,7875 | 4,9416 |

Por tanto, calculamos la recta de regresión de $Y'$ sobre $X$:

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{6} x_i n_{i.} = \frac{100}{6} = 16.\bar{6}$$

$$\bar{y'} = \frac{1}{n}\sum_{j=0}^{6} y'_j n_{.j} = \frac{1}{n}\sum_{j=0}^{6} \ln(y_j) n_{.j} = \frac{1}{n}\ln\left(\prod_{j=0}^{6} y_j\right) = \frac{\ln 14112 \cdot 10^7}{6} = 4{,}2788$$

$$\begin{aligned}\sigma_{xy'} &= \frac{1}{n}\sum_{i,j=0}^{6} x_i y'_j n_{ij} - \bar{x}\bar{y'} = \frac{1}{n}\sum_{i,j=0}^{6} x_i \ln y_j - \bar{x}\bar{y'} = \frac{1}{n}\sum_{i,j=0}^{6} \ln(y_j^{x_i}) - \bar{x}\bar{y'} \\ &= \frac{1}{n}\ln\left(\prod_{i,j=0}^{6} y_j^{x_i}\right) - \bar{x}\bar{y'} = \frac{449{,}89945}{6} - \bar{x}\bar{y'} = 3{,}6699\end{aligned}$$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=0}^{6} n_{ij} x_i^2 - \bar{x}^2 = \frac{2058}{6} - \bar{x}^2 = 65.\bar{2}$$

$$\sigma_{y'}^2 = \frac{1}{n}\sum_{j=0}^{6} n_{ij}(y'_j)^2 - \bar{y'}^2 = \frac{111{,}4638}{6} - \bar{y'}^2 = 0{,}26918$$

Por tanto, recta de regresión de $Y'$ sobre $X$ es:

$$y' - \bar{y'} = \frac{\sigma_{xy'}}{\sigma_x^2}(x - \bar{x}) \Longrightarrow y' = 0{,}05627x + 3{,}341$$

Deshaciendo el cambio de variable:

$$\ln y = 0{,}05627x + 3{,}341 \Longrightarrow y = e^{0{,}05627x+3{,}341} = e^{3{,}341}e^{0{,}05627x} = 28{,}2475e^{0{,}05627x}$$

$$y = 28{,}2475e^{0{,}05627x} \Longrightarrow y = 28{,}2475 \cdot 1{,}05788^x$$

Para estudiar la bondad del ajuste, calculamos la varianza residual teniendo en cuenta que el ajuste **no** es lineal en los parámetros:

$$\begin{aligned}\sigma_{ry}^2 &= \frac{1}{n}\sum_{i,j=1}^{6} n_{ij}(y_j - f(x_i))^2 = \\ &= \frac{(30-46{,}87137)^2 + (50-49{,}5843)^2 + (70-55{,}4903)^2}{6} + \\ &+ \frac{(80-65{,}69406)^2 + (120-97{,}4049)^2 + (140-170{,}9799)^2}{6} = \\ &= \frac{2170{,}2999}{6} = 361{,}71667\end{aligned}$$

Estos resultados los podemos ver en el diagrama de dispersión de la figura 8.27.

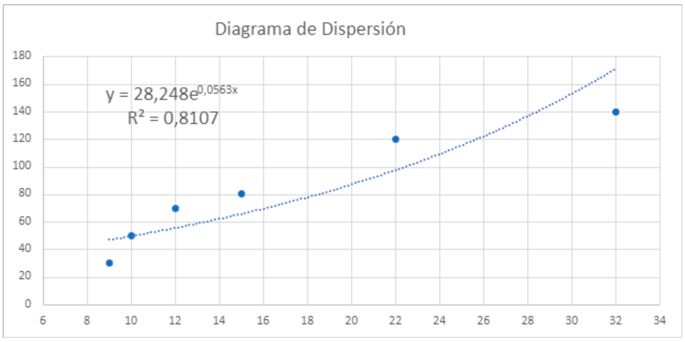


Figura 8.27: Diagrama de dispersión del ejercicio 8.2.14.4

¿Qué ajuste es más adecuado?

Como el menor valor de $\sigma^2_{ry}$ es para el ajuste hiperbólico, concluimos que este es el ajuste más adecuado, ya que deja sin explicar la menor cantidad de datos. En ese caso, vemos que $r^2 = 0{,}98102$, por lo que explica adecuadamente más del 98 % de los casos.

## 8.3. Espacios de Probabilidad

**Ejercicio 8.3.1.** Durante un año, las personas de una ciudad utilizan 3 tipos de transportes: metro (M), autobús (A), y coche particular (C). Las probabilidades de que durante el año hayan usado unos u otros transportes son:

M: 0.3; A: 0.2; C: 0.15; M y A: 0.1; M y C: 0.05; A y C: 0.06; M, A y C: 0.01

Calcular las probabilidades siguientes:

1. Que una persona viaje en metro y no en autobús.

$$P(M \cap \bar{A}) = P(M - A) = P(M) - P(A \cap M) = 0{,}3 - 0{,}1 = 0{,}2$$

2. Que una persona tome al menos dos medios de transporte.

El suceso descrito es $S = (M \cap A) \cup (M \cap C) \cup (A \cap C) \cup (A \cap M \cap C)$. Como son sucesos incompatibles, tenemos que:

$$\begin{aligned} P(S) &= P[(M \cap A) \cup (M \cap C) \cup (A \cap C) \cup (A \cap M \cap C)] = \\ &= P[(M \cap A) \cup (M \cap C)] + P[(A \cap C) \cup (A \cap M \cap C)] \\ &\qquad - P[[(M \cap A) \cup (M \cap C)] \cap [(A \cap C) \cup (A \cap M \cap C)]] = \\ &= P(M \cap A) + P(M \cap C) - P(M \cap A \cap C) + P(A \cap C) - P(A \cap C \cap M) = \\ &= 0{,}1 + 0{,}05 - 0{,}01 + 0{,}06 - 0{,}01 = 0{,}19 \end{aligned}$$

donde he usado que:

$$(A \cap C) \cup (A \cap M \cap C) = A \cap C, \text{ ya que } (A \cap M \cap C) \subset (A \cap C)$$

$$[(M \cap A) \cup (M \cap C)] \cap [(A \cap C) \cup (A \cap M \cap C)] = [M \cap (A \cup C)] \cap (A \cap C) = A \cap C \cap M$$

3. Que una persona viaje en metro o en coche, pero no en autobús.

$$\begin{aligned} P[(M \cup C) \cap \bar{A}] &= P[(M \cup C) - A] = P(M \cup C) - P[(M \cup C) \cap A] = \\ &= P(M) + P(C) - P(M \cap C) - P[(M \cup C) \cap A] = \\ &= 0{,}3 + 0{,}15 - 0{,}05 - 0{,}15 = 0{,}25 \end{aligned}$$

donde he tenido que usar que:

$$\begin{aligned} P[(M \cup C) \cap A] &= P[(M \cap A) \cup (C \cap A)] = \\ &= P(M \cap A) + P(C \cap A) - P[(M \cap A) \cap (C \cap A)] = \\ &= P(M \cap A) + P(C \cap A) - P(A \cap M \cap C) = \\ &= 0{,}1 + 0{,}06 - 0{,}01 = 0{,}15 \end{aligned}$$

4. Que viaje en metro, o bien en autobús y en coche.

$$\begin{aligned} P[M \cup (A \cap C)] &= P(M) + P(A \cap C) - P(A \cap C \cap M) = \\ &= 0{,}3 + 0{,}06 - 0{,}01 = 0{,}35 \end{aligned}$$

5. Que una persona vaya a pie.

   Se presupone que ir a pie es la única alternativa a los tres medios de transporte descritos. En ese caso,

$$\begin{aligned}
P(\bar{A}\cap\bar{M}\cap\bar{C}) &= P\left(\overline{A\cup M\cup C}\right) = 1 - P(A\cup M\cup C) = \\
&= 1 - [P(A) + P(M\cup C) - P[A\cap(M\cup C)]] = \\
&= 1 - [P(A) + P(M) + P(C) - P(M\cap C) - P[A\cap(M\cup C)]] = \\
&= 1 - [0{,}2 + 0{,}3 + 0{,}15 - 0{,}05 - 0{,}15] = 1 - 0{,}45 = 0{,}55
\end{aligned}$$

**Ejercicio 8.3.2.** Sean $A, B$ y $C$ tres sucesos de un espacio probabilístico $(\Omega, \mathcal{A}, P)$, tales que $P(A) = 0{,}4$, $P(B) = 0{,}2$, $P(C) = 0{,}3$, $P(A\cap B) = 0{,}1$ y $(A\cup B)\cap C = \emptyset$. Calcular las probabilidades de los siguientes sucesos:

1. Sólo ocurre $A$,

$$\begin{aligned}
P((A-B)-C) = P(A-B) - \cancelto{0}{P((A-B)\cap C)} &= P(A-B) = \\
&= P(A) - P(A\cap B) = 0{,}4 - 0{,}1 = 0{,}3
\end{aligned}$$

   donde he empleado que $(A-B)\cap C = \emptyset$. Esto se debe a que:

$$\emptyset = (A\cup B)\cap C = (A\cap C)\cup(B\cap C) \Longrightarrow A\cap C = B\cap C = \emptyset$$

   Como $A\cap C = \emptyset$ y $A - B\subset A$, entonces $[(A-B)\cap C]\subset[A\cap C] = \emptyset$.

2. Ocurren los tres sucesos,

   Teniendo en cuenta que $B\cap C = \emptyset$,

$$P(A\cap B\cap C) = P(A\cap\emptyset) = P(\emptyset) = 0$$

3. Ocurren A y B pero no C,

$$P((A\cap B) - C) = P(A\cap B) - \cancelto{0}{P(A\cap B\cap C)} = P(A\cap B) = 0{,}1$$

4. Por lo menos dos ocurren,

   Teniendo en cuenta que $A\cap C = B\cap C = \emptyset$, tenemos que:

$$P[(A\cap B)\cup(A\cap C)\cup(B\cap C)\cup(A\cap B\cap C)] = P(A\cap B) = 0{,}1$$

5. Ocurren dos y no más,

   Sea $S$ el suceso,

$$\begin{aligned}
P(S) &= P\left[((A\cap B)\cup \cancelto{\emptyset}{(A\cap C)} \cup \cancelto{\emptyset}{(B\cap C)})\cap(\overline{A\cap B\cap C})\right] \\
&= P[A\cap B\cap(\overline{A\cap B\cap C})] = P(A\cap B\cap\overline{\emptyset}) = \\
&= P(A\cap B\cap\Omega) = P(A\cap B) = 0{,}1
\end{aligned}$$

6. No ocurren más de dos,

   Esto suceso equivale a que no ocurran los tres, es decir, $\overline{A\cap B\cap C}$.

   $$P\left(\overline{A\cap B\cap C}\right) = 1 - P(A\cap B\cap C) = 1 - P(\emptyset) = 1$$

7. Ocurre por lo menos uno,

   Cabe destacar que este suceso es:

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

   $$\begin{aligned} P(A\cup B\cup C) &= P(A\cup B) + P(C) - \cancelto{0}{P[(A\cup B)\cap C]} = \\ &= P(A) + P(B) + P(C) - P(A\cap B) = \\ &= 0{,}4 + 0{,}2 + 0{,}3 - 0{,}1 = 0{,}8 \end{aligned} \tag{8.1}$$

8. Ocurre sólo uno,

   Sea $S_i$ el suceso en el que sólo ocurre el suceso $i$, es decir,

   $$S_a = A\cap\bar{B}\cap\bar{C} \qquad S_b = \bar{A}\cap B\cap\bar{C} \qquad S_c = \bar{A}\cap\bar{B}\cap C$$

   Además, como $A\cap C = B\cap C = \emptyset$, tenemos que:

   $$S_a = A\cap\bar{B} \qquad S_b = \bar{A}\cap B \qquad S_c = C$$

   Como además tenemos que dichos sucesos son incompatibles, ya que no pueden ocurrir a la vez, tenemos que:

   $$\begin{aligned} P(S_a\cup S_b\cup S_c) &= P(S_a) + P(S_b) + P(S_c) = \\ &= P(A\cap\bar{B}) + P(\bar{A}\cap B) + P(C) = P(A-B) + P(B-A) + P(C) = \\ &= P(A) + P(B) + P(C) - P(A\cap B) - P(B\cap A) = \\ &= 0{,}4 + 0{,}2 + 0{,}3 - 2\cdot 0{,}1 = 0{,}7 \end{aligned}$$

9. No ocurre ninguno.

   $$P(\bar{A}\cap\bar{B}\cap\bar{C}) = P(\overline{A\cup B\cup C}) = 1 - P(A\cup B\cup C) \overset{Ec.\ 8{,}1}{=} 1 - 0{,}8 = 0{,}2$$

**Ejercicio 8.3.3.** Se sacan dos bolas sucesivamente sin devolución de una urna que contiene 3 bolas rojas distinguibles y 2 blancas distinguibles.

1. Describir el espacio de probabilidad asociado a este experimento.

   Sea el experimento aleatorio sacar dos bolas sucesivamente sin devolución de una urna que contiene 3 bolas rojas distinguibles y 2 blancas distinguibles.

   Sean $r_i$, $(i = 1,2,3)$ sacar una de las tres bolas rojas y $b_j$, $(j = 1,2)$ sacar una de las dos bolas blancas. El espacio de probabilidad viene dado por:

   $$\begin{aligned} \Omega = \{r_ir_j \mid i,j = 1,2,3; i\neq j\} &\cup \{r_ib_j \mid i = 1,2,3; j = 1,2\}\cup \\ &\cup \{b_ir_j \mid i = 1,2; j = 1,2,3\} \cup \{b_ib_j \mid i,j = 1,2; i\neq j\} \end{aligned}$$

2. Descomponer en sucesos elementales los sucesos: *la primera bola es roja*, *la segunda bola es blanca* y calcular la probabilidad de cada uno de ellos.

   El suceso *la primera bola es roja* se descompone como:

$$1^a\text{roja} = \{r_ir_j \mid i,j=1,2,3; i\neq j\} \cup \{r_ib_j \mid i=1,2,3; j=1,2\}$$

   La cantidad de sucesos totales de mi experimento son:

$$V_{5,2} = \frac{5!}{(5-2)!} = 5\cdot 4 = 20$$

   La cantidad de sucesos posibles en los que las dos bolas son rojas son:

$$V_{3,2} = \frac{3!}{(3-2)!} = 6$$

   La cantidad de sucesos posibles de la forma $\{r_ib_j\} \mid i=1,2,3,\ j=1,2$ son:

$$6 \Longrightarrow (r_1b_1, r_2b_1, r_3b_1, r_1b_2, r_2b_2, r_3b_2)$$

   Por tanto, usando la Ley de Laplace y sabiendo que todos los sucesos elementales son incompatibles, tenemos que:

$$\begin{aligned}P(1^a\text{roja}) &= P[\{r_ir_j \mid i,j=1,2,3; i\neq j\} \cup \{r_ib_j \mid i=1,2,3; j=1,2\}] = \\ &= P[\{r_ir_j \mid i,j=1,2,3; i\neq j\}] + P[\{r_ib_j \mid i=1,2,3; j=1,2\}] = \\ &= \frac{6}{20}+\frac{6}{20} = \frac{6}{10} = \frac{3}{5} = 0{,}6\end{aligned}$$

   Por otro lado, el suceso *la segunda bola es blanca* se descompone como:

$$2^a\text{blanca} = \{b_ib_j \mid i,j=1,2; i\neq j\} \cup \{r_ib_j \mid i=1,2,3; j=1,2\}$$

   La cantidad de sucesos posibles en los que las dos bolas son blancas son:

$$V_{2,2} = \frac{2!}{(2-2)!} = 2$$

   Por tanto, usando la Ley de Laplace y sabiendo que todos los sucesos elementales son incompatibles, tenemos que:

$$\begin{aligned}P(2^a\text{blanca}) &= P[\{b_ib_j \mid i,j=1,2; i\neq j\} \cup \{r_ib_j \mid i=1,2,3; j=1,2\}] = \\ &= P[\{b_ib_j \mid i,j=1,2; i\neq j\}] + P[\{r_ib_j \mid i=1,2,3; j=1,2\}] = \\ &= \frac{2}{20}+\frac{6}{20} = \frac{8}{20} = \frac{2}{5} = 0{,}4\end{aligned}$$

**Ejercicio 8.3.4.** Una urna contiene $a$ bolas blancas y $b$ bolas negras. ¿Cuál es la probabilidad de que al extraer dos bolas simultáneamente sean de distinto color?

Sea $N$ sacar una bola negra y $B$ sacar una bola blanca. Tenemos que:

$$\Omega = \{NN, BB, NB\}$$

En primer lugar, tenemos en cuenta no influye el orden en el que sacan las bolas. Además, como una misma bola no se puede sacar dos veces, estamos ante combinaciones sin repetición.

Los casos totales, sabiendo que tengo que hacer grupos de 2 de un total de $(a+b)$ elementos, son:

$$C_{2,a+b} = \binom{a+b}{2} = \frac{(a+b)!}{2!(a+b-2)!} = \frac{(a+b)(a+b-1)}{2}$$

Calculamos en primer lugar la probabilidad de los siguientes sucesos:

1. dos bolas negras,

   En este caso, el suceso es $\{NN\} \in \mathcal{A}$.

   Los casos favorables son las combinaciones de 2 elementos entre un total de $b$ bolas negras, es decir,

   $$C_{2,b} = \binom{b}{2} = \frac{b!}{2! \cdot (b-2)!} = \frac{b(b-1)}{2}$$

   Por tanto, usando la regla de Laplace, tenemos que:

   $$P(NN) = \frac{C_{2,b}}{C_{2,a+b}} = \frac{b(b-1)}{(a+b)(a+b-1)}$$

2. dos bolas blancas,

   En este caso, el suceso es $\{BB\} \in \mathcal{A}$.

