



UNIVERSIDAD DE VALLADOLID

Dpto. Estadística e I.O.

La enseñanza de la estadística con herramientas didácticas como “R”

ANEXO 1:A111 1º ESO

**Trabajo Final del Máster Universitario de Profesor en Educación
Secundaria Obligatoria y Bachillerato. Especialidad de Matemáticas.**

Alumno: Julián Rodríguez Vaca.

Tutor: Dr. David Conde del Río.

Valladolid, Junio 2018.

Índice general.

Índice general.....	3
Capítulo 1. Introducción.....	5
Capítulo 2. Contenidos y estándares oficiales.....	7
Capítulo 3. Introducción a “R” Statistics.....	9
Capítulo 4. Contenidos y estándares oficiales con “R” Statistics.	13
Capítulo 5. Estándares de aprendizaje evaluables.....	31
Capítulo 6. Bibliografía	33

Capítulo 1. Introducción.

El objetivo del presente trabajo es plantear y ofrecer una propuesta para la mejora de la didáctica de la estadística mediante el empleo de un potente software, destinado hasta hoy a estudios superiores.

Por un motivo ético, se ha elegido Software Libre, con el que se pretende fomentar este tipo de herramientas en el aula.

Este software, además de ahorrar una gran cantidad de tiempo, permitirá hacer más dinámica esta parte, gracias al manejo de grandes volúmenes de información, la realización de gráficos estadísticos de manera automática, y permitiendo el análisis de los datos de una forma más adecuada.

Con esto el alumno comenzará a tomar contacto con un software de programación y un lenguaje de alto nivel, lo que le mostrará puertas por abrir, y le aportará una buena ventaja sobre todo si se plantea estudios superiores.

El trabajo se presenta en forma de memoria, donde se recopila cada punto del temario en su versión más extensa, y donde aparecen más de 50 ejemplos de cómo resolver los ejercicios de forma tradicional, y con “R”. Incorpora nueve anexos con el temario preparado para cada uno de los cursos, que el profesor puede proporcionar a sus alumnos. Tanto en la memoria como en los anexos, aparece todo el código utilizado en la elaboración del trabajo. Las versiones de introducción o de repaso de cada punto del temario se han dejado en cada uno esos anexos, para evitar la duplicación de los contenidos en la memoria.

Puesto que el bloque de estadística se presenta en todos los cursos en el último bloque de la asignatura de matemáticas, sufre los retrasos de todos los bloques precedentes, dejando en la mayoría de las ocasiones un tiempo muy reducido para el desarrollo del mismo. Con el uso de este método, no se trata de evitar que el alumno trabaje el tan necesario cálculo mental y manual. Sin embargo, si el grupo llega hasta este punto con retraso, uno de los motivos puede

ser precisamente el llevar trabajando cerca de ocho meses en esta línea. Por ello, se trata de optimizar el poco tiempo del que disponga el profesor, evitando pérdidas en la representación de gráficos a mano, nubes de puntos, o tablas de contingencia.

El BOCYL establece, en sus Ordenes EDU 362 y 363 del 4 de mayo de 2015, que el quinto bloque, «Estadística y probabilidad», es de suma importancia.

Esto no sólo es cierto, sino que además, en la era de la información y de la competitividad, el futuro de las empresas y de los países no dependerá tanto del volumen de información de que dispongan, sino de la mejor explotación que hagan de la misma.

Independientemente de su elección tras acabar la ESO o el Bachillerato, el alumno adquirirá los conceptos y el vocabulario necesarios para poder aplicarlos de manera prácticamente autónoma en su futura profesión.

Así, al finalizar sus estudios será capaz de realizar análisis críticos de una mayor cantidad de información mediante tablas y gráficas, con la ayuda de “R”.

Será capaz de recopilar datos por sí mismo, organizarlos, resumirlos, estudiarlos y explotarlos, lo que le será de gran utilidad en su ámbito profesional.

El contenido del trabajo está adaptado a la comunidad de Castilla y León, según las órdenes EDU 362 y 363 de 2015 por las que se establecen los currículos y se regulan la implantación, evaluación y desarrollo de la educación secundaria obligatoria y del bachillerato en la Comunidad de Castilla y León:

ORDEN EDU/362/2015, de 4 de mayo, por la que se establece el currículo y se regula la implantación, evaluación y desarrollo de la educación secundaria obligatoria en la Comunidad de Castilla y León.

ORDEN EDU/363/2015, de 4 de mayo, por la que se establece el currículo y se regula la implantación, evaluación y desarrollo del bachillerato en la Comunidad de Castilla y León.

Así, establecen los temas para el bloque de estadística que veremos a continuación.

Capítulo 2. Contenidos y estándares oficiales.

1º ESO.

- 1E1.- Población e individuo. Muestra. Variables estadísticas.
- 1E2.- Variables cualitativas y cuantitativas discretas.
- 1E3.- Frecuencias absolutas y relativas.
- 1E4.- Organización en tablas de datos recogidos en una experiencia.
- 1E5.- Diagramas de barras, y de sectores. Polígonos de frecuencias.
- 1E6.- Medidas de tendencia central.
- 1E7.- Medidas de dispersión.

Capítulo 3. Introducción a “R” Statistics.

