

Estadística

Práctica 4

Informe de estadística descriptiva

Contenido

1. Introducción	3
2. Organizar los datos.....	4
3. Calcular medidas estadísticas	5
3.1 Medidas de centralización	5
3.2 Medidas de dispersión.....	5
3.3 Medidas de localización	6
3.4 Medidas de forma.....	6
3.5 Otras medidas.....	7
4. Obtener tablas de frecuencias	8
5. Dibujar diagramas.....	10
5.1 Diagramas de tarta.....	10
5.2 Histogramas y diagramas de caja	11
6. Escribir el informe	12
6.1 Título	12
6.2 Introducción.....	12
6.3 Metodología	13
6.4 Resultados	13
6.5 Análisis de resultados	16
6.6 Conclusiones.....	18
6.7 Referencias.....	18
7. Referencias	19

1. Introducción

Con esta práctica se trata de aprender a realizar un informe descriptivo de una variable estadística de forma independiente y también en combinación con otra variable. Aunque se podría realizar un informe sobre el total de observaciones de la variable en una muestra, la utilidad se pone de manifiesto si se comparan los resultados en dos grupos de la muestra, y así se hará en esta práctica.

Se utilizarán los mismos datos que en la primera parte, es decir los disponibles en el archivo "[encuesta.csv](#)".

Para elaborar un informe sobre un análisis descriptivo de datos hay que seguir los siguientes pasos:

- Plantear el objetivo o pregunta de investigación. Se trata de definir qué es lo que queremos analizar y sobre qué variable estadística tratará el informe.
- Recoger datos. Hay que conseguir una muestra de valores de la variable estadística a analizar y de otras posibles variables relacionadas que sirvan por ejemplo para realizar agrupaciones de datos y comparaciones.
- Organizar los datos. En muchas ocasiones no es posible trabajar directamente con los datos obtenidos, es necesario realizar una organización de estos, por ejemplo, cambiar de formato, agrupar datos, eliminar datos con valores extraños, etc.
- Realizar los cálculos estadísticos. Una vez preparados los datos hay que utilizar una herramienta, como R, para calcular medidas estadísticas y realizar diagramas.
- Escribir el informe. Se trata de recoger en un documento los resultados del análisis

2. Organizar los datos

Vamos a analizar la variable estadística “Nota de acceso a la universidad” ya utilizada en la práctica anterior, cuyas observaciones se encuentran en la columna NOTA del archivo “encuesta.csv”. Pero también vamos a usar otra variable TURNO que represente el turno al que pertenece cada alumno, que puede ser “Mañana” o “Tarde”. El turno puede obtenerse a partir de la variable GRUPO, ya que los grupos A1 y A2 son de tarde y los grupos B1 y B2 son de mañana.

Podrían leerse los datos del archivo encuesta.csv en un dataset y añadir una nueva columna TURNO y rellenarla a partir de la columna GRUPO usando comandos de R como rep() y rbind()¹, pero es más sencillo modificar directamente el archivo encuesta.csv, añadir una nueva columna y rellenarla manualmente hasta obtener una nueva versión del archivo, por ejemplo llamado [encuesta_turno.csv](#).

GRUPO	GRADO	NOTA	VIAJE	DORMIR	MOVIL	SO	TURNO
A1	Quinta opcion	8,5	75	7	Movistar/O2/T	Android	Mañana
A1	Quinta opcion	7,1	90	7	Movistar/O2/T	iOS	Mañana
A1	Segunda opcio	8,63	60	6	Digi Mobil	Android	Mañana
A1	Segunda opcio	8,62	35	8	Movistar/O2/T	iOS	Mañana

Por lo que para esta práctica se usará este archivo, disponible en encuesta_turno.csv.

