MaMaEuSch

Management **Ma**thematics for
European **Sch**ools
http://www.mathematik.uni-kl.de/~ mamaeusch

Introducción al Cálculo de Probabilidades a través de casos reales

Paula Lagares Barreiro[*]
Federico Perea Rojas-Marcos[*]
Justo Puerto Albandoz[*]

MaMaEuSch[**]
Management Mathematics for European Schools
94342 - CP - 1 - 2001 - 1 - DE - COMENIUS - C21

[*]Universidad de Sevilla

[**]Este proyecto ha sido llevado a cabo con ayuda parical de la Comunidad Europea en el marco del programa Sócrates. El contenido del proyecto no reflejy necesariamente la posición de la Comunidad Europea, ni implica ninguna responsabilidad por su parte.

Índice general

Capítulo 1

Juegos y Azar

Vamos a jugar al mus. Se reparten las cartas y llega la hora de decidir cuánto apostamos. Hemos de tener en cuenta que no jugamos solos, sino que competimos con otros participantes. Nerviosos vamos observando nuestras cartas una a una, según las vayan repartiendo. ¿Qué cartas nos tocarán? ¿Mejores que las de nuestros rivales?

Antes de nada, vamos a plantear los objetivos que se cubrirán en este manual y a dar unas reglas para jugar al mus.

1.1. Objetivos

- Comprender el concepto de experimento aleatorio y distinguirlo de uno determinista.
- Identificar los sucesos aleatorios tras un experimento y diferenciar un suceso simple de un suceso compuesto.
- Encontrar algunos sucesos especiales: el suceso seguro y el suceso imposible.
- Operar con sucesos aleatorios e interpretar los sucesos resultantes tras efectuar uniones, intersecciones y diferencias.
- Asignar probabilidades a los sucesos aleatorios sencillos de dos formas diferentes: mediante la regla de Laplace y a través de las frecuencias relativas.
- Entender el concepto de probabilidad condicionada y su utilidad.
- Comprender la independencia de sucesos y su uso para el cálculo de probabilidades.
- Manejar el teorema de la Probabilidad Total y la Regla de Bayes, sus diferencias y su aplicabilidad en el cálculo de probabilidades.

1.2. El juego del mus

Vamos a jugar a un juego sencillo de cartas. En él, participarán dos equipos, formados por dos jugadores cada uno, que competirán entre sí. La pareja ganadora, recibirá el placer de ganar la partida y haber disfrutado de un buen rato con sus amigos, no se jugará con dinero. Se juega con una baraja de 40 cartas, distribuidas de la siguiente forma:

- 8 ases o unos,
- 4 cuatros,
- 4 cincos,
- 4 seises,
- 4 sietes,
- 4 sotas o dieces,
- 4 caballos o onces,
- 8 reyes o doces.

Se repartirán 4 cartas a cada jugador, elegidas aleatoriamente de la baraja, y tras ese reparto se pueden obtener las siguientes jugadas:

- Tener dos cartas iguales y las otras dos desiguales entre sí y con respecto a las dos primeras, eso es tener una *pareja*. Por ejemplo, la jugada (cuatro, rey, diez, cuatro) es una pareja de cuatros.
- Tener tres cartas iguales y la tercera desigual es una jugada que llamaremos *media*. Si obtenemos unas cartas así: (rey, as, rey, rey) es que tienes una media de reyes.
- Tener dos parejas entre las cuatro cartas, iguales o diferentes entre sí, es un *dúplex*. Por ejemplo, una jugada que sea (cuatro diez diez cuatro) es un dúplex de dieces-cuatros, mientras que si tienes (as, as, as, as), eso es un dúplex de ases-ases. En ambos casos, se considerarán dúplex.

En este juego, el dúplex tiene más valor que la media, y ésta más que la pareja. En caso de haber dos dúplex, dos medias o dos parejas, ganará aquella que tenga las cartas más altas.

Las cartas ordenadas de más baja a más alta son las que puedes encontrar al principio de esta sección, es decir: as, cuatro, cinco, seis, siete, sota, caballo y rey.

Por ejemplo, un dúplex de reyes y ases ganará a un dúplex de sotas y caballos. Igualmente, una pareja de caballos ganará a una pareja de sotas.

En caso de empate, ganará aquel jugador que tenga cartas más altas acompañando a la pareja o a la media. En caso de que las cuatro cartas sean iguales, ganará aquel jugador que vaya de mano, es decir, aquel jugador que haya recibido las cartas en primer lugar.

Cuatro amigos pasan mucho tiempo jugando a este juego, y resulta que se han dado cuenta de que la pareja de reyes, la media de reyes o ases y un dúplex cualquiera salen aproximadamente el mismo número de veces. Están discutiendo a ver cuál de esas tres jugadas sale con mayor frecuencia. ¿Podrías ayudarles a averiguar cuál es?

Creemos que al terminar este manual, podremos responder a esa pregunta, y lo haremos. Ahora comenzaremos a introducir los diferentes conceptos que nos serán de utilidad para contestar a la pregunta del ejemplo y a muchas otras.

1.3. Experimentos Aleatorios

Ejemplo 1.3.1 *Imagina la siguiente situación: se reparten las cartas a los participantes. ¿Sabremos antes de verlas que cartas vamos a tener?*

Como ves, no tenemos ninguna certeza de las cartas que vamos a recibir hasta que no las veamos. Podemos obtener tres reyes y un as ó cuatro sotas. Ambas posibilidades y muchas más pueden darse en nuestro reparto de cartas. Al hecho de no tener seguridad sobre el resultado tras el reparto le llamamos azar.

En nuestro caso, tenemos un experimento: sacar cuatro cartas de la baraja. Como tras la realización del experimento se pueden dar varios resultados, decimos que se trata de un experimento aleatorio.

Si por el contrario supiéramos el resultado del experimento de antemano, diríamos que se trata de un experimento determinista. Por ejemplo, si dejamos caer una piedra desde nuestra mano, sabemos que ésta caerá al suelo. Aquí no hay posibilidad de resultados diferentes, sólo uno: la piedra caerá al suelo.

Así pues, podemos decir que la gran diferencia entre un experimento aleatorio y uno determinista, es que el primero puede dar lugar a diferentes resultados, y no podemos predecir cuál de ellos será el que ocurra, mientras que en el segundo sólo tenemos una posibilidad, y esa es la que ocurrirá.

Ejercicio 1.3.1 *Plantea dos experimentos aleatorios y dos deterministas.*

Definición 1.3.1 (Experimento Aleatorio) *Decimos que un experimento es aleatorio si no podemos predecir su resultado.*

1.4. Sucesos Aleatorios y Espacio Muestral

Una vez comprendida la idea de experimento aleatorio, con toda naturalidad nos surgirán preguntas sobre qué pasará.

Se puede ver que tras el reparto de cartas, podemos obtener un gran número de combinaciones de cartas diferentes. Nos pueden salir manos como (as, rey, as, sota) o (cuatro, rey, cinco, siete). A cada uno de esos posibles repartos, los llamaremos sucesos aleatorios.

Es decir, en el experimento planteado anteriormente, las cartas (as, siete, sota, seis) forman un suceso aleatorio.

Al conjunto de todos los sucesos de un experimento aleatorio, lo llamaremos Espacio Muestral, y lo denotaremos por E. En nuestro experimento, el espacio muestral es el conjunto formado por todos los posibles repartos.

Ejercicio 1.4.1 *Imagina el siguiente experimento aleatorio: sacar de una baraja de mus una carta al azar.*

Describe el espacio muestral de ese experimento mediante la enumeración de todos sus sucesos aleatorios.

Definición 1.4.1 (Espacio Muestral y suceso aleatorio) *Al conjunto de todos los resultados que pueden obtenerse al realizar un experimento aleatorio, se le llama espacio muestral, y lo denotaremos por E.*
Cualquier subconjunto del espacio muestral es un suceso.

1.4.1. Sucesos elementales y compuestos

Dentro de los sucesos aleatorios, vamos a distinguir, de momento, dos tipos: los sucesos aleatorios elementales y los sucesos aleatorios compuestos.

Imagina que observas la primera carta de tu reparto, y es un as. Plantéate ahora estos dos sucesos aleatorios:

- es un as,
- es menor que siete.

Puedes notar que entre esos dos sucesos aleatorios hay una gran diferencia. En el primero especificamos que carta es: un as. Mientras que en el segundo dejamos abierta la posibilidad a varios tipos de cartas: un as, un cuatro, un cinco ó un seis.

Es decir, que el segundo suceso aleatorio descrito (la carta es menor que siete) engloba a su vez varios sucesos aleatorios más.

Diremos entonces que el primer suceso aleatorio descrito es un suceso elemental, mientras que el segundo es un suceso compuesto.

Definición 1.4.2 (sucesos elementales) *Decimos que un suceso es elemental cuando consta de un solo elemento del espacio muestral. En caso contrario, diremos que es un suceso compuesto.*

Ejemplo 1.4.1 *El experimento aleatorio es el siguiente: sacar de una baraja española completa una carta al azar. Tomaremos como elementos del espacio muestral: "as de oros", "dos de oros", "tres de oros",...,"rey de bastos".*
El suceso "sacar el rey de oros" es elemental. Sin embargo, si elegimos el suceso "sacar un rey", se trata de un suceso compuesto, ya que está formado por varios sucesos elementales: "sacar el rey de oros", "sacar el rey de copas", "sacar el rey de espadas" y "sacar el rey de bastos".

1.4.2. Sucesos compatibles e incompatibles

Volvamos a la situación del reparto de cartas. Nos van a dar cuatro. Planteemos estos dos sucesos aleatorios:

- suceso A = "De entre las cuatro cartas, dos son reyes",
- suceso B = "De entre las cuatro cartas, dos son ases".

¿Se pueden dar los sucesos A y B a la vez? Es decir, ¿se pueden obtener 2 reyes y 2 ases en la misma mano? Claro que sí. Lo que tendremos es un bonito dúplex de reyes y ases.

Como el suceso A y el B se pueden dar a la vez, decimos que son sucesos compatibles.

Sin embargo, no siempre es así, hay parejas de sucesos que no se pueden dar simultáneamente. Como ejemplo, imagina estos dos nuevos sucesos:

- suceso $C =$ "De entre las cuatro cartas, exactamente tres son reyes",
- suceso $D =$ "De entre las cuatro cartas, exactamente dos son ases".

Y nos volvemos a preguntar: ¿Se pueden dar los sucesos C y D a la vez? Es decir, ¿se pueden obtener 3 reyes y 2 ases en la misma mano? En este caso no, pues si tuviéramos 3 reyes y 2 ases en la misma jugada, eso significaría que tenemos 5 cartas, cuando en este juego sólo se reparten cuatro. Es decir, es imposible.

Así, como los sucesos C y D no se pueden dar en el mismo experimento, decimos que son sucesos incompatibles.

Ejercicio 1.4.2 *Encontrar una pareja de sucesos compatibles y otra de sucesos incompatibles, diferentes a las dadas antes, en el mismo experimento.*

Definición 1.4.3 *Dados dos sucesos de un experimento aleatorio, diremos que son compatibles si se pueden dar los dos al mismo tiempo, y diremos que son incompatibles si no se pueden dar los dos a la vez.*

1.4.3. El suceso seguro

Imagina que cogemos la baraja de cartas y la dividimos en dos grupos, por un lado los ases (8 cartas en total) y por el otro las demás cartas (32 cartas). Cogemos el grupo de los 8 ases, y de él elegimos una carta al azar: ¿podemos asegurar algo? Sí, que la carta elegida será un as. Así de simple es el suceso seguro, aquel suceso aleatorio que se da siempre.

Ejercicio 1.4.3 *En el reparto de cartas del mus, plantea un suceso seguro en tu reparto, o sea, en tus cuatro cartas.*

Definición 1.4.4 *El suceso seguro es aquel suceso aleatorio de un experimento que se da siempre. Se puede ver que el suceso seguro es el espacio muestral E.*

1.4.4. El suceso imposible

Sin embargo, en el mismo experimento podemos asegurar que la carta no será un seis, ya que no hay ningún seis entre las cartas del montón elegido. Decimos entonces que el suceso "la carta es un seis" es un suceso imposible.

Ejercicio 1.4.4 *En el reparto de cartas del mus, plantea un suceso imposible en tu reparto, o sea, en tus cuatro cartas.*

Definición 1.4.5 *Decimos que un suceso es imposible cuando no puede darse en el experimento. Se denota por $\emptyset$ a cualquier suceso imposible.*

1.4.5. Sucesos complementarios

Supón ahora que la baraja ha sido dividida en dos grupos: en el primero sólo están los reyes y los ases (16 cartas) y en el segundo están las restantes 24 cartas. Elegimos una carta al azar de entre las del primer grupo. Observa estos dos sucesos:

- $A =$ "La carta elegida es un as",
- $B =$ "La carta elegida es un rey".

Como característica especial de estos sucesos, podemos decir que si no se da el suceso A entonces se da el B, y si no se da el B se da el A. Es decir, siempre se da alguno de los dos sucesos. También puedes observar que son sucesos incompatibles, es decir, no se pueden dar los dos a la vez.

Estas dos características son las que han de cumplir un par de sucesos aleatorios para poder decir que son sucesos complementarios o contrarios.

Ejercicio 1.4.5 *Obtener una pareja de sucesos complementarios en el experimento de sacar una carta al azar de la baraja del mus completa.*

Definición 1.4.6 *Dos sucesos son complementarios si son incompatibles (no se pueden dar los dos a la vez) y, si se da uno, siempre se cumple alguno de los dos. Si denotamos por A a un suceso, su complementario será denotado $\overline{A}$ ó A^c.*

1.5. Operaciones con Sucesos Aleatorios

De igual forma que operamos con los números (los sumamos, restamos, multiplicamos, etc.) podemos hacer operaciones con los Sucesos Aleatorios. Esta vez las operaciones son diferentes a las usadas con los números, así hablaremos de uniones e intersecciones de sucesos.

1.5.1. Unión de sucesos

Imagina que volvemos a repartir las cartas en el juego del mus, es decir, cuatro cartas a cada jugador. Por tanto, tenemos el experimento de sacar cuatro cartas y observarlas. Planteemos ahora dos sucesos aleatorios resultantes de ese experimento:

- $A =$ "Sacar dos reyes",
- $B =$ "Sacar un as".

Imagina que después del reparto, nuestras cartas son: (as sota siete siete). A la vista de este reparto, ¿se ha cumplido el suceso A? No, pues no tenemos dos reyes, ni tan siquiera tenemos uno. Y el suceso B, ¿se ha cumplido? Este sí, dado que hemos sacado un as. Así, diremos que se ha cumplido el suceso A unión B, y lo representaremos $A \cup B$.

Así pues, si tenemos dos sucesos aleatorios y tras el experimento se cumple uno de los dos o los dos, entonces diremos que se ha cumplido el suceso $A \cup B$.

Ejercicio 1.5.1 *Estamos otra vez con el experimento anterior. Repartimos cuatro cartas. Considera los dos sucesos:*

- A = *"Sacamos tres reyes",*
- B = *"Sacamos tres ases".*

A partir de esto describe dos repartos en los que se cumpla el suceso $A \cup B$*.*

Definición 1.5.1 (Unión) *Dados los sucesos* A *y* B*, se define el suceso* A *unión* B*, y lo denotaremos* $A \cup B$*, como el suceso consistente en que se cumpla al menos uno de los dos. Notar, que si se cumplen los dos a la vez, también se cumple* $A \cup B$*.*

1.5.2. Intersección

Planteemos ahora dos nuevos sucesos aleatorios tras la realización del experimento consistente en repartir cuatro cartas de la baraja del mus.

- A = "Sacar 2 ases",
- B = "Sacar 1 siete".

Imagina que tras el reparto de cartas, nos encontramos con esta combinación: (as, rey, siete, as). Tras observarlo, ¿se ha cumplido el suceso A? Sí, pues entre las cuatro cartas, tenemos dos ases. ¿Y el suceso B? También, pues la tercera carta es un siete. Como se han cumplido los sucesos A y B en el mismo reparto, decimos entonces que en ese reparto se ha dado el suceso A intersección B, y lo denotamos por $A \cap B$

Ejercicio 1.5.2 *Considera los siguiente sucesos aleatorios tras sacar las cuatro cartas de la baraja:*

- A = *"Sacar dos reyes",*
- B = *"Sacar dos ases".*

Plantea un reparto en el que se cumpla el suceso $A \cap B$ *y otro en el que no.*

Definición 1.5.2 (Intersección) *Dados los sucesos* A *y* B*, se define el suceso* A *intersección* B*, y lo denotaremos* $A \cap B$*, como el suceso consistente en que se cumplan los dos sucesos simultáneamente.*

Date cuenta, que si la intersección de dos sucesos es el suceso imposible ($\emptyset$), los sucesos son incompatibles, (recuerda la definición dada de *sucesos incompatibles* en el apartado anterior). Si la intersección no es el suceso imposible, entonces los sucesos son compatibles.

1.5.3. Diferencia de sucesos

Planteemos ahora dos nuevos sucesos tras el reparto de las cuatro cartas de la baraja del mus:

- $A =$ "Sacar tres sotas",
- $B =$ "Sacar un rey".

Después de repartir nos han salido las siguientes cartas: (sota, as, sota, sota).

¿Se ha cumplido el suceso A? Claramente sí, pues tenemos entre las cuatro cartas exactamente tres sotas. ¿Y el suceso B? En este caso no, pues no tenemos ningún rey entre nuestras cartas. Entonces diremos que se ha cumplido el suceso diferencia A menos B, y lo denotaremos por $A \setminus B$.

Así pues, siempre que se dé un suceso y otro no, diremos que se ha dado el suceso diferencia entre el primero y el segundo.

Ejercicio 1.5.3 *Estamos otra vez con el experimento anterior. Considera los dos sucesos:*

- $A =$ *"Sacamos dos ases",*
- $B =$ *"Sacamos dos sotas".*

A partir de esto, plantea un posible reparto en el que se cumpla el suceso $A \setminus B$ y otro en el que se cumpla el suceso $B \setminus A$.

Definición 1.5.3 (Diferencia) *Dados los sucesos A y B, se define el suceso diferencia $A \setminus B$, como el suceso consistente en que se cumpla el suceso A pero no el B.*

1.5.4. Propiedades de las operaciones con sucesos

Las propiedades más utilizadas son las que describiremos ahora. Tener en cuenta, que el suceso E es el suceso seguro, el suceso $\emptyset$ es el suceso imposible, los sucesos A, B y C son tres sucesos cualesquiera, subconjuntos del espacio muestral y que A^c es el suceso contrario o complementario del suceso A.

Unión:

$$A \cup B = B \cup A, \quad A \cup E = E, \quad A \cup \emptyset = A, \quad A \cup A^c = E.$$

Intersección:

$$A \cap B = B \cap A, \quad A \cap E = A, \quad A \cap \emptyset = \emptyset, \quad A \cap A^c = \emptyset.$$

Diferencia:

$$A \setminus B = A \cap B^c.$$

Leyes de De Morgan:

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c.$$

Otras propiedades:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \qquad A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Capítulo 2

Probabilidad

2.1. Introducción

En el capítulo anterior, habíamos dicho que al realizar un experimento aleatorio, no hay seguridad sobre el resultado que obtendremos: en otras palabras, hay incertidumbre. Pues bien, utilizaremos para 'medir' esa aleatoriedad o incertidumbre, un número que asociaremos a cada suceso, y llamaremos probabilidad.

Por ejemplo, si elegimos una carta al azar de entre las de la baraja del mus, no sabríamos predecir con seguridad qué carta saldrá. Aunque como sabemos que hay más ases que sietes, sería lógico pensar que tiene más posibilidades de salir un as que un siete. Por eso diremos que, tras el experimento, el suceso aleatorio "la carta es un as" tiene más posibilidades de darse que el otro suceso "la carta es un siete".