3.1 Sobre “R”:

R es un lenguaje y entorno para el procesamiento y representación de datos estadísticos. Es un proyecto de GNU similar al lenguaje y al entorno S, que fue desarrollado en Bell Laboratories, por John Chambers y su equipo. Hay algunas diferencias importantes, pero gran parte del código escrito para S corre inalterado bajo R.

R proporciona una amplia variedad de técnicas estadísticas y gráficos, y es altamente extensible mediante la creación de librerías por los usuarios, al ser una herramienta de código abierto.

Uno de los puntos fuertes de R es la facilidad con la que se pueden crear gráficos de calidad, incluyendo símbolos matemáticos y fórmulas si es necesario.

R está disponible como Software Libre bajo los términos de la Licencia Pública General GNU de la Free Software Foundation en forma de código fuente. Se compila y se ejecuta en una amplia variedad de plataformas UNIX y sistemas similares (incluidos FreeBSD y Linux), Windows y MacOS.

Fuente: <https://www.r-project.org/about.html>

3.2 Descarga e instalación de RStudio:

Rstudio es un software gratuito que podemos descargar de forma totalmente legal y sin coste ni publicidad, de la siguiente página:

<https://www.rstudio.com/products/rstudio/download/>

Pulsando en la tecla download, que aparece en la columna Free (gratis), nos dirige a la zona para elegir nuestro sistema operativo:

Installers	Size	Date	MD5
RStudio 1.1.453 - Windows Vista/7/8/10	85.8 MB	2018-05-16	bf287e385aef53829204023087e98735
RStudio 1.1.453 - Mac OS X 10.6+ (64-bit)	74.5 MB	2018-05-16	00a0088424ed06ac434f7a966f602b9c
RStudio 1.1.453 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-05-16	6cfd86770c7b6dbc13e66f4f59c299ce
RStudio 1.1.453 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	97.4 MB	2018-05-16	63e36e8138e369d19f9aaf4b0e995bbc
RStudio 1.1.453 - Ubuntu 16.04+/Debian 9+ (64-bit)	64.4 MB	2018-05-16	85b3e76c9fad4613bc9cf0de1f34b183
RStudio 1.1.453 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.1 MB	2018-05-16	37cade7e162eab62483e6556e39dedee
RStudio 1.1.453 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	90.6 MB	2018-05-16	44cddd285bc31c41e4eaeed174b8eebb

Dependiendo del sistema operativo de nuestro PC, descargamos el que corresponda.

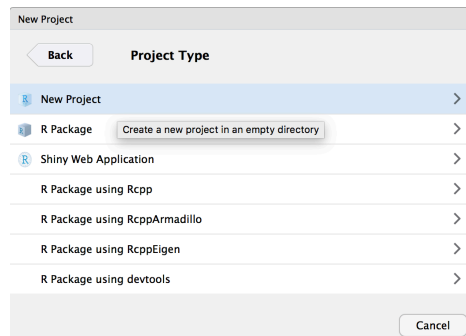
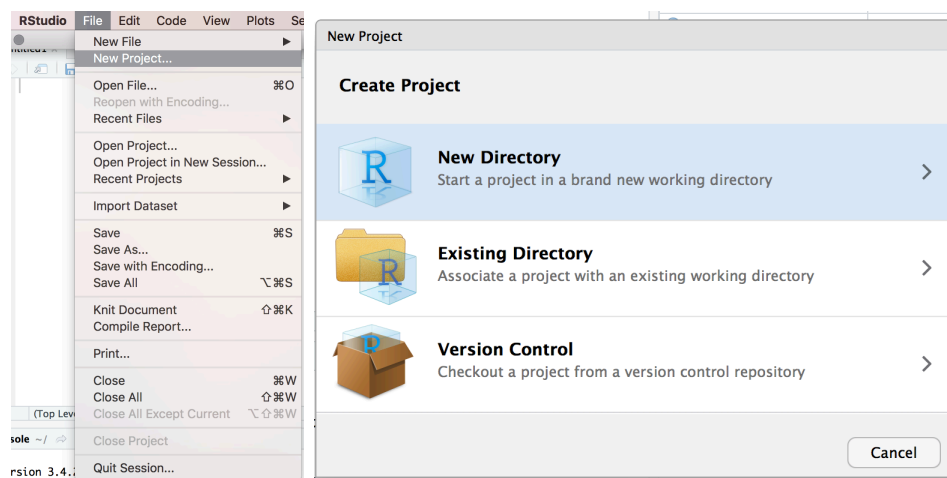
Una vez descargado, se ejecuta el programa instalador, y se van siguiendo los pasos del asistente de instalación, como en cualquier otro programa.

3.3.- Creación de un nuevo proyecto:

1.- Vamos a crear nuevo proyecto con el nombre población y muestra, ubicado en el escritorio del PC.

