

Chapter 3 Probabilidad

En este capítulo introduciremos el concepto de probabilidad a partir de frecuencias relativas.

Definiremos los eventos como los elementos sobre los que se aplica la probabilidad. Los eventos compuestos se definirán usando álgebra de conjuntos.

Luego discutiremos el concepto de probabilidad condicional derivado de la probabilidad conjunta de dos eventos.

3.1 Experimentos aleatorios

Recordemos el objetivo básico de la estadística. La estadística se ocupa de los datos que se presentan en forma de observaciones.

- Una **observación** es la adquisición de un número o una característica de un experimento

Las observaciones son realizaciones de **resultados**.

- Un **resultado** es una posible observación que es el resultado de un experimento.

Al realizar experimentos, a menudo obtenemos resultados diferentes. La descripción de la variabilidad de los resultados es uno de los objetivos de la estadística.

- Un **experimento aleatorio** es un experimento que da diferentes resultados cuando se repite de la misma manera.

La pregunta filosófica detrás es ¿Cómo podemos conocer algo si cada vez que lo miramos cambia?

3.2 Probabilidad de medición

Nos gustaría tener una medida para el resultado de un experimento aleatorio que nos diga **cuán seguros** estamos de observar el resultado cuando realicemos un **futuro** experimento aleatorio.

Llamaremos a esta medida la probabilidad del resultado y le asignaremos valores:

- 1, cuando estamos seguros de que la observación sucederá.

3.3 Probabilidad clásica

Siempre que un experimento aleatorio tenga M resultados posibles que son todos **igualmente probables**, la probabilidad de cada resultado i es

$$P_i = \frac{1}{M}$$

La probabilidad clásica fue defendida por Laplace (1814).

Dado que cada resultado es **igualmente probable** en este tipo de experimento, declaramos una completa ignorancia y lo mejor que podemos hacer es distribuir equitativamente la misma probabilidad para cada resultado.

- No observamos P_i
- Deducimos P_i de nuestra razón y no necesitamos realizar ningún experimento para conocerla.

Ejemplo (dado):

¿Cuál es la probabilidad de que obtengamos 2 en el lanzamiento de un dado?

$$P_2 = 1/6 = 0.166666.$$

3.4 Frecuencias relativas

¿Qué sucede si tenemos experimentos aleatorios para que sus posibles resultados **no** son igualmente probables?

¿Cómo podemos entonces definir las probabilidades de los resultados?

Ejemplo (experimento aleatorio)

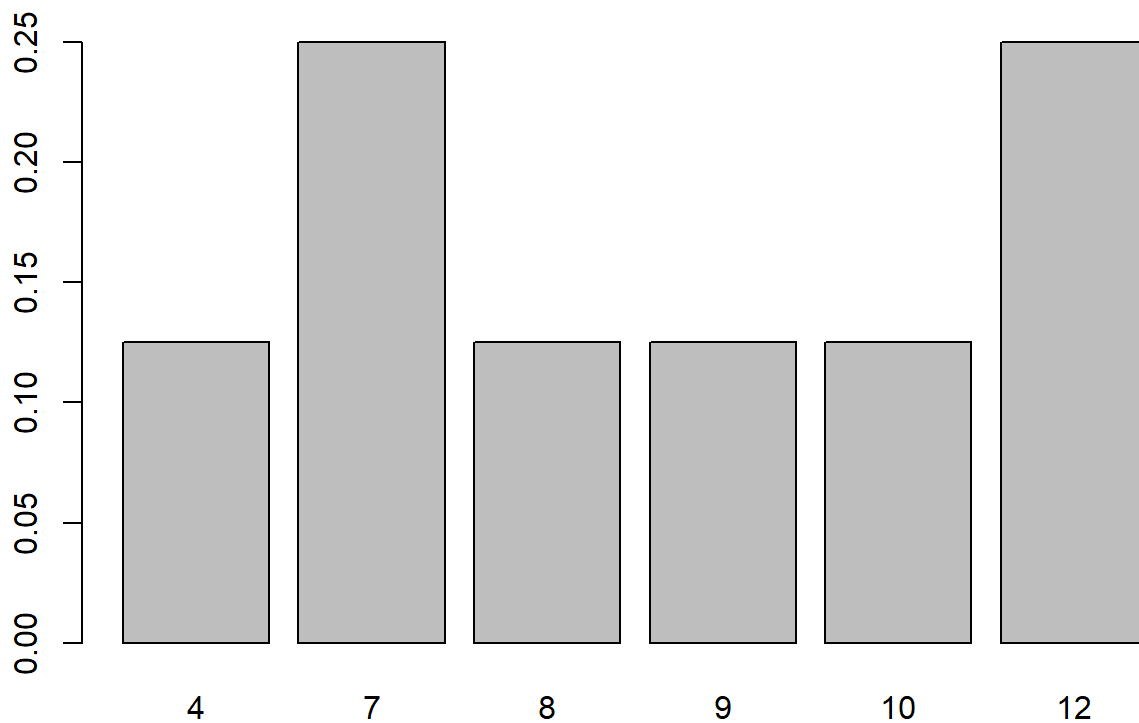
Imaginemos que repetimos un experimento aleatorio 8 veces y obtenemos las siguientes observaciones

