Módulo de Matemáticas Académicas II

Nivel II de ESPAD

Unidad 8
Estadística y Probabilidad

Este documento ha sido realizado por la profesora Carmen de la Fuente para el alumnado que cursa el Ámbito Científico Tecnológico (en las modalidades “enseñanzas académicas” y “enseñanzas aplicadas”) del nivel II de Educación Secundaria para Personas Adultas, concretando y desarrollando el currículo establecido para la obtención del título de Graduado en Educación Secundaria Obligatoria por personas adultas en la **Comunidad de Madrid** (BOCM de 16 de mayo de 2017).

La autora del documento agradece al equipo de Matemáticas de **Marea Verde** por compartir los archivos de sus apuntes. Para la elaboración de este documento se han utilizado partes del siguiente capítulo de los textos elaborados por el equipo de Matemáticas de **Marea Verde** **(www.apuntesmareaverde.org.es):**

7 Estadística. Azar y probabilidad. 4ºA ESO. Autores: María Molero y Andrés García Mirantes.

También se han tomado ideas y ejemplos del libro “THE CARTOON GUIDE TO STATISTICS” de LARRY GONICK & WOOLLCOTT SMITH.

ÍNDICE

1. ESTADÍSTICA

1.1. Estudios estadísticos. Población y muestra

Si queremos hacer un estudio estadístico tenemos que:

a) Recoger los datos.

b) Describir esos datos con tablas y gráficas, cálculo de parámetros estadísticos…

c) Extraer conclusiones.

Para recoger los datos y determinar los valores de la variable se puede utilizar a toda la población, todo el universo sobre el que se realiza el estudio, o seleccionar una muestra. En muchas ocasiones no es conveniente recoger valores de toda la población, porque es complicado o demasiado costoso, o incluso porque es imposible, como en el caso de un control de calidad, en que se destruya el objeto a analizar. La parte de la Estadística que se ocupa de cómo seleccionar adecuadamente las muestras se denomina **Teoría de Muestras**.

Población o universo es todo el conjunto de individuos sobre el que se realiza el estudio estadístico.

Una **muestra** es un subconjunto representativo de esa población.

Cada uno de los elementos de la población es un **individuo**.

Las características de la población que se estudian se denominan **variables estadísticas**, que se clasifican en **cuantitativas** y **cualitativas** según que los valores que tomen sean o no numéricos. Las variables cuantitativas que toman valores aislados se denominan **variables discretas** y las que pueden tomar cualquier valor de un intervalo de la recta real, **variables continuas**.

La parte de la Estadística que ordena, analiza y representa un conjunto de datos para describir sus características se denomina **Estadística Descriptiva**. Para extraer conclusiones se utilizan probabilidades y la parte de la Estadística que se ocupa de ello es la **Inferencia Estadística**.

Ejemplos:

- Si queremos conocer las preferencias en deportes del alumnado del nivel II de Educación Secundaria para Personas Adultas,, es posible preguntar a toda la población (alumnado del nivel II de Educación Secundaria para Personas Adultas,), aunque es adecuado elegir una muestra representativa, seleccionando a algunos estudiantes. En este estudio sobre preferencias deportivas, la variable utilizada es cualitativa (los valores de la variable pueden ser: fútbol, baloncesto, tenis…)
- Para conocer la intención de voto ante unas elecciones se utilizan muestras, pues preguntar a toda la población sería muy costoso (y eso ya se hace en las elecciones). La variable en este caso también es cualitativa (los valores posibles son los distintos partidos políticos que se presentan).
- Si una fábrica quiere conocer las horas de vida útil de una bombilla, una nevera, un camión… no puede poner a funcionar a toda la población, (todas las bombillas o neveras o camiones…) hasta que se estropeen pues se queda sin producción. En este caso es imprescindible seleccionar una muestra. La variable en este caso es cuantitativa, y el tiempo puede tomar cualquier valor real, es una variable cuantitativa continua.
- Si se hace un estudio sobre el número de personas que habitan las viviendas de una ciudad, se trata de una variable cuantitativa discreta (los valores posibles son números naturales).

1.2. Variable discreta. Tablas y gráficos

Al hacer un estudio estadístico o realizar un experimento aleatorio la información obtenida se resume en una tabla o distribución de frecuencias.

Posibles resultados	Frecuencia absoluta
Les gusta	28
No les gusta	12
Total	40

Ejemplo:

- *Preguntamos a 40 personas si les gusta, o no, el fútbol. En la tabla del margen reflejamos los resultados. Es una tabla de frecuencias absolutas.*

Al dividir la frecuencia absoluta entre el número total de individuos de la muestra tenemos la frecuencia relativa, así la frecuencia relativa de los que les gusta el fútbol es 28/40 = 0,7 o el porcentaje, 70 %, y la de los que no les gusta el futbol es 12/40 = 3/10 = 0,3 o el porcentaje 30 %.

La **frecuencia absoluta** de un resultado es el número de veces que se ha obtenido ese resultado.

Posibles resultados	Frecuencias relativas	Porcentaje
Les gusta	0,7	70 %
No les gusta	0,3	30 %
Suma total	1	100 %

La **frecuencia relativa** se obtiene dividiendo la frecuencia absoluta entre el número total de datos.
La suma de las frecuencias relativas es siempre igual a 1 (100 %).

Actividad resuelta

- *Se han obtenido los datos sobre el número de visitas que se han hecho a la web de los Textos Marea Verde de Matemáticas en los meses indicados, y se han reflejado en una tabla. Haz una tabla de frecuencias absolutas, relativas y porcentajes, de frecuencias acumuladas absolutas y de frecuencias relativas acumuladas.*

Marea verde	Frecuencias absolutas	Frecuencias relativas	Porcentajes	Frecuencias acumuladas absolutas	Frecuencias acumuladas relativas
Septiembre	1834	0,51	51	1834	0,52
Octubre	956	0,26	26	2790	0,77
Noviembre	432	0,12	12	3222	0,89
Diciembre	389	0,11	11	3611	1
TOTAL	3611	1	100		

Observa que las **frecuencias acumuladas** se obtienen sumando la frecuencia anterior e indica, en este ejemplo, el número de visitas hasta ese momento.

Actividades propuestas

Resultados	Frecuencias absolutas
1	17
2	12
3	17
4	15
5	21
6	14

1. Copia y completa la siguiente tabla de frecuencias absolutas de los valores obtenidos al tirar un dado, con las frecuencias relativas y porcentajes, y con frecuencias acumuladas absolutas y frecuencias acumuladas relativas.

Las **representaciones gráficas** ayudan a comprender el significado de los datos.

Dada una tabla de frecuencias (absolutas, relativas, porcentajes, acumuladas absolutas o acumuladas relativas) para representar un **diagrama de rectángulos o de barras** se traza para cada valor de la variable un rectángulo o barra de altura proporcional a la frecuencia que se esté representando.

Si se unen los puntos medios de los extremos superiores de las barras tenemos un **polígono de frecuencias relativas o diagrama de líneas**.

En un **diagrama de sectores** se dibuja un círculo que se divide en sectores de amplitudes proporcionales a las frecuencias.

Actividad resuelta

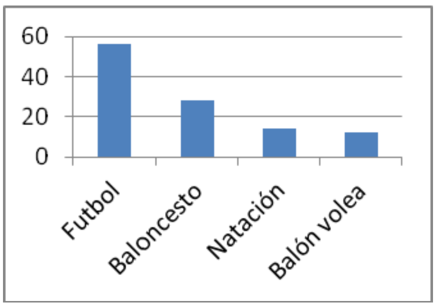
Deporte	Frecuencia Absoluta
Fútbol	56
Baloncesto	28
Natación	14
Balón volea	12
TOTAL	110

- *Tenemos un estudio estadístico sobre las preferencias deportivas del alumnado* del nivel II de Educación Secundaria para Personas Adultas, *de un determinado centro. Represéntalos en un diagrama de barras de frecuencias absolutas, en un polígono de frecuencias relativas y en un diagrama de sectores (calcula previamente la medida angular en grados de cada sector)*

El número total de datos es 110. Calculamos las frecuencias relativas y las multiplicamos por 360° para obtener los grados de cada sector:

Deporte	Fútbol	Baloncesto	Natación	Balón volea
Ángulo del sector	$\frac{56}{110}\cdot 360° \approx 183°$	$\frac{28}{110}\cdot 360° \approx 92°$	$\frac{14}{110}\cdot 360° \approx 46°$	$\frac{12}{110}\cdot 360° \approx 39°$

Diagrama de barras de frecuencias absolutas



Polígono de frecuencias relativas o diagrama de líneas

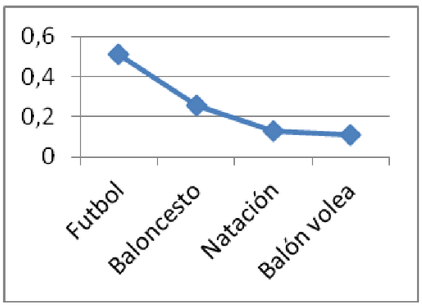
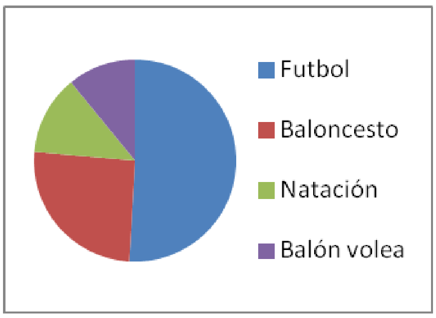


Diagrama de sectores



1.3. Parámetros de centralización

Los parámetros de centralización dan información sobre el "*centro*" de un conjunto de datos cuantitativos. Estudiamos la media aritmética, la moda y la mediana.

Para calcular la **media** ($\bar{x}$) de $x_1, x_2, \ldots, x_n$, se suman todos y se divide por el número total de datos (n).

$$\textbf{Media} = \bar{x} = (x_1 + x_2 + \ldots + x_n)/n$$

La **moda** (mo) de una distribución de frecuencias es el valor más frecuente.

La **mediana** (me) es el valor central que deja por debajo el mismo número de valores de la variable que por encima (con datos ordenados de menor a mayor).

Actividad resuelta

- *Nieves ha tenido en Matemáticas las siguientes notas: 8, 4, 6, 10 y 10. Calcula su media, su moda y su mediana.*

Su nota media se calcula sumando todas las notas: 8 + 4 + 6 + 10 + 10 = 38, y dividiendo la suma entre el número total de notas que es 5: 38/5 = 7,6.

La moda es 10 pues es el valor más frecuente.

Una forma de calcular la mediana es ordenar los valores de menor a mayor, y si el número de datos es impar, el valor central es la mediana. Si el número de datos es par, la mediana es la media de los dos datos centrales.

En nuestro caso: $4 \leq 6 \leq \mathbf{8} \leq 10 \leq 10$, por lo que la mediana es 8.

Actividades propuestas

2. Dadas las temperatura en una ciudad a una hora determinada el día 1 de cada mes se tiene la siguiente tabla:

MES	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Temperatura (° C)	–2	5	8	9	11	13	27	33	21	14	9	4

Calcula la temperatura media, la moda y la mediana.

3. Calcula la media, la mediana y la moda de las distribuciones siguientes:

a) 2, 3, 4, 5, 7, 9, 9, 1000   b) 2, 3, 4, 5, 7, 9, 9, 10   c) 0, 0, 4, 5, 7, 9, 9, 100, 200

Observa en cada caso cómo influyen los valores extremos. ¿Influyen en la moda? ¿Y en la mediana? ¿Y en la media?

Actividad resuelta

- *En una clase de 40 alumnos las calificaciones han sido:*

x_i	0	1	2	3	4	5	6	7	8	9	10	**Suma**
f_i	1	2	0	1	2	8	7	6	6	4	3	**40**

A cada nota la llamamos x_i y a la frecuencia absoluta de esa nota: f_i. Esto significa que ha habido un cero, dos unos, ningún 2… y 3 dieces.

Para calcular la media aritmética añadimos a la tabla una fila con los productos $x_i \cdot f_i$ y sumamos esa fila:

$x_i \cdot f_i$	0	2	0	3	8	40	42	42	48	36	30	**251**

Al ser 40 el número total de estudiantes la media es: **Media** = $\overline{x}$ = 251 / 40 = 6,275.