En este capítulo descubriremos varias técnicas para medir la frecuencia con la que se dan los sucesos aleatorios, es decir, introduciremos varios métodos para asignar probabilidades.

Ejercicio 2.1.1 *Describe dos sucesos en el reparto de cartas del mus que creas que tienen muy diferenciadas sus probabilidades o posibilidades de ocurrir y explica por qué.*

2.1.1. Definición de probabilidad a partir de frecuencias relativas: probabilidad empírica.

Decíamos en el apartado anterior, que en un experimento aleatorio, la probabilidad es un número que asignábamos a cada suceso, y con el que queremos indicar la frecuencia con la que se da dicho suceso.

Una sencilla manera de obtener la probabilidad de un suceso aleatorio es a través de la tabla de frecuencias relativas de ese experimento. A esa probabilidad la llamamos probabilidad empírica o concepto frecuentista de probabilidad, porque se obtiene una vez realizado el experimento. Así, si hemos realizado el experimento un número de veces n y observado sus resultados, y nos damos cuenta de que en k de esas ocasiones se ha verificado el suceso que queremos analizar, llamémosle

A, decimos que la probabilidad de que ocurra el suceso A, y lo denotamos $P(A)$, es $\frac{k}{n}$, es decir:

$$P(A) = \frac{k}{n}.$$

Ejemplo 2.1.1 *Supongamos que vamos sacando una a una 200 cartas de la baraja, y las vamos dejando otra vez en la baraja después de observarlas. Así, obtenemos la siguiente tabla de frecuencias relativas para este experimento:*

Carta	*Frecuencia Relativa*
as	$\frac{38}{200}$
cuatro	$\frac{17}{200}$
cinco	$\frac{21}{200}$
seis	$\frac{24}{200}$
siete	$\frac{21}{200}$
sota	$\frac{23}{200}$
caballo	$\frac{18}{200}$
rey	$\frac{38}{200}$

A partir de esta tabla, podríamos decir que si sacamos una carta al azar de la baraja del mus, obtendremos cada uno de los diferentes naipes con las siguientes probabilidades:

- *P(la carta es un as)=*$\frac{38}{200}$,
- *P(la carta es un cuatro)=*$\frac{17}{200}$,
- *P(la carta es un cinco)=*$\frac{21}{200}$,
- *P(la carta es un seis)=*$\frac{24}{200}$,
- *P(la carta es un siete)=*$\frac{21}{200}$,
- *P(la carta es una sota)=*$\frac{23}{200}$,
- *P(la carta es un caballo)=*$\frac{18}{200}$,
- *P(la carta es un rey)=*$\frac{38}{200}$.

Imagina ahora que llevamos a cabo 1000 extracciones, y nos salen las siguientes frecuencias relativas:

Carta	*Frecuencia Relativa*
as	$\frac{192}{1000}$
cuatro	$\frac{111}{1000}$
cinco	$\frac{109}{1000}$
seis	$\frac{85}{200}$
siete	$\frac{87}{1000}$
sota	$\frac{116}{1000}$
caballo	$\frac{91}{1000}$
rey	$\frac{209}{1000}$

A partir de esta tabla, podríamos decir que si sacamos una carta al azar de la baraja del mus, obtendremos cada uno de los diferentes naipes con las siguientes nuevas probabilidades:

- *P(la carta es un as)*=$\frac{192}{1000}$,
- *P(la carta es un cuatro)*=$\frac{111}{1000}$,
- *P(la carta es un cinco)*=$\frac{109}{1000}$,
- *P(la carta es un seis)*=$\frac{85}{1000}$,
- *P(la carta es un siete)*=$\frac{87}{1000}$,
- *P(la carta es una sota)*=$\frac{116}{1000}$,
- *P(la carta es un caballo)*=$\frac{91}{1000}$,
- *P(la carta es un rey)*=$\frac{209}{1000}$.

A la vista de estas probabilidades, y una vez comprendido que la probabilidad es un número que asignamos a los sucesos aleatorios para medir la frecuencia con la que se dan, podríamos decir que el suceso "la carta elegida es un rey" es más probable*, es decir, que tiene más posibilidades de darse que el suceso "la carta elegida es un siete", dado que P(la carta es un rey) > P(la carta es un siete).*

Con el mismo razonamiento, diremos que el suceso "la carta es una sota" es más probable que el suceso "la carta es un caballo", etc.

Nota 2.1.1 *Este método de asignar probabilidades se basa en la 'ley de frecuencias relativas', que viene a decir que la frecuencia relativa de un suceso, cuando el número de realizaciones del experimento se va haciendo grande, se aproxima a su verdadera probabilidad. Cuanto más grande sea el número de repeticiones del experimento, mayor será la fiabilidad de la probabilidad que obtengamos.*

Es decir, en el ejemplo anterior, si en lugar de haber sacado 200 cartas hubiéramos sacado sólo 100, la fiabilidad de las probabilidades obtenidas para cada suceso, sería menor que la que obtenemos con las 200 extracciones y al revés, si hubiéramos sacado 10000 cartas, las probabilidades habrían sido más fiables que con las 1000 cartas que hemos sacado en el experimento.

En cualquier caso, la tabla obtenida tras la realización de las 1000 extracciones es más fiable que la obtenida tras sacar 200 cartas.

Ejercicio 2.1.2 *Realiza el siguiente experimento: efectúa 20 repartos de mus, es decir, cuatro cartas de la baraja, y anota si en cada reparto se ha obtenido: una pareja, una media, un dúplex o ninguno de ellos.*

Construye la tabla de frecuencias relativas de ese experimento y después asigna probabilidades a esos sucesos. ¿Crees que son fiables esas probabilidades? ¿Por qué?

En la siguiente sección, vamos a asignar probabilidades a los sucesos de otra forma. No necesitaremos realizar el experimento para conocer dichas probabilidades por lo que en algunos casos, será un método bastante más sencillo que el explicado anteriormente.

2.1.2. Regla de Laplace: probabilidad teórica

Como has podido comprobar, resulta bastante tedioso asignar probabilidades a partir de las frecuencias relativas, pues es necesario realizar el experimento una gran cantidad de veces para conseguir una buena aproximación de la verdadera probabilidad de un suceso y, aún así, nunca estaremos seguros de conseguirla.

Por esa razón es necesario introducir un método alternativo para el cálculo de probabilidades que sea más manejable.

Imaginemos el ejemplo de antes: tenemos la baraja del mus y vamos a sacar una carta. Queremos conocer las diferentes probabilidades de todos los posibles sucesos.

Bien, es lógico pensar que la baraja está bien hecha y, por tanto, sacaremos una carta con igual probabilidad que otra. Es decir, no hay cartas más grandes que otras, ni cartas dobladas, etc., en otras palabras, la baraja no está trucada y por tanto podemos sacar cualquiera de las cuarenta cartas con las mismas posibilidades. Se dice en este caso que son sucesos equiprobables. Otros sucesos equiprobables pueden ser el número que sale tras el lanzamiento de un dado (con igual probabilidad sale uno, dos,..., seis) o el hecho de salir cara o cruz en el lanzamiento de una moneda (con igual probabilidad sale cara o cruz), siempre y cuando no estén trucados.

Volvamos a nuestro ejemplo con la baraja del mus. Tenemos cuarenta cartas, todas ellas iguales en peso, forma, etc. Entre ellas hay 8 ases, por tanto parece lógico pensar que de cuarenta extracciones de cartas, aproximadamente en ocho saldrá un as. Esto es algo teórico, pues puedes comprobar que no siempre ha de ser así en la práctica (si sacas cuarenta cartas, lo mismo te salen 3 ases que te salen 12). Pero el hecho de que haya 8 ases, nos da una idea de cuan probable es que salga esa carta.

Así pues, diremos que la probabilidad de que salga un as es $\frac{8}{40}$. Es una probabilidad teórica, repetimos, en la práctica no siempre se dará que de cuarenta extracciones, en ocho de ellas saquemos un as. En este experimento, como hay cuarenta cartas en total, diremos que en este experimento hay cuarenta *casos posibles* (podemos sacar cuarenta cartas diferentes) y ocho *casos favorables* porque el suceso que estamos analizando ("sacar un as") tiene ocho oportunidades de darse.

Ahora, una vez introducidos estos conceptos, podemos enunciar la regla de Laplace para el cálculo de probabilidades, que dice:

Definición 2.1.1 (regla de Laplace) *Si todos los sucesos elementales de un experimento son equiprobables, y tenemos un suceso cualquiera A de dicho experimento, entonces se tiene que*

$$P(A) = \frac{N^o \;\; total \;\; de \;\; casos \;\; favorables \;\; al \;\; suceso}{N^o \;\; total \;\; de \;\; casos \;\; posibles},$$

donde el "Nº total de casos favorables al suceso" son las posibilidades reales de obtener ese suceso.

Esta fue la primera definición formal de probabilidad que se dio en la historia, y lo hizo Pierre Simon de Laplace.

Ahora, estás en disposición de hacer correctamente el siguiente ejercicio:

Ejercicio 2.1.3 *Calcula las probabilidades, utilizando la regla de Laplace, de obtener cada una de las diferentes cartas en el experimento de sacar una carta al azar de la baraja del mus.*

2.2. Extracciones con y sin reemplazamiento. Diagramas de árbol.

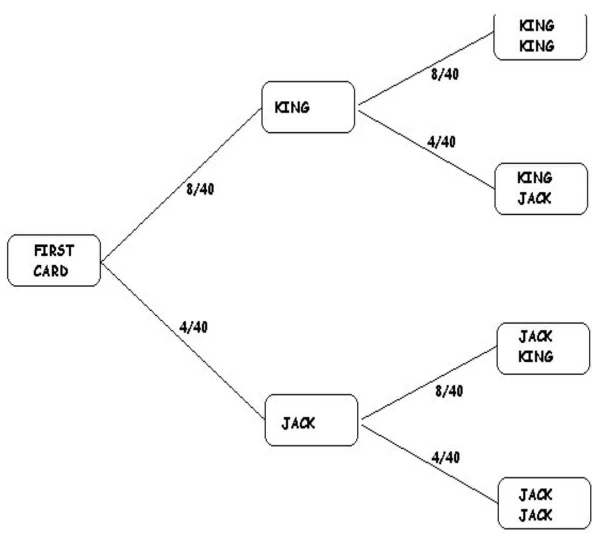
Vamos a plantear en esta sección unos nuevos experimentos algo más complejos, en lugar de sacar una sola carta, vamos a sacar varias. Tras el estudio de esta sección, podremos analizar en profundidad las diferentes jugadas del mus.

2.2.1. Extracciones con reemplazamiento

Comencemos por una situación sencilla, sacamos dos cartas de la baraja del mus, de una en una y devolviendo la carta a la baraja una vez observada. A este proceso lo llamaremos extracción con reemplazamiento.

Si llamamos A_1 al suceso "La primera carta es un rey" y A_2 al suceso "La segunda carta es una sota", nos podríamos preguntar cuál es la probabilidad de que se den esos dos sucesos a la vez, es decir, que la primera carta elegida sea un rey y la segunda una sota.

Para calcular la probabilidad de este suceso, dibujamos el siguiente diagrama, llamado diagrama de árbol:



La probabilidad de que la primera carta sea un rey y la segunda una sota, es el producto de las probabilidades del camino hasta llegar al resultado (la regla del producto), es decir, $\frac{8}{40} \cdot \frac{4}{40} = \frac{1}{50}$.

Sin embargo, si lo único que queremos es que las dos cartas sean un rey y una sota sin importar el orden en el que aparezcan, nos tenemos que plantear que salgan (rey, sota) por ese orden, o el orden inverso (sota, rey).

Si llamamos B_1 al suceso "La primera carta es una sota" y B_2 al suceso "La segunda carta es un rey", para que se de la combinación (sota, rey) se tiene que dar el suceso $B_1 \cap B_2$.

Por el mismo razonamiento de antes, y con un árbol similar al anterior, obtenemos que la probabilidad de obtener (sota, rey) es $\frac{4}{40} \cdot \frac{8}{40} = \frac{1}{50}$.

Por tanto, para obtener la probabilidad de obtener una sota y un rey sin importar el orden, sumamos las probabilidades anteriores "(rey, sota) + (sota,rey)" (regla de la suma), y obtenemos $\frac{1}{50} + \frac{1}{50} = \frac{1}{25}$.

Ejercicio 2.2.1 *Calcula en el experimento anterior la probabilidad de que las dos cartas sean ases, dibujando su árbol respectivo.*

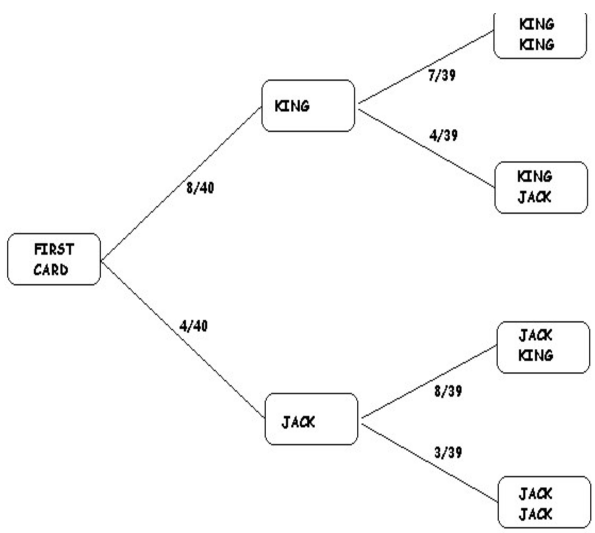
En el siguiente apartado veremos otro tipo de extracciones en las que no se devuelven las cartas una vez observadas.

2.2.2. Extracciones sin reemplazamiento

En los experimentos realizados antes, devolvíamos las cartas a la baraja una vez observadas pero, ¿qué pasaría si no las devolvemos? Bien, la cosa cambia pero el razonamiento es más o menos el mismo, lo único que varían son las segundas probabilidades, es decir, las probabilidades referentes a la segunda extracción de cartas. Es lógico, ya que si en la primera extracción tenemos cuarenta cartas, en la segunda sólo tendremos 39, pues la primera que cogimos no es devuelta.

Entonces, en el ejemplo de antes, si queremos que la primera carta sea un rey y la segunda sea una sota el árbol es el mismo, sólo cambian las probabilidades en la segunda extracción.

Vamos a calcular de nuevo las probabilidades de obtener un rey y una sota en ese orden, pero esta vez con unas extracciones diferentes, sin devolver las cartas a la baraja una vez observadas. Son las llamadas *extracciones sin reemplazamiento*. El árbol de este experimento es muy parecido al anterior, veámoslo.



En este caso tenemos que la probabilidad de obtener (rey, sota) es $\frac{8}{40} \cdot \frac{4}{39}$.

Notar que en este caso el segundo factor es $\frac{4}{39}$ porque al no devolver la primera carta que cogimos, nos quedarán sólo 39, entre las que hay cuatro sotas.

Igualmente calcularíamos la probabilidad de obtener la combinación (sota, rey) en dos extracciones sin reemplazamiento, y saldría $\frac{4}{40} \cdot \frac{8}{39}$.

Y de nuevo, si queremos calcular la probabilidad de obtener sota y rey sin importar el orden, en dos extracciones sin reemplazamiento, solo tenemos que aplicar la regla de la suma, por lo que nos queda que es $\frac{4}{40} \cdot \frac{8}{39} + \frac{4}{40} \cdot \frac{8}{39} = 2 \cdot \frac{4}{40} \cdot \frac{8}{39}$.

Ejercicio 2.2.2 *Haz lo mismo que en el ejercicio anterior pero suponiendo que hacemos extracciones sin reemplazamiento.*

2.3. Definición axiomática de probabilidad (1º BACH)

Ahora, en esta sección, vamos a introducir una definición más abstracta de la probabilidad. Lo haremos a partir de unos principios que aceptaremos como evidentes (a los que llamamos axiomas). Estos axiomas son:

1. Para cada suceso A, su probabilidad es un número entre 0 y 1, es decir,
$$0 \leq P(A) \leq 1.$$
2. $P(E) = 1$, donde E es el suceso seguro.
3. Si A y B son dos sucesos incompatibles, se tiene que
$$P(A \cup B) = P(A) + P(B).$$

A partir de estos axiomas, podemos deducir una gran cantidad de propiedades que cumple la probabilidad, las más importantes son:

1. Si denotamos por A^c el suceso complementario de A, se tiene que
$$P(A^c) = 1 - P(A).$$
Ejercicio 2.3.1 *Demostrar esta propiedad a partir de los axiomas anteriores.*
2. Si tenemos un conjunto de sucesos $A_1, A_2, ..., A_n$, que son incompatibles dos a dos ($A_i \cap A_j = \emptyset \; \forall i \neq j$), se tiene que
$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$
Como caso particular, podemos estudiar el caso en el que el conjunto de sucesos $A_1, A_2, ..., A_n$ cumple también que $\bigcup_{i=1}^{n} A_i = E$, donde E representa el suceso seguro. En este caso decimos que el conjunto de sucesos $A_1, A_2, ..., A_n$ es un *sistema completos de sucesos*, y se tiene que $\sum_{i=1}^{n} P(A_i) = 1$.

3. Si el espacio muestral se puede descomponer en n sucesos elementales, pongamos $E = \{x_1, ..., x_n\}$, entonces se tiene que

$$P(x_1) + P(x_2) + ... + P(x_n) = \sum_{i=1}^{n} P(x_i) = 1.$$

Como caso particular, si la probabilidad de cada suceso elemental es la misma, es decir, $P(x_i) = 1/n$, y A es un suceso compuesto por k sucesos elementales, se tiene que $P(A) = k/n$, que es nuevamente la regla de Laplace.

4. Si A y B son dos sucesos cualesquiera, se cumple que

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

Esta propiedad se puede extender al caso de tres sucesos, cumpliéndose que

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

Ejemplo 2.3.1 *Supón que realizamos el siguiente experimento: escogemos cuatro cartas de la baraja del mus con reemplazamiento. En cada elección apuntamos si la carta elegida es figura (sota, caballo o rey) o no lo es y al final anotamos todas las figuras que nos han salido.*
Consideramos como espacio muestral el número de figuras que han salido, (0,1,2,3 ó 4).
a)Describir los sucesos elementales y calcular sus probabilidades.
Los sucesos elementales son
A_i = *"han salido i figuras"*, $i = 0, 1, 2, 3, 4$.
Cómo hay 16 figuras en total se tiene que, aplicando la regla de Laplace, la probabilidad de sacar una figura al azar en la baraja del mus es $\frac{16}{40} = \frac{2}{5}$ *y la probabilidad de sacar una carta que no sea figura es* $\frac{24}{40} = \frac{3}{5}$. *A partir de eso y haciendo cálculos como los realizados anteriormente en este apartado, se puede ver que las probabilidades de estos sucesos son, respectivamente:*

$$P(A_0) = \left(\frac{3}{5}\right)^4, \; P(A_1) = 4 \cdot \frac{2}{5} \cdot \left(\frac{3}{5}\right)^3, \; P(A_2) = 6 \cdot \left(\frac{2}{5}\right)^2 \cdot \left(\frac{3}{5}\right)^2, \; P(A_3) = 4 \cdot \left(\frac{2}{5}\right)^3 \cdot \frac{3}{5}, \; P(A_4) = \left(\frac{2}{5}\right)^4.$$

b) Si B = *"Ha salido alguna figura", calcula* $P(B)$.
Tenemos que $B = A_0^c$, *así pues,* $P(B) = P(A_0^c) = 1 - P(A_0) = 1 - \frac{3}{5}^4 = 0'1296$.
c) Si C = *"Han salido tres ó más figuras", calcula* $P(C)$.
$C = A_3 \cup A_4$, *y además* $A_3 \cap A_4 = \emptyset$. *Entonces tenemos que* $P(C) = P(A_3 \cup A_4) = P(A_3) + P(A_4) = 4 \cdot \left(\frac{2}{5}\right)^3 \cdot \frac{3}{5} + \left(\frac{2}{5}\right)^4 = 0'1792$.