Para ello, abrir Rstudio y pulsar:

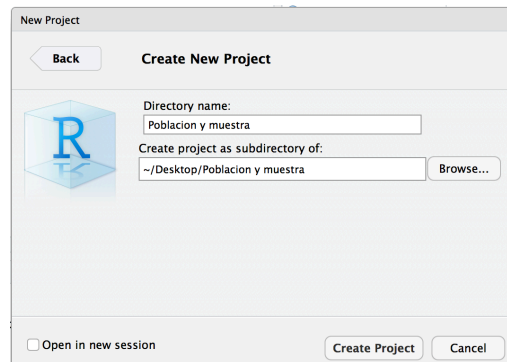
File/New project.../New directory/New Project



En la siguiente ventana, escribimos:

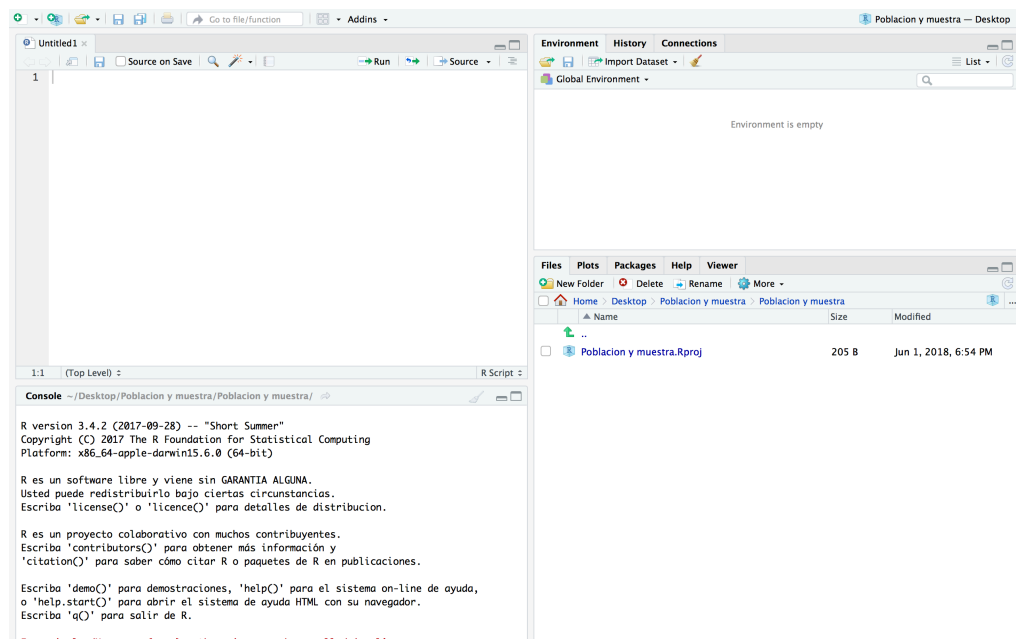
Directory name: Población y muestra.

Create project as subdirectory of: Click en browse... y creamos una carpeta en nuestro escritorio (desktop) que se llame Población y muestra.



Nota: las carpetas podemos crearlas tanto en el escritorio como en un USB o donde queramos, y luego localizarla usando la tecla browse.

Con esto, se nos abre el entrono de “R”, listo para empezar. Tendrá la siguiente pinta:



Los comandos se escriben en la zona inferior izquierda, y los gráficos se mostrarán en la ventana inferior derecha. Las ventanas superiores son para la selección y visualización de tablas y otras variables. Con esto el sistema está listo para comenzar a trabajar

Capítulo 4. Contenidos y estándares oficiales con “R” Statistics.

DeSeCo (2003) define competencia como «la capacidad de responder a demandas complejas y llevar a cabo tareas diversas de forma adecuada».

La competencia «supone una combinación de habilidades prácticas, conocimientos, motivación, valores éticos, actitudes, emociones, y otros componentes sociales y de comportamiento que se movilizan conjuntamente para lograr una acción eficaz».

Se contemplan, pues, como conocimiento en la práctica, es decir, un conocimiento adquirido a través de la participación activa en prácticas sociales y, como tales, se pueden desarrollar tanto en el contexto educativo formal, a través del currículo, como en los contextos educativos no formales e informales.

Fuente: <https://www.mecd.gob.es/>

En este trabajo se ha buscado contribuir a las competencias en:

Comunicación lingüística, mediante el fomento de un uso del vocabulario apropiado, de la lectura y sobre todo de la interpretación de los enunciados, que contribuyen finalmente a expresarse y comunicarse con propiedad.

Competencia matemática, mediante el análisis matemático del comportamiento de las variables de estudio de la población, extrayendo conclusiones en función de la regresión lineal y correlación de los datos de las variables, y bajo la interpretación conjunta de parámetros estadísticos.

Competencia digital, mediante el fomento de un uso ético, cívico y crítico de las nuevas tecnologías, y mediante el empleo de una herramienta Software de alto nivel.

Competencias sociales y cívicas, mediante el análisis de datos de nuestro entorno, como PIB, IPC, extrayendo conclusiones de posibles desigualdades salariales en poblaciones, o identificando las malas prácticas de las presentaciones de datos de forma interesada.

Competencia cultural, representando e interpretando la información con relación a ejemplos de plantas y otros datos del entorno.

La competencia aprender a aprender, mediante el ejemplo de la búsqueda de información para mejorar la utilización del software R de manera casi autodidacta.

Sentido de la iniciativa y espíritu emprendedor, mostrando al alumno el inicio de un camino, que con su propia iniciativa podrá recorrer hasta donde le lleve su curiosidad científica. Por su potencia y escasa inversión, el alumno será capaz de imaginar escenarios de emprendimiento, donde con un ordenador y este software como herramientas podrá realizar estudios de alto valor a nivel profesional.

1E1.- Población e individuo. Muestra. Variables estadísticas.

La Estadística es una ciencia que se ocupa del estudio de los métodos y procedimientos para recoger, clasificar, resumir y analizar datos observados sobre una población de individuos.