8 4 12 7 10 7 9 12

- ¿Qué tan seguro estamos de obtener el resultado 12 en la siguiente observación?

La tabla de frecuencias es

##	outcome	n_i	f_i
## 1	4	1	0.125
## 2	7	2	0.250
## 3	8	1	0.125
## 4	9	1	0.125
## 5	10	1	0.125
## 6	12	2	0.250



La **frecuencia relativa** f_i parece una medida de probabilidad razonable porque

- es un número entre 0 y 1.
- mide la proporción del total de observaciones que observamos de un resultado particular.

Como $f_{12} = 0.25$ entonces estaríamos un cuarto seguros, una de cada 4 observaciones, de obtener 12.

Pregunta: ¿Qué tan bueno es f_i como medida de certeza del resultado i ?

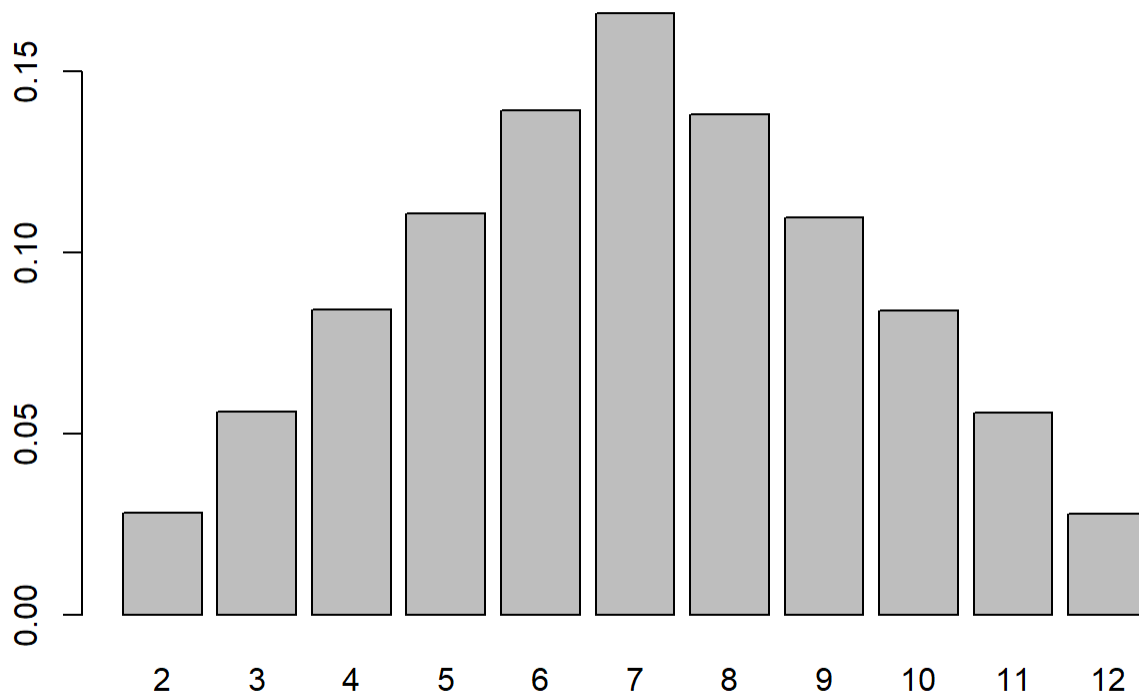
Ejemplo (experimento aleatorio con mas repeticiones)