Empezamos cargando en sendos vectores los valores de NOTA y TURNO.

```
> encuesta.con.turno = read.csv2("encuesta_turno.csv")

> (nota=encuesta.con.turno$NOTA)
[1] 8.50 7.10 8.63 8.62 8.20 8.70 7.21 7.63 8.40 8.30
[11] 8.20 9.10 9.79 10.11 8.02 7.31 7.50 8.71 8.34 9.21
[21] 7.80 10.30 7.99 6.90 7.80 10.00 8.59 7.00 8.05 10.80
[31] 7.99 8.55 7.34 6.75 9.56 7.42 6.94 7.21 7.68 10.28
[41] 7.86 10.26 7.27 5.80 7.30 7.14 8.60 7.50 8.00 7.54
[51] 7.29 7.83 6.75 9.81 6.80 6.44 6.65 7.80 10.27 7.60
[61] 7.87 7.00 7.08 7.48 8.07 5.82 6.50 9.90 7.50 6.50
[71] 9.46 8.00 7.80 7.65

> (turno=encuesta.con.turno$TURNO)
[1] "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana"
[11] "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana"
[21] "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana"
[31] "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana" "Mañana"
[41] "Mañana" "Mañana" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde"
[51] "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde"
[61] "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde" "Tarde"
[71] "Tarde" "Tarde" "Tarde" "Tarde"
```

¹ Si se creara una columna TURNO con R a partir de GRUPO, se podría hacer con los siguientes comandos:

```
encuesta = read.csv2("encuesta.csv")
encuesta.mañana=encuesta[ (( encuesta$GRUPO=="A1" ) | ( encuesta$GRUPO=="A2" ) ) , ]
encuesta.tarde=encuesta[ (( encuesta$GRUPO=="B1" ) | ( encuesta$GRUPO=="B2" ) ) , ]
encuesta.mañana$TURNO=rep(c("Mañana"), times=nrow(encuesta.mañana))
encuesta.tarde$TURNO=rep(c("Tarde"), times=nrow(encuesta.tarde))
encuesta.con.turno=rbind(encuesta.mañana, encuesta.tarde)
```

3. Calcular medidas estadísticas

Debemos realizar los cálculos de las medidas estadísticas que analizaremos en el informe, tanto a de la nota a nivel de todos los alumnos, como las medidas condicionadas de la nota según el valor de la variable turno.

Comenzando por el tamaño de la muestra:

```
> #Tamaño
> length(nota)
[1] 74
> tapply(nota, turno, length)
Mañana Tarde
    42    32
```

3.1 Medidas de centralización

```
> #Media
> round(mean(nota), 2)
[1] 8.02
> round(tapply(nota, turno, mean), 2)
Mañana Tarde
    8.35    7.59

> #Mediana
> round(median(nota), 2)
[1] 7.81
> round(tapply(nota, turno, median), 2)
Mañana Tarde
    8.2    7.5
```

3.2 Medidas de dispersión

```
> #Mínimo
> min(nota)
[1] 5.8
> tapply(nota, turno, min)
Mañana Tarde
    6.75    5.80

> #Máximo
> max(nota)
[1] 10.8
> tapply(nota, turno, max)
Mañana Tarde
    10.80    10.27

> #Rango
> max(nota) - min(nota)
[1] 5
> tapply(nota, turno, max) - tapply(nota, turno, min)
Mañana Tarde
```

4.05 4.47

```
> #Varianza (muestral)
> round(var(nota),2)
[1] 1.28
> round(tapply(nota,turno,var),2)
Mañana Tarde
  1.15  1.17

> #Desviación estándar
> round(sd(nota),2)
[1] 1.13
> round(tapply(nota,turno,sd),2)
Mañana Tarde
  1.07  1.08

> #Coeficiente de variación
> round(sd(nota)/mean(nota),2)
[1] 0.14
> round(tapply(nota,turno,sd)/tapply(nota,turno,mean),2)
Mañana Tarde
  0.13  0.14
```

3.3 Medidas de localización

```
> #Cuartiles
> round(summary(nota),2)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.80  7.28   7.82   8.02   8.60  10.80
> tapply(nota,turno,quantile)
$Mañana
  0%    25%    50%    75%   100%
 6.7500 7.5325 8.2000 8.7075 10.8000
$Tarde
  0%    25%    50%    75%   100%
 5.8000 6.9500 7.5000 7.9025 10.2700

> #Rango intercuartílico
> round(IQR(nota),2)
[1] 1.32
> round(tapply(nota,turno,IQR),2)
Mañana Tarde
  1.17  0.95
```