El objetivo de cualquier estudio estadístico es obtener información acerca de las características de los individuos de cierto colectivo, llamado población estadística.

Población: Es el conjunto total de individuos sobre los que se quieren estudiar unos datos determinados.

Individuo: Cada uno de los componentes de la población. Pueden ser personas, animales, plantas, u objetos.

Cuando la población o colectivo sea muy grande, se hará difícil el estudio de la misma. Estos inconvenientes pueden ser superados mediante la elección de muestras.

Muestra: Es una parte de la población representativa de la misma. Tiene por tanto características similares. Ha de elegirse al azar. Se utiliza cuando la población es muy grande, o difícil de estudiar.

Variable estadística: Es el dato o característica que se quiere estudiar. Por ejemplo: la estatura, la nota de Matemáticas, el sexo de una persona, o su peso.

Ejemplo. Población.

Vamos a crear un archivo de población con los alumnos de la clase, y su sexo. Para crear la población, nos vamos a ayudar de otro programa informático: LibreOffice Calc.

Lo abrimos, y vamos colocando en la primera fila, uno encima del otro, los nombres de toda la clase por orden de izquierda a derecha, y de adelante a atrás según estemos sentados en el aula. En la segunda columna, colocaremos el primer apellido de cada compañero. En la tercera columna, colocaremos el segundo apellido de cada compañero.

Marta	Hernández	Poza
Paula	Fernández	Mate
Mario	Rodríguez	Crespo
Alberto	Camino	Arranz
Elsa	Moral	Rodríguez
David	Castro	Martin
Álvaro	Martínez	Lobato
Alberto	Martín	Sueña
Sofía	González	Manuel
Claudia	Álvarez	Alonso
María	Lozano	Alonso
Manuel	Martínez	Lobato
Esteban	González	Aguado
Jesús	González	Ortega
Iván	Cano	Hernández
Luis	Mateo	Sánchez
Jimena	Sacristán	Blas
Marga	Domingo	Garzo
Eva	Viñas	Zamora
Ana	Merino	Valdezate
Francisco	Rodríguez	Álvarez
Andrea	Moreno	Jiménez
Carla	Rojo	Domínguez

Lo guardamos con el nombre de Población.xlsx en la carpeta del proyecto Población y muestra. Ya tenemos un archivo con la población de la

clase. Esto se conoce como censo. En la antigüedad, los reyes pedían hacer censos constantemente. Incluso en la biblia aparecen referencias a censos.

En el libro de Números, en su capítulo 1, se puede leer lo siguiente:

Censo de las tribus de Israel:

El primer día del segundo mes del segundo año desde la salida de los israelitas de Egipto, el Señor, dio las siguientes instrucciones a Moisés, que se encontraba en el santuario, en el desierto del Sinaí.

«Haz un censo de todos los hombres, mayores de veinte años, capaces de ir a la guerra. En la lista anota la tribu y familia a la que pertenezcan.»

Ejemplo 2. Muestra:

Tomar una muestra representativa de la población que acabamos de crear.

Vamos a tomar como muestra a 10 compañeros. Los 10 que queramos:

Marta	Hernández	Poza
Paula	Fernández	Mate
Mario	Rodríguez	Crespo
Alberto	Camino	Arranz
Elsa	Moral	Rodríguez
David	Castro	Martin
Álvaro	Martínez	Lobato
Alberto	Martín	Suaña
Sofía	González	Manuel
Claudia	Álvarez	Alonso

1E2.- Variables cualitativas y cuantitativas.

Variable cualitativa: Describen cualidades que no puede expresarse por números. Por ejemplo, provincias españolas, colores favoritos, qué libro lectura prefieren los adolescentes, o el coche más vendido.

Variable cuantitativa: Las variables cuantitativas toman valores numéricos. Por ejemplo, la estatura, o la nota de una asignatura. Pueden ser:

Variable cuantitativa discreta: Los valores de la variable son números enteros 1, 2, 3, 4, 5. Por ejemplo el número de compras de un producto en un mes, no puede ser 4,8.

Variable cuantitativa continua: Pueden tomar todos los valores dentro de un intervalo. La temperatura, la humedad, o la estatura, son ejemplos de variables cuantitativas continuas. Los valores razonables de la variable temperatura en una persona pueden valer desde 34,5°C hasta 42°C, pudiendo tomar cualquier valor intermedio.

Ejemplo 3. Variable cualitativa.

Añadir a la tabla anterior, una variable cualitativa a estudiar: color del pelo. La llamaremos Pelo, para no poner espacios ni caracteres especiales.

Nombre	Apellido1	Segundo2	Pelo
Marta	Hernández	Poza	castaño
Paula	Fernández	Mate	rubio
Mario	Rodríguez	Crespo	moreno
Alberto	Camino	Arranz	rubio
Elsa	Moral	Rodríguez	castaño
David	Castro	Martin	rubio
Álvaro	Martínez	Lobato	moreno
Alberto	Martín	Suaña	castaño
Sofía	González	Manuel	rubio
Claudia	Álvarez	Alonso	castaño
María	Lozano	Alonso	castaño
Manuel	Martínez	Lobato	castaño
Esteban	González	Aguado	moreno
Jesús	González	Ortega	castaño

Iván	Cano	Hernández	castaño
Luis	Mateo	Sánchez	moreno
Jimena	Sacristán	Blas	moreno
Marga	Domingo	Garzo	pelirrojo
Eva	Viñas	Zamora	castaño
Ana	Merino	Valdezate	moreno
Francisco	Rodríguez	Álvarez	rubio
Andrea	Moreno	Jiménez	moreno
Carla	Rojo	Domínguez	pelirrojo

Ejemplo. Variable cuantitativa:

Añadir a la tabla anterior, una columna con una variable cuantitativa discreta y otra continua.