Digamos que repetimos el experimento 100000 veces más:

La tabla de frecuencias es ahora

##	outcome	ni	fi
## 1	2	2807	0.02807
## 2	3	5607	0.05607
## 3	4	8435	0.08435
## 4	5	11070	0.11070
## 5	6	13940	0.13940
## 6	7	16613	0.16613
## 7	8	13806	0.13806
## 8	9	10962	0.10962
## 9	10	8402	0.08402
## 10	11	5581	0.05581
## 11	12	2777	0.02777

y el gráfico de barras es



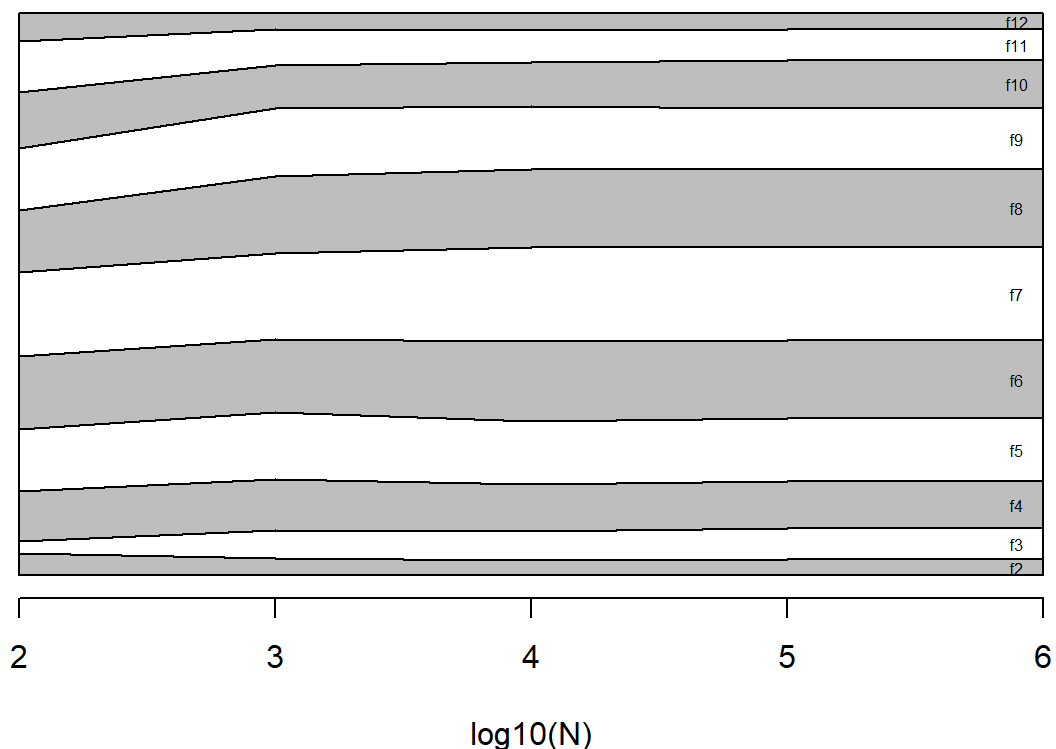
Aparecieron nuevos resultados y f_{12} ahora es solo 0.027, ahora estamos 3% seguros de obtener 12 en el próximo experimento. Las probabilidades medidas por f_i cambian con N .

3.5 Frecuencias relativas en el infinito

Una observación crucial es que si medimos las probabilidades de f_i en valores crecientes de N ¡convergen!

En este gráfico cada sección vertical da la frecuencia relativa de cada observación. Vemos que después de $N = 1000$ ($\log_{10}(N) = 3$) las proporciones apenas varían con mas N .

Proportion of relative frequencies at a given N



Encontramos que cada una de las frecuencias relativas f_i converge a un valor constante

$$\lim_{N \rightarrow \infty} f_i = P_i$$

3.6 Probabilidad frecuentista

Llamamos **Probabilidad** P_i al límite cuando $N \rightarrow \infty$ de la **frecuencia relativa** de observar el resultado i en un experimento aleatorio.

Defendida por Venn (1876), la definición frecuentista de probabilidad se deriva de datos/experiencia (empírica).