3.4 Medidas de forma

```
> install.packages("e1071")
> library(e1071)

> #Coeficiente de asimetría
> round(skewness(nota),2)
[1] 0.57
> round(tapply(nota,turno,skewness),2)
Mañana Tarde
```

0.59 0.80

```
> #Coeficiente de apuntamiento o curtosis
> round(kurtosis(nota),2)
[1] -0.28
> round(tapply(nota,turno,kurtosis),2)
Mañana Tarde
-0.64 0.31
```

3.5 Otras medidas

```
> #Valores inferiores a la media
> (nota.menores=length(nota[nota<mean(nota)]))
[1] 45
> round(100*nota.menores/length(nota))
[1] 61
> nota.mañana=encuesta.con.turno[encuesta.con.turno$TURNO=="Mañana",]$NOTA
> nota.tarde=encuesta.con.turno[encuesta.con.turno$TURNO=="Tarde",]$NOTA
> (nota.mañana.menores=length(nota.mañana[nota.mañana<mean(nota.mañana)]))
[1] 24
> round(100*nota.mañana.menores/length(nota.mañana))
[1] 57
> (nota.tarde.menores=length(nota.tarde[nota.tarde<mean(nota.tarde)]))
[1] 18
> round(100*nota.tarde.menores/length(nota.tarde))
[1] 56

> #Valores superiores a la media
> (nota.mayores=length(nota[nota>mean(nota)]))
[1] 29
> round(100*nota.mayores/length(nota))
[1] 39
> (nota.mañana.mayores=length(nota.mañana[nota.mañana>mean(nota.mañana)]))
[1] 18
> round(100*nota.mañana.mayores/length(nota.mañana))
[1] 43
> (nota.tarde.mayores=length(nota.tarde[nota.tarde>mean(nota.tarde)]))
[1] 14
> round(100*nota.tarde.mayores/length(nota.tarde))
[1] 44
```

4. Obtener tablas de frecuencias

Hay que decidir los intervalos de clase en los que agrupar los datos. Podemos agrupar en un número de intervalos o clases cercano al recomendando por la regla de Sturges, que en este caso es 8:

```
> #Regla de Sturges para una muestra de tamaño 74
> ceiling(1+3.322*log10(74))
[1] 8
```

Un número razonable de intervalos en este caso serían 6 intervalos iguales de 1 punto cada uno, desde la nota 5 hasta 11.

La tabla de contingencia para nota y turno con frecuencias absolutas se obtendría con los siguientes comandos:

```
> #Tabla de contingencia con 6 clases
> L=c(5,6,7,8,9,10,11)
> nota.agrupada=cut(nota, breaks=L, right=FALSE, include.lowest = TRUE)
> addmargins(table(nota.agrupada,turno))
```

nota.agrupada	turno		Sum
	Mañana	Tarde	
[5,6)	0	2	2
[6,7)	3	6	9
[7,8)	15	16	31
[8,9)	14	4	18
[9,10)	4	3	7
[10,11]	6	1	7
Sum	42	32	74

```
> round(addmargins(prop.table(table(nota.agrupada,turno))),2)
```

nota.agrupada	turno		Sum
	Mañana	Tarde	
[5,6)	0.00	0.03	0.03
[6,7)	0.04	0.08	0.12
[7,8)	0.20	0.22	0.42
[8,9)	0.19	0.05	0.24
[9,10)	0.05	0.04	0.09
[10,11]	0.08	0.01	0.09
Sum	0.57	0.43	1.00