Si la variable estadística toma los valores $x_1, x_2, \ldots, x_n$, con una frecuencia absoluta $f_1, f_2, \ldots, f_n$, para calcular la **media** se multiplica cada valor por su frecuencia absoluta, se suman dichos productos y se divide por n el total de valores de la variable:

$$\textbf{media} = \overline{x} = (x_1 \cdot f_1 + x_2 \cdot f_2 + \ldots + x_n \cdot f_n) / (f_1 + f_2 + \ldots + f_n)$$

La formula anterior se resume, utilizando el símbolo sumatorio ∑ en la siguiente:

$\overline{x} = \dfrac{\sum_i x_i \cdot f_i}{N}$ donde $N = \sum_i f_i$ es el tamaño de la muestra o número total de individuos.

La **moda** es la nota más frecuente, que es mo = 5 pues es la de mayor frecuencia.

La **moda** es el valor de la variable estadística que tiene la frecuencia más alta.

Puede ocurrir que una distribución de frecuencias tenga más de una moda. Por ejemplo, la distribución:

x_i	1	2	3	4	5	6
f_i	10	9	10	8	7	10

tiene 3 modas, 1, 3 y 6, ya que el valor más alto de la frecuencia absoluta es 10 en los tres casos.

Para calcular la **mediana** añadimos una nueva fila, la de las frecuencias acumuladas:

x_i	0	1	2	3	4	5	6	7	8	9	10
Frecuencias acumuladas	1	3	3	4	6	14	21	27	33	37	40

La mitad de los datos es 40/2 = 20, y como 14 < 20 < 21, la mediana es 6.

Para obtener la **mediana** se calculan las frecuencias acumuladas y se busca el valor de la variable que ocupa el lugar central N/2.

Actividades propuestas

4. Se ha lanzado un dado 100 veces y se ha confeccionado la siguiente tabla de frecuencias absolutas:

x_i	1	2	3	4	5	6
f_i	18	16	14	16	16	20

a) Calcula la media, moda y mediana.

b)

5. Lanzamos 2 dados y sumamos los valores obtenidos. Repetimos el experimento 1000 veces y obtenemos las siguiente tabla de frecuencias absolutas.

x_i	2	3	4	5	6	7	8	9	10	11	12
f_i	24	65	73	81	158	204	148	79	68	59	41

a) Calcula la media, la mediana y la moda.

6. Calcular la media, la mediana y la moda de la siguiente tabla de frecuencias absolutas, que indica el número de hijos que tienen 200 familias entrevistadas:

x_i	0	1	2	3	4	5	6	7	8	9	10
f_i	14	65	73	27	9	6	2	1	0	2	1

1.4. Parámetros de dispersión

Los parámetros de dispersión muestran cómo están de dispersos o alejados del centro los datos de un estudio estadístico (en general ese centro es la media aritmética de la distribución).

El **recorrido** es la diferencia del valor máximo menos el valor mínimo. Aunque es una medida muy fácil de calcular, en la mayoría de los casos no es una medida muy significativa.

El parámetro más utilizado como medida de la dispersión de los datos es la desviación típica (o desviación estándar) que se define a partir de otro parámetro llamado varianza. La **varianza** es la media de los cuadrados de las desviaciones de los valores de la variable estadística respecto a la media.

Varianza = $\frac{\sum (x_i - \overline{x})^2 \cdot f_i}{N}$ o también **Varianza** = $\frac{\sum x_i^{\,2} \cdot f_i}{N} - \overline{x}^2$

Desviación típica = $\sqrt{\text{var ianza}}$

Las dos fórmulas para calcular la varianza son equivalentes. La primera es la definición de la varianza como suma de los cuadrados de las desviaciones respecto a la media. La segunda es más cómoda de usar para realizar los cálculos, sobre todo cuando los valores de la variable son enteros y el valor de la media no.

Ya sabes que la mediana nos indica el valor de la variable que ocupa el lugar central. Se denomina **primer cuartil (Q1)** al valor de la variable que deja menores o iguales que él a la cuarta parte de los datos, (o un 25 %), (siendo por tanto las tres cuartas partes mayores o iguales que él). La mediana es el segundo cuartil, que deja por debajo la mitad de los datos o un 50 %. El **tercer cuartil (Q3)** es el valor de la variable que deja menores o iguales que él las tres cuartas partes de los datos o un 75 % (y mayores o iguales la cuarta parte). Se llama **intervalo intercuartil** (o recorrido intercuartílico) a la distancia entre el tercer y el primer cuartil (**Q3 – Q1**). Por lo que hemos dicho, en ese intervalo están la mitad de los datos.

Actividad resuelta

- *Nieves ha tenido en Matemáticas las siguientes notas: 8, 4, 6, 10 y 10. Calcula su recorrido, la varianza, la desviación típica, los cuartiles y el intervalo intercuartil.*

La mayor calificación ha sido un 10 y la menor un 4, luego el **recorrido** es 10 – 4 = 6.

Recorrido = Máximo – Mínimo.

La media ya la hemos calculado y es 7’6. Queremos analizar cómo las observaciones se separan de la media. Si a cada valor le restamos la media, unos salen positivos y otros negativos, y si sumamos todos, se compensan, por lo que sale 0. Es posible superar esa dificultad calculando esas diferencias en valor absoluto, o elevándolas al cuadrado. Si las elevamos al cuadrado, sumamos todo y dividimos por el número total de valores de la variable menos 1, obtenemos la varianza.

x_i	x_i – media	(x_i – media)^2
8	0'4	0'16
4	-3'6	12'96
6	-1'6	2'56
10	2'4	5'76
10	2'4	5'76
Media = 7'6		Suma = 27'2

x_i	x_i^2
8	64
4	16
6	36
10	100
10	100
Suma = 38	Suma = 316

Si después calculamos la raíz cuadrada, se obtiene la desviación típica. Estamos evaluando la distancia de los valores de la variable a la media.

Si dividimos 27'2 entre 5 (N) se obtiene la **varianza = 5,44**.

Calculamos la raíz cuadrada de la varianza: **desviación típica = 2,33**.

En la tabla de la derecha se han dispuesto los cálculos para usar la fórmula alternativa para la varianza y comprobar que se obtiene el mismo resultado.

Varianza = (316/5) – (7'6)^2 = 63'2 – 57'76 = 5'44.

Para calcular los cuartiles debemos ordenar los datos; 4 ≤ 6 ≤ 8 ≤ 10 ≤ 10.

1º	2º	3º	4º	5º
4	**6**	8	**10**	10

El primer cuartil deja por debajo la cuarta parte o el 25 % de los datos. Hay 5 datos y 5/4 = 1'25, como 1 < 1'25 < 2, el primer cuartil es 6. Q1 = 6. El tercer cuartil deja por debajo las tres cuartas partes o el 75 % de los datos: 3(5/4) = 3'75. Como 3 < 3'75 < 4, entonces Q3 = 10.

El intervalo intercuartil = Q3 – Q1 = 10 – 6 = 4.

Intervalo intercuartil = Q3 – Q1.

Actividades propuestas

7. Dadas las temperatura en una ciudad de un ejercicio anterior:

Meses	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Temperatura	–2	5	8	9	11	13	27	33	21	14	9	4

Calcula el recorrido, la varianza, la desviación típica, los cuartiles y el intervalo intercuartil.

8. Calcula el recorrido, la varianza, la desviación típica, los cuartiles y el intervalo intercuartil de las distribuciones siguientes:

a) 2, 3, 4, 5, 7, 9, 9, 1000 b) 2, 3, 4, 5, 7, 9, 9, 10 c) 0, 0, 4, 5, 7, 9, 9, 100, 200

1.5. Variable continua: intervalos y marcas de clase. Histogramas

Recuerda que las variables pueden ser cualitativas, si no son numéricas, o cuantitativas, que a su vez pueden ser discretas o continuas.

Por ejemplo: Si se hace un estudio estadístico sobre la población de estudiantes, se puede preguntar sobre la profesión de sus padres y madres, que es una variable cualitativa, sobre el número de hermanos, que es una variable cuantitativa discreta (nadie tiene 3,7 hermanos), o sobre el peso, la estatura, la calificación media… que son variables cuantitativas continuas.

Con las variables cuantitativas continuas, o las variables que toman muchos valores diferentes, tiene sentido agrupar los valores en intervalos o clases, para que sean más fáciles de manejar e interpretar.

Al valor central o punto medio del intervalo se le denomina **marca de clase**. Es el valor x_i que representa a todo el intervalo para el cálculo de la media y de la varianza.

La representación gráfica más adecuada es el **histograma** que es un diagrama de rectángulos en el que el área de cada rectángulo es proporcional a la frecuencia.

Actividad resuelta

- *Para realizar un estudio estadístico se han recogido los datos de los pesos de 40 estudiantes de un centro escolar y se han agrupado en intervalos como se muestra en la siguiente tabla de frecuencias absolutas:*

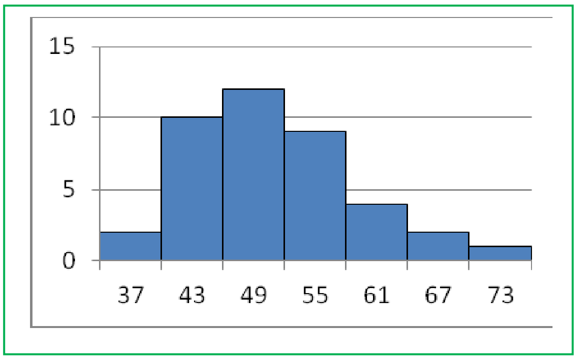
Peso (kg)	[34, 40)	[40, 46)	[46, 52)	[52, 58)	[58, 64)	[64, 70)	[70, 76)
Estudiantes	2	10	12	9	4	2	1

La tabla nos dice que hay 2 estudiantes cuyo peso es mayor o igual a 34 kg y menor que 40 kg, 10 estudiantes cuyo peso es mayor o igual a 40 kg y menor que 46 kg, …

Calculamos las marcas de clase, buscando el punto medio de cada intervalo, que se obtiene sumando los extremos y dividiendo la suma entre 2: (40 + 34)/2 = 37. Todos los intervalos en este ejemplo tienen una longitud de 6. Escribimos la tabla con las marcas de clase y las frecuencias absolutas:

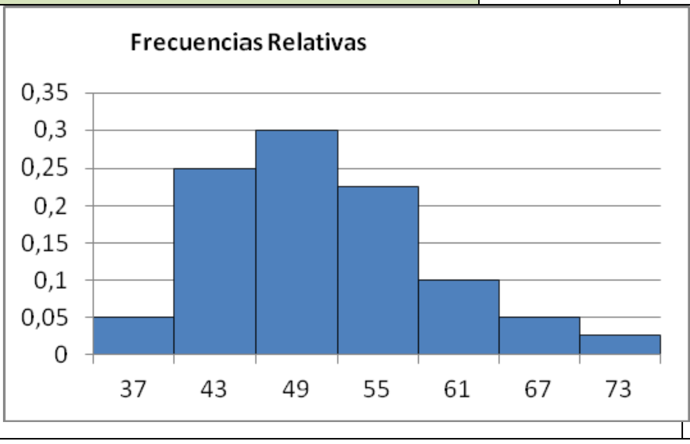
x_i	37	43	49	55	61	67	73
f_i	2	10	12	9	4	2	1

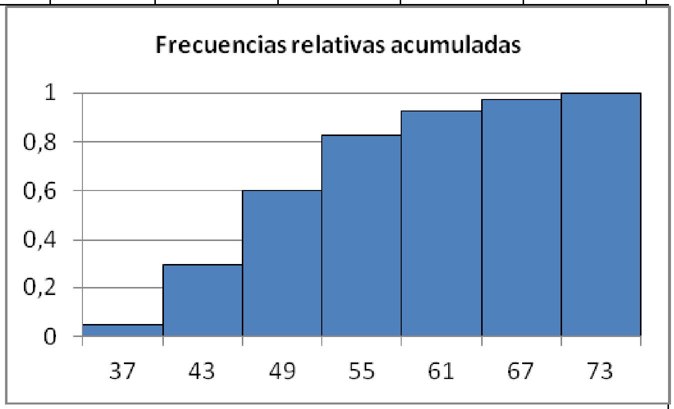
En este caso el histograma de las frecuencias absolutas es muy sencillo pues todos los intervalos tienen igual longitud. Si no fuera así, habría que calcular con cuidado las alturas de los rectángulos para que las áreas fueran proporcionales a las frecuencias.