- No observamos P_i , observamos f_i
- **Estimamos** P_i con f_i (normalmente cuando N es grande), escribimos:

$$\hat{P}_i = f_i$$

Similar a la relación entre **observación** y **resultado**, tenemos la relación entre **frecuencia relativa** y **probabilidad** como un valor concreto de una cantidad abstracta.

3.7 Probabilidades clásicas y frecuentistas

Tenemos situaciones en las que se puede usar la probabilidad clásica para encontrar el límite de frecuencias relativas.

- Si los resultados son **igualmente probables**, la probabilidad clásica nos da el límite:

$$P_i = \lim_{N \rightarrow \infty} \frac{n_i}{N} = \frac{1}{M}$$

- Si los resultados en los que estamos interesados pueden derivarse de otros resultados **igualmente probables**; Veremos más sobre esto cuando estudiemos los modelos de probabilidad.

Ejemplo (suma de dos dados)

Nuestro ejemplo anterior se basa en la suma de dos dados. Si bien realizamos el experimento muchas veces, anotamos los resultados y calculamos las frecuencias relativas, podemos conocer el valor exacto de probabilidad.

Esta probabilidad se deduce del hecho de que el resultado de cada dado es igualmente probable. A partir de esta suposición, podemos encontrar que (Ejercicio 1)

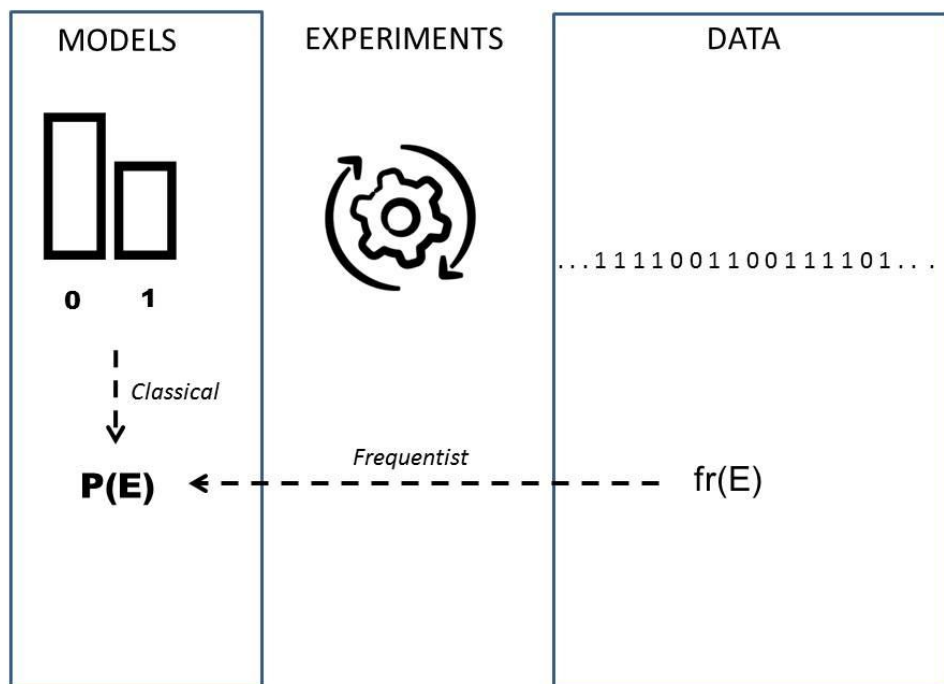
$$P_i = \frac{i - 1}{36}$$

para $\{2, 3, 4, 5, 6, 7\}$ y

$$P_i = \frac{13 - i}{36}$$

para $\{8, 9, 10, 11, 12\}$

La motivación de la definición frecuentista es **empírica** (datos) mientras que la de la definición clásica es **racional** (modelos). A menudo combinamos ambos enfoques para conocer las probabilidades de nuestro experimento aleatorio.



3.8 Definición de probabilidad

Una probabilidad es un número que se asigna a cada resultado posible de un experimento aleatorio y satisface las siguientes propiedades:

1. cuando los resultados E_1 y E_2 son mutuamente excluyentes; es decir, solo uno de ellos puede ocurrir, entonces la probabilidad de observar $E_1 \cup E_2$, escrito como $E_1 \cup E_2$, es su suma:

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

2. cuando S es el conjunto de todos los resultados posibles, entonces su probabilidad es 1 (al menos se observa algo): $P(S) = 1$
3. La probabilidad de cualquier resultado está entre 0 y 1 $P(E) \in [0, 1]$

Propuesto por Kolmogorov's hace menos de 100 años (1933)

3.9 Tabla de probabilidades

Las propiedades de Kolmogorov son las reglas básicas para construir una tabla de probabilidad, de manera similar a la tabla de frecuencia relativa.