Pero también podría ser interesante comprobar la distribución de frecuencias por nivel de calificación, con tres intervalos: [5,7) para Aprobado, [7,9) para Notable y [9,11] para Sobresaliente.

```
> #Tabla de contingencia con 3 clases
> L=c(5,7,9,11)
> nota.agrupada=cut(nota, breaks=L, right=FALSE, include.lowest = TRUE)
> addmargins(table(nota.agrupada,turno))
```

nota.agrupada	turno		Sum
	Mañana	Tarde	
[5,7)	3	8	11
[7,9)	29	20	49
[9,11]	10	4	14
Sum	42	32	74


```
> round(addmargins(prop.table(table(nota.agrupada, turno))), 2)
      turno
nota.agrupada Mañana Tarde  Sum
[5,7)         0.04  0.11 0.15
[7,9)         0.39  0.27 0.66
[9,11]        0.14  0.05 0.19
Sum           0.57  0.43 1.00
```

5. Dibujar diagramas

5.1 Diagramas de tarta

Se puede utilizar un diagrama de tarta para mostrar gráficamente el porcentaje de alumnos que pertenecen al turno de mañana y de tarde.

```
> #Diagrama de tarta de turnos
> par(mfrow=c(1,1))
> ft=table(turno)
> etiquetas.turno = c("Mañana", "Tarde")
> propt = round(ft/sum(ft)*100)
> etiquetas.turno = paste(etiquetas.turno, propt)
> etiquetas.turno = paste(etiquetas.turno,"%",sep="")
> pie(f,labels = etiquetas.turno, col=rainbow(2))
```

Podría ser interesante mostrar gráficamente los porcentajes de las calificaciones de Aprobado, Notable y Sobresaliente.

```
> #Diagramas de tarta de calificaciones
> par(mfrow=c(1,3))
>
> L=c(5,7,9,11)
> fn=table(nota.agrupada)
> etiquetas <- c("Aprobado", "Notable", "Sobresaliente")
> prop <- round(fn/sum(fn)*100)
> etiquetas= paste(etiquetas, prop)
> etiquetas = paste(etiquetas,"%",sep="")
> pie(fn,labels = etiquetas, col=rainbow(3), main="Calificaciones total alumnos")
>
> nota.agrupada.mañana=cut(nota.mañana, breaks=L, right=FALSE, include.lowest =
TRUE)
> fm=table(nota.agrupada.mañana)
> etiquetas <- c("Aprobado", "Notable", "Sobresaliente")
> prop <- round(fm/sum(fm)*100)
> etiquetas= paste(etiquetas, prop)
> etiquetas = paste(etiquetas,"%",sep="")
> pie(fm,labels = etiquetas, col=rainbow(3), main="Calificaciones turno Mañana")
>
> nota.agrupada.tarde=cut(nota.tarde, breaks=L, right=FALSE, include.lowest =
TRUE)
> ft=table(nota.agrupada.tarde)
> etiquetas <- c("Aprobado", "Notable", "Sobresaliente")
> prop <- round(ft/sum(ft)*100)
> etiquetas= paste(etiquetas, prop)
> etiquetas = paste(etiquetas,"%",sep="")
> pie(ft,labels = etiquetas, col=rainbow(3), main="Calificaciones turno Tarde")
```