2.4. Cálculo de probabilidades en casos complejos

2.4.1. Probabilidad condicionada

Volvemos a repartir las cartas. Son repartidas de uno en uno y nuestro turno es el cuarto, es decir, el último. Al primer jugador le ha tocado un rey, al segundo un as y al tercero una sota. ¿Cuál es la probabilidad de que nos toque a nosotros un as? Aplicando la regla de Laplace, podemos decir que la probabilidad de que nos toque un as es $\frac{7}{37}$, debido a que ya han sido repartidas tres cartas y por tanto en la baraja solo quedan 37, y uno de los ases fue a parar al segundo jugador, por lo que solo quedan siete en la baraja. Imagina ahora, que ninguno de ellos hubiese recibido un as, ¿cuál sería entonces la probabilidad de que nos tocara a nosotros uno? En este caso, quedarían aun los ocho ases en la baraja, y por tanto ahora la probabilidad sería $\frac{8}{37}$. Pero, ¿y si dos de ellos tuvieran un as? ¿Con qué probabilidad nos tocaría a nosotros otro? En este caso, sería $\frac{6}{37}$. Como puedes ver el valor de las probabilidades de que nos toque un as, varía según las cartas que tengan nuestros contrincantes. Es decir, la probabilidad de un suceso puede depender de la información que tengamos antes de realizar el experimento. En este caso, la información que tenemos previa son las cartas de nuestros contrincantes, es decir, sabemos que cartas van a faltar en la baraja cuando nos toque nuestro turno.

En estos casos, es muy sencillo calcular las probabilidades, pero hay otros en los que nos tendremos que apoyar en la fórmula de la probabilidad condicionada.

Volvamos al ejemplo introducido antes. Si llamamos A al suceso ”Mi carta será un as” y B al suceso ”Los tres primeros jugadores han recibido: rey, as y sota respectivamente”, nosotros queremos calcular $P(A)$, pero sabiendo las cartas que tienen los otros, es decir, sabiendo que se ha dado el suceso B. Es decir, queremos calcular la probabilidad del suceso A condicionado a B, y lo denotaremos A/B. Para ello podemos aplicar la fórmula de la probabilidad condicionada, que dice:

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

Por tanto, tendremos que calcular $P(A \cap B)$ y $P(B)$. Para calcular $P(B)$ aplicamos lo aprendido antes en las extracciones sin reemplazamiento, y con un razonamiento similar obtenemos que

$$P(B) = \frac{8}{40} \cdot \frac{8}{39} \cdot \frac{4}{38},$$

mientras que para que se de $A \cap B$ tiene que ocurrir que los cuatro jugadores reciban un rey, un as, una sota y un as, es decir que

$$P(A \cap B) = \frac{8}{40} \cdot \frac{8}{39} \cdot \frac{4}{38} \cdot \frac{7}{37}.$$

Por tanto, aplicando la fórmula de la probabilidad condicionada, tenemos que

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{8}{40} \cdot \frac{8}{39} \cdot \frac{4}{38} \cdot \frac{7}{37}}{\frac{8}{40} \cdot \frac{8}{39} \cdot \frac{4}{38}} = \frac{7}{37},$$

cómo ya sabíamos antes.

En este caso, podríamos haber resuelto la cuestión sin recurrir a la fórmula de la probabilidad condicionada, pero en otros casos es necesario utilizarla.

Definición 2.4.1 *La probabilidad de que se dé un suceso A condicionado al suceso B, al que denotamos A/B, es*

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

2.4.2. Independencia de sucesos

Volvamos por un momento un poco atrás en nuestras explicaciones y pensemos en el ejemplo que vimos en el apartado de los diagramas de árbol. Si recuerdas, tuvimos que hacer muchas operaciones para calcular las probabilidades pedidas, y eso que ese era un caso de los más sencillos. Imagina que si en lugar de dos posibles resultados en cada extracción (rey o no rey en la primera extracción y sota o no sota en la segunda) tuviésemos tres, habría en total 9 posibilidades, si hubiera cuatro casos en cada extracción, tendríamos en total 16 posibilidades tras la segunda extracción. En general, si tenemos n posibles resultados en cada extracción, después de dos extracciones tendríamos que analizar n^2 casos, lo cual es bastante. Y esto es sólo si hablamos de dos extracciones, si fuesen tres, tendríamos n^3, si fuesen cuatro n^4, etc. Lo que queremos decir es que la técnica del diagrama de árbol, sólo es útil en casos muy sencillos, en cuanto los números se vayan haciendo un poco más grandes, el árbol se hace casi imposible de dibujar.

¿Habría alguna forma más sencilla de calcular la probabilidad de este tipo de sucesos? Pues sí, pero antes tenemos que estudiar un concepto nuevo: la independencia de sucesos aleatorios.

Definición 2.4.2 *Dado un experimento aleatorio, y dos sucesos cualesquiera de ese experimento, llamémosles A y B, decimos que esos dos sucesos son independientes si no importa que se dé uno de ellos para que se cumpla el otro, es decir*

$$P(A/B) = P(A) \quad y \quad P(B/A) = P(B).$$

De las fórmulas anteriores, se puede deducir que si dos sucesos son independientes, se tiene que

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

Este resultado es de gran utilidad para el cálculo de probabilidades en la repetición de sucesos aleatorios. Así, si repetimos n veces un experimento, siendo el resultado en cada ocasión independiente de las anteriores, y queremos calcular la probabilidad de que ocurra el suceso A_i en cada repetición, $\forall i = 1, ..., n$, tendremos que la probabilidad de que se den todos esos sucesos, que será el suceso intersección de todos ellos $A_1 \cap A_2 \cap ... \cap A_n$, es

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1) \cdot P(A_2) \cdot ... \cdot P(A_n).$$

Ejemplo 2.4.1 *Si elegimos dos cartas de la baraja del mus al azar con reemplazamiento, ¿cuál será la probabilidad de que hayan sido elegidas dos figuras? ¿Y dos cartas que no sean figuras? ¿Y una de cada tipo?*

Si llamamos A al suceso "la primera carta es una figura" y B al suceso "La segunda carta es una figura", tenemos que el suceso "Las dos cartas son figuras", es el suceso $A \cap B$. Claramente son sucesos independientes, ya que hemos hecho extracciones con reemplazamiento, y por tanto las condiciones en ambas extracciones son las mismas. Así, tenemos que

$$P(A \cap B) = P(A) \cdot P(B) = \frac{16}{40} \cdot \frac{16}{40}.$$

El suceso "Las dos cartas son no figuras", se representa en función de los sucesos A y B, de la siguiente manera $A^c \cap B^c$. También son sucesos independientes. Por tanto, se puede calcular la probabilidad de este suceso así

$$P(A^c \cap B^c) = P(A^c) \cdot P(B^c) = (1 - P(A)) \cdot (1 - P(B)) = \frac{24}{40} \cdot \frac{24}{40}.$$

En el suceso "Una carta es figura y la otra no", se representa en función de los sucesos A y B, de la siguiente manera $A^c \cap B$ y $A \cap B^c$, ya que puede ser que la primera sea figura y la segunda no y al revés. Así, el suceso que queremos estudiar, será la unión de ambos, $(A^c \cap B) \cup (A \cap B^c)$. Por tanto, como son sucesos incompatibles, (disjuntos ya que $(A^c \cap B) \cap (A \cap B^c) \subset A \cap A^c = \emptyset \Rightarrow (A^c \cap B) \cap (A \cap B^c) = \emptyset$),se puede calcular la probabilidad de este suceso así

$$P((A^c \cap B) \cup (A \cap B^c)) = P(A^c \cap B) + P(A \cap B^c) = P(A^c) \cdot P(B) + P(A) \cdot P(B^c) = \frac{24}{40} \cdot \frac{16}{40} + \frac{16}{40} \cdot \frac{24}{40}.$$

Puedes ver cuál es el suceso más probable de los tres.

Ejercicio 2.4.1 *Considera el siguiente experimento: Elige al azar dos cartas de nuestra baraja con reemplazamiento, es decir, devolviendo a la baraja la primera carta una vez observada y antes de realizar la segunda extracción. Suma el valor numérico que le da el mus a ambas cartas. Responde a las siguientes preguntas:*

a) Describe el espacio muestral, el suceso seguro y un suceso imposible del experimento.

b) Calcula la probabilidad de que la suma de las cartas sea 20.

c) Calcula la probabilidad de que la suma sea menor o igual que seis.

Ejercicio 2.4.2 *Haz lo mismo que en el ejercicio anterior pero suponiendo que elegimos al azar cartas SIN reemplazamiento.*

2.4.3. Probabilidad total

Imagina que seleccionamos dos cartas cualesquiera de la baraja, sin reemplazamiento. Observamos la primera carta, y después la segunda. ¿Cuál es la probabilidad de que la segunda carta sea un rey? Con los conocimientos que ya tenemos, fácilmente podemos decir que, si la primera carta fue un rey, la probabilidad de que la segunda también lo sea es $\frac{7}{39}$. Sin embargo, si la primera no lo fue, entonces tenemos que la probabilidad de que la segunda carta sea un rey es $\frac{8}{39}$. Como ves, depende de que tipo de carta fue la primera para poder asegurar algo sobre la segunda. En el apartado de la probabilidad condicionada, partíamos con la ventaja de que lo sabíamos, sabíamos qué tipo de carta era la primera. Pero ahora no. ¿Cómo solucionamos ese problema? Pues teniendo en cuenta ambas posibilidades, que la primera carta sea un rey o que no lo sea. Veamos como lo resolvemos:

Consideremos los siguientes sucesos aleatorios:

1. A_1 = "La primera carta es un rey",
2. A_2 = "La segunda carta es un rey".

Queremos calcular $P(A_2)$. Pues bien, vamos a tener en cuenta si se da A_1 o no se da. ¿Cómo? Dividiendo $P(A_2)$ en varias probabilidades que son más sencillas de calcular. Para ello vamos a recurrir a las propiedades de las operaciones con los sucesos.

Si consideramos a $\overline{A}_1$ como el suceso complementario de A_1, claramente tenemos que

$$A_1 \cup \overline{A}_1 = E, \qquad A_1 \cap \overline{A}_1 = \emptyset.$$

Así pues, tenemos que

$$A_2 = A_2 \cap E \Longrightarrow A_2 = (A_2 \cap A_1) \cup (A_2 \cap \overline{A}_1).$$

Además, como

$$(A_2 \cap A_1) \cap (A_2 \cap \overline{A}_1) = \emptyset,$$

tenemos que

$$P(A_2) = P(A_2 \cap A_1) + P(A_2 \cap \overline{A}_1).$$

Aplicando la fórmula de la probabilidad condicionada, obtenemos

$$P(A_2 \cap A_1) = P(A_2/A_1) \cdot P(A_1),$$

y

$$P(A_2 \cap \overline{A}_1) = P(A_2/\overline{A}_1) \cdot P(\overline{A}_1).$$

Y esas dos probabilidades sí las podemos calcular fácilmente, aplicando la regla de Laplace y las técnicas vistas para las extracciones sin reemplazamiento. Así, tendremos que:

$$P(A_2/A_1) \cdot P(A_1) = \frac{7}{39} \cdot \frac{8}{40},$$

y

$$P(A_2/\overline{A}_1) \cdot P(\overline{A}_1) = \frac{8}{39} \cdot \frac{32}{40} = ... = 0'2.$$

Como ves, hemos dividido la probabilidad en dos sumando: un primer sumando en donde suponemos que la primera carta es un rey, A_1, y otro en el que suponemos que la primera carta no es un rey, $\overline{A}_1$.

Los sucesos A_1 y $\overline{A}_1$ tienen dos características muy especiales

$$A_1 \cap \overline{A}_1 = \emptyset \quad \text{y} \quad A_1 \cup \overline{A}_1 = E.$$

Esta técnica la podemos aplicar como regla general:

si tenemos un conjunto $A_1, A_2, ..., A_n$ de sucesos incompatibles dos a dos ($A_i \cap A_j = \emptyset \quad \forall i \neq j$), cumpliendo que $A_1 \cup A_2 \cup ... \cup A_n = E$, (si cumplen esas dos condiciones, decimos que ese conjunto es un sistema completo de sucesos), entonces la probabilidad de un suceso $S \subset E$ es

$$P(S) = P(A_1) \cdot P(S/A_1) + P(A_2) \cdot P(S/A_2) + ... + P(A_n) \cdot P(S/A_n),$$

que es la llamada fórmula de la probabilidad total.

La parte más difícil a la hora de aplicar esta fórmula, es escoger bien el sistema completo de sucesos, ya que una mala elección sólo nos provocaría más dificultades en la resolución del problema. Se tendrán que estudiar cuáles son los sucesos que nos convienen, porque una mala elección de el sistema completo de sucesos no nos ayudaría a resolver el problema.

Ejercicio 2.4.3 *Halla la probabilidad, en el mismo experimento que planteamos al principio de la sección, de que la segunda carta no sea una figura.*

Ejercicio 2.4.4 *Planteemos el siguiente ejercicio: repartimos tres cartas de la baraja del mus. Calcula, eligiendo un adecuado sistema completo de sucesos y aplicando la fórmula de la probabilidad total, la probabilidad de que la tercera carta sea un as.*

¿Cuál es la probabilidad de que la tercera carta no sea un as?

Indicación: escoger como sistema completo de sucesos el número de ases que hayan salido entre las dos primeras cartas.

2.4.4. Teorema de Bayes

Pongámonos en la situación planteada anteriormente como ejemplo. Sacábamos dos cartas seguidas sin reemplazamiento. Se nos puede plantear una nueva pregunta: sabiendo que la segunda carta fue un rey, ¿cuál es la probabilidad de que la primera también lo fuera? Esta pregunta puede parecer similar a la que planteábamos en la anterior sección, pero tiene una gran diferencia: en este caso, ya hemos realizado el experimento completo (hemos visto la segunda carta) y nos preguntamos qué carta era la primera. Es decir, si llamamos a los sucesos A_1 y A_2 como antes, queremos calcular

$$P(A_1/A_2).$$

Aplicando la fórmula de la probabilidad condicionada, obtenemos que

$$P(A_1/A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}.$$

Y desarrollando el denominador según la fórmula de la probabilidad total y aplicando de nuevo la fórmula de la probabilidad condicionada en el numerador, obtenemos que

$$P(A_1/A_2) = \frac{P(A_2/A_1) \cdot P(A_1)}{P(A_2/A_1) \cdot P(A_1) + P(A_2/\overline{A}_1) \cdot P(\overline{A}_1)}.$$

El cálculo será sencillo a partir de los que realizamos en la anterior sección, así, tenemos que

$$P(A_1/A_2) = \frac{\frac{7}{39} \cdot \frac{8}{40}}{\frac{1}{5}} = \frac{7}{39}.$$

En general, la fórmula de Bayes se obtiene de la siguiente forma:

dado un Sistema Completo de Sucesos $A_1, A_2, ..., A_n$ y un suceso cualquiera S, queremos calcular la probabilidad de que se de el suceso A_i, sabiendo que al hacer el experimento, se dio el suceso S, es decir, calcular $P(A_i/S)$. Por el mismo razonamiento de antes se tiene que:

$$P(A_i/S) = \frac{P(A_i \cap S)}{P(S)} = \frac{P(S/A_i) \cdot P(A_i)}{\sum_{i=i}^{n} P(S/A_i) \cdot P(A_i)},$$

donde $P(A_i)$ es la probabilidad a priori del suceso A_i (se sabe antes de realizar el experimento) y $P(A_i/S)$ su probabilidad a posteriori, pues se calcula una vez realizado el experimento.

Al igual que dijimos en la sección dedicada a la probabilidad condicionada, para aplicar correctamente la regla de Bayes, hay que elegir un adecuado sistema completo de sucesos, que será la parte más difícil del problema.

Ejercicio 2.4.5 *Al sacar dos cartas de la baraja del mus sin reemplazamiento, calcula la probabilidad de que la primera carta fuera un siete sabiendo que la segunda es una sota.*

Ejercicio 2.4.6 *Sacamos cuatro cartas al azar. Sabemos que la carta que salió en cuarto lugar es un as. ¿Cuál es la probabilidad de que la primera también lo fuera? ¿Y de que la primera fuera una sota?*

Nota: hay que elegir dos sistemas completos de sucesos diferentes para responder a cada pregunta.

2.5. Resolución de la pregunta inicial

Decíamos que los cuatro amigos habían echado una partida, y habían obtenido 8 parejas de reyes, seis medias de ases o reyes y cinco dúplex. Discutían cuál de esas tres jugadas era más probable que ocurriera y no se ponían de acuerdo. Por eso, vamos a ayudarles. Denotaremos a esos tres sucesos así:

- $RR =$ ”Sacar una pareja de reyes”.
- $M =$ ”Sacar una media de ases o de reyes”.
- $D =$ ”Sacar un dúplex cualquiera”.

Vamos a calcular sus diferentes probabilidades y veremos cuál es más posible que salga.

Comencemos por calcular la probabilidad de obtener una pareja de reyes, es decir, $P(RR)$. Veamos, para obtener esa pareja tenemos que tener dos reyes (obviamente) y de las otras dos, ninguna puede ser un rey pues si no tendríamos una media o un dúplex de reyes, ni tampoco pueden ser iguales entre sí, pues si no tendríamos un dúplex de reyes y las otras cartas.

Tendremos que distinguir dos casos, uno en el que una de las cuatro cartas es un as y otro en el que no, debido a que el número de ases es diferente al número de las otras cartas.

Así, si denotamos el suceso $B =$ ”Sacar un as”, aplicando el teorema de la probabilidad total tenemos que

$$P(RR) = P(RR/B) \cdot P(B) + P(RR/\overline{B}) \cdot P(\overline{B}) = P(RR \cap B) + P(RR \cap \overline{B}).$$

Si denotamos por A un as y a C y C' cartas distintas de reyes o ases y también entre sí, tenemos que una pareja de reyes podría ser:

$$(R, R, A, C),$$

si sale algún as o

$$(R, R, C, C'),$$

si no tenemos ningún as, y que las cartas (R, R, A, C) y todas sus posibles variaciones de orden, representan al suceso $RR \cap B$ mientras que la combinación (R, R, C, C') y todas sus variaciones de orden representan al suceso $RR \cap \overline{B}$.

Comencemos calculando la probabilidad de que se de el reparto de cartas (R, R, A, C) en ese mismo orden, después calcularemos cuantas posibles reordenaciones hay, y como son todas equiprobables, solo tendremos que multiplicar la probabilidad obtenida por el número de variaciones.

Calculemos $P(R, R, A, C)$. La probabilidad de que la primera carta sea un rey es $\frac{8}{40}$, de que la segunda también lo sea es $\frac{7}{39}$. La probabilidad de que la tercera carta sea un as es $\frac{8}{38}$, mientras que la probabilidad de que la cuarta carta sea una diferente a rey o as es $\frac{24}{37}$. Por tanto tenemos que

$$P(R, R, A, C) = \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{8}{38} \cdot \frac{24}{37}.$$

Calculemos la probabilidad de una variación de orden en esas cartas, para ver que en efecto son equiprobables. Por ejemplo,

$$P(A,R,C,R)=\frac{8}{40}\cdot\frac{8}{39}\cdot\frac{24}{38}\cdot\frac{7}{37}=\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{8}{38}\cdot\frac{24}{37}=P(R,R,A,C).$$

Lo mismo se podría hacer con todas las variaciones, pero creemos que queda suficientemente claro que son equiprobables.