Ejemplo (Dado)

La tabla de probabilidad para el lanzamiento de un dado

resultado	Probabilidad
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
$P(1 \cup 2 \cup \dots \cup 6)$	1

1. Donde $1 \cup 2$ es, por ejemplo, el **evento** de lanzar un 1 o un 2. Entonces
 $P(1 \cup 2) = P(1) + P(2) = 2/6$
2. Como $S = \{1, 2, 3, 4, 5, 6\}$ se compone de resultados mutuamente excluyentes, entonces

$$P(S) = P(1 \cup 2 \cup \dots \cup 6) = P(1) + P(2) + \dots + P(n) = 1$$

3. Las probabilidades de todos los resultados están entre 0 y 1.

3.10 Espacio muestral

El conjunto de todos los resultados posibles de un experimento aleatorio se denomina **espacio muestral** y se denota como S .

El espacio muestral puede estar formado por resultados categóricos o numéricos.

Por ejemplo:

- temperatura humana: $S = (36, 42)$ grados Celsius.

- niveles de azúcar en humanos:
- el tamaño de un tornillo de una línea de producción:
- número de correos electrónicos recibidos en una hora:
- el lanzamiento de un dado:

3.11 Eventos

Un **evento** es un **subconjunto** del espacio muestral. Es una **colección** de resultados.

Ejemplos de eventos:

- El evento de una temperatura saludable: grados Celsius
- El evento de producir un tornillo con un tamaño:
- El evento de recibir más de 4 emails en una hora:
- El evento de obtener un número menor o igual a 3 en la tirada de a dice:

Un evento se refiere a un posible conjunto de **resultados**.

3.12 Álgebra de eventos

Para dos eventos A y B , podemos construir los siguientes eventos derivados utilizando las operaciones básicas de conjuntos:

- Complemento A^c : el evento de **no** A
- Unión $A \cup B$: el evento de **o** A o B
- Intersección $A \cap B$: el evento de **y** A y B

Ejemplo (dado)

Lancemos un dado y veamos los eventos (conjunto de resultados):

- un número menor o igual a tres
- un número par

Veamos como podemos construir nuevos eventos con las operaciones de conjuntos:

- un número no menor de tres:
- un número menor o igual a tres **o** par:
- un número menor o igual a tres **y** par

3.13 Resultados mutuamente excluyentes

Los resultados como tirar 1 y 2 en un dado son eventos que no pueden ocurrir al mismo tiempo. Decimos que son **mutuamente excluyentes**.

En general, dos eventos denotados como A y B son mutuamente excluyentes cuando

Ejemplos:

- El resultado de tener una gravedad de misofonía de 1 y una gravedad de 2 .
- Los resultados de obtener 1 y 2 al sumar el lanzamiento de dos dados.

De acuerdo con las propiedades de Kolmogorov, solo los resultados **mutuamente excluyentes** se pueden organizar en **tablas de probabilidad**, como en las tablas de frecuencias relativas.

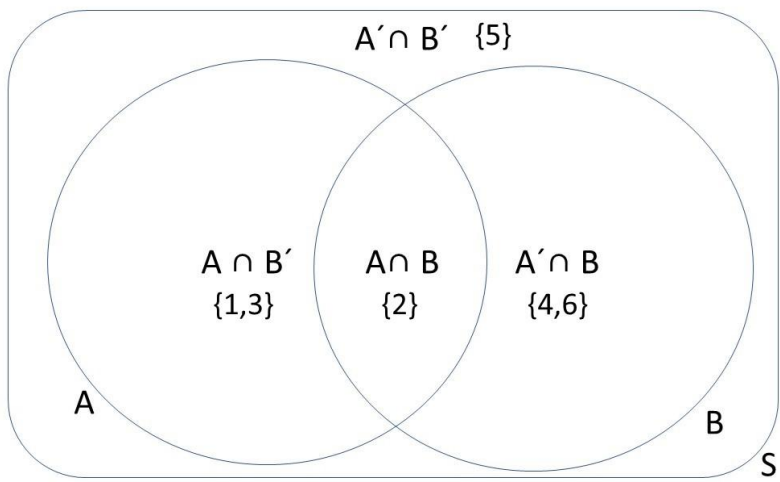
3.14 Probabilidades conjuntas

La **probabilidad conjunta** de A y B es la probabilidad de A y B . Eso es

o $P(A \cap B)$.

