

Horas de uso diarias de los portátiles

Mario Ramos Salsón y Sergio Verde Llorente

INTRODUCCIÓN Y OBJETIVOS

Se pretende estudiar distintos datos sobre la gente que utiliza un portátil para ver como influyen en las horas de uso. Entre estos encontramos la edad del usuario, la ocupación del encuestado, la marca del portátil y en que emplean su tiempo de uso.

Descripción de los datos:

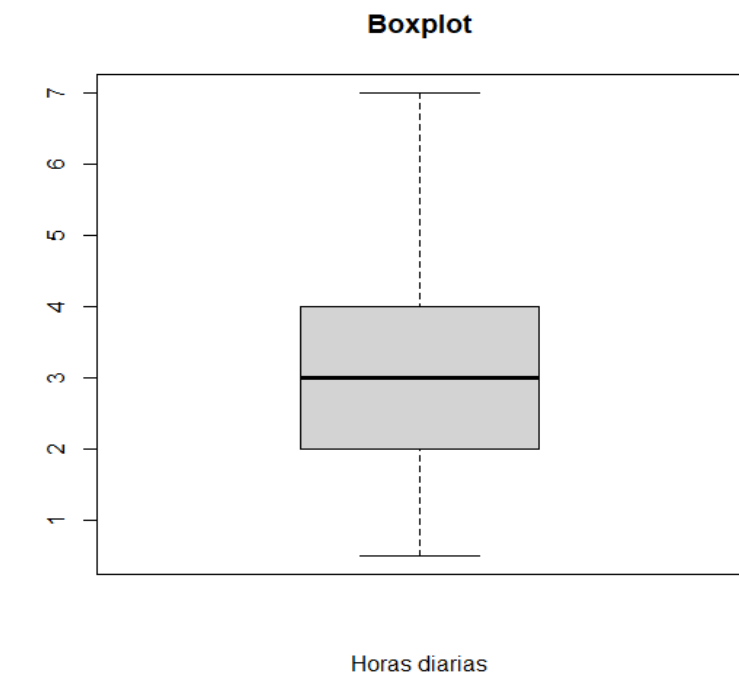
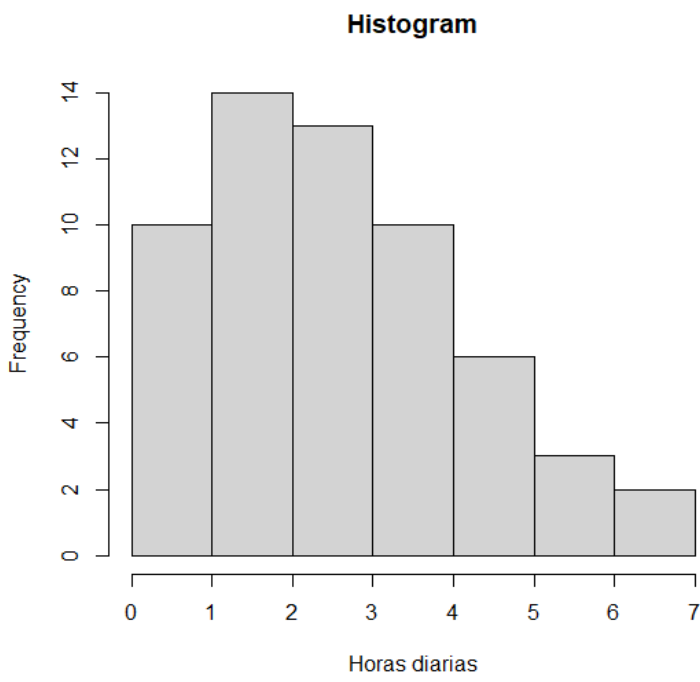
- Número de observaciones muestrales: 58
- Fuente: Encuesta realizada por nosotros a través de Google Forms

Variables:

- Variable principal:**
 - Horas de uso diario
- Variables secundarias:**
 - Edad
 - Ocupación
 - Marca
 - Actividades para las que se usa

ESTUDIO DESCRIPTIVO

RESUMEN VARIABLE PRINCIPAL	
N	58
Media	3,06
Desviación típica	1,64
Q1	2
Mediana(Q2)	3
Q3	4
Rango intercuartico	2
Max	7
Min	0,5
Rango	6,5
CV	0,54(54%)