Una vez hallada la probabilidad de uno de ellos, calculemos el número total de variaciones de orden de esa combinación de cartas.

Tenemos una variación de cuatro cartas en las que se repiten dos. Recordemos que la fórmula general de las variaciones con repetición de n elementos en los que hay n_1 elementos de una clase, n_2 elementos de otra,... y n_k de otra es

$$V^n_{n_1\ldots n_k}=\frac{n!}{n_1!\cdot n_2!\cdot\ldots\cdot n_k!}.$$

Por tanto, en nuestro caso tenemos que

$$P(R,R,A,C)=\frac{4!}{2!1!1!}=\frac{24}{2}=12.$$

Así pues, tenemos que

$$P(RR\cap B)=12\cdot\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{8}{38}\cdot\frac{24}{37}.$$

Para obtener $P(RR\cap\overline{B})$ vamos a seguir un proceso análogo. Lo primero que vamos a hacer es calcular $P(R,R,C,C')$. Utilizando como antes la regla de Laplace carta a carta, obtenemos que

$$P(R,R,C,C')=\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{24}{38}\cdot\frac{20}{37}.$$

De nuevo todas sus variaciones son equiprobables y también otra vez, tenemos que hay 12 posibles reordenaciones, ya que vuelven a ser variaciones con repetición de 4 elementos de los que se repiten dos. Así pues, vuelve a haber 12 posibles reordenaciones, todas ellas equiprobables y por tanto, la probabilidad de obtener una pareja de reyes sin ningún as es

$$P(RR\cap\overline{B})=12\cdot\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{24}{38}\cdot\frac{20}{37}$$

Así pues, aplicando la fórmula de la probabilidad total tenemos que

$$P(RR)=P(RR\cap B)+P(RR\cap\overline{B})=12\cdot\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{8}{38}\cdot\frac{24}{37}+12\cdot\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{24}{38}\cdot\frac{20}{37}=0'255.$$

Calculemos ahora $P(M)$. La media puede ser de reyes o de ases. Es decir, la primera carta puede ser un rey o un as. La segunda tendrá que ser un rey si la primera también lo era o un as si la anterior era otro as. Con la tercera pasa lo mismo mientras que la cuarta carta no puede ser igual a las anteriores, pues si no tendríamos un dúplex y no una media. Así pues, la probabilidad de obtener una media de ases o reyes es

$$\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{6}{38}\cdot\frac{32}{37}+\frac{8}{40}\cdot\frac{7}{39}\cdot\frac{6}{38}\cdot\frac{32}{37}=\frac{16}{40}\cdot\frac{7}{39}\cdot\frac{6}{38}\cdot\frac{32}{37},$$

y son todas equiprobables. Notar que el primer sumando corresponde a la probabilidad de obtener una media de reyes y el segundo a la probabilidad de obtenerla de ases. Además son sucesos incompatibles, por lo que se pueden sumar sus probabilidades para obtener el resultado.

Para averiguar el número de posibles reordenaciones, recurrimos otra vez a la fórmula de las variaciones con repetición. En este caso, tenemos cuatro elementos en los que se repiten tres del mismo tipo y el otro es desigual. Es decir tenemos $V^4_{3,1} = \frac{4!}{3!} = 4$ posibles reordenaciones de las mismas cartas.

Y multiplicando la probabilidad de una de ellas por todas sus posibles reordenaciones, tenemos que

$$P(M) = 4\frac{16}{40} \cdot \frac{7}{39} \cdot \frac{6}{38} \cdot \frac{32}{37} = 0'039.$$

Por último, vamos a calcular la probabilidad de obtener un dúplex cualquiera en el reparto del mus. Para ello, vamos a dividir el suceso sacar un dúplex cualquiera en los diferentes dúplex que podemos obtener según si tienen reyes, ases o los dos.

Así pues podemos obtener los siguientes dúplex:

1. Reyes-Ases (R, R, A, A). $P(R, R, A, A) = \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{8}{38} \cdot \frac{7}{37}$. Tenemos variaciones de cuatro elementos que son iguales dos a dos, es decir, $V^4_{2,2} = \frac{4!}{2!2!} = 6$ posibles reordenaciones. Por tanto, la probabilidad de obtener un dúplex de reyes ases es

$$6 \cdot \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{8}{38} \cdot \frac{7}{37}.$$

2. Un dúplex de reyes y otra carta que no sea ni rey ni as, (R, R, C, C). $P(R, R, C, C) = \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{24}{38} \cdot \frac{3}{37}$. En este caso volvemos a tener $V^4_{2,2} = \frac{4!}{2!2!} = 6$ variaciones, por lo que la probabilidad de obtener este tipo de dúplex es

$$6 \cdot \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{24}{38} \cdot \frac{3}{37}.$$

3. Un dúplex de ases y otra carta que no sea ni rey ni as, (A, A, C, C). $P(A, A, C, C) = \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{24}{38} \cdot \frac{3}{37}$. En este caso volvemos a tener $V^4_{2,2} = \frac{4!}{2!2!} = 6$ variaciones, por lo que la probabilidad de obtener este tipo de dúplex es

$$6 \cdot \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{24}{38} \cdot \frac{3}{37}.$$

4. Reyes-Reyes (R,R,R,R). Tenemos que $P(R, R, R, R) = \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{6}{38} \cdot \frac{5}{37}$ y como todas las cartas son iguales, tenemos una única posible ordenación, por lo que la probabilidad de obtener un dúplex de reyes-reyes es

$$\frac{8}{40} \cdot \frac{7}{39} \cdot \frac{6}{38} \cdot \frac{5}{37}.$$

5. Ases-Ases (A,A,A,A). Tenemos que $P(A, A, A, A) = \frac{8}{40} \cdot \frac{7}{39} \cdot \frac{6}{38} \cdot \frac{5}{37}$ y como todas las cartas son iguales, tenemos una única posible ordenación, por lo que la probabilidad de obtener un dúplex de ases-ases es

$$\frac{8}{40} \cdot \frac{7}{39} \cdot \frac{6}{38} \cdot \frac{5}{37}.$$

6. El suceso ”obtener un dúplex de cartas que no sean ni ases ni reyes y distintas entre sí” se puede representar como (C,C,C’,C’), y $P(C,C,C',C') = \frac{24}{40}\cdot\frac{3}{39}\cdot\frac{20}{38}\cdot\frac{3}{37}$. En cuanto a las posibles reordenaciones, volvemos a tener $V^4_{2,2} = \frac{4!}{2!2!} = 6$. Por tanto, la probabilidad de obtener este tipo de dúplex es
$$6\cdot\frac{24}{40}\cdot\frac{3}{39}\cdot\frac{20}{38}\cdot\frac{3}{37}.$$

7. El último dúplex posible es el formado por cuatro cartas iguales y distintas de ases o reyes, es decir, una combinación como (C,C,C,C). La probabilidad de que se de es $\frac{24}{40}\cdot\frac{3}{3}\cdot\frac{2}{38}\cdot\frac{1}{37}$ y como las cuatro cartas son iguales, solo tenemos una posible ordenación. Por tanto, la probabilidad de obtener este dúplex es
$$\frac{24}{40}\cdot\frac{3}{3}\cdot\frac{2}{38}\cdot\frac{1}{37}.$$

Estos sucesos que hemos enumerado, los diferentes dúplex, son todos ellos disjuntos y su unión da lugar al suceso $D =$ ”sacar un dúplex cualquiera”. Por tanto, $P(D)$ será igual a la suma de las probabilidades obtenidas para los dúplex anteriores. Efectuando cálculos, resulta que
$$P(D) = 0'035.$$

En definitiva, tenemos que
$$P(RR) = 0'255, \qquad P(M) = 0'039, \qquad P(D) = 0'035,$$
es decir, claramente la pareja de reyes es el suceso más probable de los tres, mientras que los sucesos ”sacar una media de reyes o ases” y ”sacar un dúplex cualquiera” están bastante igualados, aunque con un poco más de probabilidad se dará el primero de ellos.

Con estos cálculos, hemos podido dar respuesta a la pregunta que se hacían los amigos al terminar su partida. Como ves, aunque es mucho más probable sacar una pareja de reyes que un dúplex, esa proporción no se ha visto reflejada en la realidad, esto es debido a que cuando decimos que un suceso es más probable que otro, no queremos decir que siempre se va a dar en más ocasiones que el menos probable, sino que teóricamente tiene más posibilidades de ocurrir. Esa es la gran diferencia existente entre teoría y práctica.

Capítulo 3

Distribuciones de probabilidad unidimensionales

3.1. Objetivos

- Conocer el concepto de distribución de probabilidad y su motivación.
- Calcular la función de densidad y la función de distribución de una distribución de probabilidad.
- Entender el concepto de moda, esperanza y varianza de una distribución de probabilidad discreta, así como poder calcularlas a partir de la función de densidad o de la función de distribución.

3.2. El ejemplo

Un inversor bursátil tiene a su disposición 1.000.000 de euros para invertir en bolsa. Se está planteando dos opciones:

- invertirlo todo a plazo fijo, con lo que se aseguraría un beneficio fijo del 16 %,
- arriesgarse a un plan de inversiones.

Según el estudio de un analista de bolsa, ese plan de inversiones da los siguientes beneficios expresados en tanto por ciento:

Beneficio (%)	Probabilidad
30	0'15
25	0'2
20	0'25
15	0'15
10	0'1
5	0'1
0	0'05

El inversor tiene que decidir que hacer con su dinero. Se pueden plantear dos preguntas:

El inversor se fía mucho del concepto de ganancia esperada (el dinero que ganaría en promedio si invirtiera muchas veces), ¿qué decidir teniendo en cuenta la ganancia esperada del plan de inversiones, y en pro de obtener un mejor beneficio?

Otra opción es: el inversor no se quiere arriesgar mucho y dice que quiere obtener al menos el 16 % de beneficio con el 70 % de probabilidad. ¿Qué debe hacer en ese caso?

Resolveremos este problema, apoyándonos en las distribuciones de probabilidad, concepto que introduciremos previamente.

3.3. Introducción. Variable aleatoria discreta y distribuciones de probabilidad

Anteriormente, hemos estudiado cómo asignar probabilidades a distintos sucesos aleatorios, pero todos ellos de forma particular. En este capítulo, veremos métodos generales que luego particularizaremos para cada caso.

Como puedes comprobar, en el ejemplo nos han dado una tabla con las probabilidades de obtener los diferentes beneficios tras el plan de inversiones. Con ayuda de esa tabla vamos a decidir qué hacer con el dinero. Antes de nada, tenemos que recordar que las conclusiones que vamos a obtener son teóricas. Puede que decidamos algo y después el azar, porque las preguntas enunciadas en el ejemplo dependen del azar, nos quite la razón y eche por tierra todos nuestros cálculos. Aún así, la probabilidad nos dará una idea aproximada de cómo irá la ganancia.

Comencemos a analizar el problema.

Definamos el concepto de variable aleatoria discreta.

En nuestro problema tenemos planteado un experimento aleatorio, el beneficio que obtenemos al meter el dinero en el plan de inversiones. Tras el experimento aleatorio podemos obtener varios beneficios diferentes en cada caso. Bien, esos posibles beneficios serán los resultados de nuestro experimento aleatorio.

Así pues, si denotamos por

$$X = \text{"beneficio obtenido al llevar a cabo el plan de inversiones"},$$

decimos que X es una variable aleatoria discreta y los posibles valores que puede tomar esta variable son los resultados.

En este caso la variable aleatoria X puede tomar siete resultados o valores (los diferentes beneficios) que denotaremos con la letra x y un subíndice en cada caso. Es decir, tenemos que la variable aleatoria discreta X puede tomar los valores $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ donde

1. $x_1 = 0 \Longrightarrow P(X = x_1) = 0'05$,
2. $x_2 = 5 \Longrightarrow P(X = x_2) = 0'1$,
3. $x_3 = 10 \Longrightarrow P(X = x_3) = 0'1$,
4. $x_4 = 15 \Longrightarrow P(X = x_2) = 0'15$,
5. $x_5 = 20 \Longrightarrow P(X = x_5) = 0'25$,
6. $x_6 = 25 \Longrightarrow P(X = x_6) = 0'2$,
7. $x_7 = 30 \Longrightarrow P(X = x_7) = 0'15$.

Al conjunto de los valores que puede tomar una variable aleatoria discreta y sus respectivas probabilidades, lo llamamos distribución de probabilidad discreta. En nuestro ejemplo tenemos la siguiente distribución de probabilidad:

$$P(X = x_1) = 0'05, \quad P(X = x_2) = 0'1, \quad P(X = x_3) = 0'1, \quad P(X = x_4) = 0'15,$$

$$P(X = x_5) = 0'2, \quad P(X = x_6) = 0'25, \quad P(X = x_7) = 0'15.$$

Lógicamente, la suma de todas las probabilidades tiene que ser igual a 1. Esto deriva del hecho de que si una variable aleatoria discreta toma sus valores en el conjunto $\{x_1, ..., x_n\}$, es decir la variable aleatoria tomará SEGURO uno de esos valores, es decir,

$$P(X \in \{x_1, ..., x_n\}) = 1,$$

pues pase lo que pase la variable cogerá uno de esos valores y eso es precisamente lo que definíamos como suceso seguro, y como son sucesos disjuntos, es decir, la variable aleatoria X no puede tomar dos valores diferentes a la vez, tenemos que

$$P(X = x_1) + ... + P(X = x_n) = P(X = x_1 \ \cup \ ... \ \cup X = x_n) = P(X \in \{x_1, ..., x_n\}) = 1.$$

Así pues, definiremos una variable aleatoria discreta y su distribución de probabilidad como sigue:

Definición 3.3.1 *Una variable aleatoria es una aplicación que toma sus valores según el resultado de un experimento aleatorio.*

Decimos que X es una variable aleatoria discreta, si puede tomar sus valores dentro de un conjunto $x_1, x_2, ..., x_n, ...$ que se pueda contar (finito o infinito), con probabilidades respectivas

$P(X = x_1), P(X = x_2), ..., P(X = x_n), ...$ *y cumpliéndose que* $0 \leq P(X = x_n) \leq 1,\ \forall n$ *y* $\sum_{i=1}^{\infty} P(X = x_i) = 1$.

Las probabilidades dadas anteriormente forman una distribución de probabilidad *para la variable* X*, es decir, la distribución de probabilidad de la variable aleatoria* X *es la función que asigna a cada uno de sus posibles valores la probabilidad con la que éste es tomado.*

En nuestro estudio de variables aleatorias discretas, profundizaremos en aquellas variables que sólo pueden tomar un conjunto finito de valores, como por ejemplo la situación que planteamos al principio del capítulo, solo siete posibilidades.

Como resumen de esta sección recordemos qué variable aleatoria tenemos en el problema y cuál es su distribución de probabilidad:

la variable aleatoria es $X =$ "el beneficio tras llevar a cabo el plan de inversiones" y su distribución de probabilidad es:

$$P(X = x_1) = 0'05, \quad P(X = x_2) = 0'1, \quad P(X = x_3) = 0'1, \quad P(X = x_4) = 0'15,$$

$$P(X = x_5) = 0'2, \quad P(X = x_6) = 0'25, \quad P(X = x_7) = 0'15,$$

donde $x_1 = 0$, $x_2 = 5$, $x_3 = 10$, $x_4 = 15$, $x_5 = 20$, $x_6 = 25$ y $x_7 = 30$.

3.4. Función de Distribución

Si te has dado cuenta, antes hemos asociado a los sucesos $x_1, ..., x_7$ con los diferentes beneficios de una forma especial, en un orden creciente, es decir, al x_1 le asociamos el más pequeño, el 0, al x_2 le asociamos el 5,... y al x_7 le asociamos el más grande, el 30.

Eso se ha hecho así no por casualidad, sino porque nos va a servir para definir una nueva función relacionada con la variable aleatoria X y que la caracterizará unívocamente, la función de distribución que definiremos más adelante.

Antes de eso vamos a responder a la segunda pregunta del problema. El inversor quiere que la probabilidad de obtener al menos el 16 % de beneficio sea superior a 0'7 en el plan de inversiones. Si eso no es así, meterá todo su dinero a plazo fijo donde se asegurará el 16 % de beneficios.

Es decir, en términos de nuestra variable aleatoria, quiere que $P(X > 16) \geq 0'7$. ¿Cómo calculamos esto? Muy fácil, cogemos los posibles valores que puede tomar la variable aleatoria discreta X y sumamos sus probabilidades. En nuestro problema tenemos que

$$P(X > 16) = P(X = 20) \ + \ P(X = 25) \ + \ P(X = 30)$$

$$= P(X = x_5) \ + \ P(X = x_6) \ + \ P(X = x_7) = 0'25 + 0'2 + 0'15 = 0'6.$$

Por tanto, en nuestro problema de inversiones, tenemos que la probabilidad de obtener un beneficio superior al 16 % es sólo del 0'6, frente al 0'7 que se quería asegurar el inversor. Por tanto deberíamos aconsejar al inversor que no llevara a cabo el plan de inversiones y metiera el dinero a plazo fijo donde sí obtendría seguro ese 16 %.

Vamos a definir ahora esa función de la que hablábamos antes y que nos caracterizará nuestra variable aleatoria de forma unívoca. Además, con esa función también podremos dar respuesta a la pregunta que respondíamos antes. Esa función es la función de distribución (FDD). Pasemos a definirla:

Definición 3.4.1 *Si tenemos los posibles valores que puede tomar una variable aleatoria discreta X ordenados de forma creciente, $x_1, ..., x_n$ definimos la función*

$$F : \mathbb{R} \longrightarrow [0, 1],$$

de la siguiente manera

$$F(x) = P(X \leq x), \quad \forall x \in \mathbb{R}.$$

Para un valor posible de X, x_i, tenemos que $P(X \leq x_i) = P(X = x_1) + P(X = x_2) + ... + P(X = x_i)$, por tanto $P(X \leq x)$ para un x cualquiera de $\mathbb{R}$, tendremos que $P(X \leq x)$ será la suma de las probabilidades de todos los x_i menores o iguales que x, es decir,

$$P(X \leq x) = \sum_{x_i \leq x} P(X = x_i).$$

Así pues, en el ejemplo del plan de inversiones tenemos que F es:

- Si $x < 0$, $\quad F(x) = 0$, pues no hay ningún $x_i < 0$.
- Si $x \in [0, 5)$, $\quad F(x) = P(X = x_1) = 0'05$.
- Si $x \in [5, 10)$, $\quad F(x) = P(X = x_1) + P(X = x_2) = 0'05 + 0'1 = 0'15$.
- Si $x \in [10, 15)$, $\quad F(x) = P(X = x_1) + P(X = x_2) + P(X = x_3) = 0'05 + 0'1 + 0'1 = 0'25$.
- Si $x \in [15, 20)$, $\quad F(x) = P(X = x_1) + P(X = x_2) + P(X = x_3) + P(X = x_4) = 0'05 + 0'1 + 0'1 + 0'15 = 0'4$.
- Si $x \in [20, 25)$, $\quad F(x) = P(X = x_1) + P(X = x_2) + P(X = x_3) + P(X = x_4) + P(X = x_5) = 0'05 + 0'1 + 0'1 + 0'15 + 0'2 = 0'6$.
- Si $x \in [25, 30)$, $\quad F(x) = P(X = x_1) + P(X = x_2) + P(X = x_3) + P(X = x_4) + P(X = x_5) + P(X = x_6) = 0'05 + 0'1 + 0'1 + 0'15 + 0'2 + 0'25 = 0'85$.
- Si $x \geq 30$, $\quad F(x) = P(X = x_1) + P(X = x_2) + P(X = x_3) + P(X = x_4) + P(X = x_5) + P(X = x_6) + P(X = x_7) = 0'05 + 0'1 + 0'1 + 0'15 + 0'2 + 0'25 + 0'15 = 1$.