   Los casos favorables son las combinaciones de 2 elementos entre un total de $a$ bolas blancas, es decir,

   $$C_{2,a} = \binom{a}{2} = \frac{a!}{2! \cdot (a-2)!} = \frac{a(a-1)}{2}$$

   Por tanto, usando la regla de Laplace, tenemos que:

   $$P(BB) = \frac{C_{2,a}}{C_{2,a+b}} = \frac{a(a-1)}{(a+b)(a+b-1)}$$

Por tanto, calculamos ahora la probabilidad de que las bolas extraídas sean de distinto color, es decir, $P(NB)$. Como los sucesos del espacio muestral son mutuamente excluyentes, tenemos que:

$$1 = P(\Omega) = P(NN \cup BB \cup NB) = P(NN) + P(BB) + P(NB) \Longrightarrow$$
$$\Longrightarrow P(NB) = 1 - \frac{a(a-1)}{(a+b)(a+b-1)} - \frac{b(b-1)}{(a+b)(a+b-1)} =$$
$$= \frac{(a+b)(a+b-1) - a(a-1) - b(b-1)}{(a+b)(a+b-1)} = \frac{2ab}{(a+b)(a+b-1)}$$

Por tanto, tengo que:

$$P(NB) = \frac{2ab}{(a+b)(a+b-1)}$$

*Observación.* Como podemos ver, las probabilidades coinciden con las calculadas en el ejercicio 8.3.5.

**Ejercicio 8.3.5.** Una urna contiene 5 bolas blancas y 3 rojas. Se extraen 2 bolas simultáneamente. Calcular la probabilidad de obtener los siguientes sucesos.

Sea $R$ sacar una bola roja y $B$ sacar una bola blanca. Tenemos que:

$$\Omega = \{RR, BB, RB\}$$

En primer lugar, tenemos en cuenta no influye el orden en el que sacan las bolas. Además, como una misma bola no se puede sacar dos veces, estamos ante combinaciones sin repetición.

Los casos totales, sabiendo que tengo que hacer grupos de 2 de un total de 8 elementos, son:

$$C_{2,8} = \binom{8}{2} = \frac{8!}{2!(8-2)!} = \frac{8 \cdot 7 \cdot 6!}{2 \cdot 6!} = 4 \cdot 7 = 28$$

1. dos bolas rojas,

   En este caso, el suceso es $\{RR\} \in \mathcal{A}$.

   Los casos favorables son las combinaciones de 2 elementos entre un total de 3 bolas rojas, es decir,

   $$C_{2,3} = \binom{3}{2} = \frac{3!}{2! \cdot 1!} = 3$$

   Por tanto, usando la regla de Laplace, tenemos que:

   $$P(RR) = \frac{C_{2,3}}{C_{2,8}} = \frac{3}{28} \approx 0{,}1071$$

2. dos bolas blancas,

   En este caso, el suceso es $\{BB\} \in \mathcal{A}$.

   Los casos favorables son las combinaciones de 2 elementos entre un total de 5 bolas blancas, es decir,

   $$C_{2,5} = \binom{5}{2} = \frac{5!}{2! \cdot 3!} = \frac{5 \cdot 4}{2} = 10$$

   Por tanto, usando la regla de Laplace, tenemos que:

   $$P(BB) = \frac{C_{2,5}}{C_{2,8}} = \frac{10}{28} = \frac{5}{14} \approx 0{,}357$$

3. una blanca y otra roja.

   En este caso, se pide la probabilidad de $P(RB)$. Como los sucesos del espacio muestral son mutuamente excluyentes, tenemos que:

$$1 = P(\Omega) = P(RR \cup BB \cup RB) = P(RR) + P(BB) + P(RB) \Longrightarrow$$
$$\Longrightarrow P(RB) = 1 - P(RR) - P(BB) = 1 - \frac{3}{28} - \frac{5}{14} = \frac{15}{28}$$

   Por tanto, tengo que:

$$P(RB) = \frac{15}{28} \approx 0{,}5357$$

**Ejercicio 8.3.6.** En una lotería de 100 billetes hay 2 que tienen premio.

1. ¿Cuál es la probabilidad de ganar al menos un premio si se compran 12 billetes?

   Estamos ante una situación en la que no importa el orden de los billetes y es sin repetición, ya que a comprar un billete ya hay uno menos disponible. Por tanto, estamos ante combinaciones.

   El número total de combinaciones posibles es:

$$C_{100,12} = \binom{100}{12} = \frac{100!}{12! \cdot 88!}$$

   Sea $A$ el suceso de ganar al menos con un billete. Por tanto, $\bar{A}$ es el suceso de no haber comprado ningún billete premiado.

   Las combinaciones no premiadas son:

$$C_{98,12} = \binom{98}{12} = \frac{98!}{12! \cdot 86!}$$

   Por tanto,

$$P(A) = 1 - P(\bar{A}) = 1 - \frac{C_{98,12}}{C_{100,12}} = 1 - \frac{98! \cdot 12! \cdot 88!}{100! \cdot 12! \cdot 86!} = 1 - \frac{88 \cdot 87}{100 \cdot 99} =$$
$$= 1 - \frac{58}{75} = \frac{17}{75} = 0{,}22\bar{6}$$

2. ¿Cuantos billetes habría que comprar para que la probabilidad de ganar al menos un premio sea mayor que $4/5$?

   Sea $x$ el número de billetes comprado. Las combinaciones totales son:

$$C_{100,x} = \frac{100!}{x!(100-x)!}$$

   El número de combinaciones no premiadas son:

$$C_{98,x} = \frac{98!}{x!(98-x)!}$$

Sea $A$ el suceso de ganar al menos un premio, por lo que sea $\bar{A}$ no ganar ningún premio. Como $P(A) \geqslant \frac{4}{5} = 0{,}8$, tenemos que:

$$P(A) \geqslant 0{,}8 \Longleftrightarrow 1 - P(\bar{A}) \geqslant 0{,}8 \Longleftrightarrow P(\bar{A}) \leqslant 0{,}2$$

Además, por la Regla de Laplace, tenemos que:

$$P(\bar{A}) = \frac{C_{98,x}}{C_{100,x}} = \frac{98! \cdot x!(100-x)!}{100! \cdot x!(98-x)!} = \frac{(100-x)(99-x)}{100 \cdot 99} \leqslant 0{,}2 \Longleftrightarrow$$
$$\Longleftrightarrow 100 \cdot 99 - 100x - 99x + x^2 \leqslant 0{,}2(100 \cdot 99) \Longleftrightarrow x^2 - 199x + 7920 \leqslant 0$$

Las raíces de esa parábola son $x_1 = 144$, $x_2 = 55$. Además, como se trata de una parábola cóncava hacia arriba, tenemos que es $\leqslant 0 \Longleftrightarrow x \in [55, 144]$.

No obstante, como tenemos que el número máximo de billetes disponible es de 100, tenemos que:

$$P(A) \geqslant 0{,}8 \Longleftrightarrow x \geqslant 55$$

Por tanto, habría que comprar, como mínimo, 55 billetes.

**Ejercicio 8.3.7.** Se consideran los 100 primeros números naturales. Se sacan 3 al azar.

1. Calcular la probabilidad de que en los 3 números obtenidos no exista ningún cuadrado perfecto.

   Llamemos a este suceso el *suceso* $A$. Además, tenemos en cuenta que hay 10 cuadrados perfectos en $\mathbb{N} \cap [1, 100]$.

   Como el orden en el que se sacan no importa, estamos ante combinaciones. Es necesario distinguir entre los casos en los que hay repetición y los que no:

   - Si no hay repetición:
     Los casos posibles son:

     $$C_{100,3} = \binom{100}{3} = \frac{100!}{3! \cdot 97!}$$

     Los casos en los que no hay cuadrados perfectos son las combinaciones posibles de 3 elementos de un conjunto de 90 elementos $(100 - 10)$ sin repetición:

     $$C_{90,3} = \binom{90}{3} = \frac{90!}{3! \cdot 87!}$$

     Por tanto, usando la Ley de Laplace, tenemos que:

     $$P(A) = \frac{C_{90,3}}{C_{100,3}} = \frac{90! \cdot 3! \cdot 97!}{100! \cdot 3! \cdot 87!} = \frac{90 \cdot 89 \cdot 88 \cdot 87! \cdot 97!}{100 \cdot 99 \cdot 98 \cdot 97! \cdot 87!} =$$
     $$= \frac{90 \cdot 89 \cdot 88}{100 \cdot 99 \cdot 98} = \frac{178}{245} \approx 0{,}72652$$

- Si sí hay repetición:
  Los casos posibles son:

$$CR_{100,3} = \binom{100+3-1}{3} = \binom{102}{3} = \frac{102!}{3! \cdot 99!}$$

  Los casos en los que no hay cuadrados perfectos son las combinaciones posibles de 3 elementos de un conjunto de 90 elementos $(100-10)$ con repetición:

$$CR_{90,3} = \binom{90+3-1}{3} = \binom{92}{3} = \frac{92!}{3! \cdot 89!}$$

  Por tanto, usando la Ley de Laplace, tenemos que:

$$P(A) = \frac{CR_{90,3}}{CR_{100,3}} = \frac{92! \cdot 3! \cdot 99!}{102! \cdot 3! \cdot 89!} = \frac{92 \cdot 91 \cdot 90 \cdot 89! \cdot 99!}{102 \cdot 101 \cdot 100 \cdot 99! \cdot 89!} = \frac{92 \cdot 91 \cdot 90}{102 \cdot 101 \cdot 100} = \frac{6279}{8585} \approx 0{,}73139$$

2. Calcular la probabilidad de que exista al menos un cuadrado perfecto.

   Llamemos a este suceso el *suceso* $B$. Como tenemos que $B = \bar{A}$, tenemos que:

$$P(B) = P(\bar{A}) = 1 - P(A)$$

   Por tanto, deberemos distinguir de nuevo si hay o no repetición. Esto es, deberemos usar el resultado del apartado anterior calculado con o sin repetición.

   - Si no hay repetición:

$$P(B) = P(\bar{A}) = 1 - P(A) = 0{,}27348$$

   - Si sí hay repetición:

$$P(B) = P(\bar{A}) = 1 - P(A) = 0{,}26860$$

3. Calcular la probabilidad de que exista un sólo cuadrado perfecto, de que existan dos, y la de que los tres lo sean.

   Denotemos ahora por 1C, 2C y 3C el resultado de obtener 1, 2 ó 3 cuadrados perfectos, respectivamente. Nuevamente debemos distinguir si hay o no repetición.

   - Si no hay repetición:
     Los casos posibles son:

$$C_{100,3} = \binom{100}{3} = \frac{100!}{3! \cdot 97!} = \frac{100 \cdot 99 \cdot 98}{6}$$

Los casos en los que sólo hay un cuadrado perfecto son las combinaciones posibles de 2 elementos de un conjunto de 90 elementos $(100-10)$ y de 1 elemento de 10 elementos sin repetición:

$$C_{90,2} \cdot C_{10,1} = \binom{90}{2} \cdot \binom{10}{1} = \frac{90!}{2! \cdot 88!} \cdot 10$$

Por tanto, usando la Ley de Laplace, tenemos que:

$$P(1C) = \frac{C_{90,2} \cdot C_{10,1}}{C_{100,3}} = \frac{90! \cdot 3! \cdot 97! \cdot 10}{100! \cdot 2! \cdot 88!} = \frac{90 \cdot 89 \cdot 88! \cdot 3 \cdot 2! \cdot 97! \cdot 10}{100 \cdot 99 \cdot 98 \cdot 97! \cdot 2! \cdot 88!} =$$
$$= \frac{90 \cdot 89 \cdot 3 \cdot 10}{100 \cdot 99 \cdot 98} = \frac{90 \cdot 89 \cdot 3}{10 \cdot 99 \cdot 98} = \frac{267}{1078} \approx 0{,}24768$$

Los casos en los que sólo hay dos cuadrados perfectos son las combinaciones posibles de 1 elemento de un conjunto de 90 elementos $(100-10)$ y de 2 elemento de 10 elementos sin repetición:

$$C_{90,1} \cdot C_{10,2} = \binom{90}{1} \cdot \binom{10}{2} = 90 \cdot \frac{10!}{2! \cdot 8!} = 90 \cdot \frac{10 \cdot 9}{2} = 4050$$

Por tanto, usando la Ley de Laplace, tenemos que:

$$P(2C) = \frac{C_{90,1} \cdot C_{10,2}}{C_{100,3}} = \frac{4050 \cdot 6}{100 \cdot 99 \cdot 98} \approx 0{,}025046$$

Los casos en los que hay tres cuadrados perfectos son las combinaciones posibles de 3 elemento de un conjunto de 10 elementos sin repetición:

$$C_{10,3} = \binom{10}{3} = \frac{10!}{3! \cdot 7!} = \frac{10 \cdot 9 \cdot 8}{6} = 120$$

Por tanto, usando la Ley de Laplace, tenemos que:

$$P(3C) = \frac{C_{10,3}}{C_{100,3}} = \frac{120 \cdot 6}{100 \cdot 99 \cdot 98} \approx 0{,}000742$$

- Si sí hay repetición:

  Los casos posibles son:

$$CR_{100,3} = \binom{100+3-1}{3} = \binom{102}{3} = \frac{102!}{3! \cdot 99!} = 171700$$

Los casos en los que sólo hay un cuadrado perfecto son las combinaciones posibles de 2 elementos de un conjunto de 90 elementos $(100-10)$ y de 1 elemento de 10 elementos con repetición:

$$CR_{90,2} \cdot CR_{10,1} = \binom{90+2-1}{2} \cdot \binom{10}{1} = \binom{91}{2} \cdot 10 = \frac{91!}{2! \cdot 89!} \cdot 10$$

Por tanto, usando la Ley de Laplace, tenemos que:

$$P(1C) = \frac{CR_{90,2} \cdot CR_{10,1}}{CR_{100,3}} = \frac{91! \cdot 3! \cdot 99! \cdot 10}{102! \cdot 2! \cdot 89!} = \frac{91 \cdot 90 \cdot 89! \cdot 3 \cdot 2! \cdot 99! \cdot 10}{102 \cdot 101 \cdot 100 \cdot 99! \cdot 2! \cdot 89!} =$$
$$= \frac{91 \cdot 90 \cdot 3 \cdot 10}{102 \cdot 101 \cdot 100} = \frac{91 \cdot 90 \cdot 3}{102 \cdot 101 \cdot 10} = \frac{819}{3434} \approx 0{,}238497$$

Los casos en los que hay dos cuadrados perfectos son las combinaciones posibles de 1 elemento de un conjunto de 90 elementos $(100 - 10)$ y de 2 elemento de 10 elementos con repetición:

$$CR_{90,1} \cdot CR_{10,2} = 90 \cdot \binom{11}{2} = 90 \cdot \frac{11!}{2! \cdot 9!} = 4950$$

Por tanto, usando la Ley de Laplace, tenemos que:

$$P(2C) = \frac{CR_{90,1} \cdot CR_{10,2}}{CR_{100,3}} = \frac{4950}{171700} \approx 0{,}028829$$

Los casos en los que hay tres cuadrados perfectos son las combinaciones posibles de 3 elementos de un conjunto de 10 elementos con repetición:

$$CR_{10,3} = \binom{12}{3} = \frac{12!}{3! \cdot 9!} = 220$$

Por tanto, usando la Ley de Laplace, tenemos que:

$$P(3C) = \frac{CR_{10,3}}{CR_{100,3}} = \frac{220}{171700} \approx 0{,}0012813$$

**Ejercicio 8.3.8.** En una carrera de relevos cada equipo se compone de 4 atletas. La sociedad deportiva de un colegio cuenta con 10 corredores y su entrenador debe formar un equipo de relevos que disputará el campeonato, y el orden en que participarán los seleccionados.

1. ¿Entre cuántos equipos distintos habría de elegir el entrenador si los 10 corredores son de igual valía? (Dos equipos con los mismos atletas en orden distinto se consideran diferentes)

   En este caso, el orden sí afecta a la hora de hacer los grupos. Además, como un atleta no puede ser seleccionado dos veces, tenemos que no hay repetición. Por tanto, estamos ante variaciones sin repetición.

   $$V_{10,4} = \frac{10!}{(10-4)!} = 10 \cdot 9 \cdot 8 \cdot 7 = 5040$$

   Por tanto, habría 5040 equipos distintos.

2. Calcular la probabilidad de que un alumno cualquiera sea seleccionado.

   Dado un alumno, la cantidad de grupos distintos que se pueden hacer sin él son:

   $$V_{9,4} = \frac{9!}{(9-4)!} = 9 \cdot 8 \cdot 7 \cdot 6 = 3024$$

Por tanto, la cantidad de grupos que se pueden hacer con él son:

$$V_{10,4} - V_{9,4} = 2016$$

Por tanto, por la regla de Laplace, tenemos que la probabilidad de que un alumno cualquiera sea elegido es:

$$\frac{2016}{V_{10,4}} = \frac{2}{5} = 0{,}4$$

**Ejercicio 8.3.9.** Una tienda compra bombillas en lotes de 300 unidades. Cuando un lote llega, se comprueban 60 unidades elegidas al azar, rechazándose el envío si se supera la cifra de 5 defectuosas. ¿Cuál es la probabilidad de aceptar un lote en el que haya 10 defectuosas?

Analizamos en primer lugar el problema:

- No se repiten las bombillas, pues no se pueden revisar 2 bombillas iguales.
- El orden es indiferente, ya que solo importa el número de bombillas defectuosas.

Por tanto, se trata de combinaciones sin repeticiones. La cantidad de lotes posibles de 60 bombillas a analizar son:

$$C_{300,60} = \binom{300}{60} = \frac{300!}{60! \cdot 240!}$$

Sea $D_n$ el suceso de que, en el lote a revisar, haya exactamente $n$ bombillas defectuosas. Como los sucesos $D_i, D_j$ $(i \neq j)$ son incompatibles, ya que no puede haber a la vez dos números distintos de bombillas, tenemos que la probabilidad de aceptar el envío es:

$$P\left(\bigcup_{n=0}^{5} D_n\right) = \sum_{n=0}^{5} P(D_n)$$

Sabiendo que de las 300 bombillas hay 10 defectuosas y que se hacen lotes de 60 bombillas, vemos cuántos lotes se pueden hacer con $n$ bombillas defectuosas.