5.2 Histogramas y diagramas de caja

Es interesante mostrar en una sola figura los histogramas y debajo los diagramas de caja (boxplots) para todos los alumnos y por separado para cada turno, utilizando el número de intervalo más próximo al recomendado por la regla de Sturges, que sería el caso de 6 intervalos.

```
> #Histogramas y diagramas de caja
> par(mfrow=c(3,3))
>
> L=c(5,6,7,8,9,10,11)
> hist(nota, breaks = L, main="", ylab = "Frecuencia", xlab = "Nota")
> abline(v=median(nota), lwd=3)
> abline(v=mean(nota), lwd=2, lty=2, col="darkred")
>
> hist(nota.mañana, breaks = L, main="", ylab = "Frecuencia", ylim=c(0,30), xlab =
"Nota turno Mañana")
> abline(v=median(nota.mañana), lwd=3)
> abline(v=mean(nota.mañana), lwd=2, lty=2, col="darkred")
>
> hist(nota.tarde, breaks = L, main="", ylab = "Frecuencia", ylim=c(0,30), xlab =
"Nota turno Tarde")
> abline(v=median(nota.tarde), lwd=3)
> abline(v=mean(nota.tarde), lwd=2, lty=2, col="darkred")
>
> boxplot(nota, horizontal = TRUE, ylim=c(5,11))
> segments(x0 = mean(nota), y0 = 0.8, x1 = mean(nota), y1 = 1.2, col = "darkred",
lwd = 2, lty=2)
>
> boxplot(nota.mañana, horizontal = TRUE, ylim=c(5,11))
> segments(x0 = mean(nota.mañana), y0 = 0.8, x1 = mean(nota.mañana), y1 = 1.2, col
= "darkred", lwd = 2, lty=2)
>
> boxplot(nota.tarde, horizontal = TRUE, ylim=c(5,11))
> segments(x0 = mean(nota.tarde), y0 = 0.8, x1 = mean(nota.tarde), y1 = 1.2, col
= "darkred", lwd = 2, lty=2)
```

6. Escribir el informe

Se debe crear un documento con los apartados habituales en este tipo de informes: Título, Introducción, Metodología, Resultados, Análisis de resultados, Conclusiones y Referencias.

6.1 Título

Debe elegirse un título que refleje claramente sobre qué trata el informe.

Ejemplo:

Análisis de las notas de acceso de los estudiantes del Grado en Ingeniería en Sistemas de Información de la Universidad de Alcalá del curso 2021-22.

6.2 Introducción

En informe comienza situando el trabajo en su contexto, y a continuación se plantea el objetivo del mismo y su alcance.

Ejemplo:

Contexto

En la Universidad de Alcalá se imparten diferentes estudios de grado relacionados con las Tecnologías de la Información y las Comunicaciones, entre los que se encuentra el Grado en Ingeniería en Sistemas de Información. Dado el elevado número de alumnos existente, los estudios se organizan en dos turnos: Mañana y Tarde, y en cada turno los alumnos asisten a un grupo de teoría de cada asignatura, pero se reparten en dos grupos diferentes para realizar las prácticas.

Objetivo

El objetivo de la investigación llevada a cabo ha sido “realizar un análisis estadístico descriptivo de las notas de acceso a la Universidad de Alcalá de los alumnos del Grado en Ingeniería en Sistemas de Información del curso 2021-22, comprobando si hay diferencias entre los alumnos del turno de mañana y de tarde”.

Alcance

La investigación se limita a los alumnos del curso 2021-22 matriculados en la asignatura Estadística de primero del Grado en Ingeniería en Sistemas de Información, y se ha utilizado una muestra de 74 alumnos de los 108 matriculados en la asignatura.

6.3 Metodología

Hay que describir el origen de los datos, las herramientas utilizadas y el diseño de la investigación, que incluye la definición de la variable estadística y si se ha realizado manipulaciones previas para organizar los datos como, por ejemplo, descarte de valores, agrupamiento o división de los datos en clases o categorías.

Ejemplo:

Origen de los datos

Se ha trabajado con un archivo “encuesta.csv” que incluye las notas de la muestra, así como la información sobre el grupo al que pertenece cada alumno, siendo A1 y A2 los grupos del turno de mañana, y B1 y B2 los del turno de tarde. Se ha realizado una encuesta a toda la población de alumnos mediante una encuesta elaborada por el profesor durante la primera semana del curso.

Herramientas

Se han procesado los datos utilizando la aplicación RStudio para Windows, versión 2022.07.1 y los paquetes de funciones: “modeest” para calcular la moda, y “e1071” para calcular los coeficientes de asimetría y apuntamiento.