Nombre	Apellido1	Segundo2	Pelo	Sexo	Hermanos	Nota_1EV
Marta	Hernández	Poza	castaño	Chica	0	4,71
Paula	Fernández	Mate	rubio	Chica	2	1,75
Mario	Rodríguez	Crespo	moreno	Chico	2	5,52
Alberto	Camino	Arranz	rubio	Chico	2	1,97
Elsa	Moral	Rodríguez	castaño	Chica	0	4,81
David	Castro	Martin	rubio	Chico	2	3,50
Álvaro	Martínez	Lobato	moreno	Chico	1	6,40
Alberto	Martín	Suaña	castaño	Chico	0	4,80
Sofía	González	Manuel	rubio	Chica	2	1,06
Claudia	Álvarez	Alonso	castaño	Chica	1	5,00
María	Lozano	Alonso	castaño	Chica	1	3,58
Manuel	Martínez	Lobato	castaño	Chico	2	5,03
Esteban	González	Aguado	moreno	Chico	1	6,62
Jesús	González	Ortega	castaño	Chico	2	5,00
Iván	Cano	Hernández	castaño	Chico	1	4,86
Luis	Mateo	Sánchez	moreno	Chico	0	9,18
Jimena	Sacristán	Blas	moreno	Chica	3	7,53
Marga	Domingo	Garzo	pelirrojo	Chica	1	0,70
Eva	Viñas	Zamora	castaño	Chica	1	5,28
Ana	Merino	Valdezate	moreno	Chica	1	6,80
Francisco	Rodríguez	Álvarez	rubio	Chico	2	1,99
Andrea	Moreno	Jiménez	moreno	Chica	1	7,72
Carla	Rojo	Domínguez	pelirrojo	Chica	3	0,95

1E3.- Frecuencias absolutas y relativas.

Frecuencia absoluta (n_i): Es el número de veces que aparece cada valor (x_i) de la variable.

La suma de las frecuencias absolutas es el número total de datos (N).

En “R”, usaremos la función `table` para crear tablas de frecuencias absolutas:

```
> FA_Pelo<-table(Población$Pelo)
> FA_Pelo
```

castaño	moreno	pelirrojo	rubio
9	7	2	5

Frecuencia relativa (f_i): es el resultado de dividir la frecuencia absoluta entre el número total de datos (N):

$$f_i = \frac{n_i}{N}$$

En “R”, usaremos la tabla de frecuencias absolutas, y la aplicaremos la función `prop.table`:

```
> prop.table(FA_Pelo)
```

castaño	moreno	pelirrojo	rubio
0.39130435	0.30434783	0.08695652	0.21739130

Estos ejercicios se realizan mucho más rápido utilizando el ordenador, como vamos a ver a continuación.

Frecuencias acumuladas

La frecuencia absoluta acumulada (N_i) de un valor X_i del conjunto (X_1, X_2, \dots, X_N) es la suma de las frecuencias absolutas de los valores menores o iguales a X_i , es decir: $N_i = n_1 + n_2 + \dots + n_i$.

En “R”, usaremos la tabla de frecuencias absolutas, y la aplicaremos la función `cumsum`:

```
> cumsum(FA_pelo)
  castaño    moreno pelirrojo    rubio
      9      16      18      23
```

Ejercicio. Frecuencia Absoluta:

Escribe la frecuencia absoluta de los alumnos rubios, contando uno a uno cada uno de ellos.

Solución: 5

Ejercicio. Frecuencia Relativa:

Escribe la frecuencia relativa de tener 2 hermanos sin la ayuda de la calculadora.

Solución: 0,35

Reflexiona: ¿Qué pasaría si quisiéramos escribir la frecuencia relativa de todos los habitantes rubios de Valladolid? Habría que contar todos los habitantes rubios, y dividir por el número de habitantes.

1E4.- Organización en tablas de datos recogidos en una experiencia.

Es muy común encontrarnos los datos colocados en tablas y gráficas para ayudar a su interpretación. Estos datos pueden haber sido obtenidos mediante encuestas, de muy diversos tipos.

Encuesta: Una encuesta es un procedimiento que permite obtener los datos para hacer un estudio de ellos. Puede ser oral o escrita. En función de las preguntas realizadas, se construye una tabla donde se representan los datos.

La tabla Población.xlsx que hemos construido es una encuesta. Podríamos seguir insertando columnas, con las preguntas que se nos ocurriesen.

Hemos visto que si una encuesta se efectúa sobre una población muy grande, puede resultar muy complicado extraer conclusiones. Vamos a aprender a utilizar las herramientas informáticas para analizar estos datos.

Ejemplo. Tablas en "R":

Ya tenemos los datos de todos los alumnos de la clase en una tabla. Ahora vamos a empezar a estudiar los comportamientos de las variables con la ayuda del software "R".

1.- Crea nuevo proyecto con el nombre población y muestra, ubicado en el escritorio del PC.