De un total de 10 bombillas defectuosas, se han de elegir $n$. Por tanto, se pueden hacer $C_{10,n}$ combinaciones para las $n$ defectuosas.

Además, de un total de $300-10 = 290$ bombillas correctas, se han de elegir $60-n$, por lo que se pueden hacer $C_{290,60-n}$ combinaciones para las $60 - n$ defectuosas.

Por tanto, como cada grupo de bombillas defectuosas puede estar con un grupo de bombillas correctas distinto, tenemos que el número total de lotes con $n$ bombillas defectuosas es:

$$C_{10,n} \cdot C_{290,60-n}$$

Por tanto, por la regla de Laplace tenemos que la probabilidad de hacer un lote de $n$ bombillas defectuosas es:

$$P(D_n) = \frac{C_{10,n} \cdot C_{290,60-n}}{C_{300,60}} \qquad n = 0, 1, \ldots, 10$$

Calculamos por tanto la probabilidad de encontrar $n = 0, 1, \dots, 5$ bombillas defectuosas:

$$P(D_0) = \frac{C_{10,0} \cdot C_{290,60}}{C_{300,60}} = \frac{\binom{10}{0} \cdot \binom{290}{60}}{\binom{300}{60}} = \frac{10! \cdot 290! \cdot 60! \cdot 240!}{300! \cdot 10! \cdot 0! \cdot 60! \cdot 230!} = \frac{240 \cdot \dots \cdot 231}{300 \cdot \dots \cdot 291} \approx 0{,}1033$$

$$P(D_1) = \frac{C_{10,1} \cdot C_{290,59}}{C_{300,60}} = \frac{\binom{10}{1} \cdot \binom{290}{59}}{\binom{300}{60}} = \frac{10! \cdot 290! \cdot 60! \cdot 240!}{300! \cdot 9! \cdot 1! \cdot 59! \cdot 231!} = \frac{240 \cdot \dots \cdot 232 \cdot 10 \cdot 60}{300 \cdot \dots \cdot 291} \approx 0{,}26838$$

$$P(D_2) = \frac{C_{10,2} \cdot C_{290,58}}{C_{300,60}} = \frac{\binom{10}{2} \cdot \binom{290}{58}}{\binom{300}{60}} = \frac{10! \cdot 290! \cdot 60! \cdot 240!}{300! \cdot 8! \cdot 2! \cdot 58! \cdot 232!} = \frac{240 \cdot \dots \cdot 233 \cdot 10 \cdot 9 \cdot 60 \cdot 59}{300 \cdot \dots \cdot 291 \cdot 2!} \approx 0{,}3071$$

$$P(D_3) = \frac{C_{10,3} \cdot C_{290,57}}{C_{300,60}} = \frac{\binom{10}{3} \cdot \binom{290}{57}}{\binom{300}{60}} = \frac{10! \cdot 290! \cdot 60! \cdot 240!}{300! \cdot 7! \cdot 3! \cdot 57! \cdot 233!} = \frac{240 \cdot \dots \cdot 234 \cdot 10 \cdot \dots \cdot 8 \cdot 60 \cdot \dots \cdot 58}{300 \cdot \dots \cdot 291 \cdot 3!} \approx 0{,}2039$$

$$P(D_4) = \frac{C_{10,4} \cdot C_{290,56}}{C_{300,60}} = \frac{\binom{10}{4} \cdot \binom{290}{56}}{\binom{300}{60}} = \frac{10! \cdot 290! \cdot 60! \cdot 240!}{300! \cdot 6! \cdot 4! \cdot 56! \cdot 234!} = \frac{240 \cdot \dots \cdot 235 \cdot 10 \cdot \dots \cdot 7 \cdot 60 \cdot \dots \cdot 57}{300 \cdot \dots \cdot 291 \cdot 4!} \approx 0{,}08691$$

$$P(D_5) = \frac{C_{10,5} \cdot C_{290,55}}{C_{300,60}} = \frac{\binom{10}{5} \cdot \binom{290}{55}}{\binom{300}{60}} = \frac{10! \cdot 290! \cdot 60! \cdot 240!}{300! \cdot 5! \cdot 5! \cdot 55! \cdot 235!} = \frac{240 \cdot \dots \cdot 234 \cdot 10 \cdot \dots \cdot 6 \cdot 60 \cdot \dots \cdot 56}{300 \cdot \dots \cdot 291 \cdot 5!} \approx 0{,}02485$$

Por tanto, la probabilidad de que se acepte el envío (es decir, que tenga entre 0 y 5 bombillas defectuosas) es:

$$P\left(\bigcup_{n=0}^{5} D_n\right) = \sum_{n=0}^{5} P(D_n) \approx 0{,}99444$$

**Ejercicio 8.3.10.** Una secretaria debe echar al correo 3 cartas; para ello, introduce cada carta en un sobre y escribe las direcciones al azar. ¿Cuál es la probabilidad de que al menos una carta llegue a su destino?

En este caso, tenemos que considerar 2 conjuntos, el conjunto $C$ de las cartas y conjunto $D$ de los destinos. Sean los siguientes sucesos:

- $A_1$: la carta 1 ($C_1$) llega a su destino ($D_1$).

- $A_2$: la carta 2 ($C_2$) llega a su destino ($D_2$).
- $A_3$: la carta 3 ($C_3$) llega a su destino ($D_3$).

Tenemos que la probabilidad del suceso pedido es:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1)+P(A_2)+P(A_3)-P(A_1 \cap A_2)-P(A_1 \cap A_3)-P(A_2 \cap A_3)+P(A_1 \cap A_2 \cap A_3)$$

Para calcular las probabilidades, empleamos combinatoria. Como buscamos las formas en las que se pueden organizar las direcciones, buscamos las distintas permutaciones. Como no se pueden repetir, son **permutaciones sin repetición**.

Para los casos posibles, como hay 3 direcciones, tenemos que son:

$$P_3 = 3!$$

En el suceso $A_i$ de que la carta $C_i$ llegue a su destino $D_i$, tenemos que se ha fijado dicha dirección ($C_i$ con $D_i$), por lo que quedan 2 direcciones que se pueden ordenar como se desee. Las permutaciones favorables son, por tanto,

$$P_2 = 2!$$

Por la Regla de Laplace, tengo que:

$$P(A_i) = \frac{2!}{3!} = \frac{1}{3} \qquad i = 1,2,3$$

En el suceso $A_i \cap A_j$ de que las carta $C_i, C_j$ lleguen a su destino $D_i, D_j$, tenemos que se han fijado dos direcciones, por lo que queda una dirección que se pueden ordenar como se desee. Las permutaciones favorables son, por tanto,

$$P_1 = 1! \Longrightarrow P(A_i \cap A_j) = \frac{1!}{3!} = \frac{1}{6} \qquad i,j = 1,2,3 \quad i \neq j$$

Además, en el suceso $A_1 \cap A_2 \cap A_3$ las tres direcciones están fijadas, por lo que solo hay una posibilidad. Es decir,

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{6}$$

Por tanto, usando las probabilidades ya obtenidas, tengo que:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - \\ &\qquad - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3) = \\ &= 3 \cdot \frac{1}{3} - 3 \cdot \frac{1}{6} + \frac{1}{6} = \frac{2}{3} \end{aligned}$$

## 8.4. Probabilidad Condicionada e Independencia de Sucesos

**Ejercicio 8.4.1.** En una batalla naval, tres destructores localizan y disparan simultáneamente a un submarino. La probabilidad de que el primer destructor acierte el disparo es 0,6, la de que lo acierte el segundo es 0,3 y la de que lo acierte el tercero es 0,1. ¿Cuál es la probabilidad de que el submarino sea alcanzado por algún disparo?

Sea $A$ el suceso de que el primer destructor acierte, $P(A) = 0{,}6$.
Sea $B$ el suceso de que el segundo destructor acierte, $P(B) = 0{,}3$.
Sea $C$ el suceso de que el tercer destructor acierte, $P(C) = 0{,}1$.

Tenemos, además, que $A, B, C$ son independientes; ya que el hecho de que uno acierte no influye en que lo hagan los otros o no.

Por tanto, tenemos que la probabilidad de que el submarino sea alcanzado por algún disparo es:

$$\begin{aligned}P(A\cup B\cup C) &= P(A) + P(B\cup C) - P(A\cap(B\cup C)) = \\ &= P(A) + P(B) + P(C) - P(B\cap C) - P[(A\cap B)\cup(A\cap C)] = \\ &= P(A) + P(B) + P(C) - P(B\cap C) - P(A\cap B) - P(A\cap C) + P(A\cap B\cap C)\end{aligned}$$

Como tenemos que son sucesos independientes:

$$P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(B)P(C)-P(A)P(C)+P(A)P(B)P(C) = 0{,}748$$

Por tanto, tenemos que la probabilidad de que el submarino haya sido alcanzado es de 0,748.

**Ejercicio 8.4.2.** Un estudiante debe pasar durante el curso 5 pruebas selectivas. La probabilidad de pasar la primera es 1/6. La probabilidad de pasar la $i-$ésima, habiendo pasado las anteriores es $1/(7-i)$. Determinar la probabilidad de que el alumno apruebe el curso.

Sea $A_i$ el suceso de aprobar la $i$-ésima prueba. El enunciado que la probabilidad de aprobar la prueba $n$-ésima condicionada a haber aprobado las $n-1$ anteriores es de $\frac{1}{7-n}$:

$$P\left[A_n \middle| \bigcap_{i=1}^{n-1} A_i\right] = \frac{1}{7-n} \qquad n = 1, \dots, 5$$

Por tanto, por el Teorema de la Probabilidad Compuesta,

$$P\left[\bigcap_{i=1}^{5} A_i\right] = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|(A_1\cap A_2))\cdot P\left(A_4|\cap_{j=1}^{3} A_j\right)\cdot P\left(A_5|\cap_{j=1}^{4} A_j\right) =$$

$$= \prod_{i=1}^{5} \frac{1}{7-i} = \frac{1}{6}\cdot\frac{1}{5}\cdot\frac{1}{4}\cdot\frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6!} = \frac{1}{720} = 0{,}0013\overline{8}$$

Por tanto, tenemos que la probabilidad de aprobar el curso de dicho estudiante es de $0{,}0013\overline{8}$.

**Ejercicio 8.4.3.** En una ciudad, el 40 % de las personas tienen pelo rubio, el 25 % tienen ojos azules y el 5 % el pelo rubio y los ojos azules. Se selecciona una persona al azar. Calcular la probabilidad de los siguientes sucesos:

1. tener el pelo rubio si se tiene los ojos azules,

   Sea $A$ tener los ojos azules y $R$ tener el pelo rubio. Tenemos que:

   $$P(R) = 0{,}4 \qquad\qquad P(A) = 0{,}25 \qquad\qquad P(A \cap R) = 0{,}05$$

   Por definición de probabilibad condicionada:

   $$P(R|A) = \frac{P(A \cap R)}{P(A)} = \frac{1}{5} = 0{,}2$$

2. tener los ojos azules si se tiene el pelo rubio,

   Por definición de probabilibad condicionada:

   $$P(A|R) = \frac{P(A \cap R)}{P(R)} = \frac{1}{8} = 0{,}125$$

3. no tener pelo rubio ni ojos azules,

   $$P(\bar{R} \cap \bar{A}) = P(\overline{A \cup R}) = 1 - P(A \cup R) = 1 - P(A) - P(R) + P(A \cap R) = 0{,}4$$

4. tener exactamente una de estas características.

   $$\begin{aligned} P[(A \cap \bar{R}) \cup (\bar{A} \cap R)] &= P(A \cap \bar{R}) + P(\bar{A} \cap R) - \cancelto{0}{P[(A \cap \bar{R}) \cap (\bar{A} \cap R)]} = \\ &= P(A - R) + P(R - A) = P(A) + P(R) - 2 \cdot P(A \cap R) = 0{,}55 \end{aligned}$$

**Ejercicio 8.4.4.** En una población de moscas, el 25 % presentan mutación en los ojos, el 50 % presentan mutación en las alas, y el 40 % de las que presentan mutación en los ojos presentan mutación en las alas.

1. ¿Cuál es la probabilidad de que una mosca elegida al azar presente al menos una de las mutaciones?

   Sea $O$ tener una mutación en los ojos y $A$ tener una mutación en las alas. Tenemos que:

   $$P(O) = 0{,}25 \qquad\qquad P(A) = 0{,}5 \qquad\qquad P(A|O) = 0{,}4$$

   Por la definición de probabilidad condicionada, tenemos que:

   $$P(A|O) = \frac{P(A \cap O)}{P(O)} \Longrightarrow P(A \cap O) = P(A|0) \cdot P(O) = 0{,}1$$

   Por tanto, tenemos que la probabilidad de tener al menos una de las mutaciones es:

   $$P(A \cup O) = P(A) + P(O) - P(A \cap O) = 0{,}65$$

2. ¿Cuál es la probabilidad de que presente mutación en los ojos pero no en las alas?
$$P(O \cap \bar{A}) = P(O - A) = P(O) - P(A \cap O) = 0{,}15$$

**Ejercicio 8.4.5.** Una empresa utiliza dos sistemas alternativos, $A$ y $B$, en la fabricación de un artículo, fabricando por el sistema $A$ el 20 % de su producción. Cuando a un cliente se le ofrece dicho artículo, la probabilidad de que lo compre es $2/3$ si éste se fabricó por el sistema $A$ y $2/5$ si se fabricó por el sistema $B$. Calcular la probabilidad de vender el producto.

Sea $A$ el suceso de que un producto sea fabricado usando el sistema $A$, y $B$ en el caso contrario. Se considera también el suceso $C$, que es que el cliente compra el producto. Entonces, por las condiciones del enunciado tenemos que:

$$P(A) = 0{,}2 = \frac{1}{5} \qquad P(B) = 0{,}8 = \frac{4}{5}$$
$$P(C|A) = \frac{2}{3} \qquad P(C|B) = \frac{2}{5}$$

Por el teorema de la probabilidad total, tenemos que:

$$P(C) = P(A)P(C|A) + P(B)P(C|B) = \frac{2}{15} + \frac{8}{25} = \frac{34}{75} \approx 0{,}4533$$

**Ejercicio 8.4.6.** Se consideran dos urnas: la primera con 20 bolas, de las cuales 18 son blancas, y la segunda con 10 bolas, de las cuales 9 son blancas. Se extrae una bola de la segunda urna y se deposita en la primera; si a continuación, se extrae una bola de ésta, calcular la probabilidad de que sea blanca.

Sean las bolas blancas de la primera urna notadas como $1_B$, y $1_O$ las que son de otro color. Respecto de la segunda urna, sean estas bolas $2_B, 2_O$ respectivamente.

Usando la regla de la Probabilidad total, tenemos que:

$$P(1_B) = P(2_B) \cdot P(1_B|2_B) + P(2_O) \cdot P(1_B|2_O)$$

Por tanto, usando la Ley de Laplace, tenemos que:

$$P(B) = \frac{9}{10} \cdot \frac{19}{21} + \frac{1}{10} \cdot \frac{18}{21} = \frac{9}{10} = 0{,}9$$

**Ejercicio 8.4.7.** Se dispone de tres urnas con la siguiente composición de bolas blancas y negras:

$U_1$: 5B y 5N $\qquad$ $U_2$: 6B y 4N $\qquad$ $U_3$: 7B y 3N

Se elige una urna al azar y se sacan cuatro bolas sin reemplazamiento.

1. Calcular la probabilidad de que las cuatro sean blancas.

   Por el teorema de la Probabilidad Total, tenemos que:

   $$P(4B) = P(U_1)P(4B|U_1) + P(U_2)P(4B|U_2) + P(U_3)P(4B|U_3)$$

   Usando la ley de Laplace, tenemos que:

   $$P(4B) = \frac{1}{3}\left[\frac{5}{10} \cdot \frac{4}{9} \cdot \frac{3}{8} \cdot \frac{2}{7} \cdot + \frac{6}{10} \cdot \frac{5}{9} \cdot \frac{4}{8} \cdot \frac{3}{7} \cdot + \frac{7}{10} \cdot \frac{6}{9} \cdot \frac{5}{8} \cdot \frac{4}{7} \cdot\right] = \frac{1}{3} \cdot \frac{1320}{5040} = \frac{11}{126} \approx 0{,}0873$$

2. Si en las bolas extraídas sólo hay una negra, ¿cuál es la probabilidad de que la urna elegida haya sido $U_2$?

   Se pide $P(U_2|1N)$. Por la regla de Bayes, tenemos que:

$$P(U_2|1N) = \frac{P(U_2)P(1N|U_2)}{P(U_1)P(1N|U_1) + P(U_2)P(1N|U_2) + P(U_3)P(1N|U_3)}$$

   Como tenemos que $P(U_i) = \frac{1}{3}\ \forall i$, tenemos que:

$$P(U_2|1N) = \frac{P(1N|U_2)}{P(1N|U_1) + P(1N|U_2) + P(1N|U_3)}$$

   Para calcular cada probabilidad, empleamos combinatoria. Tenemos que se trata de combinaciones sin remplazamineto, por lo que:

$$P(1N|U_1) = \frac{C_{5,1}C_{5,3}}{C_{10,4}} = \frac{5}{21} \qquad P(1N|U_2) = \frac{C_{4,1}C_{6,3}}{C_{10,4}} = \frac{8}{21}$$

$$P(1N|U_3) = \frac{C_{3,1}C_{7,3}}{C_{10,4}} = \frac{1}{2}$$

   Por tanto, tenemos que:

$$P(U_2|1N) = \frac{P(1N|U_2)}{P(1N|U_1) + P(1N|U_2) + P(1N|U_3)} = \frac{16}{47} \approx 0{,}3404$$

**Ejercicio 8.4.8.** La probabilidad de que se olvide inyectar el suero a un enfermo durante la ausencia del doctor es 2/3. Si se le inyecta el suero, el enfermo tiene igual probabilidad de mejorar que de empeorar, pero si no se le inyecta, la probabilidad de mejorar se reduce a 0.25. Al regreso, el doctor encuentra que el enfermo ha empeorado. ¿Cuál es la probabilidad de que no se le haya inyectado el suero?