Vamos a representar también el histograma de las frecuencias relativas y de las frecuencias relativas acumuladas:

x_i	37	43	49	55	61	67	73
Frecuencias relativas	0’05	0’25	0’3	0’225	0’1	0’05	0’025
Frecuencias relativas acumuladas	0’05	0’3	0’6	0’825	0’925	0’975	1





Cálculo de la media y la desviación típica

Procedemos de la forma que ya conocemos, calculando el producto de las marcas de clase por las frecuencias absolutas:

x_i	37	43	49	55	61	67	73	**Suma**
f_i	2	10	12	9	4	2	1	40
$x_i \cdot f_i$	74	430	588	495	244	134	73	2038

La **media** es igual a 2038/40 = 50’95.

Para calcular la **desviación típica** restamos a cada marca de clase, la media, elevamos al cuadrado y multiplicamos por la frecuencia relativa:

x_i	37	43	49	55	61	67	73	Suma
f_i	2	10	12	9	4	2	1	40
$x_i - \bar{x}$	–13’95	–7’95	–1’95	4’05	10’05	16’05	22’05	
$(x_i - \bar{x})^2$	194’60	63’2025	3’8025	16’4025	101’0025	257’6025	486’2025	1122’8175
f_i	2	10	12	9	4	2	1	40
$(x_i - \bar{x})^2 \cdot f_i$	389’20	632’025	45’63	147’62	404’01	515’205	486’2025	2619’9

La suma de los productos de los cuadrados de las desviaciones (diferencias de cada valor menos la media) por las frecuencias relativas correspondientes es 2619’9. Ahora dividimos entre N, 40, y se obtiene 65’5 que es la varianza. Calculamos la raíz cuadrada. La desviación típica es 8’09.

Si usamos la otra fórmula, $\text{var}ianza = \frac{\sum_i x_i \cdot f_i}{N} - \bar{x}^2$ organizamos los cálculos de la siguiente forma:

x_i	37	43	49	55	61	67	73	Suma
f_i	2	10	12	9	4	2	1	40
x_i^2	1369	1849	2401	3025	3721	4489	5329	22183
$x_i^2 \cdot f_i$	2738	18490	28812	27225	14884	8978	5329	106456

Varianza = (106456/40) – (50'95)^2 = 2661'4 – 2595'9 = 65'5 y *desviación típica* = 8'09.

Cálculo de la mediana y los cuartiles

Calculamos las frecuencias absolutas acumuladas y vemos a qué intervalos corresponden las frecuencias absolutas acumuladas $n/2$ para la mediana, $n/4$ para el primer cuartil, y $3n/4$ para el tercero. En nuestro caso $n/2$= 20, $n/4$= 10 y $3n/4$=30.

Intervalos	[34, 40)	[40, 46)	[46, 52)	[52, 58)	[58, 64)	[64, 70)	[70, 76)
f_i	2	10	12	9	4	2	1
F_i	2	12	24	33	37	39	40

Observamos que 12 < 20 < 24, por lo que la mediana está en el intervalo [46, 52) cuya marca de clase es 49; como 2 < 10 < 12, el primer cuartil en el intervalo [40, 46) cuya marca de clase es 43, y como 24<30<33, el tercer cuartil está en el intervalo [52, 58) cuya marca de clase es 55.

Actividades propuestas

9. Se conocen las cantidades de residuos sólidos recogidos en m^3/semana durante 12 semanas de una urbanización: 23, 27, 30, 34, 38, 21, 30, 33, 36, 39, 32, 24. Escribe una tabla de frecuencias absolutas con cuatro intervalos: [20, 25), [25, 30), [30, 35) y [35, 40). Calcula las marcas de clase. Dibuja el histograma de frecuencias absolutas. Calcula la media y la desviación típica. Calcula los intervalos de clase de la mediana y de los cuartiles.

2. AZAR Y PROBABILIDAD

2.1. Experimento aleatorio y sucesos

Se llama **experimento aleatorio** a todo experimento del que no se puede predecir el resultado porque depende del azar.

- Son ejemplos de experimentos aleatorios:

a) Lanzar una moneda y anotar si sale cara o cruz.

b) Lanzar un dado y anotar el número de la cara superior.

c) Si en una urna hay bolas blancas y rojas, sacar una al azar y anotar el color.

d) Sacar una carta de una baraja y anotar de qué palo es.

e) Abrir un libro y anotar la página por la que se ha abierto.

Sin embargo, calcular el coste de una mercancía, sabiendo el peso y el precio por kg, no es un experimento aleatorio. Tampoco lo es calcular el coste del recibo de la luz sabiendo el gasto.

Se llama **suceso elemental** a cada uno de los posibles resultados de un experimento aleatorio.

El conjunto de los posibles resultados de un experimento aleatorio se denomina **espacio muestral**, E.

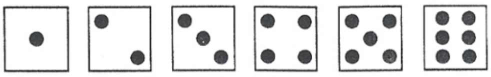
Un **suceso** es un subconjunto del espacio muestral.

Cuando el suceso es un subconjunto de E formado por un solo resultado elemental, se llama suceso elemental. Cuando el suceso se compone de varios sucesos elementales se llama **suceso compuesto**.

Al realizar el experimento aleatorio queda determinada la ocurrencia o no ocurrencia de un suceso. Como veremos más adelante, a cada suceso se le asigna una probabilidad, un número entre 0 y 1, que mide "las posibilidades" de ocurrencia del suceso antes de realizar el experimento.

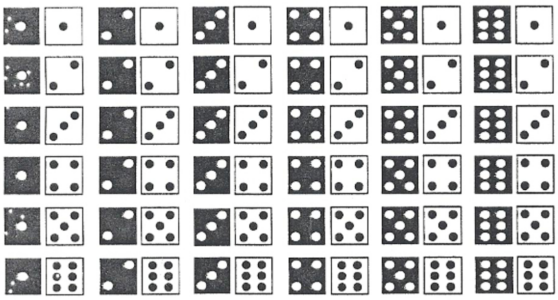
Ejemplos:

- *Los posibles resultados al tirar una moneda son que salga cara o salga cruz. El conjunto de sucesos elementales es* $E = \{cara, cruz\}$.
- *El conjunto de posibles resultados del experimentos aleatorio que consiste en sacar una bola, y anotar el color, de una bolsa en la que hay 9 bolas blancas y 5 negras es* $E = \{blanca, negra\}$.
- *El espacio muestral del experimento aleatorio que consiste en lanzar un dado y anotar el número de la cara superior es* $E = \{1, 2, 3, 4, 5, 6\}$.

  El suceso obtener par es $\{2, 4, 6\}$*, el suceso obtener impar es* $\{1, 3, 5\}$*, el suceso obtener múltiplo de 3 es* $\{3, 6\}$*, sacar un número menor que 3 es* {1, 2}.
- *Para el lanzamiento de un par de dados, supongamos que, por ejemplo, un dado es negro y el otro blanco, y en cada realización del experimento anotamos los resultados del negro y del blanco por este orden. El espacio muestral tiene 36 sucesos elementales.*

- *En el experimento "extraer una carta de una baraja española" el espacio muestral está formado por las cuarenta cartas, cada una de las cuales es un suceso elemental. El suceso "es un as" está compuesto por cuatro sucesos elementales y el suceso "es un oros" por diez.*



E ={(1,1), (2,1), (3,1), (4,1), (5,1), (6,1),
(1,2), (2,2), (3,2), (4,2), (5,2), (6,2),
(1,3), (2,3), (3,3), (4,3), (5,3), (6,3),
(1,4), (2,4), (3,4), (4,4), (5,4), (6,4),
(1,5), (2,5), (3,5), (4,5), (5,5), (6,5),
(1,6), (2,6), (3,6), (4,6), (5,6), (6,6)}

Algunos sucesos compuestos del experimento aleatorio de tirar dos dados son:

A = *"los dados suman 3"*	*A* = {(1,2), (2,1)}	*suceso formado por 2 resultados elementales*
B = *"los dados suman 6"*	*B* = {(1,5), (2,4), (3,3), (4,2), (5,1)}	*formado por 5 resultados elementales*
C = *"el dado blanco es 1"*	*C* = {(1,1), (2,1), (3;1), (4,1), (5,1), (6,1)}	*formado por 6 resultados elementales*
D = *"el dado negro es 1"*	*D* = {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)}	*formado por 6 resultados elementales.*

- *Al lanzar dos monedas escribimos* el conjunto de posibles resultados de la siguiente forma: E = {(C, C), (C, +), (+, C), (+, +)} (C significa cara y + significa cruz). El suceso *sacar cero caras* es {(+, +)}, *sacar una cara* es el suceso {(C, +), (+, C)} y *sacar dos caras* {(C, C)}.

El conjunto de todos los sucesos que pueden definirse para el espacio muestral de un experimento aleatorio recibe el nombre de **Espacio de Sucesos**, y lo designaremos por $\boldsymbol{S}$. Si el espacio muestral, $\boldsymbol{E}$, tiene n elementos, $\boldsymbol{S}$ está formado por 2^n sucesos, entre los que hay que incluir dos sucesos muy especiales:

Suceso seguro. Se llama así al suceso que siempre se verifica en la realización del experimento. Es evidente que tiene que estar formado por todos los resultados elementales y, por tanto, coincide con el propio espacio muestral, $\boldsymbol{E}$.

Suceso imposible. Se llama así al suceso que no se verifica nunca en la realización del experimento y se designa por **Ø** (conjunto vacío porque no tiene ningún elemento).

Actividades propuestas

10. Escribe el espacio muestral del experimento aleatorio: "*Escribir en cinco tarjetas cada una de las vocales y sacar una al azar*".

11. En el juego de lotería, escribe el espacio muestral e indica dos sucesos distintos de los elementales respecto a la cifra de las unidades del primer premio.

2.2. Operaciones con sucesos

Lo bueno de usar sucesos, en lugar de resultados elementales, es que podemos combinar los sucesos para construir otros sucesos, usando operaciones lógicas. Las palabras clave son ***Y, O, NO*** equivalentes a las operaciones con conjuntos: intersección ($\cap$), unión ($\cup$) y complementario ($\bar{\ }$). Es decir, dados dos sucesos A y B, podemos definir nuevos sucesos mediante las siguientes operaciones con ellos:

Intersección de sucesos: A ***Y*** $B = A \cap B$ = "Los sucesos A y B ocurren a la vez".

El suceso $A \cap B$ (se lee "A intersección B") es el subconjunto de elementos del espacio muestral que pertenecen simultáneamente a A y a B). El suceso $A \cap B$ ocurre cuando lo hacen a la vez A y B.

Unión de sucesos: A ***O*** $B = A \cup B$ = "Ocurre el suceso A o el suceso B o ambos".

El suceso $A \cup B$ (se lee "A unión B") está formado por los elementos que pertenecen a A o a B o a ambos). El suceso $A \cup B$ ocurre cuando lo hacen A o B o ambos.

Suceso contrario o complementario: ***NO*** $A = \overline{A} = E - A$= "El suceso A no ocurre" (La diferencia de conjuntos $E - A$ es el conjunto de elementos de E que no pertenecen a A).

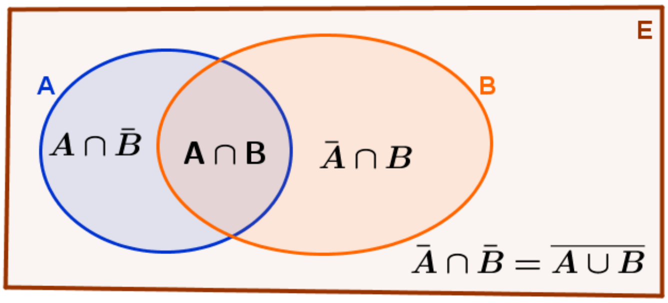
El suceso seguro y el suceso imposible son sucesos contrarios: $\overline{E} = \phi$ y $\overline{\phi} = E$.

El suceso contrario del contrario de A es, evidentemente, el propio suceso A: $\overline{\overline{A}} = A$.