Diseño

El resumen del diseño es el siguiente:

- *Variables estadísticas: Se ha analizado la variable estadística “Nota de acceso”, de tipo cuantitativo discreto, con valores positivos que pueden ser mayores o iguales a 5. También se ha utilizado la variable estadística “Turno”, de tipo cualitativo, con valores “Mañana” o “Tarde”.*
- *Población: 108 alumnos matriculados en la asignatura Estadística del Grado en Ingeniería en Sistemas de Información de la Universidad de Alcalá en el curso 2021-22.*
- *Muestra: 74 alumnos, de los cuales 42 son del grupo de mañana y 32 del grupo de tarde.*
- *Organización de datos: Se han eliminado del archivo “encuesta.csv” los datos erróneos, resultando un total de 74 observaciones válidas. Se ha modificado el archivo añadiendo una nueva columna TURNO con el valor “Mañana” en el caso de los alumnos de los grupos A1 y A2, y el valor “Tarde” en los alumnos de los grupos B1 y B2.*

6.4 Resultados

En este apartado se deben presentar los resultados obtenidos, como medidas de centralización, dispersión, localización y forma; las tablas de frecuencias; y los diagramas. Tanto de la variable estadística para la muestra completa como para los diferentes grupos en los que se ha organizado.

Ejemplo:

Medidas estadísticas

En la tabla 1 se muestran los resultados de las medidas estadísticas de centralización, dispersión, localización y forma más relevantes, para el total de alumnos de la muestra, y para los alumnos de los turnos de mañana y tarde.

Tabla 1. Medidas estadísticas descriptivas

Medida	Todos	Mañana	Tarde
Tamaño muestra	74	42	32
Medidas de centralización			
Media	8.02	8.35	7.59
Mediana	7.81	8.20	7.50
Medidas de dispersión			
Mínimo	5.80	6.75	5.80
Máximo	10.80	10.80	10.27
Rango	5.00	4.05	4.47
Varianza	1.28	1.15	1.17
Desviación estándar	1.13	1.07	1.08
Coefficiente de variación	0.14	0.13	0.14
Medidas de localización			
Primer cuartil (Q1)	7.27	7.53	6.95
Tercer cuartil (Q3)	8.60	8.71	7.90
Rango intercuartílico (IQR)	1.32	1.17	0.95
Medidas de forma			
Coefficiente de asimetría	0.57	0.59	0.80
Coefficiente de apuntamiento (curtosis)	-0.28	-0.64	0.31
Otras			
Valores inferiores a la media	45 (61%)	24 (57%)	18 (56%)
Valores superiores a la media	29 (39%)	18 (43%)	14 (44%)

Tablas de contingencia

En la tabla 2 se muestra la tabla de contingencia para las variables nota y turno en el caso de agrupar las notas en 6 intervalos. En la tabla se representa con f las frecuencias absolutas y con h las frecuencias relativas.

Tabla 2. Tabla de contingencia para nota y turno (con 6 intervalos)

NOTA/TURNO	Mañana	Tarde	f_{NOTA}	h_{NOTA}
[5, 6)	0	2	2	0.03
[6,7)	3	6	9	0.12
[7, 8)	15	16	31	0.42
[8,9)	14	4	18	0.24
[9,10)	4	3	7	0.09
[10,11]	6	1	7	0.09
f_{TURNO}	42	32	74	
h_{TURNO}	0.57	0.43		1

En la tabla 3 se muestra la tabla de contingencia para las variables nota y turno agrupando la nota en 3 intervalos, diferenciando la calificación de aprobado (nota entre 5 y 7 puntos), de notable (entre 7 y 9 puntos) y de sobresaliente (más de 9 puntos).

Tabla 3. Tabla de contingencia para nota y turno (con 3 intervalos)

NOTA/TURNO	Mañana	Tarde	f_{NOTA}	h_{NOTA}
[5, 7) Aprobado	3	8	11	0.15
[7, 9) Notable	29	20	49	0.66
[9, 11] Sobresaliente	10	4	14	0.19
f_{TURNO}	42	32	74	
h_{TURNO}	0.57	0.43		1

Diagramas

En la figura 1 se muestra el porcentaje de alumnos por turno, el número total de alumnos es 74.