Tenemos los siguientes suceso:

- $A \longrightarrow$ El paciente mejora.
- $\bar{A} \longrightarrow$ El paciente empeora.
- $B \longrightarrow$ Se le ha inyectado el suero.
- $\bar{B} \longrightarrow$ No se le ha inyectado el suero.

El enunciado afirma que $P(\bar{B}) = \frac{2}{3}$. Por tanto,

$$P(B) = 1 - P(\bar{B}) = 1 - \frac{2}{3} = \frac{1}{3}$$

Además, tenemos que $P(A|B) = P(\bar{A}|B)$. Por tanto:

$$P(A|B) = P(\bar{A}|B) \Longleftrightarrow \frac{P(A \cap B)}{P(B)} = \frac{P(\bar{A} \cap B)}{P(B)} \Longleftrightarrow$$
$$\Longleftrightarrow P(A \cap B) = P(B - A) = P(B) - P(A \cap B) \Longleftrightarrow 2P(A \cap B) = P(B) \Longleftrightarrow$$
$$\Longleftrightarrow \frac{P(A \cap B)}{P(B)} = \frac{1}{2} = P(A|B) = P(\bar{A}|B)$$

También sabemos por el enunciado que $P(A|\bar{B}) = \frac{1}{4}$. Usando la regla de la probabilidad total, tenemos que:

$$P(\Omega) = 1 = P(A|\bar{B}) + P(\bar{A}|\bar{B}) \Longrightarrow P(\bar{A}|\bar{B}) = 1 - P(A|\bar{B}) = \frac{3}{4} = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{B})} \Longrightarrow$$
$$\Longrightarrow P(\bar{A} \cap \bar{B}) = \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$$

Se pide calcular $P(\bar{B}|\bar{A})$. Usando la definición de probabilidad y la regla de Bayes, tenemos que:

$$P(\bar{B}|\bar{A}) = \frac{P(\bar{B} \cap \bar{A})}{P(\bar{A})} = \frac{P(\bar{B} \cap \bar{A})}{P(B)P(\bar{A}|B) + P(\bar{B})P(\bar{A}|\bar{B})} = \frac{\frac{1}{2}}{\frac{1}{3} \cdot \frac{1}{2} + \frac{2}{3} \cdot \frac{3}{4}} = \frac{3}{4}$$

Por tanto, la probabilidad de que no se le haya inyectado el suero y por ello haya empeorado es de 0,75.

**Ejercicio 8.4.9.** $N$ urnas contienen cada una 4 bolas blancas y 6 negras, mientras otra urna contiene 5 blancas y 5 negras. De las $N+1$ urnas se elige una al azar y se extraen dos bolas sucesivamente, sin reemplazamiento, resultando ser ambas negras. Sabiendo que la probabilidad de que queden 5 blancas y 3 negras en la urna elegida es $1/7$, encontrar $N$.

En este caso, tenemos dos experimentos aleatorios. Respecto al primer experimento, que consiste en elegir una urna, sea su espacio muestral $\Omega_1 = \{U_A, U_B\}$, con:

- $U_A$: sacar una urna del primer tipo. Hay $N$ urnas de este tipo, cada una con 4 bolas blancas y 6 negras.
- $U_B$: sacar una urna del segundo tipo. Hay 1 urna de este tipo, con 5 bolas blancas y 5 negras.

Respecto al segundo experimento aleatorio, que consiste en elegir dos bolsas sin reemplazamiento de la urna elegida, sea su espacio muestral $\Omega_2 = \{BB, NN, BN\}$, con:

- $BB$: sacar dos bolas blancas.
- $NN$: sacar dos bolas negras.
- $BN$: sacar una bola blanca y una bola negra.

Usando la regla de Laplace, tenemos que:

$$P(U_A) = \frac{N}{N+1} \qquad\qquad P(U_B) = \frac{1}{N+1}$$

Consideramos ahora el caso de que, tras extraer dos bolas negras, queden 5 blancas y 3 negras en la urna elegida, cuya probabilidad sabemos que es 1/7. Originalmente había en la urna 5 bolas de cada tipo, por lo que la urna es del tipo $U_B$. Por tanto, este suceso es elegir la urna $U_B$ sabiendo que hemos sacado dos bolas negras, es decir, $U_B|(NN)$. Por tanto,

$$P(U_B|NN) = \frac{1}{7} = \frac{P(U_B \cap NN)}{P(NN)}$$

Usando la regla de la multiplicación, tenemos que:

$$P(U_B \cap NN) = P(U_B) \cdot P(NN|U_B) = \frac{1}{N+1} \cdot \frac{5}{10} \cdot \frac{4}{9} = \frac{2}{9(N+1)}$$

Usando el teorema de la Probabilidad total, tenemos que:

$$\begin{aligned} P(NN) &= P(U_A) \cdot P(NN|U_A) + P(U_B) \cdot P(NN|U_B) \\ &= \frac{N}{N+1} \cdot \frac{6}{10} \cdot \frac{5}{9} + \frac{1}{N+1} \cdot \frac{5}{10} \cdot \frac{4}{9} \\ &= \frac{N}{3(N+1)} + \frac{2}{9(N+1)} = \frac{3N+2}{9(N+1)} \end{aligned}$$

Por tanto, como $P(U_B|NN) = \frac{1}{7}$, tenemos que:

$$P(U_B|NN) = \frac{1}{7} = \frac{\frac{2}{9(N+1)}}{\frac{3N+2}{9(N+1)}} = \frac{2}{3N+2} \Longrightarrow 3N+2 = 14 \Longrightarrow N = 4$$

Por tanto, tenemos que hay 4 urnas del tipo $U_A$.

**Ejercicio 8.4.10.** Se dispone de 6 cajas, cada una con 12 tornillos; una caja tiene 8 buenos y 4 defectuosos; dos cajas tienen 6 buenos y 6 defectuosos y tres cajas tienen 4 buenos y 8 defectuosos. Se elige al azar una caja y se extraen 3 tornillos con reemplazamiento, de los cuales 2 son buenos y 1 es defectuoso. ¿Cuál es la probabilidad de que la caja elegida contuviera 6 buenos y 6 defectuosos?

En este caso, tenemos dos experimentos aleatorios. Respecto al primer experimento, que consiste en elegir una caja, sea su espacio muestral $\Omega_1 = \{C_1, C_2, C_3\}$, con:

- $C_1$: elegir una caja del primer tipo. Solo hay una caja de este tipo, que tiene 8 tornillos buenos y 4 defectuosos.
- $C_2$: elegir una caja del segundo tipo. Hay dos cajas de este tipo, cada una con 6 tornillos buenos y 6 defectuosos.

- $C_3$: elegir una caja del tercer tipo. Hay tres cajas de este tipo, cada una con 4 tornillos buenos y 8 defectuosos.

Respecto al segundo experimento aleatorio, se repite tres veces y consiste en elegir cada vez un tornillo con reemplazamiento de la caja elegida. Sea su espacio muestral $\Omega_2 = \{B, D\}$, con:

- $B$: elegir un tornillo bueno.
- $D$: elegir un tornillo defectuoso.

El enunciado nos pide:

$$P[C_2|(BBD)] = \frac{P[C_2 \cap (BBD)]}{P(BBD)}$$

Por la regla de la multiplicación, tenemos que:

$$P[C_2 \cap BBD] = P(C_2) \cdot P(B|C_2) \cdot P[B|(B \cap C_2)] \cdot P[B|(B \cap B \cap C_2)]$$

usando la regla de Laplace en cada caso, tenemos que:

$$P[C_2 \cap BBD] = \frac{2}{6} \cdot \frac{6}{12} \cdot \frac{6}{12} \cdot \frac{6}{12} = \frac{1}{3} \cdot \frac{1}{2^3} = \frac{1}{24}$$

Usando de nuevo la regla de la multiplicación y la regla de Laplace, junto con la regla de la probabilidad total, tenemos que:

$$\begin{aligned}
P(BBD) &= P(C_1)P(BBD|C_1) + P(C_2)P(BBD|C_2) + P(C_3)P(BBD|C_3) = \\
&= \frac{1}{6} \cdot \frac{8}{12} \cdot \frac{8}{12} \cdot \frac{4}{12} + \frac{2}{6} \cdot \frac{6}{12} \cdot \frac{6}{12} \cdot \frac{6}{12} + \frac{3}{6} \cdot \frac{4}{12} \cdot \frac{4}{12} \cdot \frac{8}{12} = \\
&= \frac{2}{3^4} + \frac{1}{2^3 \cdot 3} + \frac{1}{3^3} = \frac{67}{648} \approx 0{,}103395
\end{aligned}$$

Por tanto, tenemos que la probabilidad de que la caja elegida contuviera 6 buenos y 6 defectuosos es:

$$P[C_2|(BBD)] = \frac{P[C_2 \cap (BBD)]}{P(BBD)} = \frac{\frac{1}{24}}{\frac{67}{648}} = \frac{27}{67} \approx 0{,}4029$$

**Ejercicio 8.4.11.** Se seleccionan $n$ dados con probabilidad $p_n = \frac{1}{2^n}$, $n \in \mathbb{N}$. Si se lanzan estos $n$ dados y se obtiene una suma de 4 puntos, ¿Cuál es la probabilidad de haber seleccionado 4 dados?

Sea $B$ el suceso de que sumen 4 puntos, y suponemos cada dado como un dado estándar (6 caras, equiprobables). Sea $A_n$ el suceso de haber considerado $n$ dados.

Se pide calcular $P(A_4|B)$, que calculamos empleado el Teorema de Bayes:

$$P(A_4|B) = \frac{P(A_4)P(B|A_4)}{\displaystyle\sum_{i=1}^{\infty} P(A_i)P(B|A_i)}$$

Tenemos que $P(A_i) = p_i = \frac{1}{2^n}$. Además, se tiene que $P(B|A_i) = 0 \qquad \forall i > 4$, ya que si hay más de 4 dados, la suma va a ser mayor que 4. Por tanto,

$$P(A_4|B) = \frac{\frac{1}{2^4} \cdot P(B|A_4)}{\displaystyle\sum_{i=1}^{4} \frac{1}{2^i} \cdot P(B|A_i)}$$

Calculamos ahora la probabilidad de que sumen 4 tras haber elegido $i$ dados. Los casos totales son las combinaciones de 6 números que se pueden realizar con 2 dados. Aunque aparentemente no importa el orden, de hecho sí importa porque queremos que las combinaciones $(1,3)$ y $(3,1)$ sean distintas. También hay reemplazamiento, ya que cada dado tiene siempre 6 dados, por lo que estamos ante variaciones con repetición.

- Para $n = 1$ dado:

  Tenemos que los casos totales son:

  $$VR_{6,1} = 6^1 = 6$$

  Tenemos que los casos favorables son $\{(4)\}$, por lo que solo hay un caso favorable. Por tanto,

  $$P(B|A_1) = \frac{1}{6}$$

- Para $n = 2$ dados:

  Tenemos que los casos totales son:

  $$VR_{6,2} = 6^2$$

  Tenemos que los casos favorables son $\{(1,3),(3,1),(2,2)\}$, por lo que hay tres casos favorables. Por tanto,

  $$P(B|A_2) = \frac{3}{6^2}$$

- Para $n = 3$ dados:

  Tenemos que los casos totales son:

  $$VR_{6,3} = 6^3$$

  Tenemos que los casos favorables son $\{(1,1,2),(1,2,1),(2,1,1)\}$, por lo que hay tres casos favorables. Por tanto,

  $$P(B|A_3) = \frac{3}{6^3}$$

- Para $n = 4$ dados:

  Tenemos que los casos totales son:

$$VR_{6,4} = 6^4$$

  Tenemos que los casos favorables son $\{(1,1,1,1)\}$, por lo que hay un caso favorable. Por tanto,

$$P(B|A_4) = \frac{1}{6^4}$$

Por tanto, tenemos que:

$$P(A_4|B) = \frac{\frac{1}{2^4} \cdot P(B|A_4)}{\displaystyle\sum_{i=1}^{4} \frac{1}{2^i} \cdot P(B|A_i)} = \frac{1}{2197} \approx 0{,}445 \cdot 10^{-3}$$

**Ejercicio 8.4.12.** Se lanza una moneda; si sale cara, se introducen $k$ bolas blancas en una urna y si sale cruz, se introducen $2k$ bolas blancas. Se hace una segunda tirada, poniendo en la urna $h$ bolas negras si sale cara y $2h$ si sale cruz. De la urna así compuesta se toma una bola al azar. ¿Cuál es la probabilidad de que sea negra?

En este caso, tenemos dos experimentos aleatorios. Respecto al primer experimento; que se repite dos veces, consiste en lanzar una moneda. Sea su espacio muestral $\Omega_1 = \{C, +\}$, con:

- $+$: obtener cruz.
- $C$: obtener cara.

Respecto al segundo experimento aleatorio, se saca una bola de la caja. Sea su espacio muestral $\Omega_2 = \{N, B\}$, con:

- $N$: obtener una bola negra.
- $B$: obtener una bola blanca.

Uniendo el teorema de la probabilidad total con la regla de la multiplicación, tenemos que:

$$\begin{aligned} P(N) = P(C) \cdot P(C|C) \cdot P(N|(CC)) + P(C) \cdot P(+|C) \cdot P(N|(C+)) + \\ + P(+) \cdot P(C|+) \cdot P(N|(+C)) + P(+) \cdot P(+|+) \cdot P(N|(++)) \end{aligned}$$

Como tenemos que las dos repeticiones del primer experimento son independientes; es decir, el resultado del primer lanzamiento no influye en el resultado del segundo lanzamiento; tenemos que:

$$\begin{aligned} P(N) = P(C) \cdot P(C) \cdot P(N|(CC)) + P(C) \cdot P(+) \cdot P(N|(C+)) + \\ + P(+) \cdot P(C) \cdot P(N|(+C)) + P(+) \cdot P(+) \cdot P(N|(++)) \end{aligned}$$

Por la ley de Laplace, tenemos que $P(+) = P(C) = \frac{1}{2}$. Por tanto,

$$P(N) = \frac{1}{4} \cdot P(N|(CC)) + \frac{1}{4} \cdot P(N|(C+)) + \frac{1}{4} \cdot P(N|(+C)) + \frac{1}{4} \cdot P(N|(++))$$

Veamos ahora cuántas bolas hay en cada combinación de resultados de la moneda:

$$\begin{aligned} CC &\longrightarrow k \text{ blancas y } h \text{ negras.}\\ C+ &\longrightarrow k \text{ blancas y } 2h \text{ negras.}\\ +C &\longrightarrow 2k \text{ blancas y } h \text{ negras.}\\ ++ &\longrightarrow 2k \text{ blancas y } 2h \text{ negras.}\end{aligned}$$

Por la ley de Laplace y la regla de la multiplicación, tenemos que:

$$\begin{aligned} P(N \cap C \cap C) &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{h}{k+h} & \qquad P(N \cap C \cap +) &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2h}{k+2h}\\ P(N \cap + \cap C) &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{h}{2k+h} & \qquad P(N \cap + \cap +) &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2h}{2k+2h}\end{aligned}$$

Por tanto, como $P(CC) = P(C+) = P(+C) = P(++) = \frac{1}{4}$, tenemos que:

$$\begin{aligned} P(N|(CC)) &= \frac{\frac{h}{4(k+h)}}{\frac{1}{4}} = \frac{h}{k+h} & \qquad P(N \cap C \cap +) &= \frac{2h}{k+2h}\\ P(N \cap + \cap C) &= \frac{h}{2k+h} & \qquad P(N \cap + \cap +) &= \frac{2h}{2k+2h} = \frac{h}{k+h}\end{aligned}$$

Por tanto, tenemos que:

$$P(N) = \frac{1}{4}\left[\frac{h}{k+h} + \frac{2h}{k+2h} + \frac{h}{2k+h} + \frac{h}{k+h}\right] = \frac{1}{4}\left[\frac{2h}{k+h} + \frac{2h}{k+2h} + \frac{h}{2k+h}\right]$$

## 8.5. Variables Aleatorias Unidimensionales

**Ejercicio 8.5.1.** Sea $X$ una variable aleatoria con función masa de probabilidad

$$P(X=i)=ki; \qquad i=1,\dots,20.$$

1. Determinar el valor de $k$, la función de distribución y las siguientes probabilidades:

$$P(X=4), \quad P(X<4), \quad P(3\leqslant X\leqslant 10), \quad P(3<X\leqslant 10), \quad P(3<X<10).$$

Estamos trabajando con una variable aleatoria discreta, con $|Re_X|=20$. Para que $P(X=i)$ sea una función masa de probabilidad, necesitamos que:

$$1=\sum_{i=1}^{20}P(X=i)=\sum_{i=1}^{20}ki=k\sum_{i=1}^{20}i=k\cdot\frac{20(1+20)}{2}=210k\Longrightarrow k=\frac{1}{210}$$

Por tanto, tenemos que la función masa de probabilidad es:

$$P(X=i)=\frac{1}{210}i \qquad i=1,\dots,20$$

Para calcular la función de distribución, sabemos que:

$$F_X(x)=P(X\leqslant i)=\sum_{j=1}^{i}\frac{1}{210}j=\frac{1}{210}\sum_{j=1}^{i}j=\frac{1}{210}\frac{i(i+1)}{2}=\frac{i(i+1)}{420}$$

Por tanto, la función de distribución es:

$$F_X(x)=\begin{cases}0 & \text{si} \quad x<1\\ \dfrac{i(i+1)}{420} & \text{si} \quad x\in[i,i+1[ \qquad \forall i=1,\dots,19\\ 1 & \text{si} \quad 20\leqslant x\end{cases}$$

Por tanto, las probabilidades pedidas son:

$$P(X=4)=\frac{4}{210}=\frac{2}{105}\approx 0{,}01905$$
$$P(X<4)=P(X\leqslant 3)=F_X(3)=\frac{12}{420}=\frac{1}{35}\approx 0{,}02857$$
$$P(3\leqslant X\leqslant 10)=P(X\leqslant 10)-P(X<3)=F_X(10)-F_X(2)=\frac{110-6}{420}\approx 0{,}2476$$
$$P(3<X\leqslant 10)=P(X\leqslant 10)-P(X\leqslant 3)=F_X(10)-F_X(3)=\frac{110-12}{420}\approx 0{,}2\overline{3}$$
$$P(3<X<10)=P(X<10)-P(X\leqslant 3)=F_X(9)-F_X(3)=\frac{90-12}{420}\approx 0{,}1857$$

2. Supongamos que un jugador gana 20 monedas si al observar esta variable obtiene un valor menor que 4, gana 24 monedas si obtiene el valor 4 y, en caso contrario, pierde una moneda. Calcular la ganancia esperada del jugador y

decir si el juego le es favorable.