2.- Importa la tabla con los alumnos de la clase. La llamaremos Población, escribiendo lo siguiente en la ventana inferior derecha:

```
>Población<-read_excel("~/Desktop/Población y muestra/Población.xlsx").
```

Vamos a estudiar las frecuencias relativas y absolutas de los alumnos rubios, por un lado, y las frecuencias de tener 2 hermanos por otro, sin la necesidad de hacer cuentas.

Nota: Los nombres de las Variables, escritos en cada cabecera, no deben contener espacios ni caracteres especiales. Para poner espacios, usaremos el guión bajo _.

Construimos una tabla con las frecuencias absolutas de la variable Pelo.

Para ello usamos la función `table`, escribiendo lo siguiente:

```
> FA_Pelo<-table(Población$Pelo)
```

A la tabla de frecuencias absolutas la estaremos llamando `FA_Pelo`.

Mediante el símbolo `<-`, que representa una especie de flecha, estaremos diciéndole al programa que escriba dentro de la variable lo que nosotros queramos. En este caso, como queremos escribir la tabla de frecuencias absolutas de la variable `Pelo` de nuestra población, hemos escrito a continuación de la flecha, la función que queremos utilizar, "`table`". A continuación y entre paréntesis, hemos escrito la tabla de donde tiene que buscar los datos, (`Población$Pelo`). El símbolo `dollar` indica la columna a la que queremos referirnos.

Para visualizar lo que acabamos de escribir en la tabla `FA_Pelo` escribimos:

```
> FA_Pelo
```

El programa nos muestra lo siguiente:

castaño	moreno	pelirrojo	rubio
9	7	2	5

Por lo tanto, podemos ver que hay 9 alumnos castaños, 7 morenos, 2 pelirrojos, y 5 rubios.

La **moda**, en este caso, sería tener el pelo castaño, por ser el valor más repetido.

Las frecuencias absolutas del número de hermanos serían:

```
> FA_Hermanos<-table(Población$Hermanos)
> FA_Hermanos
0 1 2 3
4 9 8 2
```

La **moda** de nuestra clase sería tener un único hermano.

Para las frecuencias relativas usaremos la tabla de frecuencias absolutas, y la aplicaremos la función `prop.table`:

```
> FR_Pelo<-prop.table(FA_Pelo)
> FR_Pelo
```

```
    castaño    moreno  pelirrojo    rubio
```

```
0.39130435 0.30434783 0.08695652 0.21739130
```

Con esta instrucción, el programa divide cada miembro de la tabla de frecuencias absolutas, por el número de miembros, y coloca el resultado en la nueva tabla.

Hallamos las frecuencias relativas de la variable Hermanos:

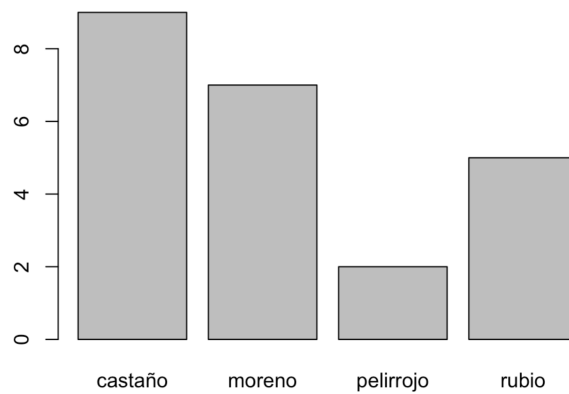
```
> FR_Hermanos<-(FA_Hermanos)/margin.table(FA_Hermanos)
> FR_Hermanos
```

```
    0      1      2      3
```

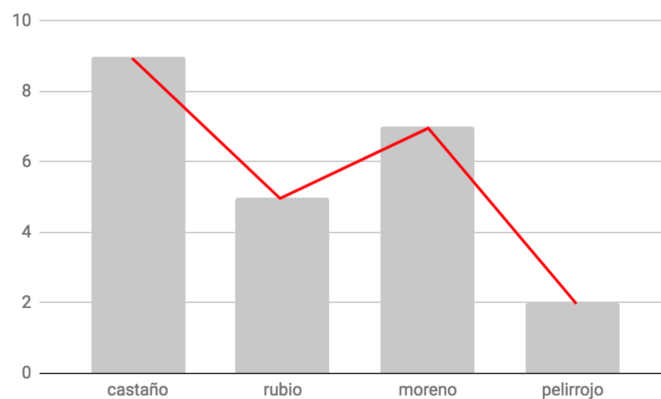
```
0.17391304 0.39130435 0.34782609 0.08695652
```


1E5.- Diagramas de barras, y de sectores. Polígonos de frecuencias.

El diagrama de barras muestra las frecuencias absolutas de los datos. Cuanto más alta es la barra más se da el valor al que corresponde.