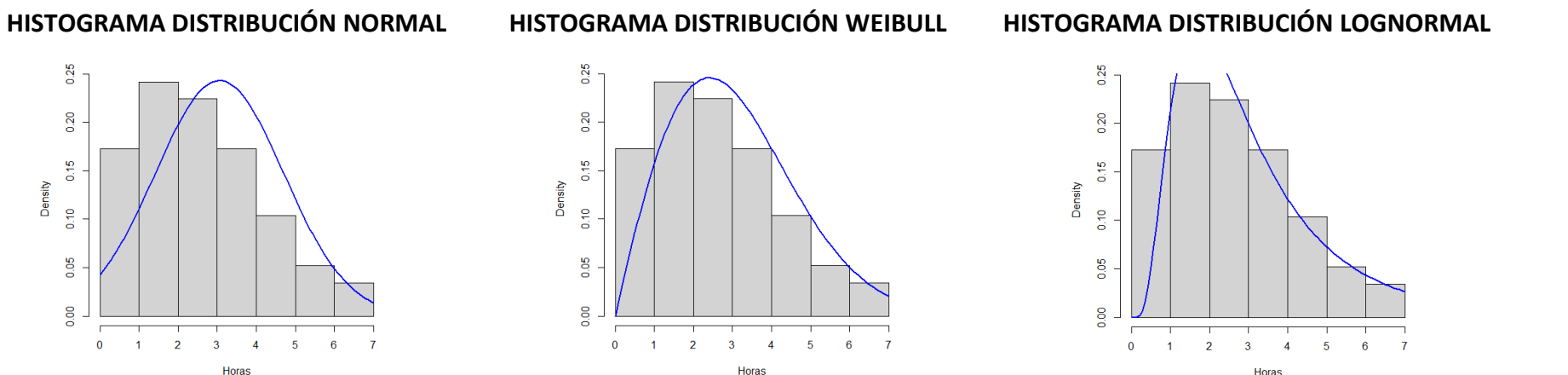


Realizando el análisis descriptivo de nuestra variable principal con 58 valores, obtenemos los valores más importantes y relevantes que agrupamos en la tabla de arriba.

Analizando el “*Boxplot*”, podemos observar que no tenemos valores atípicos, esto se debe a que las barreras internas están por encima de cualquier valor recogido, ya que $Q3 + 1.5 \cdot RI$ es mayor que nuestro valor máximo y $Q1 - 1.5 \cdot RI$ es menor que nuestro valor mínimo recogido.

Observando el histograma podemos apreciar que los valores que más se repiten están en torno al 2 y al 3 por ello la media es 3.06 y la mediana 3.

AJUSTE DE DISTRIBUCIONES



Como se puede apreciar comparando los 3 histogramas, la distribución que más se acerca a nuestra encuesta es la de Weibull.

Los valores obtenidos de dicha distribución son los siguiente: X-Squared: 7.4702 / DF = 6 / P-VALUE = 0.2795.

Haciendo uso de la función “*pchisq*” con los valores obtenidos arriba, conseguimos un P-VALUE real de 0.1130305. Estos valores si son correctos, ya que están dentro del rango establecido de 0.05. Donde todo valor que se encuentre por debajo de él, se considerará incorrecto y todo el que se encuentre por encima de él será correcto.

Como se puede observar en la tabla de la derecha donde recogemos todos los valores de las distribuciones, apreciamos que la lognormal está muy cerca de ese objetivo de 0,05, que la normal está algo más lejos y que la distribución que mejor se adecua es la de Weibull.

	X-squared	P-value
Normal	8,4482	0,2071
Lognormal	10,758	0,09616
Weibull	7,4702	0,2795

INTERVALOS DE CONFIANZA Y CONTRASTE DE HIPÓTESIS

INTERVALOS DE CONFIANZA PARA LAS HORAS DIARIAS

Intervalo de confianza para la media al 95%: UpperLimit = 3.482727
LowerLimit = 2.63793

Intervalo de confianza para la desviación típica al 95%: UpperLimit = 1.387515
LowerLimit = 2.009391

Como se puede apreciarb comparando estos intervalos con los valores de la media y la desviación típica obtenidos en el segundo apartado, observamos ambos se encuentran dentro del rango establecido.