Definimos una nueva variable aleatoria, $Y$, que indica el número de monedas obtenidas por el jugador en función del valor de $X$.

$$Y = h(X) = \begin{cases} 20 & \text{si} \quad 1 \leqslant x_i < 4 \\ 24 & \text{si} \quad x_i = 4 \\ -1 & \text{si} \quad 4 < x_i \leqslant 20 \end{cases}$$

Tenemos que $Re_y = \{-1, 20, 24\}$. Calculamos la probabilidad de cada valor de la variable $Y$:

$$\begin{aligned} P[Y = -1] = P[4 < X \leqslant 20] = P[X > 4] = 1 - P[X \leqslant 4] = \\ = 1 - P[X = 4] - P[x < 4] = 1 - \frac{1}{35} - \frac{2}{105} = \frac{20}{21} \end{aligned}$$

$$P[Y = 20] = P[1 \leqslant X < 4] = P[x < 4] = \frac{1}{35}$$

$$P[Y = 24] = P[X = 4] = \frac{2}{105}$$

Por tanto, la ganancia esperada del jugador es:

$$E[Y] = \sum_{y_i \in Re_y} y_i P[y_i] = 20 \cdot P[y = 20] + 24 \cdot P[Y = 24] - 1 \cdot P[Y = -1] = \frac{8}{105}$$

Como tenemos que $E[Y] = \frac{8}{105} > 0$, tenemos que el juego le es favorable, ya que se espera una ganancia de $\frac{8}{105}$ monedas.

**Ejercicio 8.5.2.** Sea $X$ el número de bolas blancas obtenidas al sacar dos de una urna con 10 bolas de las que 8 son blancas. Calcular:

1. Función masa de probabilidad y función de distribución.

   El experimento aleatorio tiene el siguiente espacio muestral: $\Omega = \{BB, B-, --\}$, donde $B$ representa sacar una bola blanca y $-$ se refiere a sacar una bola de otro color.

   Tenemos que $Re_X = \{0, 1, 2\}$. Calculamos su función de probabilidad:

$$P[X = 0] = P[--] = \frac{C_{2,2}}{C_{10,2}} = \frac{1}{\frac{10!}{2!8!}} = \frac{2}{9 \cdot 10} = \frac{1}{45} = 0{,}0\bar{2}$$

$$P[X = 1] = P[B-] = \frac{C_{2,1} \cdot C_{8,1}}{C_{10,2}} = \frac{16}{45} = 0{,}3\bar{5}$$

$$P[X = 2] = P[BB] = \frac{C_{8,2}}{C_{10,2}} = \frac{8! \cdot 2! \cdot 8!}{10! \cdot 2! \cdot 6!} = \frac{8 \cdot 7}{10 \cdot 9} = \frac{28}{45} = 0{,}6\bar{2}$$

   Por tanto, la función masa de probabilidad es:

$$f(x) = \begin{cases} 0{,}0\bar{2} & \text{si} \quad x = 0 \\ 0{,}3\bar{5} & \text{si} \quad x = 1 \\ 0{,}6\bar{2} & \text{si} \quad x = 2 \end{cases}$$

Por tanto, la función de distribución es:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 0{,}0\bar{2} & \text{si} \quad 0 \leqslant x < 1 \\ 0{,}3\bar{7} & \text{si} \quad 1 \leqslant x < 2 \\ 1 & \text{si} \quad 2 \leqslant x \end{cases}$$

2. Media, mediana y moda, dando la interpretación de cada una de estas medidas.

   La media de una variable aleatoria es su esperanza. Por tanto,

$$E[X] = \sum_{x_i=0}^{2} x_i f(x_i) = 0 \cdot 0{,}0\bar{2} + 1 \cdot 0{,}3\bar{5} + 2 \cdot 0{,}6\bar{2} = 1{,}6$$

   Que tenga una esperanza de 1,6 implica que se espera que tras repetir el experimento un gran número de veces, la media de bolas blancas sacadas sea 1,6. Este es el centro de gravedad de la distribución.

   Calculamos ahora la mediana.

$$P[X \leqslant 2] = 1 \geqslant \frac{1}{2} \qquad\qquad P[X \geqslant 2] = 0{,}6\bar{2} \geqslant \frac{1}{2}$$

   Por tanto, tenemos que $Me_X = 2$. Este valor deja por encima y por debajo la misma probabilidad.

   La moda es la abcisa del máximo de la función masa de probabilidad, que como podemos ver es $Mo_X = 2$. Esto implica que es el valor con mayor probabilidad.

3. Intervalo intercuartílico, especificando su interpretación.

   Calculamos en primer lugar $Q_1$:

$$P[X \leqslant 1] = 0{,}3\bar{7} \geqslant \frac{1}{4} \qquad\qquad P[X \geqslant 1] = 0{,}9\bar{7} \geqslant \frac{3}{4} = 1 - \frac{1}{4}$$

   Por tanto, tenemos que $Q_1 = 1$. Calculamos ahora $Q_3$:

$$P[X \leqslant 2] = 1 \geqslant \frac{3}{4} \qquad\qquad P[X \geqslant 2] = 0{,}6\bar{2} \geqslant \frac{1}{4} = 1 - \frac{3}{4}$$

   Por tanto, tenemos $Q_3 = 2$. De ahí concluimos que:

$$R_I = Q_3 - Q_1 = 1$$

   Por tanto, el 50 % central de la distribución se encuentra en un intervalo de amplitud 1.

**Ejercicio 8.5.3.** El número de lanzamientos de una moneda hasta salir cara es una variable aleatoria con distribución $P(X = x) = 2^{-x}; \qquad x = 1, 2, \ldots$

1. Probar que la función masa de probabilidad está bien definida.

Tenemos que $Re_X = \mathbb{N} - \{0\}$. Es necesario que:

$$\sum_{x=1}^{\infty} P(X = x) = \sum_{x=1}^{\infty} 2^{-x} = \sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^x$$

Tenemos que se trata de una serie geométrica. Estas cumplen que, dado $r$ tal que $|r| < 1$:

$$\sum_{x=1}^{\infty} r^x = \frac{r}{1-r}$$

Por tanto,

$$\sum_{x=1}^{\infty} P(X = x) = \sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^x = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$$

Además, también se cumple que la probabilidad es siempre positiva, ya que $2^{-x} \geqslant 0 \ \forall x \in \mathbb{N}$. Por tanto, tenemos que está bien definida.

2. Calcular la probabilidad de que el número de lanzamientos necesarios para salir cara esté entre 4 y 10.

$$P(4 \leqslant x \leqslant 10) = P(X \in [4, 10]) = P[X \leqslant 10] - P[X < 3] = P[X \leqslant 10] - P[X \leqslant 3] =$$

$$= \sum_{x=1}^{10} 2^{-x} - \sum_{x=1}^{3} 2^{-x} = \sum_{x=4}^{10} 2^{-x} = \frac{127}{1024} \approx 0{,}124$$

3. Calcular los cuartiles y la moda de la distribución, interpretando los valores.

Calculamos en primer lugar $Q_1$:

$$P[X \leqslant 1] = \frac{1}{2} \geqslant \frac{1}{4} \qquad P[X \geqslant 1] = 1 - P[X < 1] = 1 \geqslant \frac{3}{4} = 1 - \frac{1}{4}$$

Por tanto, $Q_1 = 1$. Calculamos ahora $Q_2$:

$$P[X \leqslant 1] = \frac{1}{2} \geqslant \frac{1}{2} \qquad P[X \geqslant 1] = 1 - P[X < 1] = 1 \geqslant \frac{1}{2} = 1 - \frac{1}{2}$$

Por tanto, $Q_2 = 1 = Me_X$. Calculamos ahora $Q_3$:

$$P[X \leqslant 2] = \frac{3}{4} \geqslant \frac{3}{4} \qquad P[X \geqslant 2] = 1 - P[X \leqslant 1] = \frac{1}{2} \geqslant \frac{1}{4} = 1 - \frac{3}{4}$$

Por tanto, $Q_3 = 2$.

Al ser una variable discreta, tenemos que tanto el primer cuarto como la mitad de la distribución se encuentra en el valor $x = 1$. El 75 % de la distribución se encuentra hasta el 2.

Como tenemos que $P(X = x) = 2^{-x}$ es estrictamente decreciente, tenemos que el máximo se alcanza en $x = 1$. Por tanto, $Mo_X = 1$. Este es el valor con mayor probabilidad.

4. Calcular la función generatriz de momentos y, a partir de ella, el número medio de lanzamientos necesarios para salir cara y la desviación típica.

   Tenemos que la función generatriz de momentos es:

$$M_X(t) = E[e^{tX}] = \sum_{x=1}^{\infty} e^{tx} P(X=x) = \sum_{x=1}^{\infty} e^{tx} \left(\frac{1}{2}\right)^x = \sum_{x=1}^{\infty} \left(\frac{e^t}{2}\right)^x$$

   Veamos para qué valores de $t$ converge esa serie geométrica:

$$\left|\frac{e^t}{2}\right| < 1 \Longleftrightarrow e^t < 2 \Longleftrightarrow t < \ln 2$$

   Por tanto, para $t < \ln 2$ tenemos que:

$$M_X(t) = E[e^{tX}] = \sum_{x=1}^{\infty} \left(\frac{e^t}{2}\right)^x = \frac{\frac{e^t}{2}}{1-\frac{e^t}{2}} = \frac{e^t}{2-e^t}$$

   Para calcular la esperanza, por las propiedades de la función generatriz de momentos, tenemos que:

$$E[X] = m_1 = M_X'(0) = 2$$

   donde he hecho uso de que:

$$M_X'(t) = \frac{e^t(2-e^t)+e^{2t}}{(2-e^t)^2} = \frac{2e^t}{(2-e^t)^2}$$

   Para calcular la desviación típica, calculo en primer la varianza. Por las propiedades de la función generatriz de momentos, tenemos que:

$$\text{Var}[X] = \mu_2 = E[X^2] - E[X]^2 = m_2 - E[X]^2 = M_X''(0) - E[X]^2 = 6-4 = 2$$

   donde he hecho uso de que:

$$M_X''(t) = \frac{2e^t(2-e^t)^2 + 2e^t(2-e^t)\cdot 2e^t}{(2-e^t)^4} = \frac{2e^t(2-e^t)+4e^{2t}}{(2-e^t)^3} = \frac{4e^t+2e^{2t}}{(2-e^t)^3}$$

   Por tanto,

$$\sigma_X = +\sqrt{\text{Var}[X]} = \sqrt{2}$$

**Ejercicio 8.5.4.** Sea $X$ una variable aleatoria con función de densidad

$$f(x) = \begin{cases} k_1(x+1) & \text{si } \ 0 \leqslant x \leqslant 4 \\ k_2 x^2 & \text{si } \ 4 < x \leqslant 6 \end{cases}$$

Sabiendo que $P(0 \leqslant X \leqslant 4) = 2/3$, determinar $k_1, k_2$, y deducir su función de distribución.

Tenemos que se trata de una variable aleatoria continua. Por tanto, tenemos que:

$$P(0 \leqslant X \leqslant 4) = P(X \in [0,4]) = \int_0^4 f(x)\,dx = \int_0^4 k_1(x+1)\,dx = k_1\left[\frac{x^2}{2}+x\right]_0^4 = 12k_1$$

Como el enunciado afirma que $P(0 \leqslant X \leqslant 4) = \frac{2}{3}$, tenemos que:

$$P(0 \leqslant X \leqslant 4) = \frac{2}{3} = 12k_1 \Longrightarrow k_1 = \frac{1}{18}$$

Además, se necesita que $\int_{-\infty}^{+\infty} f(x)\,dx = 1$, tenemos que:

$$1 = \int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_0^4 k_1(x+1)\,dx + \int_4^6 k_2x^2\,dx + \int_6^{+\infty} 0\,dx \Longrightarrow$$
$$\Longrightarrow 1 = 0 + \frac{12}{18} + k_2\left[\frac{x^3}{3}\right]_4^6 + 0 \Longrightarrow 1-\frac{2}{3} = k_2\cdot\frac{152}{3} \Longrightarrow \frac{1}{3} = \frac{152k_2}{3} \Longrightarrow k_2 = \frac{1}{152}$$

Para calcular la función de distribución, tenemos que:

$$F_X(x) = P[x \leqslant x] = \int_{-\infty}^{x} f(t)\,dt$$

- <u>Para $x < 0$</u>:
$$F_X(x) = \int_{-\infty}^{x} 0\,dt = 0$$

- <u>Para $0 \leqslant x \leqslant 4$</u>:
$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt = 0 + \int_0^x \frac{x+1}{18}\,dt = \frac{1}{18}\left[\frac{x^2}{2}+x\right]_0^x = \frac{1}{18}\left(\frac{x^2}{2}+x\right)$$

- <u>Para $4 \leqslant x \leqslant 6$</u>:
$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt = 0 + \int_0^4 \frac{x+1}{18}\,dt + \int_4^x \frac{x^2}{152}\,dt = \frac{12}{18} + \frac{1}{152}\left[\frac{x^3}{3}\right]_4^x =$$
$$= \frac{12}{18} + \frac{x^3}{3\cdot 152} - \frac{4^3}{152\cdot 3} = \frac{2}{3} + \frac{1}{456}\left(x^3-64\right)$$

Por tanto, tenemos que la función de distribución es:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \dfrac{1}{18}\left(\dfrac{x^2}{2}+x\right) & \text{si} \quad 0 \leqslant x \leqslant 4 \\ \dfrac{2}{3} + \dfrac{1}{456}\left(x^3-64\right) & \text{si} \quad 4 < x \leqslant 6 \\ 0 & \text{si} \quad 6 < x \end{cases}$$

**Ejercicio 8.5.5.** La dimensión en centímetros de los tornillos que salen de cierta fábrica es una variable aleatoria, $X$, con función de densidad

$$f(x) = \frac{k}{x^2}, \qquad 1 \leqslant x \leqslant 10.$$

1. Determinar el valor de $k$, y obtener la función de distribución.

   Para que $f$ sea una función de densidad, es necesario que:

$$1 = \int_{-\infty}^{\infty} f(x)\, dx = k\left[-\frac{1}{x}\right]_1^{10} = \frac{9k}{10} \Longrightarrow k = \frac{10}{9}$$

   Para $x \in [1, 10]$, la función de distribución es:

$$F_X(x) = \int_{-\infty}^{x} \frac{k}{t^2}\, dt = k\left[-\frac{1}{t}\right]_1^{x} = k\left(1 - \frac{1}{x}\right)$$

   Por tanto, la función de distribución queda:

$$F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{10}{9}\left(1 - \frac{1}{x}\right) & 1 \leqslant x < 10 \\ 1 & 10 \leqslant x \end{cases}$$

2. Hallar la probabilidad de que la dimensión de un tornillo esté entre 2 y 5 cm.

$$P[1 \leqslant X \leqslant 5] = \int_2^5 f(x)\, dx = k\left[-\frac{1}{x}\right]_2^5 = \frac{3k}{10} = \frac{1}{3}$$

3. Determinar la dimensión máxima del 50 % de los tornillos con menor dimensión y la dimensión mínima del 5 % con mayor dimensión.

   En primer lugar, nos piden la mediana. Al ser una variable aleatoria continua, tenemos que $Me = x \in Re_X \mid F_X(x) = \frac{1}{2}$.

$$F_X(x) = \frac{1}{2} = \frac{10}{9}\left(1 - \frac{1}{x}\right) \Longrightarrow \frac{9}{20} = 1 - \frac{1}{x} \Longrightarrow \frac{1}{x} = \frac{11}{20} \Longrightarrow x = \frac{20}{11} = 1.\overline{81}$$

   Por tanto, la dimensión máxima del 50 % de los tornillos con menor dimensión es $Me = 1.\overline{81}\ cm$.

   En segundo lugar, se pide el percentil 95. Por tanto, esto equivale a $x$ tal que:

$$0{,}95 = F_X(x) = \frac{10}{9}\left(1 - \frac{1}{x}\right) \Longrightarrow 0{,}855 = 1 - \frac{1}{x} \Longrightarrow \frac{1}{x} = 0{,}145 \Longrightarrow x = \frac{1}{0{,}145} = 6{,}897$$

   Por tanto, la dimensión mínima del 5 % con mayor dimensión es $P_{95} = 6{,}897\ cm$.

4. Si $Y$ denota la dimensión de los tornillos producidos en otra fábrica, con la misma media y desviación típica que $X$, dar un intervalo en el que tome valores la variable $Y$ con una probabilidad mínima de 0,99.

   Es necesario emplear la desigualdad de Chebyshev, que no se ha visto en clase.

**Ejercicio 8.5.6.** Sea $X$ una variable aleatoria con función de densidad

$$f(x) = \begin{cases} \dfrac{2x-1}{10} & \text{si} \quad 1 < x \leqslant 2 \\ 0{,}4 & \text{si} \quad 4 < x \leqslant 6 \end{cases}$$

1. Calcular $P(1{,}5 < X \leqslant 2)$, $P(2{,}5 < X \leqslant 3{,}5)$, $P(4{,}5 \leqslant X < 5{,}5)$, $P(1{,}2 < X \leqslant 5{,}2)$.