En resumen, podemos expresar la función F como sigue:

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 0'05 & \text{si } 0 \leq x < 5, \\ 0'15 & \text{si } 5 \leq x < 10, \\ 0'25 & \text{si } 10 \leq x < 15, \\ 0'4 & \text{si } 15 \leq x < 20, \\ 0'65 & \text{si } 20 \leq x < 25, \\ 0'85 & \text{si } 25 \leq x < 30, \\ 1 & \text{si } x \geq 30. \end{cases}$$

Como puedes ver, tenemos que definir la función de distribución de una variable aleatoria discreta a trozos, y los saltos de continuidad están en los puntos donde la probabilidad es estrictamente positiva.

La función F nos sirve para calcular la probabilidad de que una variable aleatoria tome valores menores o iguales a uno dado. Así pues podemos calcular la probabilidad que nos pedía el problema utilizando esta función y las propiedades de la probabilidad del suceso complementario de la siguiente manera:

$$P(X > 16) = 1 - P(X \leq 16) = 1 - F(16) = 1 - 0'4 = 0'6,$$

tal y como calculamos antes.

Notar que hemos utilizado que el suceso complementario de que $X > 16$ es el suceso $X \leq 16$.

La función de distribución de una variable aleatoria discreta tiene unas características especiales, que son:

1. $\lim\limits_{x \to 0} F(x) = 0.$
2. $\lim\limits_{x \to +\infty} F(x) = 0.$
3. F es creciente en todo $\mathbb{R}$.
4. F es continua a la derecha en todo $\mathbb{R}$, es decir,

   $$\lim_{x \to x^0} F(x) - F(x^0) = 0, \quad \forall x^0 \in \mathbb{R}.$$

   Además, los puntos de discontinuidad por la izquierda serán los puntos donde la variable aleatoria discreta asociada a esta función puede tomar sus valores.

Si te fijas en la función de distribución que hemos hallado, cumple todas las condiciones, además los puntos de discontinuidad son 0, 5, 10, 15, 20, 25 y 30, que son efectivamente los puntos en donde la variable aleatoria discreta X puede tomar sus valores, según el estudio que ha realizado el analista. Todo esto será teórico, pues es fácil pensar que esta variable podrá tomar otros valores.

Por otro lado, si tenemos una función que cumpla las condiciones descritas arriba, podremos decir que es una función de distribución asociada a una variable aleatoria discreta.

Si por el contrario una función no verifica alguna de las cuatro condiciones especificadas arriba, quedará automáticamente descartada como función de distribución.

3.5. La Moda

Tal y como ocurre en la vida real, la moda en una variable aleatoria es el valor que más se repite. Cuando se dice que una clase de ropa está de moda, queremos expresar que hay mucha gente que se la pone, es decir, si cogemos a una persona al azar, esta llevará puesta esa ropa con mayor probabilidad que cualquier otra prenda de vestir. Análogamente, si una variable aleatoria toma valores dentro del conjunto $\{x_1, x_2, ..., x_n\}$, la moda será aquel x_i que dé el máximo en la función de probabilidad.

En el ejemplo del inversor, el beneficio que más se repite en el plan de inversiones, es decir, el que tiene una mayor probabilidad de ocurrir, es obtener el 20 % de beneficio. Decimos entonces, que ese valor de la variable es la moda de la distribución.

3.6. La Esperanza

Ahora vamos a responder a la primera pregunta del problema. El inversor quería conocer el valor esperado de beneficios a plazo fijo para después decidir qué hacer con el dinero. Metiendo el dinero a plazo fijo, el beneficio esperado es fácil de calcular, pues como hay una sola posibilidad (ganar el 16 %) esa es la que se dará. Sin embargo, en el plan de inversión propuesto, no tenemos esa seguridad. Por eso hablaremos de esperanza, valor esperado, beneficio esperado, media, etc.

La esperanza es una medida que nos proporciona una aproximación de el promedio que cabe esperar que ocurra en un experimento tras un número grande de repeticiones. En un juego, la esperanza será el valor que 'esperamos' ganar (o perder) después de un número elevado de apuestas.

Es una media teórica que nos indica el valor promedio que cabría esperar al realizar el experimento aleatorio un número elevado de veces.

Si en una distribución de frecuencias definíamos la media aritmética como

$$\overline{x} = \sum_{i=1}^{n} x_i f_i,$$

donde x_i son los valores que puede tomar la variable y f_i son sus respectivas frecuencias relativas, en una distribución de probabilidad sustituimos la frecuencia relativa de cada valor por su probabilidad.

Definición 3.6.1 *La media o esperanza matemática de una variable aleatoria X que toma sus valores en un conjunto $\{x_1, ..., x_n\}$ con probabilidades respectivas $\{p_1, ..., p_n\}$, (es decir, $P(X = x_i) = p_i, \ \forall i \in \{1, 2, ..., n\}$) , a la que llamaremos μ ó EX, se calcula mediante la expresión*

$$EX = \mu = \sum_{i=n}^{n} x_i p_i.$$

Así pues, en nuestro problema el beneficio esperado al llevar a cabo el plan de inversiones es:

$$\mu = \sum_{i=1}^{7} x_i P(X = x_i)$$

$$= 0 \cdot 0'05 + 5 \cdot 0'1 + 10 \cdot 0'1 + 15 \cdot 0'15 + 20 \cdot 0'25 + 25 \cdot 0'2 + 30 \cdot 0'15 = 18'25.$$

Es decir, el beneficio esperado que obtendremos llevando a cabo el plan de inversiones es el $18'25$ %, superior al 16 % que nos ofrece el mercado de valores.

Claramente el inversor que confíe en el valor esperado arriesgará y llevará a cabo el plan de inversiones en pro de obtener un mayor beneficio.

3.7. La Varianza

Supongamos ahora que un inversor quiere conocer un intervalo en el cual se moverá, con una alta probabilidad, el beneficio que obtendrá en el plan de inversiones. ¿Cómo podemos calcular uno? En este apartado ofreceremos herramientas para poder construir un intervalo con esas propiedades, aunque antes hay que decir que este es muy simple y por ello no muy fiable. Para obtener intervalos más fiables habrá que estudiar inferencia estadística. Pero eso es otro tema diferente.

Hay veces en las que los posibles valores de una variable aleatoria discreta están muy esparcidos y lejos de la media. Es lógico pensar que si los valores están todos más o menos cercanos a la media, entonces una gran parte de los valores de la variable estarán en un entorno reducido del valor esperado. Gracias a esta nueva medida, la varianza, daremos un intervalo en el cual se moverá con una cierta seguridad el beneficio que obtendremos al llevar a cabo el plan de inversiones.

Para medir el grado de concentración de los valores de una variable en torno a su media, utilizaremos la varianza. La varianza es una medida que nos indica lo alejados que están los valores respecto de la media. Por tanto, será lógico pensar que utilizaremos expresiones como $(x_i - \mu)^2$, dado que este número nos indica la distancia existente entre el posible valor x_i y la media de la distribución μ. Luego sumaremos todas esas desviaciones y obtendremos una medida de la desviación total que sufren los valores de la variable.

También habrá que tener en cuenta que en la fórmula de la varianza tendremos que dar un valor ó peso a cada una de las probabilidades de los posibles valores de la variable, y que ese peso deberá ser proporcional a la distancia que separa a dicho valor de la media ó esperanza de la variable. Para aclararnos un poco veamos la definición de varianza.

Definición 3.7.1 *Dada una variable aleatoria discreta que toma valores en el conjunto* $x_1, ..., x_n$ *con probabilidades respectivas* $p_1, ..., p_n$*, se define la varianza de la variable aleatoria discreta* X*, y la denotamos por* σ^2*, como*

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 p_i.$$

Como puedes ver, al multiplicar cada uno de los cuadrados por las probabilidades p_i, estamos dando más valor en la fórmula a las distancias de los valores más probables de ser tomados por la variable X.

Hay otra manera de calcular la varianza de una distribución de probabilidad. Veamos como se obtiene a partir de la definición de varianza:

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 p_i,$$

por la definición anterior, si desarrollamos el cuadrado dentro del sumatorio, nos quedará:

$$\sum_{1=1}^{n} (x_i^2 + \mu^2 - 2x_i\mu)p_i = \sum_{i=1}^{n} x_i^2 p_i + \sum_{i=1}^{n} \mu^2 p_i - \sum_{i=1}^{n} 2x_i\mu p_i.$$

Si ahora separamos el sumatorio en tres, nos quedará

$$\sum_{i=1}^{n} x_i^2 p_i + \sum_{i=1}^{n} \mu p_i - \sum_{i=1}^{n} 2x_i\mu p_i.$$

Operando en el segundo y tercer sumatorio, podemos sacar la μ fuera, ya que es constante, y nos quedará

$$\sum_{i=1}^{n} x_i^2 p_i + \mu^2 \sum_{i=1}^{n} p_i - 2\mu \sum_{i=1}^{n} x_i p_i.$$

Como X es una variable aleatoria y los p_i son su distribución de probabilidad, tenemos que $\sum_{i=1}^{n} p_i = 1$ y, por la definición de esperanza, sabemos que $\sum_{i=1}^{n} x_i p_i = \mu$, por tanto, sustituyendo en la expresión anterior, tendremos que

$$\sum_{i=1}^{n} x_i^2 p_i + \mu^2 - 2\mu\mu = \sum_{i=1}^{n} x_i^2 p_i - \mu^2,$$

que es una fórmula más sencilla a la hora de hacer cálculos a mano.

Por tanto, podemos decir que

$$\sigma^2 = \sum_{i=1}^{n} x_i^2 p_i - \mu^2.$$

Entre las propiedades más importantes de la varianza, cabe destacar que, por su propia definición, es una medida positiva, es decir, $\sigma^2 \geq 0$ siempre.

Vamos a practicar calculando la varianza de la distribución de probabilidad del beneficio que obtendremos al invertir el dinero en el plan de inversiones descrito.

Recordemos que nuestra distribución de probabilidad era:

$$P(X = x_1) = 0'05, \quad P(X = x_2) = 0'1, \quad P(X = x_3) = 0'1, \quad P(X = x_4) = 0'15,$$

$$P(X = x_5) = 0'2, \quad P(X = x_6) = 0'25, \quad P(X = x_7) = 0'15,$$

donde $x_1 = 0$, $x_2 = 5$, $x_3 = 10$, $x_4 = 15$, $x_5 = 20$, $x_6 = 25$ y $x_7 = 30$.

Así pues, por la fórmula de la varianza tenemos que

$$\sigma^2 = \sum_{i=1}^{7} x_i^2 p_i - \mu^2 =$$

$$= 0 \cdot 0'05 + 25 \cdot 0'1 + 100 \cdot 0'1 + 225 \cdot 0'15 + 400 \cdot 0'25 + 625 \cdot 0'2 + 900 \cdot 0'15 - 333'625 = 73'1775.$$

Pero este valor no nos servirá para calcular el intervalo del que hablábamos antes, ya que es una medida que expresa las unidades de la media o esperanza al cuadrado, por lo que no se pueden restar. Para solucionar ese problema definimos otra nueva medida que será la raíz cuadrada positiva de la varianza y a la que llamaremos desviación típica.

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

En nuestro ejemplo tenemos que $\sigma = 8'55$.

Así pues, diremos que con una cierta probabilidad, mayor o menor según el caso, el valor que tome la variable aleatoria discreta X estará en el intervalo $(\mu - \sigma, \mu + \sigma)$.

En nuestro ejemplo ese intervalo es $(9'7, 26'8)$. Como ves es un intervalo bastante amplio y para nuestro ejemplo no muy útil, debido a que podíamos haber calculado un intervalo mucho más acertado a ojo y sin necesidad de hacer tantos cálculos.

Volvemos a recordar que este intervalo es sólo un primer contacto con lo que más adelante llamaremos intervalos de confianza, que serán mucho más complejos y a la vez más fiables y certeros.

También hay que citar, que al igual que en las variables estadísticas, la varianza de dos variables aleatorias no se puede comparar, ya que los valores que tomen las dos variables no tienen por qué estar expresadas en las mismas unidades. Una sencilla manera de comparar las dos variables en ese sentido, es mediante el coeficiente de variación, que se define como $CV = \frac{\sigma}{\mu},\ \mu \neq 0$, donde $\sigma = +\sqrt{\sigma^2}$ (la desviación típica) y μ es la media de la variable aleatoria. Esta medida sí se puede comparar ya que es adimensional.

3.8. Resumen del problema inicial

Simplemente recordaremos los resultados. Después del estudio realizado concluimos que:

- Si nos tenemos que basar en el valor esperado del plan de inversiones, nos decantaríamos por éste en lugar de la inversión a plazo fijo.
- Cuando nos pedían que consiguiéramos con una seguridad del 70 % un beneficio superior al 16 % en el plan de inversiones, no la podíamos dar, por lo que aconsejamos al inversor que metiera su dinero a plazo fijo.
- En cuanto al intervalo en el que se moverá el beneficio con una cierta seguridad, dimos uno muy sencillo pero que nos dará ideas para, en el futuro, construir mejores y más precisos intervalos.

Capítulo 4

Un ejemplo de distribución aleatoria discreta: la binomial

4.1. Objetivos

- Conocer los experimentos aleatorios con dos únicos posibles resultados: experimentos de Bernoulli.
- Calcular la función de densidad, función de distribución, esperanza y varianza de una variable aleatoria que se distribuya según una Bernoulli.
- Manejar la distribución Binomial y ser capaz de calcular sus características a partir de las de la Bernoulli (función de densidad, función de distribución, esperanza y varianza).
- Distinguir los fenómenos de azar que se rigen según una distribución binomial y modelarlos para que se ajusten a los parámetros de dicha distribución.

4.2. El ejemplo

La mayoría de los humanos, tienen la mano derecha mejor desarrollada para las actividades que requieren una cierta sensibilidad, como puede ser escribir, comer, etc. Son las llamadas personas diestras. Sin embargo hay otra parte de la humanidad que tiene mayor sensibilidad, desde su nacimiento, en la mano izquierda, son los zurdos. Esas personas utilizan la mano izquierda para desempeñar las actividades descritas antes (escribir, comer, etc.). Al igual que pasa con las manos, los zurdos, como regla general, manejan mejor su pierna izquierda, cosa que se puede observar en un deporte como el fútbol, en el que hay dos tipos de jugadores muy diferenciados: los zurdos y los diestros.

A pesar de que muchos millones de personas en el mundo son zurdas, siguen teniendo problemas aún no resueltos en el uso de ciertos artilugios que están pensados para el uso de personas diestras, como son cierto tipo de abrelatas, las tijeras, los bolígrafos de secado no rápido,... y otro que es el que vamos a tratar: hay muchos institutos que utilizan sillas de pala para que sus alumnos den clases o realicen exámenes, cosa que es bastante complicada para un alumno zurdo.

Por esa razón, se ha planteado el poner un cierto número de sillas de pala en cada aula, para que los alumnos zurdos puedan utilizarlas sin las dificultades que les ocasionarían las pensadas para los diestros. Así, a partir de los datos de los cuestionarios pasados en nuestra clase, vamos a dar respuesta a las siguientes preguntas:

1. ¿Cuántas sillas, como mínimo, de cada tipo habría que poner en una clase de 30 alumnos para que el número esperado de zurdos tenga a su disposición una silla y el número de diestros tenga la suya?
2. ¿Cuántas sillas, como mínimo, de zurdos habría que poner para que con probabilidad 0'9, es decir, en el 90 % de los casos, no hubiese ningún zurdo sin su correspondiente silla?
3. ¿Cuál es el porcentaje de clases de 30 alumnos en las que hay al menos 10 zurdos?

De momento, no respondas a esas preguntas, ya que serás capaz de hacerlo con exactitud al final del tema.

4.3. Introducción

Si no se ha podido rellenar la encuesta, supondremos que el 10 % de la población es zurda.

Para resolver este problema vamos a desarrollar los contenidos poco a poco, paso a paso. La primera pregunta que nos puede surgir es: ¿cuál es la probabilidad de que un alumno elegido al azar sea zurdo? Claramente podemos contestar que esa probabilidad es 0'1, en el caso de que supongamos que el 10 % de la población es zurda.

Si planteamos el experimento aleatorio "elegir un alumno de un instituto al azar y observar si es zurdo o no" podemos profundizar más en él, debido a que está muy bien estudiado.

Tras este tipo de experimentos sólo hay lugar para dos posibles resultados: ser zurdo o no.

A esos dos posibles resultados en general les llamaremos éxito (E) y fracaso (F). En nuestro ejemplo consideraremos éxito al hecho de que un alumno sea zurdo, y fracaso que sea diestro. Es decir:

- $E =$ "El alumno elegido es zurdo".
- $F =$ "El alumno elegido no es zurdo".

Como regla general, denotaremos por $p = P(E)$ y $q = P(F)$. Obviamente se tiene que $p + q = 1$, por eso a veces a q se le llamará $1 - p$.

A este tipo de experimentos con dos únicos resultados, éxito y fracaso, los llamaremos experimentos de Bernoulli y están perfectamente determinados por la probabilidad de éxito p. Un experimento Bernoulli con $P(E) = p$, será denotado por $\mathcal{B}\rceil(p)$.

Nuestro ejemplo de elegir a un niño al azar y ver si es zurdo o no será un experimento $\mathcal{B}\rceil(0'1)$. Sin embargo, el estudio de este tipo de experimentos no nos servirá para responder a las preguntas de nuestro ejemplo, aunque sí la repetición de ellos, es decir, si nos encontramos con un aula de 50 alumnos y queremos saber cuantos de ellos serán zurdos, podemos repetir el experimento Bernoulli 50 veces. Vamos mirando uno a uno a los alumnos y vemos cuántos de ellos son zurdos. Al final del recuento, veremos cuántos alumnos zurdos hay. A la variable aleatoria que nos da el número de alumnos zurdos la llamaremos Binomial. Después de analizar este problema podremos responder a las preguntas que se plantean en el ejemplo.

Veamos en general qué entendemos por una distribución binomial. Si tenemos una situación en la que:

1. Se realizan n repeticiones de un mismo experimento siempre en idénticas condiciones, y en cada uno de ellos solo hay dos posibles resultados, éxito (que denotaremos por E) y fracaso, (que denotaremos por F). Además esos sucesos son complementarios, es decir, $P(E \cup F) = 1$ y $P(E \cap F) = 0$ (siempre o se da uno o el otro, y nunca se pueden dar los dos a la misma vez).

2. La probabilidad de éxito $P(E)$ es igual en cada ensayo, llamémosla p. Por tanto, la probabilidad de fracaso también es la misma en cada ensayo, llamémosla q. Es decir, tenemos que en cada repetición del experimento se cumple:

$$P(E) = p \ , \qquad P(F) = 1 - P(E) = 1 - p = q.$$

3. Si denotamos por $X =$ "n^{o} de éxitos en n repeticiones", se tiene que X puede tomar los valores 1,...,n y que sigue una distribución binomial.

A la distribución de probabilidad que cumple estas condiciones, la llamamos distribución binomial de parámetro $p = P(E)$ con n repeticiones, y la denotaremos por $\mathcal{B}(n, p)$.

Así pues, si nuestro problema va a ser analizar el número de alumnos zurdos en un aula de 50, tendremos que plantear una variable aleatoria así:

$X =$ "n^{o} de alumnos zurdos en un aula de 50",

que como podemos suponer se distribuye según una binomial de parámetro $p = 0'1$ y $n = 50$, es decir, $X \sim \mathcal{B}(50, 0'1)$.