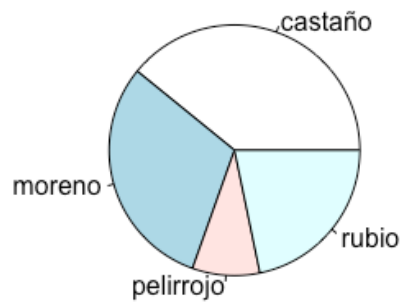
Uniando el punto medio de la parte superior de cada barra del diagrama de barras se obtiene el polígono de frecuencias (en rojo). En el ejemplo de la tabla:



En un diagrama de sectores, la amplitud de cada sector circular representa el valor de la variable:

En "R" se representa con la función `pie`

```
> pie(FA_Pelo)
```



Ejercicio: Construye un diagrama de barras de las variables Pelo y Hermanos, e intreprétalas. Para ello, responde a las siguientes preguntas:

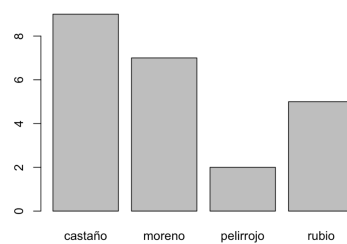
¿Quién predomina más en la clase, los rubios o los morenos?

¿Qué es más común, tener un hermano, o dos?

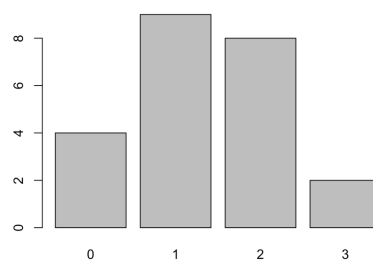
¿Qué color de pelo es el menos frecuente?

Para construir diagramas de barras, usaremos la función `barplot`.

```
> barplot(FA_Pelo)
```



```
> barplot(FA_Hermanos)
```



1E6.- Medidas de tendencia central.

En ocasiones interesa resumir la información de una muestra, en un solo valor, para hacernos una idea de cómo se comporta la variable y poder realizar comparaciones.

Las medidas de tendencia central más habituales, son la media, la mediana y la moda.

Media Aritmética: Es la suma de todos los datos dividida entre el número total de datos.

Es la medida de tendencia central más utilizada en Estadísticas y en todos los campos de la vida.

Tanto las empresas, como los países, como los medios de comunicación, constantemente hablan de medias de datos, como el gasto medio, el salario medio, la estatura media, o la media de la edad de los alumnos de primer curso de universidad.

Hay que tener cuidado, porque si no se usa debidamente, puede conducir a interpretaciones erróneas. Por ejemplo, si Pedro come dos porciones de pizza, y mi Juan ninguna, la media diría que han comido una cada uno. Y eso no es cierto.

Media Ponderada: La media ponderada (MP) de un conjunto de valores de una variable, consiste en dar a cada observación del conjunto de datos unos pesos, que indican la importancia que se quiere dar a cada uno de los valores que toma la variable.

Media Geométrica: La media geométrica de un conjunto de datos es el resultado de multiplicarlos entre si y aplicar la -ésima raíz.

Si en la media aritmética sumábamos los valores para luego dividirlos, ahora debemos multiplicarlos para luego aplicar la -ésima raíz pertinente.

La media geométrica necesita que no haya números negativos o que estos sean un número par, para evitar aplicar una raíz a un número negativo.

Mediana: La mediana divide en dos partes iguales a la distribución

estadística

Moda: Es el valor más repetido, o el valor de la variable estadística que tiene la frecuencia absoluta más alta. Pueden ser varias variables a la vez.

Ejercicios:

Media: La media la calcularemos con la función mean.

Mediana: La mediana la calcularemos con la función median.

Moda: Es el valor cuya frecuencia absoluta es más alta. Calculamos la tabla de frecuencias absolutas de las notas, y buscamos el mayor valor.

Ejercicio: Calcular la media, la mediana y la moda de las notas de la primera evaluación

Media:

```
> mean(Población$EV_P)
```

```
[1] 4.554526
```

Mediana:

```
> median(Población$EV_P)
```

```
[1] 4.862917
```

Moda:

Calculamos las frecuencias absolutas de las notas de la primera evaluación:

```
> FA_EV_P <- table(Población$EV_P)
```

```
> FA_EV_P
```

```
0.7 0.95 1.06 1.75 1.97 1.99 3.5 3.58 4.71 4.8 4.81 4.86 5 5.03 5.28 5.52
6.4 6.62
1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1
6.8 7.53 7.72 9.18
1 1 1 1
```

Como aparecen desordenadas, podemos ordenarlas de mayor a menor para que nos resulte más cómodo encontrar la moda:

```
> FA_EV_P <- sort(FA_EV_P, decreasing = TRUE)
```

```
> FA_EV_P
```

```
5 0.7 0.95 1.06 1.75 1.97 1.99 3.5 3.58 4.71 4.8 4.81 4.86 5.03 5.28 5.52
2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
6.4 6.62 6.8 7.53 7.72 9.18
1 1 1 1 1 1
```

La moda sería el valor de 5, que se repite dos veces.

1E7.- Medidas de dispersión.

Las medidas de dispersión son una serie de valores que nos informan cómo se encuentran los datos de agrupados o desagrupados.

Desviación media: mide la distancia media que hay entre todos los valores de la muestra y el valor medio.

Cuando mayor sea la desviación media quiere decir que mayor es la dispersión de la muestra, los datos están más separados, en cambio si el valor de la desviación es reducido significa que los valores de la media están muy concentrados.

Rango: mide la diferencia entre el valor mayor y el valor menor de la muestra.

Mientras mayor sea el rango más dispersos estén los valores.

Varianza: Al igual que la media, la varianza es un indicador que se utiliza para medirla dispersión (a mayor varianza, mayor dispersión), dando una información más precisa que la media.