   Tenemos que $Re_X = ]1,2] \cup ]4,6]$. Por tanto,

$$P(1{,}5 < X \leqslant 2) = \int_{1{,}5}^{2} \frac{2x-1}{10}\,dx = \left[\frac{x^2-x}{10}\right]_{1{,}5}^{2} = \frac{1}{5} - \frac{3}{40} = \frac{1}{8} = 0{,}125$$

$$P(2{,}5 < X \leqslant 3{,}5) = 0, \qquad \text{ya que } ]2{,}5,\ 3{,}5] \notin Re_X$$

$$P(4{,}5 \leqslant X < 5{,}5) = \int_{4{,}5}^{5{,}5} 0{,}4\,dx = 0{,}4\,[x]_{4{,}5}^{5{,}5} = 0{,}4$$

$$P(1{,}2 \leqslant X < 5{,}2) = \int_{1{,}2}^{2} \frac{2x-1}{10}\,dx + \int_{4}^{5{,}2} 0{,}4\,dx = \left[\frac{x^2-x}{10}\right]_{1{,}2}^{2} + 0{,}4\,[x]_4^{5{,}2} = $$
$$= \frac{22}{125} + 0{,}4(1{,}2) = \frac{82}{125} = 0{,}656$$

2. Dar la expresión general de los momentos no centrados y deducir el valor medio de $X$.

   Se definen los momentos no centrados como:

$$m_k = E[X^k] = \int_{-\infty}^{\infty} x^k f(x)\,dx = \int_1^2 x^k \cdot \frac{2x-1}{10}\,dx + \int_4^6 0{,}4x^k\,dx =$$
$$= \int_1^2 \frac{2x^{k+1}-x^k}{10}\,dx + 0{,}4\int_4^6 x^k\,dx = \frac{1}{10}\left[\frac{2x^{k+2}}{k+2} - \frac{x^{k+1}}{k+1}\right]_1^2 + 0{,}4\left[\frac{x^{k+1}}{k+1}\right]_4^6 =$$
$$= \frac{2^{k+3}-2}{10(k+2)} + \frac{-2^{k+1}+1+4\cdot 6^{k+1}-4^{k+2}}{10(k+1)}$$

   En concreto,

$$E[X] = m_1 = \frac{2^4-2}{10(3)} + \frac{-2^2+1+4\cdot 6^2-4^3}{10(2)} = \frac{259}{60} = 4{,}31\bar{6}$$

3. Calcular la función generatriz de momentos de $X$.

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \int_1^2 e^{tx}\cdot\frac{2x-1}{10}\,dx + \int_4^6 0{,}4e^{tx}\,dx =$$
$$= \frac{1}{10}\int_1^2 2xe^{tx}\,dx - \frac{1}{10}\int_1^2 e^{tx}\,dx + 0{,}4\int_4^6 e^{tx}\,dx = \left[\begin{array}{ll} u(x) = 2x & u'(x) = 2 \\ v'(x) = e^{tx} & v(x) = \frac{e^{tx}}{t} \end{array}\right] =$$
$$= \frac{1}{10}\left[\frac{2xe^{tx}}{t}\right]_1^2 - \frac{2}{10t}\int_1^2 e^{tx}\,dx - \frac{1}{10}\int_1^2 e^{tx}\,dx + 0{,}4\int_4^6 e^{tx}\,dx =$$
$$= \frac{2}{10t}\left[xe^{tx}\right]_1^2 - \frac{2}{10t^2}\left[e^{tx}\right]_1^2 - \frac{1}{10t}\left[e^{tx}\right]_1^2 + \frac{0{,}4}{t}\left[e^{tx}\right]_4^6 =$$
$$= \frac{2}{10t}\left(2e^{2t}-e^t\right) - \frac{2}{10t^2}\left(e^{2t}-e^t\right) - \frac{1}{10t}\left(e^{2t}-e^t\right) + \frac{0{,}4}{t}\left(e^{6t}-e^{4t}\right) =$$
$$= \frac{1}{10t}\left[4e^{2t} - 2e^t - \frac{2e^{2t}-2e^t}{t} - e^{2t} + e^t + 4e^{6t} - 4e^{4t}\right] =$$
$$= \frac{1}{10t}\left[3e^{2t} - e^t - \frac{2e^{2t}-2e^t}{t} + 4e^{6t} - 4e^{4t}\right]$$

Por tanto, como para $t = 0$ la función $M_X(t)$ no está definida, tenemos que no existe la función generatriz de momentos.

**Ejercicio 8.5.7.** Con objeto de establecer un plan de producción, una empresa ha estimado que la demanda de sus clientes, en miles de unidades del producto, se comporta semanalmente con arreglo a una ley de probabilidad dada por la función de densidad:

$$f(x) = \frac{3}{4}(2x - x^2), \qquad 0 \leqslant x \leqslant 2.$$

1. ¿Qué cantidad debería tener dispuesta a la venta al comienzo de cada semana para poder satisfacer plenamente la demanda con probabilidad 0,5?

   Sea $X$ una variable aleatoria que determina la demanda en miles de unidades de producto. Por tanto, se pide $\hat{x} \in [0, 2]$ tal que:

   $$0{,}5 = F_X(\hat{x}) = P[X \leqslant \hat{x}] = \int_0^{\hat{x}} f(x)\, dx = \frac{3}{4}\left[x^2 - \frac{x^3}{3}\right]_0^{\hat{x}} = \frac{3}{4}\left[\hat{x}^2 - \frac{\hat{x}^3}{3}\right]$$

   Por tanto, se busca resolver la siguiente ecuación:

   $$3\hat{x}^2 - \hat{x}^3 - 2 = 0$$

   La única solución de dicha ecuación en el intervalo $[0, 2]$ es $\hat{x} = 1$. Por tanto, han de tener dispuestas mil unidades del producto a la venta al comienzo de cada semana para poder satisfacer plenamente la demanda con probabilidad 0,5.

2. Pasado cierto tiempo, se observa que la demanda ha crecido, estimándose que en ese momento se distribuye según la función de densidad:

   $$f(y) = \frac{3}{4}(4y - y^2 - 3), \qquad 1 \leqslant y \leqslant 3.$$

   Los empresarios sospechan que este crecimiento no ha afectado a la dispersión de la demanda, ¿es cierta esta sospecha?

   En este caso, se ha denominado $Y = X$, con la diferencia en la función de densidad.

   El coeficiente de variación de Pearson se define como:

   $$C.V.(Z) = \frac{\sigma_Z}{E[Z]}$$

   Calculamos en primer lugar los siguientes valores:

   $$E[X] = \int_0^2 xf(x)\, dx = \frac{3}{4}\int_0^2 2x^2 - x^3\, dx = \frac{3}{4}\left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_0^2 = 1$$

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

$$E[Y] = \int_1^3 yf(y)\,dy = \frac{3}{4}\int_1^3 4y^2 - y^3 - 3y\,dy = \frac{3}{4}\left[\frac{4y^3}{3} - \frac{y^4}{4} - \frac{3y^2}{2}\right]_1^3 = \frac{3}{4}\cdot\frac{8}{3} = 2$$

$$E[Y^2] = \int_1^3 y^2 f(y)\,dy = \frac{3}{4}\int_1^3 4y^3 - y^4 - 3y^2\,dy = \frac{3}{4}\left[y^4 - \frac{y^5}{5} - y^3\right]_1^3 = \frac{3}{4}\cdot\frac{28}{5} = \frac{21}{5} = 4{,}2$$

Por tanto, tenemos que:

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = 0{,}2 \qquad \mathrm{Var}[Y] = E[Y^2] - E[Y]^2 = 0{,}2$$

Por tanto, la varianza no varía y la desviación típica tampoco. No obstante, como $E[X] \neq E[Y]$, la dispersión varía:

$$C.V.[X] = \frac{\sqrt{0{,}2}}{1} \neq \frac{\sqrt{0{,}2}}{2} = C.V.[Y]$$

Como podemos ver, la distribución $Y$ es más homogénea.

**Ejercicio 8.5.8.** Calcular las funciones masa de probabilidad de las variables $Y = X + 2$ y $Z = X^2$, siendo $X$ una variable aleatoria con distribución:

$$P(X = -2) = \frac{1}{5}, \quad P(X = -1) = \frac{1}{10}, \quad P(X = 0) = \frac{1}{5}, \quad P(X = 1) = \frac{2}{5}, \quad P(X = 2) = \frac{1}{10}.$$

¿Cómo afecta el cambio de $X$ a $Y$ en el coeficiente de variación?
Tenemos que $Re_X = \{-2, -1, 0, 1, 2\}$. Calculo en primer lugar la función masa de probabilidad de $Y = X + 2$, teniendo que $Re_Y = \{0, 1, 2, 3, 4\}$. Tenemos que:

$$P[Y = 0] = P[X = -2] = \frac{1}{5} \qquad P[Y = 1] = P[X = -1] = \frac{1}{10} \qquad P[Y = 2] = P[X = 0] = \frac{1}{5}$$

$$P[Y = 3] = P[X = 1] = \frac{2}{5} \qquad P[Y = 4] = P[X = 2] = \frac{1}{10}$$

Veamos si ha afectado la transformación al coeficiente de variación. Tenemos que $C.V.(X) = \dfrac{\sigma_x}{|E[X]|}$. Por ser una transformación afín, tenemos que:

$$E[Y] = E[X + 2] = E[X] + E[2] = E[X] + 2$$

$$\mathrm{Var}[Y] = \mathrm{Var}[2X + 3] = 4\,\mathrm{Var}[X] \Longrightarrow \sigma_y = 2\sigma_x$$

Por tanto, tenemos que:

$$C.V.(Y) = \frac{2\sigma_x}{|E[X] + 2|} = \frac{\sigma_x}{|E[X]|} = C.V.(X) \Longleftrightarrow \sigma_x|E[X] + 2| = 2\sigma_x|E[X]| \Longleftrightarrow$$

$$\Longleftrightarrow \left\{\begin{array}{c} \sigma_x = 0 \\ \vee \\ \pm 2E[X] = E[X] + 2 \Longleftrightarrow E[X] = 2 \quad \vee \quad E[X] = -\frac{2}{3} \end{array}\right.$$

Por tanto, tenemos que tan solo serán iguales si $\sigma_x = 0, E[X] = 2$ o $E[X] = -\frac{2}{3}$.

Calculamos ahora la función masa de probabilidad la variable $Z$, donde $Re_Z = \{0,1,4\}$.

$$P[Z=0] = P[X=0] = \frac{1}{5}$$

$$P[Z=1] = P[X=1] + P[X=-1] = \frac{1}{2}$$

$$P[Z=4] = P[X=2] + P[X=-2] = \frac{3}{10}$$

**Ejercicio 8.5.9.** Calcular las funciones de densidad de las variables $Y = 2X+3$ y $Z = |X|$, siendo $X$ una variable continua con función de densidad

$$f_X(x) = \frac{1}{4}, \qquad -2 < x < 2$$

1. $Y = 2X + 3$

   Tenemos que $Re_X = ]-2,2[$, $Re_Y = ]-7,1[$, y sea $g(x) = 2x+3$. Por el Teorema de Cambio de Variable de continua a continua, tenemos que:

   $$f_Y(y) = f_X(g^{-1}(y)) \cdot |(g^{-1})'(y)|$$

   Calculamos la inversa y su derivada:

   $$g^{-1}(y) = \frac{y-3}{2} \qquad\qquad (g^{-1})'(y) = \frac{1}{2}$$

   Por tanto, tenemos que:

   $$f_Y = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$$

2. $Z = |X|$

   Tenemos que $Re_Z = [0,2[$:

   $$h(x) = |x| = \begin{cases} -x & x < 0 \\ x & x \geqslant 0 \end{cases}$$

   En este caso, tenemos que $h$ no es inyectiva, y por tanto, hay más de una antiimagen por cada valor de $z$. En concreto, hay dos valores, el positivo y el negativo. Sean por tanto las dos inversas $h_1, h_2$.

   $$h_1^{-1}(z) = z \qquad\qquad (h_1^{-1})'(z) = 1$$

   $$h_2^{-1}(z) = -z \qquad\qquad (h_2^{-1})'(z) = -1$$

   Por tanto, por el teorema de cambio de variable, tenemos que:

   $$f_Z(z) = \sum_{k=1}^{2} f_X(h_k^{-1}(z)) \cdot |(h_k^{-1})'(z)| = \sum_{k=1}^{2} \frac{1}{4} \cdot 1 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

**Ejercicio 8.5.10.** Si $X$ es una variable aleatoria con función de densidad

$$f(x) = \frac{e^{-|x|}}{2}, \qquad -\infty < x < \infty,$$

hallar su función de distribución y las probabilidad de cada uno de los siguientes sucesos:

1. $\{|X| \leqslant 2\}$.

   En primer lugar, hallo la función de distribución:

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \frac{1}{2}\int_{-\infty}^{x} e^t = \frac{1}{2}[e^t]_{-\infty}^{x} = \frac{e^x}{2} \qquad x < 0$$

$$F_X(x) = \frac{1}{2}\int_{-\infty}^{0} e^x\,dx + \frac{1}{2}\int_{0}^{x} e^{-t}\,dt = \frac{1}{2}\left[[e^x]_{-\infty}^{0} - \left[e^{-t}\right]_0^x\right] =$$
$$= \frac{1 - e^{-x} + 1}{2} = \frac{2 - e^{-x}}{2} \qquad x \geqslant 0$$

   Por tanto, tenemos que:

$$F_X(x) = \begin{cases} \dfrac{e^x}{2} & x < 0 \\[2ex] \dfrac{2 - e^{-x}}{2} & x \geqslant 0 \end{cases}$$

   Por tanto, tenemos que:

$$P[-2 \leqslant X \leqslant 2] = P[X \leqslant 2] - P[X < -2] = P[X \leqslant 2] - P[X \leqslant -2] =$$
$$= F_X(2) - F_X(-2) = \frac{2 - e^{-2}}{2} - \frac{e^{-2}}{2} = \frac{2 - 2e^{-2}}{2} = 1 - e^{-2} \approx 0{,}8647$$

2. $\{|X| \leqslant 2 \quad \text{o} \quad X \geqslant 0\}$.

$$P[|X| \leqslant 2 \quad o \quad X \geqslant 0] = P[X \geqslant -2] = 1 - P[X < -2] = 1 - F_X(-2) = 1 - \frac{e^{-2}}{2} \approx 0{,}9323$$

3. $\{|X| \leqslant 2 \quad \text{y} \quad X \leqslant -1\}$.

$$P[|X| \leqslant 2 \quad y \quad X \leqslant -1] = P[-2 \leqslant X \leqslant -1] = F_X(-1) - F(-2) = \frac{e^{-1} - e^{-2}}{2} \approx 0{,}1163$$

4. $\{X^3 - X^2 - X - 2 \leqslant 0\}$.

   Factorizamos en primer lugar el polinomio:

$$\begin{array}{r|rrrr} & 1 & -1 & -1 & -2 \\ 2 & & 2 & 2 & 2 \\ \hline & 1 & 1 & 1 & 0 \end{array}$$

Por tanto, tenemos que $X^3 - X^2 - X - 2 = (X-2)(X^2+X+1)$.

Del polinomio restante, tenemos que $\Delta = 1 - 4 < 0$. Por tanto, no tiene soluciones. Como al evaluarlo en 0 da $1 > 0$, tenemos que es siempre positivo. Por tanto,

$$P[X^3 - X^2 - X - 2 \leqslant 0] = P[(X-2)(X^2+X+1) < 0] = P[X-2 \leqslant 0] =$$
$$= P[X \leqslant 2] = F_X(2) = \frac{2-e^{-2}}{2}$$

5. $\{X \text{ es irracional}\}$.

$$P[X \in \mathbb{R} - \mathbb{Q}] = P[X \in \{R\}] - P[X \in \{Q\}] = 1 - 0 = 0$$

La probabilidad de que esté en los reales es 1, ya que abarca todo el intervalo de definición de la variable aleatoria. En el caso de los racionales, al ser este un conjunto numerable, tenemos que su integral es nula.

**Ejercicio 8.5.11.** Sea $X$ una variable aleatoria con función de densidad

$$f(x) = 1, \qquad 0 \leqslant x \leqslant 1.$$

Encontrar la distribución de las variables:

1. $Y = \dfrac{X}{1+X}$.

Sabemos que $Re_X = [0,1]$. Además, tenemos que $h(X) = \frac{X}{1+X} = Y$. tenemos que:

$$h'(X) = \frac{(1+X) - X}{(1+X)^2} = \frac{1}{(1+X)^2} > 0$$

Por tanto, $h$ es estrictamente creciente, y tenemos que $Re_Y = \left[0, \frac{1}{2}\right]$. Como $h$ es estrictamente monótona y derivable, podemos aplicar el teorema de cambio de variable de continua a continua. Este afirma que:

$$g(y) = \begin{cases} f(h^{-1}(y))|(h^{-1})'(y)| & y \in Re_Y \\ 0 & y \notin Re_Y \end{cases}$$

Tenemos que:

$$h(X) = \frac{X}{1+X} \Longrightarrow h^{-1}(Y) = \frac{Y}{1-Y} \Longrightarrow (h^{-1})'(Y) = \frac{1-Y+Y}{(1-Y)^2} = \frac{1}{(1-Y)^2}$$

Por tanto, como $f(h^{-1}(y)) = 1$, tenemos que:

$$g(y) = \begin{cases} \frac{1}{(1-y)^2} & y \in Re_Y \\ 0 & y \notin Re_Y \end{cases}$$

Para obtener la distribución, resuelvo la integral siguiente:

$$\int_0^y \frac{1}{(1-t)^2}\,dt = \left[\frac{1}{1-t}\right]_0^y = \frac{1}{1-y} - 1$$

Por tanto,

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{1-y} - 1 & y \in Re_Y \\ 1 & y \geqslant \frac{1}{2} \end{cases}$$

2. $Z = \begin{cases} -1 & \text{si} \quad X < 3/4 \\ 0 & \text{si} \quad X = 3/4 \\ 1 & \text{si} \quad X > 3/4 \end{cases}$

   En este caso, estamos ante un cambio de variable continua a discreta. Se tiene que $Re_Z = \{-1, 0, 1\}$, y la función masa de probabilidad de $Z$ es:

$$g(-1) = P[Z = -1] = P[X < 3/4] = \int_0^{3/4} f(x)\,dx = [x]_0^{3/4} = \frac{3}{4}$$

$$g(0) = P[Z = 0] = P[X = 3/4] = \int_{3/4}^{3/4} f(x)\,dx = 0$$

$$g(1) = P[Z = 1] = P[X > 3/4] = 1 - P[X \leqslant 3/4] = \frac{1}{4}$$

   Por tanto, se tiene que la función masa de probabilidad de $Z$ es:

$$g(z) = \begin{cases} \dfrac{3}{4} & z = -1 \\[2ex] 0 & z = 0 \\[2ex] \dfrac{1}{4} & z = 1 \end{cases}$$

   Su función de dstribución es:

$$F_Z(z) = \begin{cases} 0 & z < -1 \\[2ex] \dfrac{3}{4} & -1 \leqslant z < 1 \\[2ex] 1 & 1 \leqslant z \end{cases}$$

**Ejercicio 8.5.12.** Sea $X$ una variable aleatoria simétrica con respecto al punto 2, y con coeficiente de variación 1. ¿Qué puede decirse acerca de las siguientes probabilidades?:

1. $P(-8 < X < 12)$
2. $P(-6 < X < 10)$.