Probabilidad de k éxitos

Nos puede surgir una pregunta de un modo muy intuitivo: ¿cuál es la probabilidad de que en el aula haya 2 alumnos zurdos? Es decir, para que se de ese suceso tiene que haber 2 zurdos y 48 diestros. En otras palabras, se puede decir que el primer alumno ha de ser zurdo, el 2^{o} también, el 3^{o} diestro, el cuarto diestro,..., el alumno número 50 diestro. Si denotamos por Z a un alumno zurdo y por D a uno diestro, podemos poner esta combinación:

$$\underline{Z}\ \underline{Z}\ \underbrace{\underline{D}\ \underline{D}\ ...\underline{D}\ \underline{D}}_{48 veces}.$$

Cómo la probabilidad de que un alumno sea zurdo es 0'1 y de que sea diestro es 0'9 (y son las mismas en cada observación) tenemos que la probabilidad de que se de esa combinación de alumnos es:

$$\underline{0'1}\ \underline{0'1}\ \underbrace{\underline{0'9}\ \underline{0'9}\ ...\underline{0'9}\ \underline{0'9}}_{48 veces} = (0'1)^2(0'9)^{48}.$$

Pero ahora bien, el orden en el que están los alumnos no tienen por que ser ese, es decir, los alumnos zurdos pueden ser el número 30 y el 43, o el 12 y el 17, etc. Así pues, tendremos que multiplicar la cantidad obtenida por el número total de posibles ordenaciones. ¿Cuántas hay? En total hay $\binom{50}{2} = \frac{50!}{2!(50-2)!} = 1225$. Ese número es especial, se llama número combinatorio. En general, si queremos ordenar un conjunto de n elementos en el que hay k de un tipo y $n-k$ de otro, tenemos $\binom{n}{k}$ formas de hacerlo.

Así pues, la probabilidad de que haya dos alumnos zurdos en el aula de 50, es lo mismo que decir que haya dos éxitos en nuestro experimento $\mathcal{B}(50, 0'1)$, es:

$$P(X = 2) = \binom{50}{2} \cdot (0'1)^2 (0'9)^{48} = 0'08.$$

Como puedes ver, la probabilidad es muy pequeña. En cierto modo es lógico, pues si lo piensas bien no es muy probable que haya exactamente 2 alumnos zurdos entre 50.

Veamos en general cual es la probabilidad de que se obtengan k éxitos en las n repeticiones de un experimento Bernoulli de parámetro p, es decir, en una $\mathcal{B}(n, p)$.

Si X sigue un distribución binomial $\mathcal{B}(n, p)$, tenemos que la probabilidad de obtener k éxitos en las n repeticiones viene dada por la expresión

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad \forall\ k = 0, 1, 2, ..., n,$$

donde

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}\ , \quad p = P(E)\ , \quad q = P(F), \quad p + q = 1.$$

El razonamiento para obtener esa expresión es el mismo que seguimos en el caso particular anterior de obtener 2 éxitos para nuestra variable binomial X.

Vamos a ver como casos particulares la probabilidad de que se den 0 éxitos y de que se den n éxitos. Comencemos calculando la probabilidad de que se den 0 éxitos, y por tanto n fracasos.

$$P(X = 0) = \binom{n}{0} p^0 q^{n-0} = \frac{n!}{0!n!} q^n = q^n,$$

(recordar que $0! = 1$), mientras que la probabilidad de obtener n éxitos es:

$$P(X = n) = \binom{n}{n} p^n q^{n-n} = \frac{n!}{n!0!} p^n q^0 = p^n.$$

Función de probabilidad

La función de probabilidad vendrá dada de la siguiente forma:

$$f(k) = \binom{n}{k} p^k q^{n-k}, \quad \forall\ k = 0, 1, 2, ..., n \quad \text{y 0 en el resto},$$

lo que se puede expresar de la siguiente manera:

$$f(k) = \begin{cases} \binom{n}{k} p^k q^{n-k} & \text{si } k \in \{0, 1, 2, ..., n\}, \\ 0 & \text{en caso contrario.} \end{cases}$$

Por tanto, en el ejemplo que estamos estudiando, tendremos que

$$f(k) = \begin{cases} \binom{50}{k}(0'1)^k(0'9)^{50-k} & \text{si } k \in \{0,1,2,...,50\}, \\ 0 & \text{en caso contrario.} \end{cases}$$

Función de Distribución

En la segunda pregunta nos piden calcular el número de sillas necesario para que con probabilidad $0'9$ o superior no haya zurdos sin su correspondiente silla, es decir, buscamos el primer $k \in \mathbb{Z}$ tal que $P(X \leq k) \geq 0'9$. ¿Cómo podemos calcular ese k? Pues con la ayuda de la función de distribución para esta variable aleatoria, que pasamos a calcular a continuación, ya que $P(X \leq k) = F(k)$ donde F es la FDD de la variable aleatoria X.

Para una variable aleatoria discreta X, por definición de función de distribución, tenemos que $F(x) = P(X \leq x)$, por tanto en el caso $X \sim \mathcal{B}(n,p)$ tendremos que

$$F(x) = \sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k} = \binom{n}{0} p^0 q^n + \binom{n}{1} p q^{n-1} + ... + \binom{n}{x} p^x q^{n-x}, \quad \forall x = 0,1,2,...,n.$$

Como casos particulares, podemos ver que

$$F(0) = \sum_{k=0}^{0} \binom{n}{k} p^k q^{n-k} = f(0) = q^n,$$

y aplicando la fórmula del binomio de Newton, podemos calcular

$$F(n) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = \binom{n}{0} q^n + \binom{n}{1} p q^{n-1} + ... + \binom{n}{n} p^n = (p+q)^n = 1^n = 1.$$

Algo que es lógico pensar, ya que en n repeticiones seguro que habrá un número de éxitos menor o igual que n.

En nuestro ejemplo, tendremos que la Función de Distribución es

$$F(x) = \sum_{k=0}^{x} \binom{50}{k} (0'1)^k (0'9)^{50-k}, \ \forall x \in \{0,1,...,50\}.$$

Es decir, para responder a la 2ª pregunta buscamos el primer k tal que $F(k) \geq 0'9$. Este se encuentra probando con diferentes valores para k y el cálculo es muy tedioso.

Tenemos que

$$F(k) = \sum_{i=0}^{x} \binom{50}{i} (0'1)^i (0'9)^{50-i}.$$

Probamos primero con $k = 10$:

$$F(10) = \sum_{i=0}^{10} \binom{50}{i} (0'1)^i (0'9)^{50-i} = 0'99.$$

Ésta es mayor que $0'9$, probemos con uno menor, por ejemplo $k = 7$

$$F(7) = \sum_{i=0}^{7} \binom{50}{i} (0'1)^i (0'9)^{50-i} = 0'88.$$

Nos falta un poco, vamos a probar con $k = 8$:

$$F(8) = \sum_{i=0}^{8} \binom{50}{i} (0'1)^i (0'9)^{50-i} = 0'94.$$

Como $F(7) < 0'9$, $F(8) > 0'9$ y F es creciente, tenemos que el k buscado es $k = 8$, el primero tal que $F(k) \geq 0'9$.

Así pues, para asegurarnos con una fiabilidad del 90 % que no haya zurdos sin su correspondiente silla en el aula, tendremos que poner 8 sillas para zurdos.

¿Cuántas sillas para diestros habrá que poner cómo mínimo para que con una fiabilidad del 90 % no haya diestros sin su correspondiente silla?

En este caso, tendremos que buscar el primer k en orden decreciente tal que $P(50-X \leq k) \geq 0'9$, ya que el número de diestros viene representado por la variable aleatoria $50 - X$. Vamos a poner un poco más claro cómo tenemos que calcular el k:

$$P(50 - X \leq k) \geq 0'9 \Leftrightarrow P(X \geq 50 - k) \geq 0'9 \Leftrightarrow 1 - P(X < 50 - k) \geq 0'9 \Leftrightarrow P(X < 50 - k) \leq 0'1$$

$$\Leftrightarrow P(X \leq 49 - k) \leq 0'1 \Leftrightarrow F(49 - k) = 0'1$$

Comencemos a probar:

- $k = 48, \quad F(49 - k) = F(1) = 0'03 \leq 0'1.$
- $k = 47, \quad F(49 - k) = F(2) = 0'11 \not\leq 0'1.$

Así pues, el k buscado es 47. Por tanto, para que con probabilidad $0'9$ ningún diestro se quede sin silla habrá que poner cómo mínimo 47 sillas de ese tipo.

4.3.1. La Esperanza

Ahora vamos a responder a la misma pregunta pero basándonos en el valor esperado, es decir, diremos cuántas sillas de cada tipo hay que poner para que el valor esperado de zurdos y diestros tengan una silla apropiada.

Recordemos que la esperanza de una variable aleatoria era el promedio que tiene esa variable tras un número grande de repeticiones.

Como ya vimos anteriormente, la esperanza de una variable aleatoria X que toma sus valores en un conjunto $\{x_1, ..., x_n\}$ con probabilidades respectivas $\{p_1, ..., p_n\}$, (es decir, $P(X = x_i) = p_i$, $\forall i \in \{1, 2, ..., n\}$) , a la que llamaremos μ, se calcula mediante la expresión

$$\mu = \sum_{i=n}^{n} x_i p_i.$$

Por tanto, si $X \sim \mathcal{B}(n,p)$, tendremos que su esperanza es

$$\mu = \sum_{i=1}^{n} ip(X=i) = \sum_{i=0}^{n} i\binom{n}{i}p^i q^{n-i},$$

y haciendo cálculos matemáticos, se obtiene que

$$\mu = np.$$

Así pues, la esperanza de nuestra variable aleatoria X = "nº de zurdos en el aula" es

$$E(X) = \mu = np = 50 \cdot 0'1 = 5.$$

Según el valor esperado de la variable aleatoria X tendremos 5 zurdos y 45 diestros. Por supuesto será muy arriesgado poner sólo 5 sillas de zurdos y 45 de diestros, pues no debe ser muy probable que se de exactamente ese suceso. Concretamente la probabilidad de que se de es

$$P(X=5) = \binom{50}{5}(0'1)^5(0'9)^{45} = 0'18,$$

que aunque es un suceso con probabilidad alta en comparación con el resto de los valores que puede tomar la variable X, no da mucha garantía. Sin embargo, hay que decir que la esperanza nos puede dar una idea de lo que pasará, aunque siempre es mejor dar un intervalo en donde se moverá el valor de X con una cierta fiabilidad, y eso es lo que vamos a hacer en la siguiente sección.

4.3.2. La Varianza

Ahora vamos a dar un intervalo en el cual se moverá el número de zurdos entre los 50 alumnos en función de la media y la desviación típica. Veamos la expresión de la varianza de una distribución binomial:

En general su fórmula era

$$\sigma^2 = \sum_{i=1}^{n} x_i^2 p_i - \mu^2.$$

Por tanto, si $X \sim \mathcal{B}(n,p)$ tendremos que

$$\sigma^2 = \sum_{i=0}^{n} i^2\binom{n}{i}p^i q^{n-i} - \mu^2,$$

y resolviendo ese sumatorio, resulta que

$$\sigma^2 = npq.$$

Así, en nuestro ejemplo, tendremos que

$$\sigma^2 = 50 \cdot 0'1 \cdot 0'9 = 4'5.$$

Por tanto, tenemos que

$$\sigma = +\sqrt{\sigma^2} = 2'12.$$

Entonces podremos decir que con una gran probabilidad el número de zurdos estará en el intervalo

$$(\mu-\sigma,\mu+\sigma)=(2'9,7'1),$$

y como X es una variable aleatoria discreta, podremos decir con una cierta seguridad que $X\in\{3,4,5,6,7\}$, ya que esos son los posibles valores que caen dentro del intervalo. En concreto esa probabilidad es:

$$\sum_{i=3}^{7} P(X=i)=...=0'77,$$

es decir, que en el 77 % de los casos habrá 3, 4, 5, 6 ó 7 zurdos en el aula.

Para acabar el capítulo vamos a responder a la 4ª pregunta, cuál es el porcentaje de aulas con 50 alumnos en los que hay 10 zurdos o más.

Tenemos que calcular

$$P(X\geq 10)=\sum_{i=10}^{50} P(X=i)=...=0'025.$$

Es decir, tenemos que en el $2'5$ % de las aulas de 50 alumnos habrá 10 zurdos o más.

Nos pueden surgir muchas más preguntas para este modelo. Todas ellas se pueden resolver, pero con la ayuda de un ordenador y el software necesario, ya que cálculos como $P(X\leq k)$ o $P(X\geq k)$ son muy complicados, por el gran número de cuentas a realizar, de calcular a mano. Además, a medida que va aumentando el número de repeticiones del experimento Bernoulli n el cálculo se va haciendo más y más largo, casi imposible.

Por eso, en el siguiente capítulo daremos una técnica para hacer esos cálculos más sencillos, pero para comprenderla bien hará falta manejar distribución normal, que es precisamente lo que vamos a ver en el siguiente capítulo.

Capítulo 5

Distribuciones continuas: la normal

5.1. Objetivos

- Entender el concepto de variable continua y su diferencia principal con las discretas.
- Comprender el significado de las funciones de densidad y de distribución y conocer su relación.
- Saber calcular probabilidades en intervalos de distribuciones aleatorias continuas a partir de la función de densidad, ya sea gráficamente o mediante el cálculo integral.
- Calcular probabilidades a partir de la función de distribución.
- Comprender el por qué de la importancia de la distribución Normal, conocer su función de densidad e interpretar sus parámetros (μ y σ).
- Ser capaz de ajustar unos datos a una distribución normal y comprobar si el ajuste es bueno o no.
- Tipificar una normal de media μ y desviación típica σ cualesquiera.
- Manejar la tabla de probabilidad de la Normal Estándar.
- Aproximar una distribución binomial por una normal.

5.2. El ejemplo

Dado que la altura media de la población está cambiando, en algunos institutos se están quedando pequeñas las sillas y las mesas para muchos estudiantes, pues están pensadas para alumnos de una cierta estatura. Se sabe que los alumnos que midan menos de 160 cm se sienten cómodos con las mesas y sillas del tipo A, los que tiene una estatura entre 160 y 180 cm se encuentran cómodos para estudiar en sillas y mesas del tipo B y aquellos que miden más de 180 cm, necesitan sillas y mesas del tipo C. Un director de un instituto cualquiera, quiere saber cuántas mesas de cada tipo tiene que poner de cada tipo en una clase de 30 alumnos con 30 mesas, para que la mayor parte de los alumnos se sienta cómodo en su sitio.

Dado que es un fenómeno muy estudiado, se sabe que la altura de una población se rige según una variable aleatoria continua un poco especial, la distribución normal, con media la altura media de la población y varianza la que corresponda a los datos de la muestra. La distribución normal más y mejor estudiada es la distribución normal estándar, aquella que tiene media 0 y desviación típica 1, de la cuál se disponen tablas en las que vienen tabuladas sus diferentes probabilidades.

Así pues, poco a poco podremos ir respondiendo a preguntas como:

1. ¿Cuál es la altura media según los datos de los cuestionarios? ¿Y la varianza?
2. ¿Cuál es el número de mesas de cada tipo que necesitaremos en una clase por término medio?
3. ¿Crees que la estimación que hemos realizado es acertada? ¿Por qué?

El problema de los zurdos visto para la distribución binomial, da respuesta a todas las preguntas planteadas aproximando previamente dicha distribución mediante una normal.

5.3. Introducción

Vamos a medir la altura de los alumnos de nuestra clase. Si tuviéramos un aparato para medir la altura con una precisión hasta decímetros, tendríamos datos, expresados en metros, tales como

$1'7, 1'7, 1'6, 1'5, 1'7, 1'4, 1'5, 1'8, ...$

Si por el contrario, tuviéramos una precisión hasta centímetros, tendríamos datos como estos:

$1'71, 1'75, 1'66, 1'54, 1'69, 1'48, 1'55, ...$

Imagínate la precisión que podríamos alcanzar con un aparato que midiera hasta los milímetros. Tendríamos alturas expresadas en metros de este tipo:

$1'712, 1'748, 1'663, 1'541, 1'689, 1'484, 1'552, ...$

Si pudiéramos medir con una precisión hasta 5 decimales lo haríamos, pero en ese caso, sería inútil preguntarse si alguien mide $1'56053$ m, sería más lógico preguntarse si mide más de $155'5$ centímetros y menos de $156'5$ cm y después redondearíamos y diríamos que mide 156 cm ó $1'56$ m.

Como ves, dependiendo de la precisión con la que midamos, puede que no tenga sentido ajustar esos datos a una distribución de probabilidad discreta, pues aunque el conjunto de valores sería finito, también es cierto que éste sería muy grande, excesivamente grande para que los cálculos se pudieran realizar con facilidad.

Como ya vimos antes, una distribución de probabilidad es un modelo matemático que nos ayuda a explicar un fenómeno real y a intentar predecirlo en un futuro. En el apartado de distribuciones discretas, vimos que estaban asociadas a variables aleatorias que tomaban un conjunto de valores finito. Si por el contrario estos valores pudieran ser cualquier número dentro de un intervalo de la recta real, diremos que se trata de una variable aleatoria continua.

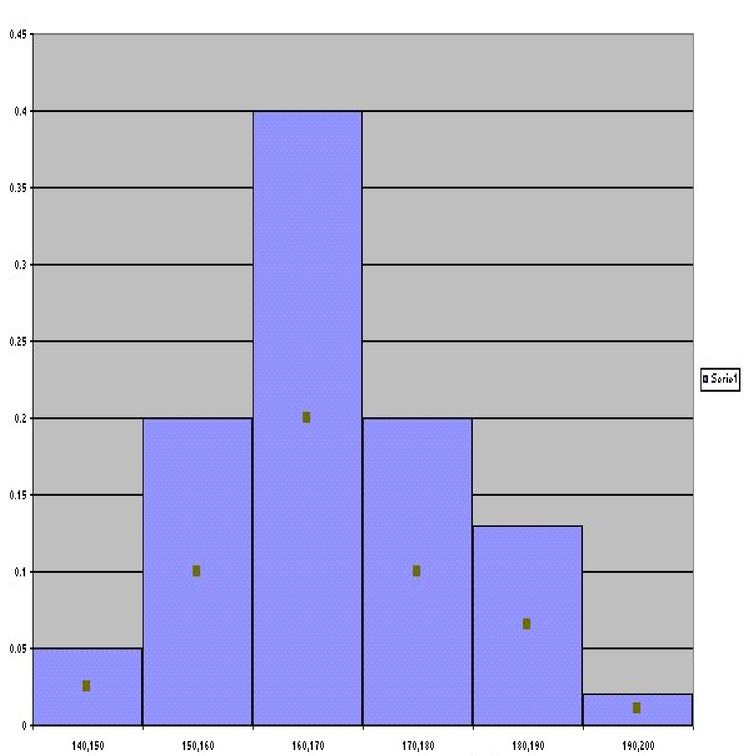
En el ejemplo que hemos visto de las alturas, podemos decir que se trata de una variable continua, pues, aunque realmente nunca conseguiremos que lo sea, si nos será más fácil de tratar de esta manera que intentarlo como si fuera una variable discreta.