Para escribir probabilidades conjuntas de eventos no mutuamente excluyentes (A y B) en una tabla de probabilidad, notamos que siempre podemos descomponer el espacio muestral en conjuntos **mutuamente excluyentes** que involucran las intersecciones:

Consideremos el diagrama de Ven para el ejemplo donde A es el evento que corresponde a sacar número menor o igual que 3 y B corresponde a un número par:



Las **marginales** de A y B son la probabilidad de A y la probabilidad de B , respectivamente:

- $P(A) = P(A \cap B') + P(A \cap B) = 2/6 + 1/6 = 3/6$
- $P(B) = P(A' \cap B) + P(A \cap B) = 2/6 + 1/6 = 3/6$

Podemos ahora escribir la **tabla de probabilidad** para las probabilidades conjuntas

Resultado	Probabilidad
$(A \cap B)$	$P(A \cap B) = 1/6$
$(A \cap B')$	$P(A \cap B') = 2/6$
$(A' \cap B)$	$P(A' \cap B) = 2/6$
$(A' \cap B')$	$P(A' \cap B') = 1/6$
suma	1

Cada resultado tiene *dos* valores (uno para la característica del tipo A y otro para el tipo B)

3.15 Tabla de contingencia

La tabla de probabilidad conjunta también se puede escribir en una **tabla de contingencia**

	B	B'	suma
A	$P(A \cap B)$	$P(A \cap B')$	$P(A)$
A'	$P(A' \cap B)$	$P(A' \cap B')$	$P(A')$
suma	$P(B)$	$P(B')$	1

Donde las marginales son las sumas en las márgenes de la tabla, por ejemplo:

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

En nuestro ejemplo, la tabla de contingencia es

	B	B'	suma
A	1/6	2/6	3/6
A'	2/6	1/6	3/6
suma	3/6	3/6	1

3.16 La regla de la suma:

La regla de la suma nos permite calcular la probabilidad de A o B , $P(A \cup B)$, en términos de la probabilidad de A y B , $P(A \cap B)$. Podemos hacer esto de tres maneras equivalentes:

1. Usando las marginales y la probabilidad conjunta

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

2. Usando solo probabilidades conjuntas

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

3. Usando el complemento de la probabilidad conjunta

$$P(A \cap B) = 1 - P(A' \cap B')$$

Ejemplo (dado)

Tomemos los eventos $A : \{1, 2, 3\}$ y $B : \{2, 4, 6\}$ un número par en el lanzamiento de un dado.

Por lo tanto:

$$1. P(A \cup B) = P(A) + P(B) - P(A \cap B) = 3/6 + 3/6 - 1/6 = 5/6$$

$$2. P(A \cup B) = P(A \cap B) + P(A \cap B') + P(A' \cap B) = 1/6 + 2/6 + 2/6 = 5/6$$

$$3. P(A \cup B) = 1 - P(A' \cap B') = 1 - 1/6 = 5/6$$

En la tabla de contingencia $P(A \cup B)$ corresponde a las casillas en negrita (o sea todas menos 1/6 de abajo a la derecha)

	B	B'
A	1/6	2/6
A'	2/6	1/6

3.17 Preguntas

Recopilamos la edad y categoría de 100 deportistas en una competición

	<i>junior</i>	<i>senior</i>
<i>1er</i>	14	12
<i>2do</i>	21	18
<i>3er</i>	22	13

1) ¿Cuál es la probabilidad estimada de que un deportista sea de 2ª categoría y senior?

a: 18/100; b: 18/43; c: 18; d: 18/39

2) ¿Cuál es la probabilidad estimada de que el atleta no esté en la tercera categoría y sea senior?

a: 35/100; b: 30/100; c: 22/100; d: 13/100

3) ¿Cuál es la probabilidad marginal de la tercera categoría?

a: 13/100; b: 35/100; c: 22/100; d: 13/22

4) ¿Cuál es la probabilidad marginal de ser senior?