Todo suceso A tiene suceso contrario $\overline{A}$ y se verifica que $A \cup \overline{A} = E$ y $A \cap \overline{A} = \phi$.

Se dice que A y B son **sucesos incompatibles** si no tienen ningún resultado elemental en común ($A \cap B = \phi$), por lo que no pueden verificarse simultáneamente en una realización del experimento. En caso contrario, es decir si $A \cap B \neq \phi$, se dice que A y B son **sucesos compatibles**.

El uso de diagramas de *Venn*, para representar los conjuntos que son los sucesos con líneas cerradas, es útil para visualizar los razonamientos en algunos problemas de probabilidad.



La imagen de la izquierda muestra dos sucesos A y B definidos en un espacio muestral E que se descompone como unión de cuatro sucesos (subconjuntos) incompatibles (disjuntos):

$A \cap \overline{B}$, $A \cap B$, $\overline{A} \cap B$ y $\overline{A \cup B}$.

Los sucesos A, B y su unión pueden descomponerse a su vez como unión de sucesos incompatibles:

$A = (A \cap B) \cup (A \cap \overline{B})$, $\quad B = (B \cap A) \cup (B \cap \overline{A})$, $\quad A \cup B = (A \cap \overline{B}) \cup (A \cap B) \cup (\overline{A} \cap B)$.

Se define la diferencia de dos suceso A y B, como el suceso que se da cuando se da A y no B, de forma que podemos escribir con la notación de la diferencia los sucesos $A \cap \overline{B} = A - B$ y $\overline{A} \cap B = B - A$.

- ***Ejemplos** con el lanzamiento de dos dados:*

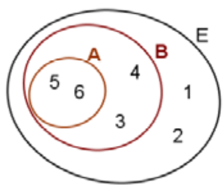
(1,1)	*(2,1)*	*(3,1)*	*(4,1)*	*(5,1)*	*(6,1)*
(1,2)	*(2,2)*	*(3,2)*	*(4,2)*	*(5,2)*	*(6,2)*
(1,3)	*(2,3)*	*(3,3)*	*(4,3)*	*(5,3)*	*(6,3)*
(1,4)	*(2,4)*	*(3,4)*	*(4,4)*	*(5,4)*	*(6,4)*
(1,5)	*(2,5)*	*(3,5)*	*(4,5)*	*(5,5)*	*(6,5)*
(1,6)	*(2,6)*	*(3,6)*	*(4,6)*	*(5,6)*	*(6,6)*

Si A es el suceso "dado blanco 1" (primera fila), y B es el suceso "dado negro 1" (primera columna):

$A \cup B$ es el área sombreada (en la que uno u otro dado es 1); $A \cap B$ es la superposición de las zonas sombreadas (en la que ambos dados son 1); $\overline{A}$ está formado por los sucesos que no están en la primera fila y $\overline{B}$ está formado por los sucesos que no están en la primera columna. $\overline{A \cup B}$ está formado por los sucesos que no están en la zona sombreada (ni en la primera fila ni en la primera columna); $\overline{A \cap B}$ está formado por todos los sucesos distintos del suceso (1,1).

- *Los sucesos C = "los dos dados suman 4" y D = "los dos dados suman 6" son incompatibles:* $C \cap D = \phi$
- *Los sucesos A = "dado blanco 1" y C = "los dos dados suman 4" son compatibles:* $A \cap C$ = {(3,1)}.

Se dice que un suceso A está **incluido** en otro suceso B, y se escribe $A \subset B$, cuando todos los elementos de A están en B, de forma que siempre que ocurre A, se da B. También se dice que A está contenido en B o que A implica B.



- **Ejemplo**: En el experimento aleatorio "lanzar un dado", el suceso A = "sacar mayor que 4" está incluido en el suceso B = "sacar mayor que 2":

$$A = \{5,6\} \qquad B = \{3, 4, 5, 6\} \qquad A \subset B.$$

2.3. Asignación de probabilidades a los sucesos elementales

Definir el concepto de probabilidad es una cuestión compleja. Entre las principales aproximaciones históricas a esta cuestión destacamos las siguientes (por su aplicación a muchos de los razonamientos que haremos en este tema):

Definición frecuentista de "probabilidad de un suceso"

Se basa en la ley de los grandes números, enunciada por *Jakob Bernouilli* (1654 - 1705) y que dice así: "La frecuencia relativa de un suceso tiende a estabilizarse en torno a un número a medida que el número de veces que se repite el experimento crece indefinidamente". Este número es a lo que se llama probabilidad del suceso.

Como comprobación experimental de la ley de los grandes números se propone realizar el experimento aleatorio de "lanzar al aire dos monedas iguales, anotando el número de caras que han salido" y repetir el experimento 100 veces. Si se va completando la siguiente tabla con las frecuencias relativas de los resultados "0 caras", "1 cara", "2 caras" después de 10 lanzamientos, después de 20 lanzamientos, etc., se puede ir viendo que cada una de estas frecuencias relativas se aproxima a un valor fijo (la probabilidad de cada resultado):

Nº de lanzamientos	10	20	30	40	50	60	70	80	90	100	$\to\infty$
Frecuencia relativa de "0 caras"											$\to$0,25
Frecuencia relativa de "1 cara"											$\to$0,5
Frecuencia relativa de "2 caras"											$\to$0,25

Usaremos esta idea intuitiva de probabilidad como frecuencia relativa en muchos razonamientos. Es decir, entendemos como probabilidad de un suceso, la frecuencia relativa con la que ocurriría dicho suceso (a la larga) si el experimento se repitiese "muchísimas veces". Como tal frecuencia relativa, la probabilidad se puede expresar en forma de fracción, como número entre 0 y 1, o como porcentaje. Así, decimos que en el lanzamiento de dos monedas la probabilidad de sacar 2 caras es 1/4, o 0,25, o del 25 % y que la probabilidad de sacar 1 cara es 1/2 = 0,5 o 50 %.

Definición clásica de "probabilidad de un suceso" o regla de Laplace

Fue enunciada por *Pierre Simon*, marqués de *Laplace* (1749 - 1827) y dice así: "La probabilidad de un suceso es el cociente entre el número de casos favorables al suceso y el número de casos posibles (dando por supuesto que todos los casos posibles tienen las mismas posibilidades de ocurrencia)".

Si indicamos la probabilidad del suceso A por p(A) la regla de *Laplace* podría escribirse mediante la siguiente fórmula:

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

Esta definición presenta un inconveniente: sólo puede aplicarse en aquellas situaciones en las que todos los casos posibles son igualmente probables (**equiprobables)**. Así, por ejemplo, nos sirve para asignar como probabilidad del suceso A = "sale 6" en el experimento de lanzar un dado, el valor p(A) = 1/6 (un caso favorable entre 6 casos posibles). Pero no nos sirve para asignar probabilidades a las distintas caras de un dado trucado. La forma de asignar probabilidades a los resultados elementales no equiprobables de este tipo de experimentos se basa en la definición frecuentista de probabilidad y consiste en realizar el experimento un número muy grande de veces, N, para asignar como valor aproximado de p(A) la frecuencia relativa de A. Cuanto mayor sea N, más preciso y fiable es el valor asignado a P(A).

$$p(A) = f_r(A) = \frac{n^{\circ}\ de\ veces\ que\ ha\ ocurrido\ A}{n^{\circ}\ de\ veces\ que\ se\ ha\ realizado\ el\ \exp erimento}.$$

Ejemplos:

- La probabilidad de que salga cara al tirar una moneda es 1/2, pues sólo hay dos casos posibles {*cara*, *cruz*}, un único caso favorable, *cara*, y suponemos que la moneda no está trucada. Si sospecháramos que la moneda estuviera trucada para asignar esa probabilidad habría que tirar la moneda un montón de veces para observar hacia qué valor se acerca la frecuencia relativa de obtener cara.
- La probabilidad de que al cruzar la calle te pille un coche NO es 1/2, aunque sólo hay dos casos posibles, que te pille el coche y que no te pille. Para calcular esa probabilidad se recogen datos de peatones atropellados y se calcula utilizando las frecuencias relativas.
- La probabilidad de sacar una bola roja de una bolsa con 7 bolas rojas y 3 bolas blancas es 7/10.
- La probabilidad de que un bebé sea niña asignada con la regla de *Laplace* es 0,5, pero al hacer el estudio con las frecuencias relativas se ha visto que es 0,49.
- Si consideramos una baraja española de 40 cartas, antes de sacar ninguna carta todas ellas son igualmente factibles, y como puede salir una cualquiera de las 40 cartas decimos que la probabilidad, de por ejemplo, *sacar el caballo de copas* es 1/40, la de *sacar un oro* es 10/40, y la de un *as* es 4/40.

Actividades propuestas

12. Calcula la probabilidad de que al sacar una carta de la baraja de 40 cartas sea una figura (sota, caballo o rey).

13. Se ha lanzado un dado trucado 1000 veces y cada cara ha salido el número de veces indicado en la siguiente tabla:

resultado obtenido	1	2	3	4	5	6
nº de veces	48	95	144	192	238	283

Asigna una probabilidad a cada uno de los resultados elementales del experimento aleatorio que consiste en lanzar este dado trucado y anotar el número obtenido.

14. Halla la probabilidad de que la suma de los puntos de las caras visibles de un dado que se lanzó al azar sea múltiplo de 5.

15. Un jugador expresó a *Galileo* su sorpresa al observar que al jugar con tres dados, la suma 10 aparece con más frecuencia que la suma 9. Explica el porqué de su sorpresa calculando la probabilidad de cada suceso.

16. Se escuchan tres discos y se vuelven a guardar al azar. ¿Cuál es la probabilidad de que al menos uno de los discos haya sido guardado en el envoltorio que le correspondía?

2.4. Probabilidad de un suceso. Propiedades de la probabilidad

A partir de la idea intuitiva de probabilidad de un suceso como frecuencia relativa con la que se produce el suceso cuando el número de realizaciones del experimento es muy grande, y teniendo en cuenta las propiedades de las frecuencias relativas, *Kolmogorov* enuncia la siguiente definición axiomática de probabilidad (que es la "oficial" en la matemática moderna):

Dado un experimento aleatorio, con espacio muestral E y espacio de sucesos S, se llama **probabilidad** a una aplicación que asocia a cada suceso A del espacio de sucesos S un número real, que llamamos probabilidad de A y representamos por p(A), que cumple los siguientes axiomas o condiciones:

1. $p(A) \geq 0$ (**No tiene sentido una probabilidad negativa**: *peor que imposible no es posible*)
2. $p(E) = 1$ (**La probabilidad del suceso seguro tiene que ser 1**)
3. **Si A y B son sucesos incompatibles ($A \cap B = \phi$), entonces** $p(A \cup B) = p(A) + p(B)$.

De estas propiedades axiomáticas se deducen otras propiedades:

Probabilidad del suceso contrario: $\boxed{p(\overline{A}) = 1 - p(A)}$

Demostración: Como $A \cup \overline{A} = E$ y $A \cap \overline{A} = \phi \Rightarrow p(A) + p(\overline{A}) = p(E) = 1 \Rightarrow p(\overline{A}) = 1 - p(A)$.

Probabilidad del suceso imposible: $\boxed{P(\phi) = 0}$ *(El suceso imposible NO puede ocurrir).*

Se demuestra a partir de la propiedad anterior porque $\overline{\phi} = E \Rightarrow p(\overline{\phi}) = 1 - p(E) = 1 - 1 = 0$.

Si $A \subset B \quad \Rightarrow \quad p(A) \leq p(B)$ (propiedad evidente a partir de la propiedad 3).

$\boxed{p(A) \leq 1}$ cualquiera que sea el suceso A (propiedad evidente a partir de la anterior).

Actividades resueltas

- *¿Cuál es la probabilidad de sacar un as en la baraja de 40 cartas? ¿Y de* ***no*** *sacar un as? ¿Y de sacar una copa? ¿Y de* ***no*** *sacar una copa?*

El suceso *no sacar un as* es el suceso **contrario** al de *sacar un as*. Cartas que no son ases hay 36, luego la probabilidad de no sacar as es 36/40 = 9/10. Se verifica que p(no as) = 1 – p(as).