Vamos a intentar explicar el concepto de variable continua a través del siguiente ejemplo:

supón que una vez recogidos los datos de la altura de todos los alumnos del instituto, y agrupándolos en grupos de 10 cm debido a que nuestro aparato solo capta decímetros, obtenemos la siguiente tabla de frecuencias relativas:

Intervalo de altura	Frecuencia relativa
[140,150)	0.05
[150,160)	0.2
[160,170)	0.4
[170,180)	0.2
[180,190)	0.13
[190,200)	0.02

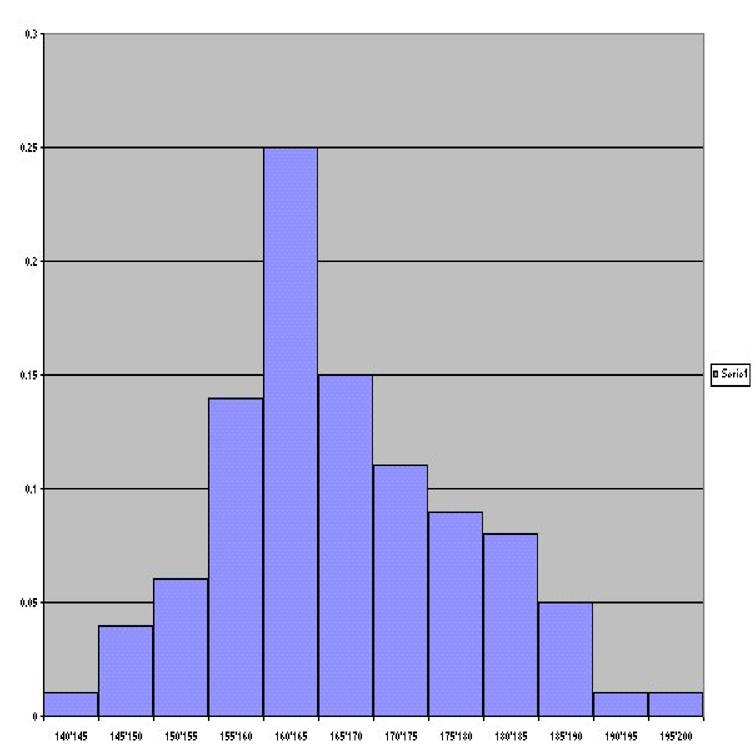
Y si representamos su histograma, obtenemos



Sin embargo, plantéate ahora, que dividimos las alturas en grupos de 5 cm, es decir, afinamos un poco más la división. Obtenemos entonces la siguiente tabla:

Intervalo de altura	Frecuencia relativa
[140,145)	0.01
[145,150)	0.04
[150,155)	0.06
[155,160)	0.14
[160,165)	0.25
[165,170)	0.15
[170,175)	0.11
[175,180)	0.09
[180,185)	0.08
[185,190)	0.05
[190,195)	0.01
[195,200)	0.01

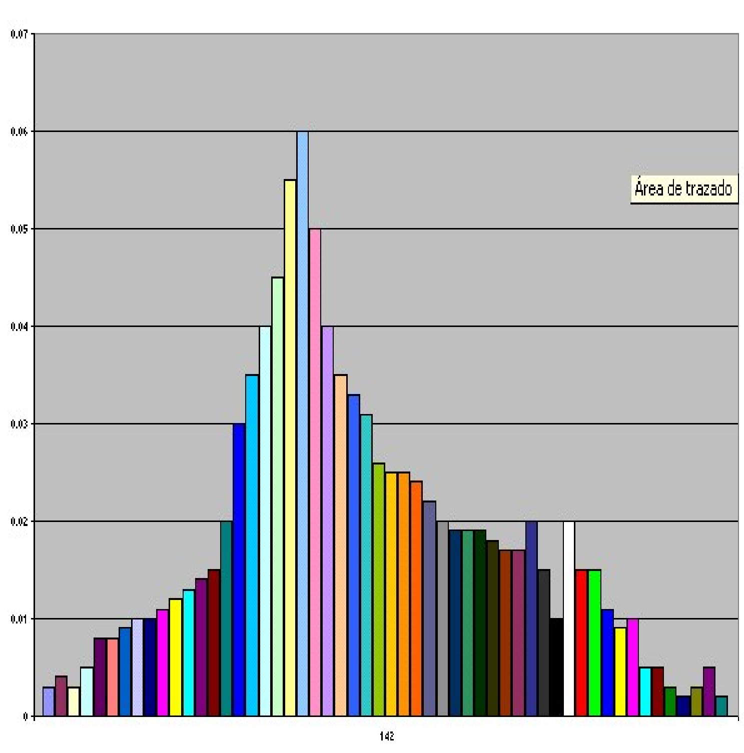
A partir de esta tabla de frecuencias, obtendríamos el siguiente histograma



Como puedes observar, es sensiblemente diferente al que obtuvimos haciendo grupos de 10 cm.
Pero, podemos afinar aún más y obtener las alturas de los alumnos expresadas en centímetros. En nuestro ejemplo obtendríamos esta tabla:

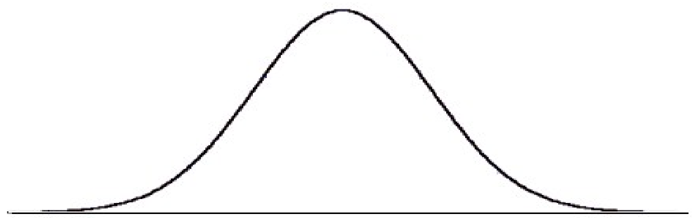
Altura	Frecuencia relativa	Altura	Frecuencia relativa
142	0.003	169	0.025
143	0.004	170	0.025
144	0.003	171	0.024
145	0.005	172	0.022
146	0.008	173	0.02
147	0.008	174	0.019
148	0.009	175	0.019
149	0.01	176	0.019
150	0.01	177	0.018
151	0.011	178	0.017
152	0.012	179	0.017
153	0.013	180	0.02
154	0.014	181	0.015
155	0.015	182	0.01
156	0.02	183	0.02
157	0.03	184	0.015
158	0.035	185	0.015
159	0.04	186	0.011
160	0.045	187	0.009
161	0.055	188	0.01
162	0.06	189	0.005
163	0.05	190	0.005
164	0.04	192	0.003
165	0.035	194	0.002
166	0.033	195	0.003
167	0.031	196	0.005
168	0.026	198	0.002

Con el siguiente histograma:

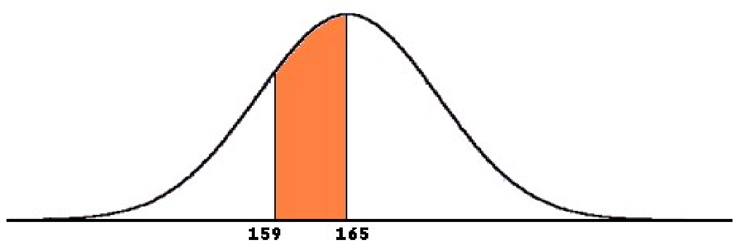


Si afináramos aún más, obtendríamos histogramas con escalones cada vez más pequeños, hasta acabar, en la idealización de la continuidad, es decir, se describiría la gráfica de una función continua, en adelante la llamaremos f. Debes darte cuenta que para ello necesitaríamos un tamaño de muestra muy grande, mucho mayor que el número de alumnos de tu instituto.

Si hiciéramos todas esas observaciones podríamos obtener una función f con una gráfica así:



Si hubiéramos hecho todas las aproximaciones necesarias hasta obtener esa función, podríamos calcular la probabilidad de que el peso de un alumno elegido al azar estuviese, por ejemplo, entre 159 y 165 cm, calculando el área de la superficie encerrada entre la función f y el eje OX, entre el 159 y el 165 en el eje de las x.



En resumen, tendremos que una variable aleatoria X es continua si cumple que:

- Tiene una distribución de probabilidad asociada que es continua y viene definida a partir de una función $f(x)$, y gracias a la cual podemos calcular probabilidades del tipo $P(x_1 \leq X \leq x_2)$, con $x_1, x_2 \in \mathbb{R},\ x_1 \leq x_2$.
- Dicha probabilidad se calculará midiendo el área encerrada entre la gráfica de la función $f(x)$ y el eje de abscisas, entre los puntos x_1 y x_2.
- En una distribución de probabilidad continua, no tiene sentido hablar de probabilidades en puntos, solo en intervalos. Esto es porque si X es una variable aleatoria continua, entonces $P(X = x) = 0,\ \forall x \in \mathbb{R}$ (en el ejemplo de antes, no tiene sentido preguntarse si alguien pesa $63'94738274658482736$ kg).

Ejemplo 5.3.1 *Consideremos la distribución de probabilidad asociada a la función $f(x)$ de la figura anterior:*

1. *Representa gráficamente las probabilidades*

$$P(60 \leq X \leq 72), \quad P(70 \geq X), \quad P(X \leq 60) y \quad P(X = 81).$$

2. *Utilizando la tabla anterior, da una aproximación numérica de sus probabilidades.*

Ejercicio 5.3.1 *Haz lo mismo para*

$$P(50 \leq X \leq 60), \quad P(61 \geq X), \quad P(X \leq 83) y \quad P(X = 76).$$

Función de densidad

La función que nos permitía calcular las probabilidades en los intervalos, a la que llamábamos f, se llama *función de densidad* de la variable aleatoria continua X.

Con la función de densidad podemos calcular todos los demás parámetros de la distribución. Es el equivalente, en las variables aleatorias continuas, de la función de probabilidad en las variables aleatorias discretas.

La función de densidad de la variable aleatoria f de una variable aleatoria X, tiene que cumplir:

1. $f(x) \geq 0, \ \forall x \in \mathbb{R}$.

2. El área total comprendida entre la gráfica de la función f y el eje OX, es 1, es decir

$$P(-\infty \leq X \leq +\infty) = 1.$$

   Algo, por otra parte, totalmente lógico. En términos de cálculo integral, se puede describir como

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

   En particular, si X sólo toma valores en el segmento delimitado por a y b, tenemos que

$$P(a \leq X \leq b) = S(a,b) = \int_a^b f(x)dx = 1,$$

   donde $S(a,b)$ denota el área de la superficie encerrada entre la gráfica de la curva $f(x)$ y el eje de abscisas, entre los puntos a y b.

3. La probabilidad de que la variable aleatoria continua X tome sus valores en el intervalo delimitado por x_1 y x_2, es decir $P(x_1 \leq X \leq x_2)$, es el área comprendida entre la función $f(x)$ y el eje OX en el intervalo x_1, x_2, es decir $S(x_1, x_2)$, para cualesquiera $x_1 \leq x_2, \ x_1, x_2 \in \mathbb{R}$. En otras palabras,

$$P(x_1 \leq X \leq x_2) = S(x_1, x_2) = \int_{x1}^{x_2} f(x)dx.$$

   Además, como consecuencia de que $P(X = x) = 0, \ \forall x \in X$, se tiene que

$$P(x_1 \leq X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X \leq x_2) = P(x_1 < X < x_2), \forall x_1 \leq x_2, \ x_1, x_2 \in \mathbb{R}.$$

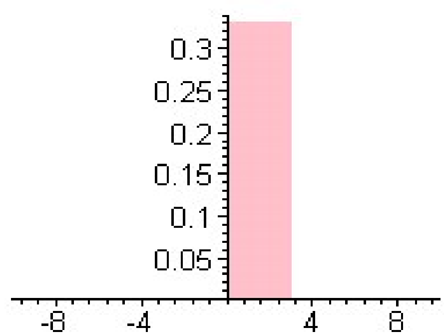
También se tienen las propiedades inversas, es decir, si tenemos una función f cumpliendo que $f(x) \geq 0$, tal que el área total entre la gráfica y el eje OX es 1, $\int_{-\infty}^{+\infty} f(x)dx = 1$, podemos definir una variable aleatoria continua asociada a ella, simplemente asignándole como función de probabilidad en cada intervalo, el área encerrada entre $f(x)$ y el eje de abscisas en dicho intervalo.

Ejemplo 5.3.2 *Veamos si la siguiente función es una función de densidad:*

$$f(x) = \begin{cases} \frac{1}{3} & \textit{si } 0 < x < 3, \\ 0 & \textit{en caso contrario.} \end{cases}$$

f tiene que cumplir:

- $f(x) \geq 0 \ \forall x$, *lo cumple claramente.*
- $\int_{-\infty}^{+\infty} f(x)dx = 1$ *Veámoslo.*
  $\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{3} f(x)dx + \int_{3}^{+\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{3} \frac{1}{3}dx + \int_{3}^{+\infty} 0dx = 0 + [\frac{x}{3}]_{x=0}^{x=3} + 0 = \frac{3}{3} - \frac{0}{3} = 1$. *Veamos esta segunda condición geométricamente.*

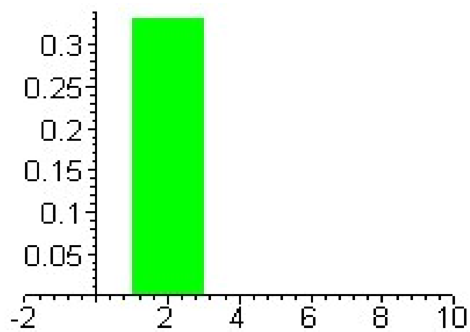


Cumple las dos condiciones, por tanto podemos asociar una variable aleatoria continua, llamémosla X, a la función de densidad f.

Vamos a utilizar la función de densidad (en adelante la denotaremos por fdd) para calcular algunas probabilidades.

$P(X \geq 1) = \int_{1}^{+\infty} f(x)dx = \int_{1}^{3} \frac{1}{3}dx = [\frac{x}{3}]_{1}^{3} = \frac{3}{3} - \frac{1}{3} = \frac{2}{3}$.

También se puede calcular mediante el área de la superficie coloreada de



Ejercicio 5.3.2 *Calcula las siguientes probabilidades, tomando como variable X la asociada a la fdd del ejemplo anterior:*

$$P(X \geq 2),\ P(0'5 \leq X < 2), P(2 < X < 4).$$

Función de Distribución

La definición de una Función de Distribución para una variable aleatoria continua, es la misma que dimos para las variables aleatorias discretas, es decir

$$F(x) = P(X \leq x),\ \forall x \in \mathbb{R}.$$

Esta función mide la probabilidad de que la variable aleatoria tome un valor menor o igual que x. En el caso discreto, lo calculábamos mediante un sumatorio finito, en el caso continuo no podremos hacer esto, así que lo haremos mediante el área encerrada o el cálculo integral, que a fin de cuentas es lo mismo.

Es decir, tenemos que

$$F(x) = S(a, x) = \int_a^x f(x)dx,$$

donde $f(x)$ es la fdd de la variable X.

La Función de Distribución (a la que a partir de ahora llamaremos FDD) tiene utilidad para el cálculo de probabilidades en intervalos, así tenemos que

$$P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1).$$

Una función F se puede considerar que es una FDD si cumple:

1. $0 \leq F(x) \leq 1,\ \forall x \in \mathbb{R}.$
2. Si $x \leq a,\ F(x) = 0$, si $x \geq b,\ F(x) = 1.$

3. $F(x)$ es monótona creciente, esto es, si $x_1 \leq x_2$ se tiene que $F(x_1) \leq F(x_2)$.

Además, tenemos que si F es la FDD de una variable aleatoria continua X y f es su fdd, entonces $F'(x) = f(x)$, y por tanto $\int_a^x f(t)dt = F(x)$.

Ejemplo 5.3.3 (Cálculo de la FDD a partir de la fdd) *Partamos de la fdd vista como ejemplo en el apartado anterior*

$$f(x) = \begin{cases} \frac{1}{3} & si\ 0 < x < 3, \\ 0 & en\ caso\ contrario. \end{cases}$$

Sabemos que $\int_a^x f(t)dt = F(x)$. *Así pues, tenemos que:*

1. *si* $x \leq 0 \ \Rightarrow F(x) = \int_{-\infty}^x 0dt = 0,$

2. *si* $x \in [0,3] \ \Rightarrow F(x) = \int_{-\infty}^0 0dt + \int_0^x \frac{1}{3}dt = \frac{x}{3},$

3. *si* $x \geq 3 \Rightarrow F(x) = \int_{-\infty}^0 0dt + \int_0^3 \frac{1}{3}dt + \int_3^x 0dt = 1.$

Es decir,

$$F(x) = \begin{cases} 0 & si\ x \leq 0, \\ \frac{x}{3} & si\ 0 < x < 3, \\ 1 & si\ x \geq 3. \end{cases}$$

Comprobemos que, en efecto, cumple las condiciones para ser Función de Distribución.

1. $0 \leq F(x) \leq 1$, *claramente lo cumple.*

2. *Si* $x \leq 0 \ \Rightarrow \ F(x) = 0.$
   Si $x \geq 3 \ \Rightarrow \ F(x) = 1.$

3. $F(x)$ *es monótona creciente. Hay que probar que si* $x_1 \leq x_2 \ \Rightarrow \ F(x_1) \leq F(x_2)$. *Veamos los diferentes casos que se pueden dar.*

   a) $x_1 < 0$ *y* $x_2 < 0$. $F(x_1) = 0 = F(x_2)$. *Se tiene.*
   b) $x_1 < 0$ *y* $x_2 \in [0,3]$. $F(x_1) = 0 \leq \frac{x_2}{3} = F(x_2)$. *Se cumple.*
   c) $x_1 < 0$ *y* $x_2 \geq 3$. $F(x_1) = 0 \leq 1 = F(x_2)$. *Se cumple.*
   d) $x_1 \in [0,3]$ *y* $x_2 \in [0,3]$. $F(x_1) = \frac{x_1}{3} \leq \frac{x_2}{3} = F(x_2)$. *Se cumple.*
   e) $x_1 \in [0,3]$ *y* $x_2 \geq 3$. $F(x_1) = \frac{x_1}{3} \leq 1 = F(x_2)$. *Se cumple.*
   f) $x_1 \geq 3$ *y* $x_2 \geq 3$. $F(x_1) = 1 \leq 1 = F(x_2)$. *Se cumple.*

Ejercicio 5.3.3 *Sea la función* F *definida como sigue*

$$F(x) = \begin{cases} 0 & si\ x \leq 0, \\ \frac{x^2}{2} & si\ 0 < x < 1, \\ 1 & si\ x \geq 1. \end{cases}$$

¿Es admisible como FDD? Si es así, calcula la fdd asociada a F *y las siguientes probabilidades:*

$$P(X \leq 0'5), \quad P(X \geq 0'8), \quad P(0'2 \leq X \leq 0'5), \quad P(0'3 \leq X \leq 1'5),$$

donde X *es la variable aleatoria continua asociada a* F.

Ejercicio 5.3.4 *Probar si la siguiente función es admisible de ser FDD o no:*

$$F(x) = \begin{cases} 0 & si\ x \leq -1, \\ \frac{x^2}{4} & si\ -1 < x < 2, \\ 1 & si\ x \geq 2. \end{cases}$$

Ejercicio 5.3.5 *Considera la siguiente función:*

$$f(x) = \begin{cases} x & si\ 2 < x < 3, \\ 0 & en\ caso\ contrario. \end{cases}$$

¿Puede ser f considerada como una fdd? ¿Por qué?

Un ejemplo: la distribución normal

Debido a su aplicabilidad práctica y teórica, la distribución de probabilidad continua más importante es, sin duda, la distribución normal. Esta aplicabilidad teórica se deriva de que en la mayoría de las situaciones, los casos raros no son *normales*, y casi todos los datos se agrupan en torno a un valor central.

Por ejemplo, lo *normal* es que un individuo varón tenga una estatura entre 170 y 180 cm (datos inventados), y será muy raro encontrar individuos que midan más de 200 cm o menos de 150 cm. Por esa razón, y por que es un ejemplo muy estudiado, veremos que la estatura de una población se distribuye según una distribución normal.

La altura es uno de los muchos ejemplos de fenómenos que se rigen por una distribución normal. Otros son el peso, el cociente intelectual, los errores en las medidas, etc.

Todos ellos tienen un denominador común: la mayoría de los datos se agrupan en torno a la media.