Es necesario emplear la desigualdad de Chebyshev, que no se ha visto en clase.

## 8.6. Modelos de Distribuciones Discretas

**Ejercicio 8.6.1.** La probabilidad de que cada enfermo de cierto hospital reaccione favorablemente después de aplicarle un calmante es 0,01. Si se aplica el calmante a 200 enfermos, determinar:

1. La distribución de probabilidad del número de enfermos que reaccionan favorablemente, la media y la varianza.

   Sea $X$ la variable aleatoria que determina el número de enfermos que reaccionan favorablemente tras aplicarle el calmante. Tenemos que $X \rightsquigarrow B(200,\ 0{,}01)$. Por tanto:
   $$P(x) = \binom{200}{x} 0{,}01^x (0{,}99)^{200-x}$$
   Por ser una distribución binomial, tenemos que:
   $$E[X] = np = 200 \cdot 0{,}01 = 2$$
   $$\mathrm{Var}[X] = np(1-p) = 200 \cdot 0{,}01 \cdot 0{,}99 = 1{,}98$$

2. Probabilidad de que a lo sumo 2 enfermos reaccionen favorablemente.
   $$P[X \leqslant 2] = P[X = 0] + P[X = 1] + P[X = 2] = 0{,}6766$$

3. Probabilidad de que más de 3 enfermos reaccionen favorablemente.
   $$P[X > 3] = 1 - P[X \leqslant 2] - P[X = 3] = 0{,}3233 - 0{,}18136 = 0{,}14196$$

**Ejercicio 8.6.2.** Cada vez que una máquina dedicada a la fabricación de comprimidos produce uno, la probabilidad de que sea defectuoso es 0,01.

1. Si los comprimidos se colocan en tubos de 25, ¿cuál es la probabilidad de que en un tubo todos los comprimidos sean buenos?

   Sea $X$ la variable aleatoria que determina el número de comprimidos defectuosos en tubos de 25 unidades. Tenemos que $X \rightsquigarrow B(25,\ 0{,}01)$. Por tanto, si todos los comprimidos son buenos, tenemos que no hay ninguno defectuoso. Por tanto, la probabilidad de que en un tubo todos sean buenos es:
   $$P[X = 0] = \binom{25}{0} 0{,}01^0 \cdot 0{,}99^{25} = 0{,}99^{25}$$

2. Si los tubos se colocan en cajas de 10, ¿cuál es la probabilidad de que en una determinada caja haya exactamente 5 tubos con un comprimido defectuoso?

   La probabilidad de que un tubo tenga un comprimido defectuoso es:
   $$P[X = 1] = \binom{25}{1} 0{,}01^1 \cdot 0{,}99^{24} = 0{,}1964$$
   Definimos $p = 0{,}9164$. Sea $Y$ una variable aleatoria que determina el número de tubos con un comprimido defectuoso en una caja de 10. Tenemos que $Y \rightsquigarrow B(10, p)$. Por tanto, la probabilidad de que haya 5 tubos con un comprimido defectuoso es:
   $$P[Y = 5] = \binom{10}{5} p^5 \cdot (1-p)^5 = 0{,}02468$$

**Ejercicio 8.6.3.** Se capturan 100 peces de un estanque que contiene 10000. Se les marca con una anilla y se devuelven al agua. Transcurridos unos días se capturan de nuevo 100 peces y se cuentan los anillados.

1. Calcular la probabilidad de que en la segunda captura se encuentre al menos un pez anillado.

   Tenemos que de una población total de $N = 10000$ se encuentran divididos en dos poblaciones, la primera de $N_1 = 100$ y la segunda de $N - N_1$.

   Sea $X$ la variable aleatoria que contabiliza la cantidad de indiviuos de $N_1$ en una muestra de $n = 100$. Tenemos que $X \rightsquigarrow H(10^4, 10^2, 10^2)$.

   Para simplificar los cálculos, como $N_1 \leqslant 0{,}1N$, aproximamos la distribución hipergeométrica a una binomial con el mismo valor de $n$ y $p = \frac{N_1}{N} = 0{,}01$. Por tanto, $X \rightsquigarrow H(10^4, 10^2, 10^2) \cong B(100, 0{,}01)$.

   Por tanto,

   $$P[X \geqslant 1] = 1 - P[X = 0] = 1 - \binom{100}{0} 0{,}01^0 \cdot 0{,}99^{100} \approx 1 - 0{,}3660 = 0{,}63397$$

2. Calcular el número esperado de peces anillados en la segunda captura.

   Para la binomial, tenemos que:

   $$E[X] = np = 100 \cdot 0{,}01 = 1$$

**Ejercicio 8.6.4.** Un comerciante de bombillas las recibe en lotes de 20 unidades. Solo acepta un lote si, al seleccionar aleatoriamente 5 bombillas del mismo, no encuentra ninguna defectuosa.

Si un determinado lote tiene dos bombillas defectuosas, calcular la probabilidad de que el comerciante lo acepte, y el número esperado de bombillas defectuosas entre las seleccionadas, en cada uno de los siguientes casos:

- Las bombillas se seleccionan con reemplazamiento.

  En este caso, tenemos la población de $N = 20$ bombillas dividida en 2 poblaciones. En primer lugar, tenemos $N_1 = 2$ bombillas defectuosas, y $N - N_1$ bombillas correctas.

  Sea $X$ la variable aleatoria que determina el número de bombillas defectuosas que hay en una muestra de 5 elegida sin reemplazamiento. Tenemos que $X \rightsquigarrow H(20, 2, 5)$. Entonces, tenemos que:

  $$p[X = 0] = \frac{\binom{2}{0}\binom{18}{5}}{\binom{20}{5}} = \frac{18! \cdot 5! \cdot 15!}{20! \cdot 13! \cdot 5!} = \frac{15 \cdot 14}{20 \cdot 19} = \frac{21}{38} \approx 0{,}5526$$

  $$E[X] = n \cdot \frac{N_1}{N} = 5 \cdot \frac{2}{20} = 0{,}5$$

  Por tanto, tenemos que la probabilidad de que no haya ninguna bombilla defectuosa en el lote y, por tanto, se acepte, es de $p[X = 0] = 0{,}5526$. Además, el número esperado de bombillas defectuosas es 0,5; es decir, 1 o 2 bombillas defectuosas.

- Las bombillas se seleccionan sin reemplazamiento.

  La probabilidad de elegir una bombilla defectuosa de las 20 viene dada por la Regla de Laplace, y es:
$$p = \frac{2}{20} = 0{,}1$$

  Sea $X$ la variable aleatoria que determina el número de bombillas defectuosas que hay en una muestra de 5 elegida con reemplazamiento, donde la probabilidad de elegir una defectuosa es $p = 0{,}1$. Tenemos que $X \rightsquigarrow B(5, 0{,}1)$. Entonces, tenemos que:
$$p[X = 0] = 0{,}5905$$
$$E[X] = np = 5 \cdot 0{,}1 = 0{,}5$$

  Por tanto, tenemos que la probabilidad de que no haya ninguna bombilla defectuosa en el lote y, por tanto, se acepte, es de $p[X = 0] = 0{,}5905$. Además, el número esperado de bombillas defectuosas es 0,5; es decir, 1 o 2 bombillas defectuosas.

  *Observación.* También se podría haber hecho con la variable $X'$ que determinase el número de bombillas correctas hasta la primera defectuosa. Entonces, tendríamos que $X' \rightsquigarrow BN(1,\ 0{,}1)$. La probabilidad pedida sería $P[X' = 5]$.

**Ejercicio 8.6.5.** Se estudian las plantas de una determinada zona donde ha atacado un virus. La probabilidad de que cada planta esté contaminada es 0,35.

1. Definir la variable que modeliza el experimento de elegir una planta al azar y comprobar si está contaminada. Dar su ley de probabilidad.

   Tenemos que $X$ es la variable aleatoria que determina si la planta está contaminada o no, de forma que:
$$X : \left\{ \begin{array}{ll} \text{“Está contaminada”} & \longmapsto 1 \\ \text{“No está contaminada”} & \longmapsto 0 \end{array} \right.$$

   Por tanto, tenemos que se trata de un experimento de Bernouilli con $p = 0{,}35$. Dar su ley de probabilidad implica describir la distribución de Bernouilli. Tenemos que:
$$P(x) = p^x(1-p)^{1-x}$$
$$m_k = E[X^k] = \sum_{x=0}^{1} x^k P[X = x] = 0^k \cdot (1-p) + 1^k \cdot p = p$$
$$\mu_k = E[(X-E[X])^k] = E[(X-p)^k] = \sum_{x=0}^{1} (x-p)^k P[X = x] = (-p)^k \cdot (1-p) + (1-p)^k \cdot p =$$
$$= p(1-p)^k + (-p)^k(1-p)$$
$$\text{Var}[X] = E[X^2] - E[X]^2 = p - p^2 = p(1-p)$$

2. ¿Cuál es el número medio de plantas contaminadas que se pueden esperar en 5 plantas analizadas?

   Sea $X$ la variable aleatoria que determina el número de plantas contaminadas en 5 plantas analizadas. Tenemos que la probabilidad se mantiene constante, por lo que $X \rightsquigarrow B(5, 0{,}35)$. En este caso, la esperanza es:

$$E[X] = np = 5 \cdot 0{,}35 = \frac{7}{4}$$

3. Calcular la probabilidad de encontrar 8 plantas contaminadas en 10 exámenes.

   Sea $X$ la variable aleatoria que determina el número de plantas contaminadas en 10 plantas analizadas. Tenemos que la probabilidad se mantiene constante, por lo que $X \rightsquigarrow B(10, 0{,}35)$. Por tanto,

$$P[X = 8] = 0{,}0043$$

4. Calcular la probabilidad de encontrar entre 2 y 5 plantas contaminadas en 9 exámenes.

   Sea $X$ la variable aleatoria que determina el número de plantas contaminadas en 9 plantas analizadas. Tenemos que la probabilidad se mantiene constante, por lo que $X \rightsquigarrow B(9, 0{,}35)$.

$$\begin{aligned} P[2 \leqslant X \leqslant 5] &= P[X = 2] + P[X = 3] + P[X = 4] + P[X = 5] = \\ &= 0{,}2162 + 0{,}2716 + 0{,}2194 + 0{,}1181 = 0{,}8253 \end{aligned}$$

5. Hallar la probabilidad de que en 6 análisis se encuentren 4 plantas no contaminadas.

   Sea $X$ la variable aleatoria que determina el número de plantas contaminadas en 6 plantas analizadas. Tenemos que la probabilidad se mantiene constante, por lo que $X \rightsquigarrow B(6, 0{,}35)$.

   Si hay 4 no contaminadas, tenemos que hay 2 contaminadas. Por tanto,

$$P[X = 2] = 0{,}3280$$

**Ejercicio 8.6.6.** Cada página impresa de un libro contiene 40 líneas, y cada línea contiene 75 posiciones de impresión. Se supone que la probabilidad de que en cada posición haya error es 1/6000.

1. ¿Cuál es la distribución del número de errores por página?

   Determinemos cuántas posiciones de impresión hay en una página:

$$1 \text{ página} \cdot \frac{40 \text{ líneas}}{1 \text{ página}} \cdot \frac{75 \text{ posiciones de impresión}}{1 \text{ línea}} = 3{\cdot}10^3 \text{ posiciones de impresión.}$$

   Sea $X$ la variable aleatoria que determina el número de errores en una página, sabiendo que la probabilidad de que haya un error en una posición es de 1/6000. Tenemos que $X \rightsquigarrow B(3 \cdot 10^3, 1/6000)$.

2. Calcular la probabilidad de que una página no contenga errores y de que contenga como mínimo 5 errores.

$$P[X=0] = \binom{3\cdot 10^3}{0} \cdot \left(\frac{1}{6000}\right)^0 \left(\frac{5999}{6000}\right)^{3\cdot 10^3} \approx 0{,}6065$$

$$\begin{aligned}P[X \geqslant 5] = 1 - P[X \leqslant 4] &= 1 - P[X=0] - P[X=1] - \\ &- P[X=2] - P[X=3] - P[X=4] = \\ &= 1 - P[X=0] - \binom{3\cdot 10^3}{1} \cdot \left(\frac{1}{6000}\right)^1 \left(\frac{5999}{6000}\right)^{2999} - \\ &- \binom{3\cdot 10^3}{2} \cdot \left(\frac{1}{6000}\right)^2 \left(\frac{5999}{6000}\right)^{2998} - \binom{3\cdot 10^3}{3} \cdot \left(\frac{1}{6000}\right)^3 \left(\frac{5999}{6000}\right)^{2997} - \\ &- \binom{3\cdot 10^3}{4} \cdot \left(\frac{1}{6000}\right)^4 \left(\frac{5999}{6000}\right)^{2996} \approx \\ &\approx 1 - 0{,}6065 - 0{,}3033 - 0{,}0758 - 0{,}0126 - 0{,}00158 = 0{,}1716 \cdot 10^{-3}\end{aligned}$$

En este caso, lo resolvemos también mediante una aproximación a la Poisson. Como $np = 0{,}5 \leqslant 5$, podemos aproximarlo a una Poisson de $\lambda = np = 0{,}5$. $X \rightsquigarrow B(np) \cong \mathcal{P}(0{,}5)$.

$$P[X=0] = 0{,}6065$$

$$\begin{aligned}P[X \geqslant 5] = 1 - P[X \leqslant 4] &= 1 - P[X=0] - P[X=1] - \\ &- P[X=2] - P[X=3] - P[X=4] = \\ &= 1 - \sum_{x=0}^{4} \frac{e^{-0{,}5} \cdot 0{,}5^x}{x!} \approx 0{,}1721 \cdot 10^{-3}\end{aligned}$$

3. ¿Cuál es la probabilidad de que un capítulo de 20 páginas no contenga ningún error?

   Sea $X$ la variable aleatoria que determina el número de errores en 20 páginas, sabiendo que la probabilidad de que haya un error en una posición es de $1/6000$. Tenemos que $X \rightsquigarrow B(6 \cdot 10^4, 1/6000)$.

   Entonces,

$$P[X=0] = \binom{6\cdot 10^4}{0} \cdot \left(\frac{1}{6000}\right)^0 \cdot \left(\frac{5999}{6000}\right)^{6\cdot 10^4} = 45{,}3621 \cdot 10^{-6}$$

**Ejercicio 8.6.7.** Se lanzan cuatro monedas 48 veces. ¿Cuál es la probabilidad de obtener exactamente 4 caras cinco veces?

Sea el experimento de Bernouilli lanzar cuatro monedas, y consideramos como éxito obtener las 4 caras. Tenemos que esa probabilidad, por ser los 4 lanzamientos independientes, es $p = \frac{1}{2^4}$.

Sea ahora $X$ la variable aleatoria que determina el número de éxitos en dicho experimento de Bernouilli en 48 repeticiones. Tenemos que $X \rightsquigarrow B(48, p)$. Por tanto:

$$P[X=5] = \binom{48}{5} p^5 (1-p)^{43} = 0{,}1018$$

**Ejercicio 8.6.8.** Un pescador desea capturar un ejemplar de sardina que se encuentra siempre en una determinada zona del mar con probabilidad 0.15. Hallar la probabilidad de que tenga que pescar 10 peces de especies distintas de la deseada antes de:

1. pescar la sardina buscada,

   Sea $X$ el número de peces de distintas especies distintas de la deseada que ha de pescar antes de pescar la sardina buscada. La probabilidad de pescar la sardina buscada es de 0,15. Tenemos que $X \rightsquigarrow BN(1,\ 0{,}15)$. Tenemos que:

   $$P[X = 10] = \binom{10}{10} 0{,}15^1 \cdot 0{,}85^{10} \approx 0{,}0295$$

2. pescar tres ejemplares de la sardina buscada.

   Sea $X$ el número de peces de distintas especies distintas de la deseada que ha de pescar antes de pescar tres ejemplares de la sardina buscada. La probabilidad de pescar la sardina buscada es de 0,15. Tenemos que $X \rightsquigarrow BN(3,\ 0{,}15)$. Tenemos que:

   $$P[X = 10] = \binom{12}{10} 0{,}15^3 \cdot 0{,}85^{10} \approx 0{,}043854$$

**Ejercicio 8.6.9.** Un científico necesita 5 monos afectados por cierta enfermedad para realizar un experimento. La incidencia de la enfermedad en la población de monos es siempre del 30 %. El científico examinará uno a uno los monos de un gran colectivo, hasta encontrar 5 afectados por la enfermedad.