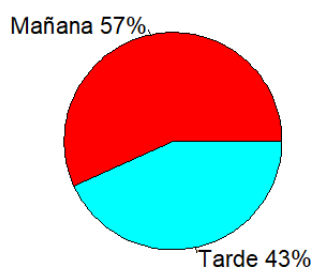


Figura 1. Diagrama con el porcentaje de alumnos por grupo

En la figura 2 se muestran los histogramas y diagramas de caja (boxplot) de la variable nota de acceso, en el caso de los histogramas, agrupada en seis intervalos, considerando todos los alumnos (izquierda), sólo los alumnos del turno de mañana (centro) y los alumnos del turno de tarde (derecha). En los diagramas aparecen tanto la mediana (línea negra continua) como la media (línea roja discontinua).

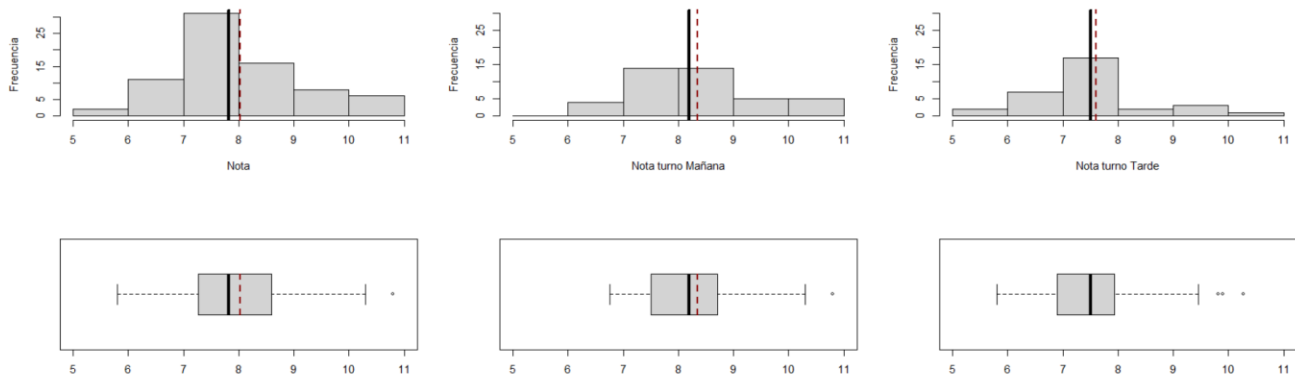


Figura 2. Histograma y diagramas de caja de la nota de todos los alumnos y según el turno

En la figura 3 se muestran los porcentajes de alumnos que han obtenido una calificación de Aprobado (entre 5 y 7 puntos), de Notable (entre 7 y 9 puntos) y de Sobresaliente (más de 9 puntos).



Figura 3. Diagramas de las calificaciones de los alumnos

6.5 Análisis de resultados

En este apartado hay que realizar la interpretación de los resultados obtenidos, resaltando las diferencias entre los grupos en los que se haya dividido la muestra. Debe hacerse referencia a todas las tablas y figuras incluidas en el apartado de resultados.

Ejemplo

Hay un mayor número de alumnos del grupo de mañana en la muestra, pues como se representa en la figura 1 y se indica en la tabla 1, un 57% de los 74 alumnos de la muestra pertenecen al turno de mañana (42 alumnos) y un 43% al turno de tarde (32 alumnos).

De los resultados obtenidos (tabla 1) y de la observación de los histogramas y diagramas de caja (figura 2), puede comprobarse que tanto la media como la mediana son superiores en el turno de mañana, donde la media es de 8.35 puntos y la mediana 8.2, mientras que en el turno de tarde la media es 7.59 y la mediana 7.5. En ambos casos la media es algo superior a la mediana, lo que indica que hay más alumnos con una nota menor que la media, lo cual se confirma con el coeficiente de asimetría, que es positivo en ambos turnos, por lo que en los dos casos hay una asimetría hacia la derecha, como se aprecia en los histogramas de la figura 2. Así, por la mañana el 57% de los alumnos tiene una nota inferior a la media, y por la tarde un 56%, que son valores muy similares.