La probabilidad de *sacar copa* es 10/40 y la de ***no*** *sacar copa* es 1 − 10/40 = 30/40 (o 3/4, o del 75 %).

Actividades propuestas

17. ¿Cuál es la probabilidad de *no* sacar un 5 al tirar un dado? ¿Y de *no* sacar un múltiplo de 3? ¿Y de *no* sacar un número menor que 2?

18. Al tirar una moneda dos veces, ¿cuál es la probabilidad de no sacar ninguna cara? ¿Y de sacar al menos una cara? Observa que sacar al menos una cara es el suceso contrario de no sacar ninguna cara.

Hemos visto que si A y B son dos sucesos incompatibles $p(A \cup B) = p(A) + p(B)$. Cuando los sucesos son compatibles hay que usar una regla más general:

PROBABILIDAD DE LA UNIÓN (regla de la suma): $\boxed{p(A \cup B) = p(A) + p(B) - p(A \cap B)}$

Esta regla general aplicada al caso particular de sucesos incompatibles se transforma en la anterior.

Ejemplos:

- *¿Cuál es la probabilidad de, en una baraja de 40 cartas, sacar una copa o un oro?*

  Hay 10 copas y 10 oros, y ninguna carta es a la vez copa y oro, luego la probabilidad es 20/40.

  Los sucesos C = "sacar una copa" y O = "sacar un oro" son incompatibles, $A \cap B = \phi$,

  $p(C) = \frac{10}{40}$, $p(O) = \frac{10}{40}$ y observamos que se verifica $p(A \cup B) = p(A) + p(B) = \frac{20}{40} = \frac{1}{2} = 0{,}5$.

- *¿Cuál es la probabilidad de, en una baraja de 40 cartas, sacar un as o un oro?*

  Hay 4 ases y hay 10 oros, pero hay el *as de oros*, luego las cartas que son o bien un as o bien un oro son 13, luego la probabilidad es 13/40.

  Los sucesos A = "sacar un as" y O = "sacar un oro" son compatibles $A \cap O =$ {as de oros}:

  $p(A) = \frac{4}{40}$, $p(O) = \frac{10}{40}$ y $p(A \cup O) = p(A) + p(O) - p(A \cap O) = \frac{4}{40} + \frac{10}{40} - \frac{1}{40} = \frac{13}{40} = 0{,}325$.

Actividades resueltas

- *Calcula la probabilidad de sacar un basto o una figura.*

  Hay 10 bastos y hay 12 figuras, pero hay 3 figuras que son a la vez bastos (sota, caballo y rey), luego p(Basto o Figura) = 10/40 + 12/40 – 3/40 = 19/40 = 0,475.

Actividades propuestas

19. Lanzamos dos dados no trucados y anotamos los números de su cara superior. Consideramos el suceso A que la suma de las dos caras sea 8, y el suceso B que esos números difieran en dos unidades. Calcula a) p(A); b) p(B); c) p($A \cap B$); d) p($A \cup B$); e) p($A \cap \overline{B}$); f) p($\overline{A} \cap B$); g) p($\overline{A} \cap \overline{B}$).

La generalización a la unión de más de dos sucesos es sencilla si son **sucesos incompatibles dos a dos**:

$$Si\ A_i \cap A_j = \phi \Rightarrow P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

La generalización de la regla de la suma cuando los sucesos no son incompatibles dos a dos no es tan sencilla, por eso es tan importante descomponer la unión en sucesos incompatibles.

Como consecuencia de la regla anterior se concluye que **la probabilidad de un suceso es la suma de las probabilidades de los sucesos elementales que lo forman**.

Para un experimento aleatorio con n sucesos elementales, E = {O_1, O_2, ..., O_n}, si P(Oi) designa la probabilidad del suceso Oi, se tiene que verificar que $P(O_1) + P(O_2) + \ldots + P(O_n) = 1$.

Cuando todos los resultados elementales del experimento aleatorio son **equiprobables** y se quiere calcular la probabilidad de un suceso compuesto, la regla de la suma se convierte en la **regla de Laplace,** ya que si todos los sucesos tienen la misma probabilidad, entonces $n\,P(O_i) = 1 \Rightarrow P(O_i) = \frac{1}{n}$.

Si un suceso A está compuesto por m de estos sucesos elementales equiprobables, entonces:

$$\mathrm{P(A)} = \frac{1}{n} + \overset{\text{m sumandos}}{\dots} + \frac{1}{n} = \frac{m}{n} = \frac{\text{nº de elementos de A}}{\text{nº de elementos de E}}$$

Pero los sucesos elementales no tienen necesariamente la misma probabilidad.

Ejemplo:

- *Si se lanza un dado cargado para que 1 salga un 25 % de las veces (a la larga) y los otros resultados sean igualmente probables, el espacio muestral es el mismo que para un dado "justo", E = {1, 2, 3, 4, 5, 6}, pero ahora P(1) = 0,25 mientras que P(2) = P(3) = P(4) = P(5) = P(6) = 0,15.*

  En este caso no podemos utilizar la regla de *Laplace* para calcular la probabilidad de un suceso.

  Por ejemplo, para calcular la probabilidad del suceso A = "sale impar", tenemos que utilizar la regla de la suma:

$$\mathrm{P(A)} = \mathrm{P(\{1\})} + \mathrm{P(\{3\})} + \mathrm{P(\{5\})} = 0{,}25 + 0{,}15 + 0{,}15 = 0{,}55$$

  (con este dado trucado el resultado sería impar el 55 % de las veces, a la larga).

2.5. Probabilidad condicionada

Nos acercaremos al concepto de probabilidad condicionada con un **ejemplo:**

- *Supongamos que cambiamos ligeramente el experimento aleatorio de lanzar dos dados (uno blanco y uno negro) y que lanzamos el dado blanco antes que el dado negro. ¿Cuál es la probabilidad de que los dos dados sumen 3, suceso al que llamaremos A?*

Antes de lanzar el primer dado $p(A) = \frac{2}{36}$ porque A = {(1, 2), (2, 1)}.

Supongamos ahora que el dado blanco ha sido 1 (Suceso C) ¿Cuál es la probabilidad de A ahora?

A esta última probabilidad la llamamos la probabilidad condicionada de que ocurra el suceso A dada la condición de que el suceso C ya se ha producido. Escribimos p(A/C) y leemos "probabilidad de A dado C" o bien "probabilidad de A condicionada por C".

Antes de que se lanzara el primer dado, el espacio muestral tenía 36 resultados elementales, pero ahora que el suceso C ha ocurrido los posibles resultados deben pertenecer al espacio muestral reducido por C. En este espacio muestral reducido, compuesto por seis resultados elementales, sólo un resultado, (2, 1), suma 3. Por lo tanto, la probabilidad condicional es 1/6.

En general, para calcular la probabilidad condicional p(A/C) miramos el suceso A/C como parte del espacio muestral reducido por C. Esta última idea se plasma en una definición formal:

La probabilidad de A condicionada por C es: $\boxed{p(A/C) = \dfrac{p(A \cap C)}{p(C)}}$

A partir de la definición se deducen algunas propiedades evidentes:

- P(A/A) = 1 (Una vez que A ocurre su ocurrencia es segura).
- Cuando A y C son incompatibles P(A/C) = 0 (Si ha ocurrido C es imposible la ocurrencia A).

En el ejemplo anterior: $$P(A/C) = \frac{P(A \cap C)}{P(C)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}.$$

Ejemplo de elaboración y análisis de una tabla de contingencia

- *Se ha producido un acalorado debate entre dos institutos próximos sobre la dificultad de aprobar el bachillerato en cada uno de ellos. Para dilucidar la cuestión, ambos institutos dan a conocer el número de estudiantes que han aprobado 2º de bachillerato en las convocatorias de Junio de los últimos 5 años y el número total de estudiantes matriculados en cada una de las modalidades "bachillerato de Ciencias Sociales" y "bachillerato de Ciencia y Tecnología":*

BACHILLERATO DE CIENCIAS SOCIALES:

INSTITUTO 1: Estudiantes matriculados: 433 Estudiantes aprobados: 141

INSTITUTO 2: Estudiantes matriculados: 695 Estudiantes aprobados: 248

BACHILLERATO DE CIENCIA Y TECNOLOGÍA:

INSTITUTO 1: Estudiantes matriculados: 597 Estudiantes aprobados: 325

INSTITUTO 2: Estudiantes matriculados: 359 Estudiantes aprobados: 207

Con estos datos en la mano, el instituto 1 afirma que en su centro aprueba 2º de bachillerato un porcentaje de estudiantes superior al que aprueba en el instituto 2 y que por lo tanto la probabilidad de aprobar es mayor en su centro. Con los mismos datos, el instituto 2 afirma que el porcentaje de estudiantes aprobados en el bachillerato de Ciencias Sociales es mayor en su centro y que también es mayor el porcentaje de estudiantes aprobados en el bachillerato de Ciencia y Tecnología. Concluye, por lo tanto, que la probabilidad de aprobar el bachillerato de Ciencias Sociales es mayor en su centro, y que también es mayor la probabilidad de aprobar el bachillerato de Ciencia y Tecnología.

¿Es posible que ambos centros tengan razón? Se puede comprobar que así es, aunque parezca paradójico, completando las siguientes tablas de contingencia y calculando después las probabilidades que intervienen en el conflicto planteado:

INSTITUTO 1	***bachillerato CS***	***bachillerato CT***	***totales***
Aprobados	*141*	*325*	*466*
Suspensos	*292*	*272*	*564*
Totales	*433*	*597*	*1030*

INSTITUTO 2	*bachillerato CS*	*bachillerato CT*	*totales*
Aprobados	*248*	*207*	*455*
Suspensos	*447*	*152*	*599*
Totales	*695*	*359*	*1054*

- *¿Cuál es la probabilidad de aprobar si se cursa el bachillerato de Ciencias Sociales en cada uno de los institutos?*

  Se pide la probabilidad condicional del suceso A = "aprobar" *condicionada por el suceso* S = "cursar la modalidad de bachillerato de Ciencias Sociales"*, es decir, la frecuencia relativa de aprobados en el grupo reducido de estudiantes que cursan esta modalidad.*

  En el INSTITUTO 1: $P(A/S) = \dfrac{N^{o}\ de\ alumnos\ del\ bachillerato\ CS\ \ y\ \ aprobados}{N^{o}\ de\ alumnos\ del\ bachillerato\ CS} = \dfrac{141}{433} = 0{,}3256$

  (Aprueban el bachillerato de Ciencias Sociales un 33 % de los alumnos, aproximadamente).

  En el INSTITUTO 2: $P(A/S) = \dfrac{N^{o}\ de\ alumnos\ del\ bachillerato\ CS\ \ y\ aprobados}{N^{o}\ de\ alumnos\ del\ bachillerato\ CS} = \dfrac{248}{695} = 0{,}3568$.

  (Aprueban el bachillerato de Ciencias Sociales un 36 % de los alumnos, aproximadamente).

  La probabilidad de aprobar el bachillerato de Ciencias Sociales es mayor en el instituto 2.

- *¿Cuál es la probabilidad de aprobar si se cursa el bachillerato de Ciencia y Tecnología en cada uno de los institutos?*

  Ahora se pide la probabilidad del suceso A = "aprobar" *condicionada por el suceso* C = "cursar la modalidad de Ciencia y Tecnología"*, es decir, la proporción de alumnos aprobados en el grupo reducido de alumnos que cursan esta modalidad.*

  En el INSTITUTO 1: $P(A/C) = \dfrac{N^{o}\ de\ alumnos\ del\ bachillerato\ CT\ \ y\ aprobados}{N^{o}\ de\ alumnos\ del\ bachillerato\ CT} = \dfrac{325}{597} = 0{,}5444$.

  (Aprueban el bachillerato de Ciencia y Tecnología un 54% de los alumnos, aproximadamente)

  En el INSTITUTO 2: $P(A/C) = \dfrac{N^{o}\ de\ alumnos\ del\ bachillerato\ CT\ \ y\ aprobados}{N^{o}\ de\ alumnos\ del\ bachillerato\ CT} = \dfrac{207}{359} = 0{,}5766$.

  (Aprueban el bachillerato de Ciencia y Tecnología un 58 % de los alumnos, aproximadamente).

  La probabilidad de aprobar el bachillerato de Ciencia y Tecnología es mayor en el instituto 2.