1. Calcular el número medio de exámenes requeridos.

   Como consideramos que es un gran colectivo y afirma explícitamente confirma que la probabilidad de que un mono esté infectado siempre es de 0,3, podemos suponer que no se trata de una dispersión hipergeométrica.

   Sea $X$ el número de monos examinados sanos antes de encontrar el 5º mono afectado por la enfermedad. Tenemos que $X \rightsquigarrow BN(5,\ 0{,}3)$. Tenemos que:

   $$E[X] = \frac{r(1-p)}{p} = \frac{5 \cdot 0{,}7}{0{,}3} = 11.\bar{6}$$

   Por tanto, el número de monos sanos examinados son, de media, $11.\bar{6}$.

   Por tanto, el número medio de exámenes requeridos será:

   $$E[X] + 5 = 16.\bar{6}$$

2. Calcular la probabilidad de que tenga que examinar por lo menos 20 monos.

   Como 5 monos serán siempre sanos, tenemos que buscamos la probabilidad de $X \geqslant 15$.

   $$P[X \geqslant 15] = 1 - P[X \leqslant 14] = 1 - \sum_{k=0}^{14} \binom{k+4}{k} 0{,}7^x \cdot 0{,}3^5 \approx 0{,}2822$$

Alternativamente, podemos definir una variable aleatoria $Y$ que determine el número de monos afectados en los 19 primeros. Para que haya como mínimo 20 exámenes, necesitamos que entre esos 19 monos haya menos de 5 monos afectados. Por tanto, como $Y \rightsquigarrow (19, 0{,}3)$, tenemos:

$$P[Y < 5] = P[Y \leqslant 4] = \sum_{k=0}^{4} \binom{19}{k} 0{,}3^k \cdot 0{,}7^{19-k} \approx 0{,}2822$$

3. Calcular la probabilidad de que encuentre 10 monos sanos antes de encontrar los 5 afectados.

   Sea $X$ el número de monos examinados sanos antes de encontrar el 5º mono afectado por la enfermedad. Tenemos que $X \rightsquigarrow BN(5,\ 0{,}3)$.

$$P[X = 10] = \binom{14}{10} 0{,}3^5 \cdot 0{,}7^{10} \approx 0{,}06871$$

**Ejercicio 8.6.10.** Para controlar la calidad de un determinado artículo que se fabrica en serie, se inspecciona diariamente el 5 % de la producción. Un día la máquina sufre una avería y, de los 1000 artículos fabricados ese día, produce $k$ defectuosos.

1. Dar la expresión de la probabilidad de no obtener más de un artículo defectuoso en la inspección de ese día.

   Tenemos que la población total es $N = 1000$, y se han producido $N_1 = k$ defectuosos. Además, tenemos que la muestra examinada es $n = 1000 \cdot 5\,\% = 50$.

   Sea $X$ una variable aleatoria que determina el número de productos defectuosos de la muestra de 50. Tenemos que $X \rightsquigarrow H(N, N_1, n)$. Tenemos que la probabilidad de no obtener más de un artículo defectuoso es:

$$P[X \leqslant 1] = \sum_{x=0}^{1} P[X = x] = \sum_{x=0}^{1} \frac{\binom{k}{x}\binom{1000-k}{50-x}}{\binom{1000}{50}}$$

2. Si $k = 90$, calcular la probabilidad de obtener menos de 6 artículos defectuosos en la inspección.

   Como $N_1 = k = 90 \leqslant 0{,}1N = 100$, podemos aproximar $X$ como una binomial de parámetro $p = \frac{N_1}{N} = 0{,}09$. Tenemos que $X \rightsquigarrow (N, N_1, n) \cong B(n, p)$. Tenemos que:

$$P[X < 6] = \sum_{x=0}^{5} \binom{50}{x} 0{,}09^x \cdot 0{,}91^{50-x} \approx 0{,}7072$$

**Ejercicio 8.6.11.** En una central telefónica de una ciudad se recibe un promedio de 480 llamadas por hora. Se sabe que el número de llamadas se distribuye según una ley de Poisson. Si la central sólo tiene capacidad para atender a lo sumo doce llamadas por minuto, ¿cuál es la probabilidad de que en un minuto determinado no sea posible dar línea a todos los clientes?

Calculamos en primer lugar cuántas llamadas hay por minuto:

$$480\ \frac{\text{llamadas}}{\text{hora}} \cdot \frac{1\ \text{hora}}{60\ \text{minutos}} = 8\ \frac{\text{llamadas}}{\text{minuto}}$$

Por tanto, sea $X$ la variable que determina las llamadas que se reciben en un minuto. Tenemos que $X \rightsquigarrow \mathcal{P}(8)$.

Para que no se pueda dar línea a todos los clientes, se han de recibir 13 o más llamadas. Por tanto,

$$P[X \geqslant 13] = 1 - P[X \leqslant 12] = 1 - e^{-8} \sum_{k=0}^{12} \frac{8^k}{k!} \approx 0{,}063797$$

**Ejercicio 8.6.12.** Cierta compañía de seguros ha determinado que una de cada 5000 personas fallecen al año por accidente laboral. La compañía tiene hechos 50000 seguros de vida en toda la nación y, en caso de accidente, debe abonar 3000 euros por póliza. ¿Cuál es la probabilidad de que la compañía tenga que pagar en un año por lo menos 36000 euros en concepto de primas?

Pagar 36000 euros en primas corresponde a $\frac{36000}{3000} = 12$ fallecimientos de asegurados al año. Sea $X$ la variable aleatoria que determina el número de fallecimientos de personas aseguradas al año. Tenemos que $X \rightsquigarrow B(5 \cdot 10^4, 1/5000)$. Tenemos que:

$$P[X \geqslant 12] = 1 - \sum_{k=0}^{11} \binom{5 \cdot 10^4}{x} \cdot \left(\frac{1}{5000}\right)^x \left(\frac{4999}{5000}\right)^{50000-x} = 0{,}3032$$

**Ejercicio 8.6.13.** Suponiendo que, en cada parto, la probabilidad de que nazca una niña es 0.51, y prescindiendo de nacimientos múltiples, calcular:

1. Probabilidad de que un matrimonio tenga tres hijos varones antes de tener una niña.

   Sea $X$ la variable aleatoria que determina el número de niños antes de tener una niña. Tenemos que $X \rightsquigarrow BN(1, 0{,}51)$.

   $$P[X = 3] = \binom{3}{3} \cdot 0{,}51^1 \cdot 0{,}49^3 \approx 0{,}06$$

2. Probabilidad de que tenga tres hijos varones antes de tener la segunda niña.

   Sea $X$ la variable aleatoria que determina el número de niños antes de tener dos niñas. Tenemos que $X \rightsquigarrow BN(2, 0{,}51)$.

   $$P[X = 3] = \binom{4}{3} \cdot 0{,}51^2 \cdot 0{,}49^3 \approx 0{,}1224$$

3. ¿Cuál es el número medio de hijos que debe tener un matrimonio para conseguir dos niñas?

   Sea $X$ la variable aleatoria que determina el número de niños antes de tener dos niñas. Tenemos que $X \rightsquigarrow BN(2, 0{,}51)$.

   $$E[X] = \frac{r(1-p)}{p} = \frac{2 \cdot 0{,}49}{0{,}51} \approx 1{,}9216$$

   Por tanto, de media se requieren 1,9216 hijos de varones antes de conseguir 2 niñas.

**Ejercicio 8.6.14.** El 60 % de los clientes de un almacén paga con dinero, el 30 % con tarjeta y el 10 % con cheque.

1. Calcular la probabilidad de que, de diez clientes, cuatro paguen con dinero.

   Sea $X$ una variable aleatoria que determina el número de personas que pagan con dinero de un total de 10 clientes. Tenemos que $X \rightsquigarrow B(10, 0{,}6)$.

   $$P[X = 4] = \binom{10}{4} 0{,}6^4 \cdot 0{,}4^6 \approx 0{,}1115$$

2. Calcular la probabilidad de que el décimo cliente sea el cuarto en pagar con dinero.

   Si el décimo cliente es el cuarto en pagar con dinero, previamente ha habido 6 que no han pagado con dinero.

   Sea $X$ la variable aleatoria que determinan el número de personas que no pagan con dinero antes de que 4 paguen con dinero. Tenemos que $X \rightsquigarrow BN(4, 0{,}6)$.

   $$P[X = 6] = \binom{9}{6} \cdot 0{,}6^4 \cdot 0{,}4^6 \approx 0{,}04459$$

**Ejercicio 8.6.15.** En un departamento de control de calidad se inspeccionan las unidades terminadas que provienen de una línea de ensamble. La probabilidad de que cada unidad sea defectuosa es 0,05.

1. ¿Cuál es la probabilidad de que la vigésima unidad inspeccionada sea la segunda que se encuentra defectuosa?

   Sea $X$ la variable aleatoria que determina el número de unidades adecuadas antes de la segunda defectuosa. Tenemos que $X \rightsquigarrow BN(2, 0{,}05)$.

   Como la vigésima unidad es la segunda defectuosa, previamente han llegado 18 correctas.

   $$P[X = 18] = \binom{19}{18} \cdot 0{,}05^2 \cdot 0{,}95^{18} \approx 0{,}018868$$

2. ¿Cuántas unidades deben inspeccionarse por término medio hasta encontrar cuatro defectuosas?

   Sea $X$ la variable aleatoria que determina el número de unidades adecuadas antes de la cuarta defectuosa. Tenemos que $X \rightsquigarrow BN(4, 0{,}05)$.

   $$E[X] = \frac{4 \cdot 0{,}95}{0{,}05} = 76$$

   Por tanto, de media se examinarán 76 unidades adecuadas, por lo que en total 80 unidades serán aproximadas.

3. Calcular la desviación típica del número de unidades inspeccionadas hasta encontrar cuatro defectuosas.

   $$\mathrm{Var}[X + 4] = \mathrm{Var}[X] = \frac{4 \cdot 0{,}95}{0{,}05^2} = 1520 \Longrightarrow \sigma_{X+4} = \sqrt{1520} \approx 38{,}987$$

**Ejercicio 8.6.16.** Se supone que la demanda de un cierto fármaco en una farmacia sigue una ley de Poisson con una demanda diaria media de 8 unidades. ¿Qué stock debe tener el farmacéutico al comienzo del día para tener, como mínimo, probabilidad 0,99 de satisfacer la demanda durante el día?

Sea $X$ la variable aleatoria que determina la demanda en la tienda en un día. Tenemos que $X \rightsquigarrow \mathcal{P}(8)$. Se pide el percentil 99.

$$P[X \leqslant P_{0,99}] \geqslant 0{,}99 \Longleftrightarrow e^{-8} \sum_{k=0}^{P_{0,99}} \frac{8^k}{k!} \geqslant 0{,}99 \Longleftrightarrow \sum_{k=0}^{P_{0,99}} \frac{8^k}{k!} \geqslant 0{,}99 \cdot e^8 \approx 2951{,}148$$

Tenemos que $P_{0,99} = 15$ cumple dicha condición (se ha determinando probando con valores naturales, a "fuerza bruta"). Comrpobemos la siguiente condición:

$$P[X \geqslant 15] \geqslant 0{,}01 \Longleftrightarrow 1-P[X < 15] \geqslant 0{,}01 \Longleftrightarrow P[X < 15] \leqslant 0{,}99 \Longleftrightarrow e^{-8} \sum_{k=0}^{14} \frac{8^k}{k!} \leqslant 0{,}99$$

Tenemos que es cierto, por lo que confirmamos que el valor pedido para poder satisfacer la demanda con una probabilidad de 0,99 es $P_{99} = 15$ fármacos.

**Ejercicio 8.6.17.** Los números $1, \dots, 10$ se escriben en diez tarjetas y se colocan en una urna. Las tarjetas se extraen una a una y sin devolución. Calcular las probabilidades de los siguientes sucesos:

1. Hay exactamente tres números pares en cinco extracciones.

   Sea $X$ una variable aleatoria que determina cuántos números pares hay en cinco extracciones. Tenemos que $X \rightsquigarrow H(10, 5, 5)$, donde la población total son los 10 números, $N_1$ son los 5 números pares y 5 son la muestra escogida. Por tanto,
   $$P[X = 3] = \frac{\binom{5}{3}\binom{5}{2}}{\binom{10}{5}} = \frac{25}{63} \approx 0{,}3968$$

2. Se necesitan cinco extracciones para obtener tres números pares.

   Sean los siguientes sucesos:

   - $A \longrightarrow$ Se extraen 4 pares y dos impares en las 4 primeras extracciones.
   - $B \longrightarrow$ Sale par en la última extracción.

   Tenemos que:
   $$P(A \cap B) = P(A) \cdot P(B|A)$$

   Por la Ley de Laplace, tenemos que:
   $$P(B|A) = \frac{(10-4)/2}{10-4} = \frac{1}{2}$$

   Para calcular la probabilidad de que se obtengan 2 números pares y 2 impares en las primeras 4 extracciones, trabajamos con la siguiente variable aleatoria. Sea $X$ una variable aleatoria que determina cuántos números pares hay en cuatro extracciones. Tenemos que $X \rightsquigarrow H(10, 5, 4)$, donde la población total son los 10 números, $N_1$ son los 5 números pares y 4 son la muestra escogida. Por tanto, la probabilidad buscada es:

$$P(A) = P[X = 2] = \frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}} = \frac{10}{21} \approx 0{,}4762$$

Por tanto, tenemos que:

$$P(A \cap B) = P(A) \cdot P(B|A) = \frac{5}{21} \approx 0{,}238$$

3. Obtener el número 7 en la cuarta extracción.

   Sean los siguientes sucesos:

   - $A \longrightarrow$ No extraer el 7 hasta la 4 extracción.
   - $B \longrightarrow$ Extraer el 7 en la cuarta extracción.

   Tenemos que:

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

   Por la Ley de Laplace, tenemos que:

   $$P(B|A) = \frac{1}{7}$$

   Para calcular la probabilidad de que no se extraiga el 7 en las primeras 3 extracciones, trabajamos con la siguiente variable aleatoria. Sea $X$ una variable aleatoria que determina cuántos 7 hay en las primeras tres extracciones. Tenemos que $X \rightsquigarrow H(10, 1, 3)$, donde la población total son los 10 números, $N_1$ es el 7 y 3 son la muestra escogida. Por tanto, la probabilidad buscada es:

   $$P(A) = P[X = 0] = \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} = \frac{7}{10}$$

   Por tanto, tenemos que:

   $$P(A \cap B) = P(A) \cdot P(B|A) = \frac{1}{10} = 0{,}1$$

**Ejercicio 8.6.18.** Supongamos que el número de televisores vendidos en un comercio durante un mes se distribuye según una Poisson de parámetro 10, y que el beneficio neto por unidad es 30 euros.

1. ¿Cuál es la probabilidad de que el beneficio neto obtenido por un comerciante durante un mes sea al menos de 360 euros?

   Beneficio neto de 360 euros equivale a $\frac{360}{30} = 12$ televisores.

   Sea $X$ una variable aleatoria que determina el número de televisores vendidos en un mes. Tenemos que $X \rightsquigarrow \mathcal{P}(10)$. Entonces,

   $$P[X \geqslant 12] = 1 - P[X \leqslant 11] = 1 - e^{-10} \sum_{k=0}^{11} \frac{10^k}{k!} = 0{,}3032$$

2. ¿Cuántos televisores debe tener el comerciante a principio de mes para tener al menos probabilidad 0,95 de satisfacer toda la demanda?

   Nos pide calcular el percentil 95.

$$P[X \leqslant P_{95}] \geqslant 0{,}95 \Longleftrightarrow e^{-10} \sum_{k=0}^{P_{99}} \frac{10^k}{k!} \geqslant 0{,}95$$

   Tenemos que $P_{95} = 15$.

$$P[X \geqslant 15] \geqslant 0{,}05 \Longleftrightarrow 1 - e^{-10} \sum_{k=0}^{14} \frac{10^k}{k!} \geqslant 0{,}05$$

   Por tanto, se confirma que el percentil 95 es $P_{95} = 15$. Deberá tener 15 televisores para cubrir adecuadamente la demanda con la probabilidad deseada.

**Ejercicio 8.6.19.** El número de accidentes que se producen semanalmente en una fábrica sigue una ley de Poisson, y se sabe que la probabilidad de que ocurran cinco accidentes en una semana es 16/15 de la probabilidad de que ocurran dos. Calcular:

1. Media del número de accidentes por semana.

   Sea $X$ una variable aleatoria que determina el número de accidentes que ocurren en dicha fábrica en una semana. Tenemos que $X \rightsquigarrow \mathcal{P}(\lambda)$. Para calcular $\lambda$, sabemos que:

$$P[X = 5] = \frac{16}{15} P[X = 2] \Longleftrightarrow e^{-\lambda} \cdot \frac{\lambda^5}{5!} = \frac{16}{15} e^{-\lambda} \cdot \frac{\lambda^2}{2!} \Longleftrightarrow \frac{1}{64} \cdot \lambda^5 = \lambda^2$$

   Como $\lambda > 0$, tenemos que esto se da si y solo si $\lambda^3 = 64 \Longleftrightarrow \lambda = 4$.

   Por tanto, tenemos que $X \rightsquigarrow \mathcal{P}(4)$. Por tanto, tenemos que:

$$E[X] = \lambda = 4$$

2. Número máximo de accidentes semanales que pueden ocurrir con probabilidad no menor que 0,9.

   En este caso, nos piden el percentil 90.

$$P[X \leqslant P_{90}] \geqslant 0{,}9 \Longleftrightarrow e^{-4} \sum_{k=0}^{P_{90}} \frac{4^k}{4!} \geqslant 0{,}9$$

   Tenemos que el valor buscado es $P_{90} = 7$. Comprobemos que cumple la segunda condición:

$$P[X \geqslant 7] \geqslant 0{,}1 \Longleftrightarrow 1 - e^{-4} \sum_{k=0}^{6} \frac{4^k}{4!} \geqslant 0{,}1$$

   Tenemos que es cierto, por lo que $P_{90} = 7$.