Se observa que el rango en el caso del turno de mañana (4,05) es menor que en el caso del turno de tarde (4,47). La desviación estándar de ambos turnos es similar, pero el coeficiente de variación es menor en el turno de mañana (0,13) que en el de tarde (0,14). Lo que evidencia que las notas del turno de mañana son algo más homogéneas que las del turno de tarde.

La mayor dispersión del turno de tarde también se puede comprobar en los diagramas de caja (figura 2), donde en el diagrama del turno de tarde hay tres datos atípicos, mientras que por la mañana sólo hay uno. A la vista del diagrama de caja, y por los valores del primer y tercer cuartil en la tabla 1, se observa que la mitad (50%) de los alumnos de la mañana tiene una nota entre 7.53 y 8,71, mientras que el 50% de los de la tarde tiene una nota entre 6,95 y 7,90; por lo que en el caso de estos valores cercanos a la media hay más dispersión en el turno de mañana, ya que su rango intercuartílico es 1,17, frente a 0,95 en el turno de tarde.

La mayor dispersión en ambos turnos ocurre en las notas altas, viéndose en el diagrama de caja que los bigotes superiores son más largos, y destacando el dato atípico de un alumno del turno de mañana con una nota muy por encima de las demás (10,80).

El coeficiente de apuntamiento del turno de mañana es negativo (-0,64) y el de la tarde es positivo (+0,31), lo que indica que la distribución de frecuencias por la mañana es “platicúrtica” o aplanada, como lo confirma su histograma (figura 2); mientras que la de la tarde es “leptocúrtica”, es decir, tiene un mayor apuntamiento. Lo que indica que por la mañana las notas están más repartidas entre los diferentes intervalos considerados en el histograma.

En relación con las calificaciones de Aprobado (notas mayores o iguales a 5 y menores que 7), Notable (mayores o iguales a 7 y menores que 9) y Sobresaliente (superiores o iguales a 9), considerando la muestra completa o por turnos, como se muestra en la figura 3, la mayoría de los alumnos tiene una calificación de Notable: 66% en el caso de todos los alumnos, un 69% en turno de mañana y un 62% en el turno de tarde. Destaca que en turno de mañana hay muy pocos alumnos con calificación de Aprobado (7%), mientras que hay muchos más con calificación de Sobresaliente (24%).

6.6 Conclusiones

En este apartado debe realizarse un resumen de la investigación explicando los hallazgos más relevantes.

Ejemplo

A partir del análisis de los resultados obtenidos, puede concluirse que las notas de acceso de los alumnos en general son elevadas, destacando el turno de mañana, con una media superior al de tarde y un mayor gran porcentaje de alumnos con calificaciones de Notable y Sobresaliente. Los datos son homogéneos, con una dispersión no muy acusada y una tendencia a valores de notas inferiores a la media, aunque existen algunos datos atípicos relevantes, destacando un alumno del turno de mañana con una nota claramente superior a la del resto.

6.7 Referencias

Si en el informe se cita algún documento o página web, en este apartado debe incluirse la referencia completa, utilizando un formado normalizado, como por ejemplo el formato [APA](#).

7. Referencias

Se recomienda consultar los siguientes documentos sobre elaboración de informes estadísticos descriptivos.

1. Examples.com (2022). 10+ Statistics Report Examples. Disponible en <https://www.examples.com/business/report/statistics-report.html>
2. Lafuente, M. (2015). El Informe Estadístico. Disponible en <https://slideplayer.es/slide/3197061/>
3. Mueller, J. (2020). How to Write a Statistical Report. Disponible en <https://www.wikihow.com/Write-a-Statistical-Report>
4. Velasco, J. (2018). Informe estadístico: Comparación de las notas obtenidas por los alumnos de metodologías y estrategias logísticas año 2017. Disponible en <https://es.slideshare.net/jorgevelascopizarro/ejemplo-informe-estadistico>
5. Caicedo, E. (2023). Proyecto Premier League. Disponible en <https://rpubs.com/Educa0698/Premier>