- *¿Cuál es la probabilidad de aprobar en cada instituto?*

  Ahora se piden la proporción de alumnos aprobados en el total de alumnos matriculados, sin diferenciar la modalidad de bachillerato.

  En el INSTITUTO 1: $P(A) = \dfrac{N^{o}\ de\ alumnos\ \ aprobados}{N^{o}\ de\ alumnos\ total} = \dfrac{466}{1030} = 0{,}4524$.

  (Aprueban 2º de bachillerato un 45 % de los alumnos, aproximadamente).

En el INSTITUTO 2: $P(A)=\dfrac{N^{o}\ de\ alumnos\ \ aprobados}{N^{o}\ de\ alumnos\ total}=\dfrac{455}{1054}=0{,}4316$.

(Aprueban 2º de bachillerato un 43 % de los alumnos, aproximadamente).

La probabilidad de aprobar 2º bachillerato es mayor en el instituto 1.

- *¿Cuál es la probabilidad de cursar el bachillerato CS y aprobar en el instituto1?*

Ahora se pide la proporción de alumnos que cursan esta modalidad y aprueban en el total de alumnos matriculados:

$$P(A\cap S)=\frac{N^{o}\ de\ alumnos\ \ del\ \ bachillerato\ CS\ \ y\ aprobados}{N^{o}\ de\ alumnos\ total}=\frac{141}{1030}=0{,}1369\ .$$

(Hacen bachillerato CS y lo aprueban un 14 % de los alumnos, aproximadamente).

- *¿Cuál es la probabilidad de cursar el bachillerato CS en el instituto1?*

Ahora se pide la proporción de alumnos de esta modalidad en el total de alumnos matriculados:

$$P(S)=\frac{N^{o}\ de\ alumnos\ \ del\ bachillerato\ CS}{N^{o}\ de\ alumnos\ total}=\frac{433}{1030}=0{,}4204\ .$$

(Hacen bachillerato CS un 42 % de los alumnos).

Comprobamos la fórmula dada como definición de probabilidad condicionada para el instituto 1:

$$\mathrm{P(A/S)}=\frac{\mathrm{P(A\cap S)}}{\mathrm{P(S)}}=\frac{\frac{141}{1030}}{\frac{433}{1030}}=\frac{141}{433}\,.$$

Actividades propuestas

20. Se está estudiando un tratamiento con un nuevo medicamento, para lo que se seleccionan 100 enfermos. A 60 se les trata con el medicamento y a 40 con un placebo. Los valores obtenidos se representan en la tabla adjunta

	Medicamento (M)	**Placebo (no M)**	**totales**
Curados (C)	50	30	80
No curados (no C)	10	10	20
totales	60	40	100

Calcula con estos datos :

a) La probabilidad de que un enfermo curado haya sido tratado con el medicamento.

b) La probabilidad de que un enfermo curado haya sido tratado con el placebo.

2.6. Probabilidad de la intersección. Independencia de sucesos

Despejando en la definición de la probabilidad condicionada , $p(C/A) = \frac{p(C \cap A)}{p(A)}$, la probabilidad de la intersección y teniendo en cuenta que $C \cap A = A \cap C$, se obtiene la <u>regla de multiplicación</u> para calcular la probabilidad de la intersección de dos sucesos:

$$\boxed{p(A \cap C) = p(A) \cdot p(C/A)}$$

Ejemplo:

- *Para calcular la probabilidad de que al sacar dos cartas de una baraja española sean ambas reyes, se hace un razonamiento como el que sigue:*

  Nombramos los sucesos R_1 = "la primera carta es rey", R_2 = "la segunda carta es rey". Nos piden la probabilidad del suceso $R_1 \cap R_2$ = "la primera y la segunda carta son reyes". Aplicamos la regla del producto: $P(R_1 \cap R_2) = P(R_1) \cdot P(R_2 / R_1) = \frac{4}{40} \cdot \frac{3}{39} = 0{,}0077$.

La regla del producto se generaliza de la siguiente manera para calcular la probabilidad de la intersección de más de dos sucesos:

$$\boxed{P(A_1 \cap A_2 \cap A_3 \cap ... \cap A_n) = P(A_1) \cdot P(A_2 / A_1) \cdot P(A_3 / A_1 \cap A_2) \cdots P(A_n / A_1 \cap A_2 \cap A_3 \cap ... \cap A_{n-1})}$$

Ejemplo:

- *La probabilidad de que al sacar cuatro cartas de una baraja española sean los cuatro reyes se calcula usando la regla del producto:*

$$P(R_1 \cap R_2 \cap R_3 \cap R_4) = P(R_1) \cdot P(R_2 / R_1) \cdot P(R_3 / R_1 \cap R_2) \cdot P(R_4 / R_1 \cap R_2 \cap R_3) = \frac{4}{40} \cdot \frac{3}{39} \cdot \frac{2}{38} \cdot \frac{1}{37} = 0{,}00001$$

INDEPENDENCIA DE SUCESOS Y REGLA ESPECIAL DEL PRODUCTO

Se dice que A y C son dos **sucesos independientes** uno del otro si la ocurrencia de uno de ellos no tiene ninguna influencia en la probabilidad de ocurrencia del otro. *Por ejemplo, en el experimento aleatorio del lanzamiento de dos dados, el lanzamiento de un dado y su resultado no tiene ningún efecto en el lanzamiento del otro (a menos que estén pegados, magnetizados, etc.)*. En caso contrario, es decir, si la ocurrencia de un suceso influye en la probabilidad de otro, se dice que son **sucesos dependientes**.

En términos de probabilidad condicionada se puede dar la siguiente definición:

A y C son dos **sucesos independientes** si $p(A/C) = P(A)$ o, lo que es equivalente, $p(C/A) = P(C)$.

Cuando A y C son independientes tenemos una regla especial de multiplicación:

$$\boxed{p(A \cap C) = p(A) \cdot p(C) \quad \Leftrightarrow \quad A \ y \ C \ son \ independientes}$$

La relación anterior se usa también para demostrar la independencia de dos sucesos A y B de los que se conocen o se pueden calcular $p(A)$, $p(B)$ y $p(A \cap B)$. En otras situaciones sabemos, por la naturaleza del experimento aleatorio, que los sucesos A y B son independientes, y entonces usamos la relación $p(A \cap B) = p(A) \cdot p(B)$ para calcular la probabilidad de la intersección de A y de B como producto de las probabilidades de cada uno de los sucesos. Hay que tener muy presente que esta relación es un caso particular de la regla general del producto que solamente es aplicable cuando los sucesos son independientes.

La regla del producto se generaliza de manera obvia para la probabilidad de la intersección de n sucesos independientes entre sí:

$$\boxed{p(A_1 \cap A_2 \cap \ldots \cap A_n) = p(A_1) \cdot p(A_2) \cdot \ldots \cdot p(A_n) \quad si \ A_1, A_2, \ldots, A_n \ \ son \ \ independientes}$$

Ejemplo:

- *Tenemos una bolsa con 3 bolas rojas y 2 bolas negras.*

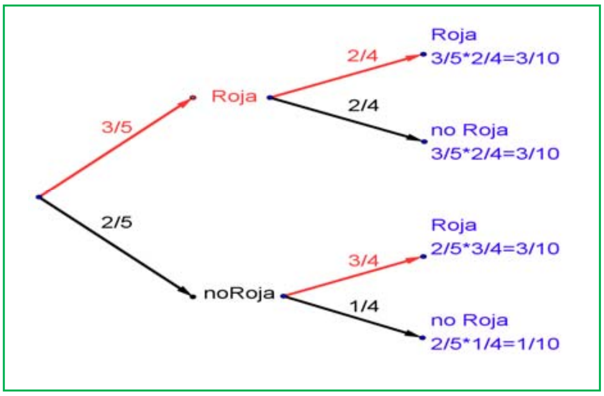
En el experimento aleatorio se sacar dos bolas de la bolsa, la probabilidad de que las dos sean rojas depende de si volvemos a meter en la bolsa la primera bola o si la dejamos fuera. En el primer caso decimos que es extracción **con reemplazamiento** y en el segundo, **sin reemplazamiento**.

La probabilidad de que la primera bola que se saca sea roja es 3/5. Si volvemos a meter la bola en la bolsa después de anotar el color, la probabilidad de sacar bola roja en la segunda extracción volverá a ser 3/5 (la probabilidad de esta segunda bola *no depende* de lo que hayamos sacado la primera vez). En este caso la probabilidad de sacar dos bolas rojas es 3/5 · 3/5 = 9/25.

Pero si la primera bola ha sido roja y la dejamos fuera, ahora en la bolsa sólo hay 4 bolas y de ellas sólo quedan 2 bolas rojas, luego la probabilidad de que esa segunda bola sea roja es 2/4 (está condicionada por lo que hemos sacado antes).

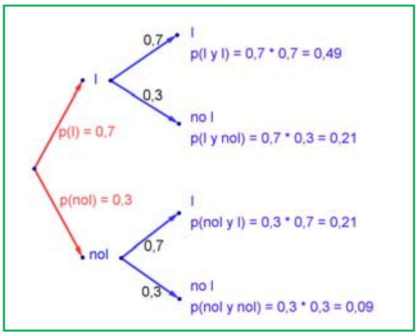
La probabilidad de sacar dos bolas rojas sin reemplazamiento es: 3/5 · 2/4 = 6/20 = 3/10.

Observa el diagrama de árbol en el que se representan todas las posibilidades para la primera y la segunda extracción y dónde se escribe cada probabilidad.



La probabilidad de sacar primero una bola roja y luego una bola negra (no roja) es 3/5 · 2/4 = 3/10 pues después de sacar una bola roja en la bolsa quedan sólo 4 bolas y de ellas 2 son negras. La probabilidad de sacar primero una bola negra y luego bola roja es 2/5·3/4 = 6/20 = 3/10, y la de sacar dos bolas negras es: 2/5·1/4 = 2/20 = 1/10.

Actividades resueltas



- *Si consideramos que la probabilidad de que un incendio sea intencionado es 0,7 ¿cuál es la probabilidad de que al considerar dos incendios, al menos uno haya sido intencionado?*

Llamamos I al suceso "ser intencionado" y noI al suceso "no ser intencionado". Representamos la situación en un diagrama de árbol. Como el que un incendio sea intencionado es independiente de cómo sea el segundo, tenemos que:

P(I y I) = 0,7 · 0,7 = 0,49 P(I y noI) = 0,7 · 0,3 = 0,21

P(noI y I) = 0,3 · 0,7 = 0,21 P(noI y noI) = 0,3 · 0,3 = 0,09

La probabilidad de que al menos uno haya sido intencionado la podemos calcular sumando las probabilidades de (I y I), (I y noI), y (noI y I) que es 0,49 + 0,21 + 0,21 = 0,91. Pero más sencillo es calcular la probabilidad del suceso contrario p(noI y noI) = 0,09 y restarla de 1:

p(al menos uno intencionado) = 1−p(ninguno intencionado9 = 1 – 0,09 = 0,91.

- *El Caballero de Mére planteó en el siglo XVII al matemático Pascal el siguiente problema:*

"¿Qué es más probable: sacar al menos un seis en cuatro lanzamientos de un dado, o sacar al menos un doble seis en 24 lanzamientos de un par de dados?".

Él pensaba erróneamente que los dos sucesos eran igualmente probables, pero había comprobado (apostando muchas veces) que perdía más frecuentemente con el segundo tipo de apuesta. Calcula ambas probabilidades para resolver el problema planteado.

Sea A el suceso de conseguir al menos un 6 en 4 lanzamientos de un dado. ¿Cuál es la probabilidad $P(A)$? Este es uno de esos sucesos para los que es más fácil calcular la probabilidad del suceso contrario o complementario: $\overline{A}$ es el suceso "no sale ningún 6 en las 4 tiradas".

Si $\overline{A_i}$ es el suceso "no se consigue un 6 en la i-ésima tirada", sabemos que $P(\overline{A_i}) = 5/6$. También sabemos que los sucesos $\overline{A_i}$ son independientes, por lo tanto:

$P(\overline{A}) = P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap \overline{A}_4) = \left(\frac{5}{6}\right)^4 = 0{,}482$ Así pues, $P(A) = 1 - P(\overline{A}) = 0{,}518$.