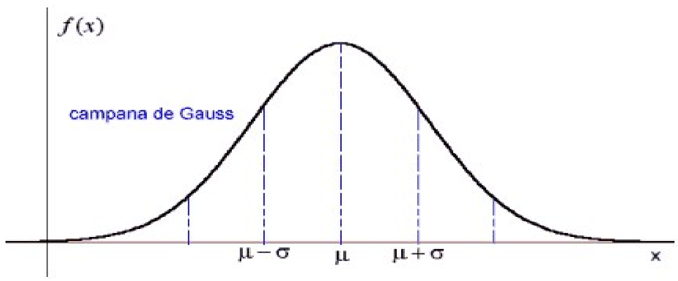
En el ejemplo de la altura de los alumnos del instituto, vamos a suponer que todos esos datos se distribuyen según una normal, de media la de los datos y desviación típica la que nos salga del cálculo en la tabla de las frecuencias relativas de las alturas.

Función de densidad normal: características

La fdd $f(x)$ de una variable aleatoria normal, con media μ y varianza σ, fue dada por Gauss. Es una función de tipo exponencial, y su expresión analítica es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}},$$

y su representación gráfica



Para probar que f es una fdd, tenemos que $f(x) \geq 0$ para cualquier x, pero también se tiene que dar que el área encerrada por la gráfica y el eje OX es 1, en otras palabras que $\int_{-\infty}^{+\infty} f(x)dx = 1$, una integral que nosotros no necesitamos saber hacer y que Laplace demostró que, efectivamente, era igual a uno.

Para responder a la primera pregunta, que nos planteamos al comienzo del capítulo, no tenemos que aprender nada nuevo, pues ya sabemos de sobra cómo calcular la media y la desviación típica de un conjunto de datos. En nuestro caso, realizando los cálculos pertinentes tenemos que

$$\mu = \sum x_i f_i = 166'59,$$

y

$$\sigma^2 = \sum x_i^2 f_i - \mu^2 = 118'29.$$

Nosotros vamos a suponer que los datos de las alturas de los alumnos del instituto se distribuyen según una normal. La función de densidad de esa distribución tiene una serie de características especiales, entre ellas que el gráfico de su fdd tiene forma de campana, es decir:

- El gráfico tiene un eje de simetría vertical, que es donde se encuentra el máximo de la fdd y la moda de la variable aleatoria correspondiente.
- La curva tiene un solo máximo, para valores de x menores que este máximo la curva $f(x)$ es decreciente y para los mayores es creciente.
- Teóricamente, puede tomar todos los valores del eje de las x, pero la curva tiene al eje OX como asíntota horizontal a ambos lados, que se acerca a la función con una gran rapidez.
- Cuanto mayor sea la varianza de la distribución, la curva que describe su fdd será más 'plana', y cuanto menor sea, será más 'picuda'. Si X es una variable aleatoria que se distribuye según una normal de media μ y desviación típica σ, lo denotaremos por $X \sim \mathcal{N}(\mu, \sigma^2)$, donde σ^2 es la varianza de X, que cumple que:
  1. $P(\mu - \sigma < X < \mu + \sigma) = 0'683$,

2. $P(\mu - 2\sigma < X < \mu + 2\sigma) = 0'954$,
3. $P(\mu - 3\sigma < X < \mu + 3\sigma) = 0'977$.

Estas últimas ecuaciones, dan una idea de lo poco probable que es obtener valores de la distribución normal lejanos a la media. Piensa que según esas igualdades, si X es una variable aleatoria que se distribuye según una normal de media 2 y varianza 1, $X \sim \mathcal{N}(2, 1)$, tendremos que el $99'7\,\%$ de las observaciones de X, estarán en el intervalo (-1,5), ¡aunque es probable que tome valores mayores a 1.000.000!

Ajustaremos los datos obtenidos para resolver el problema, a una normal cuya media μ y desviación típica σ son las obtenidas anteriormente, es decir tenemos que la altura de los alumnos del instituto se distribuye según una $\mathcal{N}(166'59, 10'88)$.

Sin embargo, esto no nos posibilitará realizar un cálculo sencillo para dar respuesta a las preguntas del problema, por lo que recurriremos a un tipo especial de distribución normal, la normal estándar o $\mathcal{N}(0, 1)$, que ya ha sido ampliamente estudiada y además, todo cálculo que necesitemos hacer para una distribución normal cualquiera, lo podemos realizar mediante ella.

La distribución $\mathcal{N}(0, 1)$

La distribución normal con media $\mu = 0$ y desviación típica $\sigma = 1$, se conoce con el nombre de distribución normal estándar, $\mathcal{N}(0, 1)$.

Su interés radica en el hecho de que los valores de sus probabilidades son perfectamente conocidos y, como veremos más adelante, podemos acudir a ella para hallar probabilidades en cualquier otra distribución normal.

Si tenemos que $Z \sim \mathcal{N}(0, 1)$, las probabilidades $P(Z \leq z)$, es decir de $F(z)$ donde F es la FDD de Z, vienen descritas en la siguiente tabla, para cualquier $z \in [0, 3]$ y eso es suficiente para calcularlas todas, porque como ya veremos, la probabilidad de que asuma valores superiores a 3 es muy pequeña, y con conocer los valores de F para $z \geq 0$, será suficiente para conocerlos todos, debido a la simetría de la fdd.

Sea $Z \sim \mathcal{N}(0, 1)$. Veamos ahora como calcular $P(Z \leq z)$ a partir de los datos de la tabla anterior, según los diferentes valores que puede tomar z. Utilizaremos para ello la simetría de la distribución respecto al cero.

- Si $Z \geq 0 \Rightarrow P(Z \leq z)$ viene descrita en la tabla.
- Si $Z < 0 \Rightarrow P(Z \leq z) = P(Z \geq -z) = 1 - P(Z \leq -z)$, que se puede obtener a partir de la tabla, ya que $-z \geq 0$.
- Si $z_1 \leq z_2$, tenemos que $P(z_1 \leq Z \leq z_2) = P(z \leq z_2) - P(z \leq z_1)$, que es una propiedad común a todas las distribuciones de probabilidad.

Ejercicio 5.3.6 *A partir de las propiedades anteriores, deducir que:*

1. $P(-z \leq Z \leq z) = 2 \cdot P(Z \leq z) - 1$,
2. $P(-Z \leq z \leq 0) = P(0 \leq Z \leq z)$,

donde nuevamente $Z \sim \mathcal{N}(0, 1)$ y z es un número real mayor o igual que cero.

	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0,1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0,2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0,3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0,4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0,5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0,6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0,7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0,8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079	0.8106	0.8133
0,9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1,0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1,1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1,2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1,3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1,4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1,5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1,6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1,7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1,8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1,9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2,0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2,1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2,2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2,3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2,4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2,5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2,6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2,7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2,8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2,9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3,0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

$$F(x)=P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Ejemplo 5.3.4 *Calculemos a partir de la tabla, siendo $Z \sim \mathcal{N}(0,1)$, las siguientes probabilidades:*

1. $P(Z \leq 0'82) = 0'7939$,
2. $P(Z \leq -1'2) = 1 - P(Z \leq 1'2) = 1 - 0'8849$.

Ejercicio 5.3.7 *Calcula a partir de la tabla, siendo $Z \sim \mathcal{N}(0,1)$, las siguientes probabilidades:*

1. $P(Z \leq 0'96)$.
2. $P(Z \leq -2'18)$.
3. $P(-2'76 \leq Z \leq -2'18)$.
4. $P(0'45 \leq Z \leq 2'31)$.

Tipificación

Como dijimos antes, hay una gran cantidad de fenómenos aleatorios que se pueden explicar a partir de una distribución normal, pero el problema es que la mayoría de ellas no tienen media 0 y desviación típica 1. Puede parecer que lo estudiado, para la normal estándar, es de escasa utilidad práctica. Sin embargo, no es así, porque toda variable normal X, $X \sim \mathcal{N}(\mu, \sigma^2)$ puede relacionarse con una $Z \sim \mathcal{N}(0,1)$ mediante el siguiente cambio de variable:

$$Z = \frac{X - \mu}{\sigma}.$$

Este procedimiento, que se llama *tipificación* de la variable X, nos permite utilizar la tabla que presentamos con las tabulaciones de la FDD de una $\mathcal{N}(0,1)$.

Este resultado, se deduce de la siguiente manera:

$$P(X \leq k) = P\left(\frac{X - \mu}{\sigma} \leq \frac{k - \mu}{\sigma}\right) = P(Z \leq \frac{k - \mu}{\sigma}).$$

Por tanto, el valor de $P(X \leq k)$ con X una normal cualquiera, se puede calcular a partir de la tabla de la normal estándar mediante la fórmula $P(Z \leq \frac{k-\mu}{\sigma})$.

Para aclarar estos conceptos, vamos a calcular la cantidad de cada tipo de mesa que habrá de comprar para un instituto con 1000 alumnos, basándonos en los datos de la tabla recogida anteriormente y suponiendo que se distribuyen según una $\mathcal{N}(166'59, 10'88)$.

Para calcular el número de mesas y sillas del tipo A tenemos que ver cuál es el porcentaje de alumnos que medirán menos de 160 cm. Es decir, si denotamos por $X =$ "Altura de un alumno de instituto" tenemos que calcular

$$P(X \leq 160).$$

Como suponemos que $X \sim \mathcal{N}(166'59, 10'88)$ tenemos que

$$P(X \leq 160) = P\left(Z \leq \frac{160 - 166'59}{10'88}\right) = P(Z \leq -0'61),$$

donde $Z \sim \mathcal{N}(0,1)$. Buscando en la tabla de la distribución normal estándar, obtenemos que $P(Z \leq -0'61) = 1 - P(Z \leq 0'61) = 1 - 0'7291 = 0'2709$.

Así pues tenemos que el $27'09\,\%$ de las mesas habrán de ser del tipo A.

Para calcular el número de mesas del tipo B tenemos que calcular $P(160 < X < 180)$, dado que en ese tipo de mesas se sienten cómodos los alumnos que tienen su estatura comprendida en ese intervalo. De nuevo tipificando la variable X obtenemos que

$$P(160 < X < 180) = P\left(\frac{160 - 166'59}{10'88} < Z < \frac{180 - 166'59}{10'88}\right) = P(-0'61 < Z < 1'23)$$

esta probabilidad se divide en dos y resulta

$$P(-0'61 < Z < 1'23) = P(Z < 1'23) - P(Z \leq -0'61) = 0'8907 - 0'2709 = 0'6198,$$

es decir, que tendremos que comprar el $61'98\,\%$ de las mesas del tipo B.

Y para ver cuántas mesas y sillas del tipo C tenemos que comprar sólo tenemos que calcular

$$P(X > 180) = P(Z > 1'23) = 1 - P(Z < 1'23) = 0'1093,$$

es decir, que el $10'93\,\%$ de las mesas y sillas serán del tipo C.

Cómo respuesta definitiva, recomendaríamos a un instituto de 1000 alumnos que comprase $1000 \cdot P(X \leq 160) = 271$ sillas del tipo A, $1000 \cdot P(160 < X \leq 180) = 620$ del tipo B y $1000 \cdot P(X \geq 160) = 109$ del tipo C.

Como ves, es muy sencillo calcular probabilidades en cualquier normal, sabiendo utilizar la tabla de la normal estándar y recordando el cambio de variable que hay que realizar.

Como ajustar otras distribuciones a una normal

En el mundo real, no todas las variables aleatorias se rigen según una distribución normal. Por tanto, antes de analizar unos datos, tenemos que comprobar si estos se ajustan o no a una distribución normal. Son los llamados *test de normalidad*. Aunque los hay muy precisos, también muy complicados, estudiaremos sólo uno, por su sencillez y porque los demás se nos escapan a nuestro nivel.

El test de normalidad que estudiaremos, procede de la siguiente manera:

1. Se extrae una muestra de n elementos de la población y se miden los valores de X, serán $x_1, x_2, ..., x_n$.
2. Se calcula la media y la desviación típica de esos datos, $\overline{x}$ y s respectivamente.
3. Contabilizamos cuántos de esos valores pertenecen a los intervalos $(\overline{x}-s, \overline{x}+s)$, $(\overline{x}-2s, \overline{x}+2s)$ y al $(\overline{x}-3s, \overline{x}+3s)$.
4. Si se tiene que aproximadamente
   el $68'3\,\%$ de los datos pertenece al intervalo $(\overline{x}-s, \overline{x}+s)$,
   el $95'4\,\%$ de los datos pertenece al intervalo $(\overline{x}-2s, \overline{x}+3s)$ y

el $99'7\,\%$ de los datos pertenece al intervalo $(\overline{x}-3s, \overline{x}+3s)$,
entonces podemos admitir que la población de la que se han obtenido los datos, se rige por una normal de media $\mu = \overline{x}$ y varianza $\sigma^2 = s^2$.

A estos niveles, diremos que la aproximación es buena si los porcentajes obtenidos no distan más del $1\,\%$ de los dados anteriormente.

Vamos a ver si los datos de la altura de los alumnos del Instituto se rigen por una distribución de probabilidad normal, tal y como supusimos en la sección anterior.

Calculamos que la media de esos datos era $\overline{x} = 166'59$ y su desviación típica era $s = 10'88$. Calculemos pues los intervalos pedidos en el test de normalidad:

- $(\overline{x}-s, \overline{x}+s) = (155'71, 177'47)$,
- $(\overline{x}-2s, \overline{x}+2s) = (144'83, 188'35)$,
- $(\overline{x}-3s, \overline{x}+3s) = (133'95, 199'23)$.

Si observamos la tabla de frecuencias relativas, tenemos que

- $P(X \in (\overline{x}-s, \overline{x}+s)) = P(X \in (155'71, 177'47)) = 0'691$,
- $P(X \in (\overline{x}-2s, \overline{x}+2s)) = P(X \in (144'83, 188'35)) = 0'965$,
- $P(X \in (\overline{x}-3s, \overline{x}+3s)) = P(X \in (133'95, 199'23)) = 1$.

Así pues tenemos que el $69'1\,\%$ de las observaciones pertenecen al intervalo $(\overline{x}-s, \overline{x}+s)$ cuando teóricamente tenía que ser el $68'3\,\%$. El $96'5\,\%$ de los datos están dentro del intervalo $(\overline{x}-2s, \overline{x}+2s)$ cuando tenía que ser el $95'4\,\%$ y el $100\,\%$ está en $(\overline{x}-3s, \overline{x}+3s)$ cuando teóricamente tenía que estar el $97'7\,\%$.

Aunque no se cumple a rajatabla el hecho de que los porcentajes empíricos no difieran más del $1\,\%$ de los teóricos, podemos decir que la aproximación es buena, ya que la suma de los errores de los tres intervalos es apenas del $5\,\%$, por lo que podemos suponer que la aproximación que hemos hecho es acertada. Además, la variable aleatoria que da la altura de una población es un fenómeno muy conocido y estudiado, y se sabe que se distribuye según una distribución normal. No obstante, si se quiere hacer un estudio exhaustivo es necesaria una mayor rigurosidad en los cálculos. Como nuestra idea es mostrar la aplicabilidad de estas herramientas, nos daremos por satisfechos si estas se entienden, no siendo necesario su absoluta rigurosidad.

La normal como aproximación de la binomial

Cuando estudiamos la distribución binomial, se nos presentaba un grave problema: el cálculo resultaba demasiado tedioso. Calcular a mano probabilidades como $P(X \leq k)$ cuando $X \sim \mathcal{B}(n, p)$ resultaba excesivamente complicado, y prácticamente sólo era posible hacerlo con un ordenador. Por eso, vamos a presentar una aproximación de distribuciones binomiales a normales, ya que estas las tenemos perfectamente controladas, para calcular sus probabilidades.

Así, si $X \sim \mathcal{B}(n,p)$ tenemos que su media es $\mu = np$ y su varianza es $\sigma^2 = npq$. Entonces, podemos decir que si ni p ni q están muy próximos a 0, entonces la distribución binomial puede aproximarse por la normal $\mathcal{N}(np, npq)$, lo que nos facilitará enormemente el cálculo como veremos a continuación. Así, usando esa aproximación y tipificando, tenemos que

$$P(X \leq k) = P\left(Z \leq \frac{k-np}{\sqrt{npq}}\right)$$

donde $Z \sim \mathcal{N}(0,1)$, por lo que podemos calcularla fácilmente.

Esta aproximación es buena, cuando np y nq son ambos mayores que cinco, y va mejorando a medida que aumenta n y p se acerca a $\frac{1}{2}$.

Todavía tenemos que resolver un problema para que esta aproximación sea válida, esto es que mientras que X es una variable discreta que toma valores enteros $X = 0, 1, 2, ..., n$, la distribución normal es continua. Eso hace que mientras que la probabilidad de que la normal tome un valor en un punto es cero, en la X no tiene por qué ser así.

Es por eso por lo que se hace la *corrección de continuidad* para calcular probabilidades del tipo $P(X \leq k)$, y se calcula en realidad $P(X \leq k + 0'5)$, para que así se capture el punto k. Para hallar $P(X < k)$, se calculará $P(X \leq k - 0'5)$ para que el punto k no caiga dentro del intervalo.

Esta corrección es imprescindible para calcular probabilidades como $P(X = k)$, y habrá que calcular $P(k - 0'5 \leq X \leq k + 0'5)$.

Con esta técnica, vamos a dar respuesta a las preguntas que planteamos en el capítulo de la distribución binomial.

La primera pregunta no necesita de esta aproximación para ser contestada fácilmente, pero las dos últimas si. Así pues, vamos a dar respuesta a esas preguntas suponiendo que nuestra variable aleatoria $X =$ "nº de zurdos en el aula de 50 alumnos" se distribuye según una $\mathcal{N}(np, npq) = \mathcal{N}(5, 4'5)$.

Así pues tenemos que para contestar a la segunda pregunta, como tenemos que calcular el primer k tal que $P(X \leq k + 0'5) \geq 0'9$ (por la corrección de continuidad), en realidad buscamos el primer entero k tal que $P\left(Z \leq \frac{k+0'5-5}{\sqrt{4'5}}\right) \geq 0'9$ donde $Z \sim \mathcal{N}(0,1)$.

$P\left(Z \leq \frac{k+0'5-5}{\sqrt{4'5}}\right) = P\left(Z \leq \frac{k-4'5}{2'12}\right)$.

El primer $x \in \mathbb{R}$ que cumple que $P(Z \leq x) \geq 0'9$ es $x = 1'29$. Así pues tenemos que encontrar el primer entero $k / \frac{k-4'5}{2'12} \geq 1'29$. Tenemos que

$$\frac{k-4'5}{2'12} \geq 1'29 \Leftrightarrow k \geq 7'24.$$

Es decir, tenemos que el k buscado es $k = 8$, que es el mismo resultado que obtuvimos en el capítulo de la distribución binomial, y esta vez no hemos necesitado ningún ordenador para calcularlo, sólo el uso de la tabla de la normal estándar y una calculadora a lo más.

Para calcular el porcentaje de grupos de 50 alumnos en los que hay al menos 10 zurdos tenemos que calcular $P(X \geq 9'5)$ y no $P(X \geq 9)$, debido a la corrección de continuidad. Es decir, tenemos que calcular

$$P\left(Z \geq \frac{9'5-5}{2'12}\right) = P(Z \geq 2'12) = 1 - P(Z \leq 2'12) = 0'0174.$$

Así pues, según esta aproximación de la binomial el $1'75\,\%$ de las aulas con 50 alumnos tendrán diez o más zurdos, cifra muy cercana al $2'25\,\%$ que calculamos siguiendo el modelo binomial puro.

Como ves, las respuestas han sido muy parecidas. Esto es por que la aproximación ha sido correcta, ya que en este caso $np = 50 \cdot 0'1 = 5$ y $nq = 50 \cdot 0'9 = 45$, ambas cifras mayores o iguales a cinco, tal y como se sugería en esta sección para dar por buena la aproximación.