a: 13/100; b: 43/100; c: 43/57; d: 57/100

5) ¿Cuál es la probabilidad de ser senior o de tercera categoría?

a: 65/100; **b:** 86/100; **c:** 78/100; **d:** 13/100

3.18 Ejercicios

3.18.0.1 Probabilidad clásica: Ejercicio 1

- Escribe la tabla de probabilidad conjunta para los resultados de lanzar dos dados; en las filas escribe los resultados del primer dado y en las columnas los resultados del segundo dado.
- ¿Cuál es la probabilidad de sacar (3, 4)? (R:1/36)
- ¿Cuál es la probabilidad de tirar 3 y 4 con cualquiera de los dos dados? (R:2/36)
- ¿Cuál es la probabilidad de tirar 3 en el primer dado y 4 en el segundo? (A:11/36)
- ¿Cuál es la probabilidad de tirar 3 o 4 con cualquier dado? (R:20/36)
- Escribe la tabla de probabilidad del resultado de lanzar dos dados y sumar sus números. Supon que el resultado de cada dado es igualmente probable. Confirma que es:

$$P_i = \frac{i-1}{36}$$

para {2, 3, 4, 5, 6, 7} y

$$P_i = \frac{13-i}{36}$$

para {8, 9, 10, 11, 12}

3.18.0.2 Probabilidad frecuentista: Ejercicio 2

El resultado de un experimento aleatorio es medir la gravedad de la misofonía y el estado de depresión de un paciente.

- Gravedad de la misofonía: $x \in \{0, 1, 2, 3, 4\}$
- Depresión: $y \in \{0, 1\}$ (no:0, si:1)

Estos son los primeros 6 pacientes:

##	Misofonia.dic	depression.dic
## 1	4	1
## 2	2	0
## 3	0	0
## 4	3	0
## 5	0	0
## 6	0	0

El estudio sobre un total de 123 pacientes mostró las frecuencias $n_{x,y}$ dadas en la tabla de contingencia:

##		Depression:0	Depression:1
##	Misophonia:4	0	9
##	Misophonia:3	25	6
##	Misophonia:2	34	3
##	Misophonia:1	5	0
##	Misophonia:0	36	5

Supongamos que N es grande y que las frecuencias **estiman** las probabilidades $f_{x,y} = \hat{P}(X, Y)$, de tal forma que la siguiente es nuestra tabla de contingencia para las probabilidades conjuntas

##		Depression:0	Depression:1
##	Misophonia:4	0.00000000	0.07317073
##	Misophonia:3	0.20325203	0.04878049
##	Misophonia:2	0.27642276	0.02439024
##	Misophonia:1	0.04065041	0.00000000
##	Misophonia:0	0.29268293	0.04065041

- ¿Cuál es la probabilidad marginal de misofonía de gravedad 3? (R/0.3)
- ¿Cuál es la probabilidad de no ser misofónico **y** no estar deprimido? (R/0.293)
- ¿Cuál es la probabilidad de ser misofónico **o** deprimido? (R/0.293)
- ¿Cuál es la probabilidad de ser misofónico **y** deprimido? (R/0.707)
- Describir en lenguaje escrito los resultados con probabilidad 0.

3.18.0.3 Ejercicio 3

Hemos realizado un experimento aleatorio 10 veces, que consiste en anotar el sexo y el estado vital de pacientes con algún tipo de cáncer después de 10 años del diagnóstico. Obtuvimos los siguientes resultados

##	A	B
## 1	male	dead
## 2	male	dead
## 3	male	dead
## 4	female	alive
## 5	male	dead
## 6	female	alive
## 7	female	dead
## 8	female	alive
## 9	male	alive
## 10	male	alive

- Crea la tabla de contingencia para el número ($n_{i,j}$) de observaciones de cada resultado (A, B)
- Crea la tabla de contingencia para la frecuencia relativa ($f_{i,j}$) de los resultados
- ¿Cuál es la frecuencia marginal de ser hombre? (R/0.6)
- ¿Cuál es la frecuencia marginal de estar vivo? (R/0.5)
- ¿Cuál es la frecuencia de estar vivo o mujer? (R/0.6)

3.18.0.4 Teoría: Ejercicio 4

- De la segunda forma de la regla de la suma, obtener la primera y la tercera forma.
- ¿Cuál es la regla de la suma de la tercera forma para la probabilidad de tres eventos $P(A \cup B \cup C)$?