(Si de Mère realizaba muchas veces este tipo de apuesta, debería ganar aproximadamente el 52 % de las mismas).

Y ahora, la otra parte: sea D el suceso "conseguir al menos un seis doble en 24 lanzamientos de dos dados". Nuevamente $NO\ D$ es más fácil de describir: $\overline{D}$ es el suceso "no se consigue un doble seis en ninguna de las 24 tiradas".

Si $\overline{D}_i$ es el suceso "no sale un seis doble en la i-ésima tirada", entonces la probabilidad de cada $\overline{D}_i$ es $P(\overline{D}_i) = 35/36$, por lo tanto: $P(\overline{D}) = P(\overline{D}_1 \cap \overline{D}_2 \cap \ldots \cap \overline{D}_{24}) = \left(\frac{35}{36}\right)^{24} = 0{,}509$.

Y podemos concluir que $P(D) = 1 - P(\overline{D}) = 0{,}491$.

(Por lo tanto, de Mére debería ganar aproximadamente el 49 % de las veces con este tipo de apuesta).

Actividades propuestas

21. Se tiene una bolsa con 10 bolas rojas y 6 negras, de la que se extraen dos bolas. Hallar la probabilidad de que ambas sean negras en cada uno de los siguientes casos:

a) La primera bola se devuelve a la bolsa antes de extraer la segunda bola;

b) La primera bola no se devuelve a la bolsa.

22. En un examen de física, un alumno sólo ha estudiado 15 temas de los 25 que contiene el cuestionario. El examen consiste en contestar a dos temas extraídos al azar del total de temas del cuestionario. Hallar la probabilidad de que los dos temas sean de los que el alumno estudió.

23. Para tratar de curar una enfermedad se ha aplicado un nuevo tratamiento a una serie de individuos, obteniéndose los resultados reflejados en la siguiente tabla que hay que completar con los totales:

Elegido un individuo al azar, halla las siguientes probabilidades:

Tratamiento	Curados	NO curados	totales
T. Nuevo	60	21	
T. Antiguo	43	36	
Totales			

a) Que se haya curado si se le ha aplicado el nuevo tratamiento.

b) Que habiendo sido sometido al nuevo tratamiento, no se haya curado.

c) Que habiendo sido sometido al tratamiento antiguo, se haya curado.

d) Que se haya curado si ha recibido el tratamiento antiguo.

e) Que se haya curado.

24. La probabilidad de que una bomba lanzada por un avión haga blanco en el objetivo es 1/3. Calcula la probabilidad de alcanzar el objetivo si se tiran tres bombas seguidas.

25. Se tienen dos urnas: la urna A contiene 6 bolas rojas y 4 bolas negras; la urna B contiene 4 bolas rojas y 8 bolas negras. Se lanza un dado y si sale un múltiplo de 3 extraemos una bola de la urna A, en caso contrario extraemos una bola de la urna B. Dibuja un diagrama en árbol para describir esta experiencia compuesta y calcula la probabilidad de que la bola extraída sea roja.

EJERCICIOS Y PROBLEMAS

Estadística

1. Las notas de alumnos de un grupo en un examen de Matemáticas son:

2, 10, 7, 8, 1, 0, 3, 5, 6, 9, 2, 4, 1, 6, 9, 10, 5, 6, 7, 8, 3, 1, 0, 1, 5, 9, 10, 9, 8, 7.

a) Haz una tabla de frecuencias absolutas, frecuencias relativas (expresadas en fracción, número porcentaje), frecuencias acumuladas absolutas y frecuencias relativas acumuladas (expresadas en fracción, número porcentaje).
b) Calcula la media, la mediana y la moda.
c) Calcula la desviación típica y los cuartiles.

2. Se ha preguntado a 40 alumnos por el número de hermanos que tenía, y se ha obtenido

Número de hermanos	**0**	**1**	**2**	**3**	**4**	**5**	**6**
Número de veces	5	15	7	6	4	2	1

Calcula la media, la mediana y la moda.

3. Se han lanzado cuatro monedas 100 veces y anotado el número de caras en cada lanzamiento. Los resultados están reflejados en la tabla siguiente:

Número de caras	**0**	**1**	**2**	**3**	**4**
Número de lanzamientos	7	25	36	26	6

a) Escribe la tabla de frecuencias absolutas, y frecuencias relativas.

b) Representa un diagrama de sectores después de calcular el ángulo de cada sector.

c) Representa un diagrama de barras para las frecuencias relativas.

d) Calcula la media y la desviación típica.

e) Calcula la mediana y los cuartiles.

4. Para conocer la distribución de un cierto país de las personas según su edad se ha recogido una muestra de diez mil personas y los valores obtenidos vienen reflejados en la tabla siguiente:

Edades	[0, 5)	[5, 10)	[10, 15)	[15, 25)	[25, 35)	[35, 45)	[45, 55)	[55, 65)	[65,105)
Número de personas	900	1000	900	1500	1300	1200	1300	900	1000

a) Representa un histograma de frecuencias absolutas. *Cuidado*: Los intervalos no son todos iguales. *Recuerda:* El área de los rectángulos debe ser proporcional a las frecuencias.
b) Calcula la media y la desviación típica.

5. Una compañía de seguros selecciona al azar a 200 propietarios y les pregunta cuántos euros han gastado en reparaciones del automóvil, con los siguientes resultados agrupados en intervalos:

Euros	[0, 100)	[100, 200)	[200, 400)	[400, 600)	[600, 800)	[800, 3000)
Nº de personas	40	30	20	40	50	20

a) Representa un histograma de frecuencias relativas. *Cuidado*: Los intervalos no son todos iguales.
b) Calcula la media y la desviación típica.

Probabilidad

6. Lanzamos dos monedas distintas, una de 50 céntimos y otra de un euro. Calcula la probabilidad de que: A) En la moneda de un euro salga cara. B) Salga una cara. C) Salga al menos una cara. D) No salga ninguna cara. E) Salga una cara y una cruz.

7. Escribe el experimento aleatorio de lanzar tres monedas y anotar si cada una de ellas ha sido cara o cruz. Calcula las probabilidades de los siguientes sucesos: A = No sale ninguna cara; B = Sale al menos una cara; C = Salen dos caras y una cruz.

8. Lanzamos dos dados y anotamos los valores de las caras superiores. Calcula las probabilidades de que la suma sea 1, sea 2, sea 3, …., sea 12.

9. Lanzamos a la vez una moneda y un dado. Calcula las probabilidades de los sucesos A = "Sale cara y un número par"; B = "Sale cruz o un número primo" y C = "sale un número primo".

10. Una moneda está trucada. La probabilidad de obtener cara es doble que la de obtener cruz. Calcula las probabilidades de los sucesos C = "obtener cara" y X =" obtener cruz" al tirar la moneda.

11. Tenemos un dado trucado de forma que los números impares tienen una probabilidad doble a la de los números pares. Calcula las probabilidades de: A) Salga un número impar. B) Salga un número primo. C) Salga un número primo impar. D) Salga un número que sea primo o sea impar.

12. En una clase hay 24 alumnos y 14 alumnas. La mitad de las alumnas y la tercera parte de los alumnos tienen los ojos azules. Se elige un estudiante al azar. A) Calcula la probabilidad de que sea chico y tenga los ojos azules. B) Calcula la probabilidad de que sea chico o tenga los ojos azules.

13. En un grupo de 12 amigas hay 3 rubias. Se eligen dos chicas al azar. Calcula la probabilidad de que: A) Ambas sean rubias. B) Ninguna sea rubia. C) Al menos una sea rubia.

14. Lanzamos dos dados y anotamos los valores de las caras superiores. Calcula las probabilidades de que: A) Los números obtenidos sean iguales. B) Los números obtenidos difieran en 3 unidades. C) Los números obtenidos sean pares.

15. Un lote de 20 artículos tiene 2 defectuosos. Se sacan 4 al azar, ¿cuál es la probabilidad de que ninguno sea defectuoso?

16. Se ha comprobado que en una ciudad están enfermos con diarrea el 60 % de los niños; con sarampión el 50 %, y el 20 % con ambas enfermedades. A) Calcular la probabilidad de que elegido un niño al azar, esté enfermo con diarrea o sarampión o ambas enfermedades. B) En un colegio con 450 niños, ¿Cuántos cabe esperar que estén enfermos con diarrea o sarampión?

17. En el lanzamiento de dos dados se definen los sucesos: A = "la suma de puntos es 5", B = "uno de los dados es 4" y C = "los dos números son iguales". Calcular A) p(A), B) p(B), C) p(C), D) p($A \cup B$), E) p($A \cap B$), F) p($A \cap C$).

18. En una empresa hay 200 empleados, de los cuales la mitad son mujeres y la mitad varones. Son fumadores 15 mujeres y 20 varones. Se elige un empleado al azar. Calcula la probabilidad de que A) sea fumador o fumadora, B) sea mujer y no fume, C) fume sabiendo que es mujer.

19. Ana tiene en su monedero 4 monedas de cinco céntimos, 3 monedas de veinte céntimos y 2 monedas de un euro. Saca dos monedas al azar. Calcula las probabilidades de que A) las dos sean de cinco céntimos, B) ninguna sea de un euro, C) saque 1,20 €.

20. Una encuesta revela que el *35 %* de los habitantes de una ciudad oyen la cadena SER, el 28 % la COPE y el 10 % ambas emisoras de radio. Se elige un ciudadano al azar: ¿Cuál es la probabilidad de que no escuche ni la SER ni la COPE?

AUTOEVALUACIÓN

Con los datos siguientes, 1, 5, 2, 8, 9, 4, 7, 7, 5, 7, calcula:

1. La media:

a) 5 b) 5'5 c) 6 d) 7

2. La mediana:

a) 5 b) 5'5 c) 6 d) 7

3. La moda:

a) 5 b) 5'5 c) 6 d) 7

4. La desviación típica:

a) 2 b) 2,3 c) 2,5 d) 2,6

5. El intervalo intercuartil

a) 3 b) 2,75 c) 4 d) 2

6. Al tirar dos dados, la probabilidad de sacar al menos un 5 es:

a) 5/6 b) 11/36 c) 25/36 d) 30/36

7. Al tirar 3 monedas, la probabilidad de sacar exactamente dos caras es:

a) 1/2 b) 3/4 c) 3/8 d) 5/8

8. Al tirar 3 monedas, la probabilidad de sacar al menos dos caras es:

a) 1/2 b) 3/4 c) 3/8 d) 5/8

9. Sacamos una carta de una baraja de 40 cartas, la probabilidad de que sea un oro o un múltiplo de 2 es:

a) 22/40 b) 19/40 c) 36/40 d) 3/4

10. Indica cuál de las afirmaciones siguientes es **siempre** correcta:

a) $P(A) + P(noA) = 1$

b) $P(A\ y\ B) = P(A) \cdot P(B)$

c) $P(A\ o\ B) = P(A) + P(B)$.

SOLUCIONES: 1b 2b 3d 4d 5b 6b 7a 8d 9a 10a

SOLUCIONES A LAS ACTIVIDADES PROPUESTAS

(1)

Resultados	**Frecuencias absolutas**	**Frecuencias relativas**			**Frecuencias absolutas acumuladas**	**Frecuencias relativas acumuladas**
		fracción	**número**	**porcentaje**		
1	**17**	**17/96**	**0,18**	**18%**	**17**	**0,18**
2	**12**	**12/96**	**0,13**	**13%**	**19**	**0,31**
3	**17**	**17/96**	**0,18**	**18%**	**46**	**0,49**
4	**15**	**15/96**	**0,15**	**15%**	**61**	**0,64**
5	**21**	**21/96**	**0,22**	**22%**	**82**	**0,86**
6	**14**	**14/96**	**0,14**	**14%**	**96**	**1**
totales	**96**	**96/96**	**1**	**100%**		

(2) Media $\overline{x} = 12{,}67°$ Moda $= 9°$ Mediana $= 10°$.

(3) Los valores extremos influyen en la media.

	Media	Mediana	Moda
a)	116,11	6	9
b)	6,11	6	9
c)	34,00	6,5	0,9

(4) Media = 3,56 Mediana = 3,625 Moda = 6.

(5) Media = 7,032 Mediana = 6,98 Moda = 7.