Se calcula como suma de las diferencias al cuadrado de cada valor respecto a la media de la muestra. Esta suma se divide entre el número de datos. La varianza se suele representar con la letra V .

Desviación típica: La desviación típica es otra medida de dispersión y se calcula como raíz cuadrada de la varianza. Es la medida de dispersión que más se utiliza.

Ejercicios:

Desviación media: Se calcula con la función $\text{sd}(x)$

Rango: Se calcula con la función $\text{range}(x)$

Varianza: Se calcula con la función $\text{var}(x)$

Desviación típica: Se calcula con la función $\text{sd}(x)$

Ejercicio: Calcular la Desviación media, el rango, la varianza y la desviación típica de las notas de la primera evaluación

```
> sd(Población$EV_P)
[1] 2.321369
> range(Población$EV_P)
[1] 0.70 9.18
> var(Población$EV_P)
[1] 5.388753
```

En este ejemplo los datos aparecen muy dispersos. Hay muchos suspensos, muchos aprobados cerca del 5, y también hay buenas notas. Por eso la desviación media es alta.

El rango nos da los extremos de las notas mínima y máxima.

.

Capítulo 5. Estándares de aprendizaje evaluables.

Estándares 1º ESO. Aplicadas

1.1. Define población, muestra e individuo desde el punto de vista de la estadística, y los aplica a casos concretos.

1.2. Reconoce y propone ejemplos de distintos tipos de variables estadísticas, tanto cualitativas como cuantitativas.

1.3. Organiza datos, obtenidos de una población, de variables cualitativas o cuantitativas en tablas, calcula sus frecuencias absolutas y relativas, y los representa gráficamente.

1.4. Calcula la media aritmética, la mediana y la moda y los emplea para resolver problemas.

2.1. Emplea la calculadora y herramientas tecnológicas para organizar datos, y calcular las medidas de tendencia central.

3.1. Identifica los experimentos aleatorios y los distingue de los deterministas.

3.2. Calcula la frecuencia relativa de un suceso mediante la experimentación.

3.3. Realiza predicciones sobre un fenómeno aleatorio a partir del cálculo exacto de su probabilidad o la aproximación de la misma mediante la experimentación.

4.1. Describe experimentos aleatorios sencillos y enumera todos los resultados posibles, apoyándose en tablas, recuentos o diagramas en árbol sencillos.

Capítulo 6. Bibliografía

Software Rstudio:

<https://cran.r-project.org/>

Manual de R:

Título: R para profesionales de los datos: una introducción

Autor: Carlos J. Gil Bellosta

Fecha: 2018-04-22

https://www.datanalytics.com/libro_r/index.html

Estructuras de datos en R:

http://www.dm.uba.ar/materias/analisis_de_datos/2009/2/practicas/TP2-2009.pdf

Creación de data.frames en R:

<http://r-econ.blogspot.com/2012/07/unir-varios-dataframes-en-un-solo-paso.html>

Curso de introducción a la Estadística:

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-00.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-02.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-04.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-05.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-06.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-07.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-08.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-09.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-10.pdf>

<http://www.postdata-statistics.com/IntroEstadistica/Tutoriales/Tutorial-11.pdf>

Contenidos y estándares oficiales:

Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato.

<https://www.boe.es/boe/dias/2015/01/03/pdfs/BOE-A-2015-37.pdf>

ORDEN EDU/362/2015, de 4 de mayo, por la que se establece el currículo y se regula la implantación, evaluación y desarrollo de la educación secundaria obligatoria en la Comunidad de Castilla y León.

<http://bocyl.jcyl.es/boletines/2015/05/08/pdf/BOCYL-D-08052015-4.pdf>

Libros de texto:

Fundamentos y métodos de Estadística, 3ª Edición,:

Autores: M. López Cachero

Editorial: Ed Piramide

ISBN 84-368-0171-7

Estadística aplicada a las ciencias de la educación:

Autores: Joan Welkowitz, Robert B. Ewen, Jacob Cohen

Editorial: Ed Santillana

ISBN: 84-294-1903-9

Libros de texto IES Cristo Rey, curso 2017-12018

Matemáticas 1º ESO:

Autores: Francisco Javier García Crespo; Ruth Martín Escanilla

Editorial: EDITEX S.A

1ª ed. (2015)

ISBN: 8490784949 ISBN-13: 9788490784945

Matemáticas 2º ESO

Autores: Fernando ... [et al.] Alcaide Guindo

Pelorroto; Juan Antonio Rocafort (il.)

Editorial: EDICIONES SM

1ª ed. (01/05/2016)

ISBN: 8467586885 ISBN-13: 9788467586886

Apuntes de Complementos de Matemáticas del Máster en profesor de educación secundaria obligatoria y bachillerato, formación profesional y enseñanzas de idiomas

http://campusvirtual2017.uva.es/pluginfile.php/415026/mod_resource/content/1/CM2017-18-Material%20Estad%C3%ADstica-2.pdf

Apuntes de estadística para Ingenieros Técnicos Industriales. Curso 2005, Escuela Universitaria Politécnica de Valladolid.

Otros:

Definiciones de Medidas de centralización y de dispersión:

Wikipedia

Referencia censos bíblicos:

<https://www.bible.com/es/bible/149/NUM.1.RVR1960?parallel=149>

Definiciones. Encuesta, censos:

http://www.ine.es/explica/explica_pasos_primera_encuesta.htm