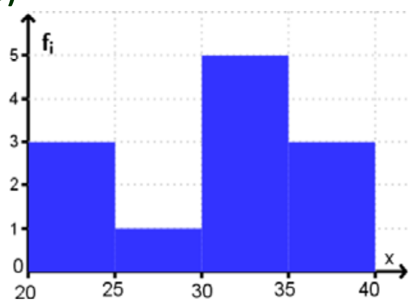
(6) Media = 2,025 Mediana = 1,78 Moda = 1.

(7) Recorrido=35 Varianza= 269 Desviación típica=16,4 Q_1=7,25 Q_3=15,75 Intervalo intercuartil=8,5.

(8)

	Recorrido	Varianza	Desviación típica	Q_1	Q_3	Intervalo intercuartil
a)	998	109870,61	331,47	4	9	5
b)	8	8,11	2,85	4	9	5
c)	200	4303,11	65,60	4,2	9	4,75

(9)



Intervalo de clase **I_i**	[20, 25)	[25, 30)	[30, 35)	[35, 40)
Marcas de clase **x_i**	22,5	27,5	32,5	37,5
Frecuencias absolutas **f_i**	3	1	5	3

Media = 30,83 m^3/semana; Desviación típica 29,53 m^3/semana. El intervalo de la mediana es [25, 30). El primer cuartil está en el intervalo [20, 25) y el tercer cuartil en [30, 35).

(10) E = {a, e, i, o, u} (11) E = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}. Respuesta abierta, por ejemplo, A = ”es múltiplo de 3” = {3, 6, 9}, B = “es mayor que 5” = {6, 7, 8, 9}.

(12) 12/40 = 0,3 o 30 %.

(13)

resultado	1	2	3	4	5	6
PROBLABILIDAD	0,048	0,095	0,144	0,192	0,238	0,283

(14) 1/3

(15) p(suma 9) = 25/216 p(suma 10) = 27/216 **(16)** *2/3* **(17)** 5/6 , 4/6, 5/6 **(18)** 1/4, 3/4

(19) a) p(A) = 5/36, b) p(B) = 8/36; c) p($A \cap B$) = 2/36, d) p($A \cup B$) = 11/36; e) p($A \cap \overline{B}$) = 3/36; f) p($\overline{A} \cap B$) = 6/36; g) p($\overline{A} \cap \overline{B}$) = 25/36.

(20) (a) p(M/C) = 50/80 = 5/8 = 0,625 (b) p(no M/C) = 30/80 = 3/8 = 0,375.

(21) a) 0,141 b)0,125 **(22)** 0,35 **(23)** a) 0,741 b) 0,259 c)0,544 d) 0,544 e) 0,644.

(24) 0,704 **(25)** 0,422.

SOLUCIONES DE EJERCICIOS Y PROBLEMAS

(1) (a)

		0	1	2	3	4	5	6	7	8	9	10	*sumas*
Frecuencias absolutas		**2**	**4**	**2**	**2**	**1**	**3**	**3**	**3**	**3**	**4**	**3**	*30*
Frecuencias relativas	**fracción**	**2/30**	**4/30**	**2/30**	**2/30**	**1/30**	**3/30**	**3/30**	**3/30**	**3/30**	**4/30**	**3/30**	*30/30*
	número	**0,07**	**0,13**	**0,07**	**0,07**	**0,03**	**0,1**	**0,1**	**0,1**	**0,1**	**0,13**	**0,1**	*1*
	porcentaje	**7%**	**13%**	**7 %**	**7 %**	**3 %**	**10%**	**10%**	**10%**	**10%**	**13 %**	**10%**	*100 %*
F. absolutas acumuladas		**2**	**6**	**8**	**10**	**11**	**14**	**17**	**20**	**23**	**27**	**30**	
Frecuencias relativas acumuladas	**fracción**	**2/30**	**6/30**	**8/30**	**10/30**	**11/30**	**14/30**	**17/30**	**20/30**	**23/30**	**27/30**	**30/30**	
	número	**0,07**	**0,2**	**0,27**	**0,34**	**0,37**	**0,47**	**0,57**	**0,67**	**0,77**	**0,9**	**1**	
	porcentaje	**7%**	**20%**	**27%**	**34%**	**37%**	**47%**	**57%**	**67%**	**77%**	**90%**	**100%**	

(b) Media = 5,4; Mediana = 6 ; Modas: 1 y 9 (c) Desviación típica = 3,25 ; Q_1 = 2 ; Q_3 = 8.

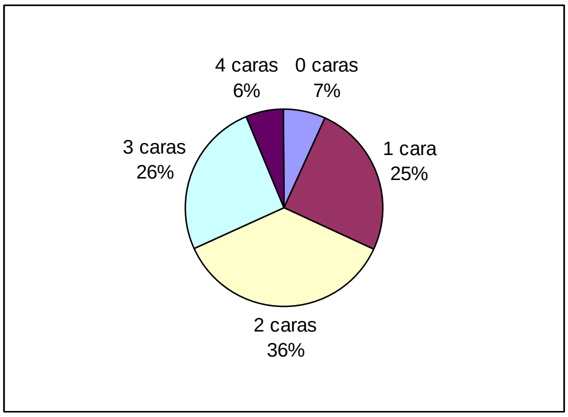
(2) Media = 1,975; Mediana = 1 Moda = 1 hermano.

(3)

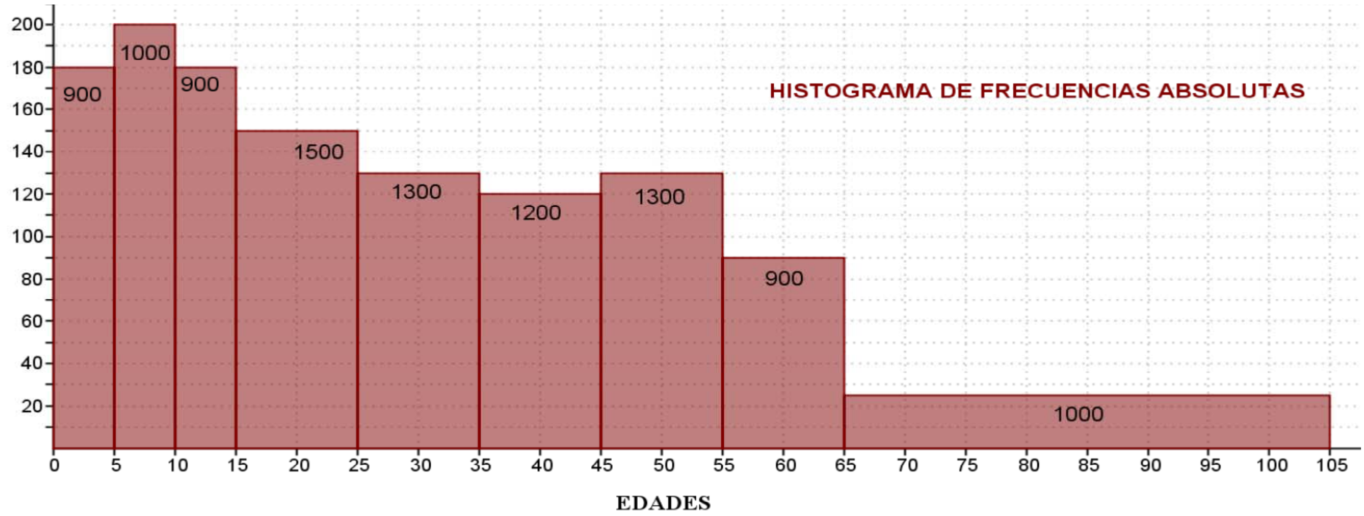
Número de caras	0	1	2	3	4
Frecuencia absoluta	7	25	36	26	6
Frecuencia relativa	0,07	0,25	0,36	0,26	0,06
Ángulo del sector	25°	90°	130°	94°	21°

Media = 1,99; Desviación típica = 1,015.

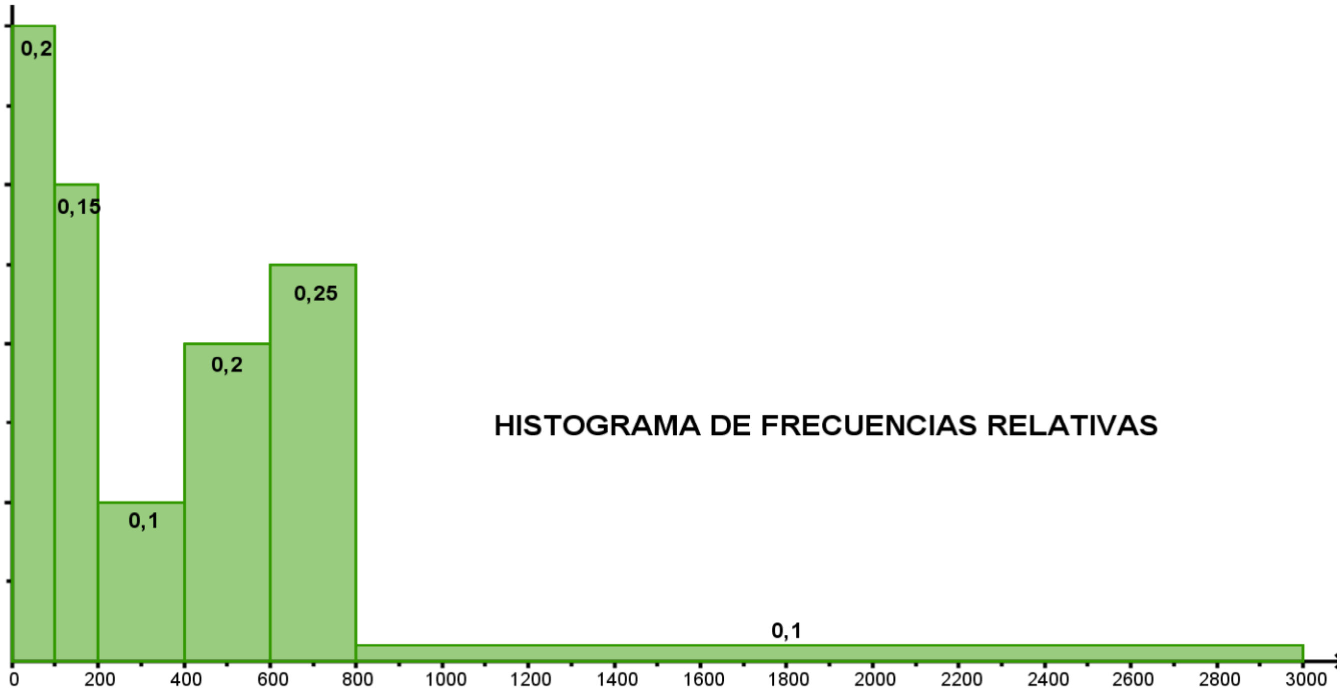
Mediana = 2; Q_1 = 1; Q_3 = 3.



(4) Media = 33,95 años; Desviación típica = 23,8 años.



(5) Media = 527,5, Desviación típica = 518.



(6) A) 1/2 B) 1/2 C) 3/4 D) 1/4 E) ½.
(7) Espacio muestral. E = {CCC, CCX, CXC, CXX, XCC, XCX, XXC, XXX}; p(A) = 1/8; p(B) = 7/8; p(C) = 3/8.
(8) p(1) = 0, p(2) = 1/36, p(3) = 2/36, p(4) = 3/36, p(5) = 4/36, p(6) = 5/36, p(7) = 6/36, p(8) = 5/36, p(9) = 4/36, p(10) = 3/36, p(11) = 2/36, p(12) = 1/36.

(9) p(A)=1/4, p(B)=3/4, p(C)=1/2 **(10)** p(C)=2/3, p(X)=1/3 **(11)** A) 6/9, B) 5/9, C) 7/9 **(12)** A)8/38, B)31/38
(13) A) 1/22 B)6/11 C)5/11. **(14)** A) 1/6, B)1/6, C)1/4. **(15)** 12/19. **(16)** A) 0,9. B)405 niños.
(17) A)1/9. B) 11/36. C) 1/6. D)13/36. E) 1/18. F)0 **(18).** A) 0,175. B) 0,425 C) 0,15.
(19) A) 1/6, B) 7/12, C) 1/6, **(20)** 0,47.