TEST DE HIPÓTESIS PARA LAS HORAS DIARIAS

El primer test realizado, hemos usado como media nuestra propia media: 3.06. Esto nos ha proporcionado una “t estadística” de 0.0016, un p-value de 0.9987 y un p-value real de 0.49936.

El segundo test realizado, hemos usado como media un valor lejos del intervalo de confianza: 1.8. Esto nos ha proporcionado una “t estadística” de 5.8483, un p-value de 2.545e-07 y un p-value real de 2.4831e-09.

Analizando estos datos, concluimos que al dar un valor dentro del rango de confianza, obtenemos un p-value correcto. Esto no sucede cuando damos un valor fuera del rango ya que obtenemos un p-value erróneo.

REGRESIÓN MÚLTIPLE

REGRESIÓN MÚLTIPLE INICIAL

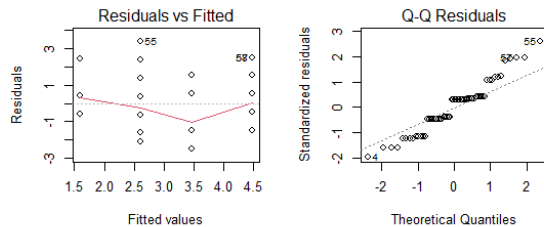
- Variable dependiente: Horas de uso
- Variable independientes:
 - Edad
 - Ocupación (Estudiante, Trabajador, Ninguno)
 - Marca (Lenovo, Dell, Asus, Acer, HP, Huawei, Apple, MSI, Otros)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.70628	1.98446	1.868	0.0683 .
edad	0.03437	0.06394	0.538	0.5935 .
ocupacion_textoEstudiante	-1.54228	0.85638	-1.801	0.0784 .
ocupacion_textoNinguno	-2.33668	2.12481	-1.101	0.2769 .
ocupacion_textoTrabajador	-0.22554	0.92405	-0.244	0.8083 .
marca_textoApple	-0.06540	0.72980	-0.090	0.9290 .
marca_textoAsus	-0.98914	0.77411	-1.278	0.2079 .
marca_textoDell	-0.55204	0.72442	-0.762	0.4500 .
marca_textoHP	-1.08633	0.61668	-1.762	0.0849 .
marca_textoHuawei	0.88311	0.92426	0.953	0.3455 .
marca_textoLenovo	-0.57362	0.61624	-0.931	0.3569 .
marca_textoMSI	0.46032	0.83184	0.553	0.5827 .
marca_textoOtros	-0.85149	1.43111	-0.595	0.5548 .

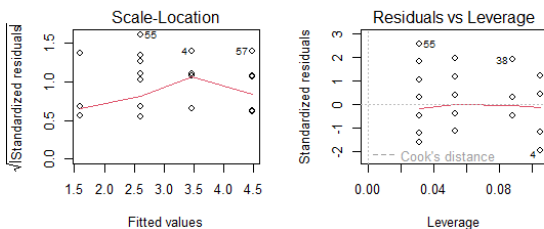
REGRESIÓN MÚLTIPLE FINAL

En la regresión inicial teníamos muchas variables secundarias no significativas. Eliminando de una en una cada variable no significativa prestando atención a si en el proceso se creaban nuevas variables significativas, nos quedamos con las siguientes:

Ocupacion1 == Estudiante (para que R lo interprete)
Marca4 == HP (para que R lo interprete)



Tras todo el proceso, nos hemos acabado quedando con que las variables que más influyen en las horas de uso del portátil son la marca HP y que el entrevistado sea un estudiante. Eso quiere decir que la gente entrevistada que tenia como portátil un HP, tenían muchas horas de uso, al igual que pasa con los entrevistados que eran estudiantes. Esto no pasa con los trabajadores, ya que estos tiene horas de uso variadas dependiendo de su oficio.



CONCLUSIONES

Como conclusión a nuestro estudio estadístico, sacamos que a la hora de analizar el uso del portátil diario en función de la ocupación, la marca, la edad y la actividad que se realice con el, no todas estas influyen.

Nos ha sorprendido que la variable “Edad” no haya estado presente entre las más significativas, ya que asociamos una edad más joven a un mayor uso de las tecnologías y los dispositivos electrónicos. Pero creemos que esto puede llegar a ser lógico, ya que las personas de mayor edad entrevistadas se movían en el sector tecnológico.

Cabe destacar que la variable estudiante si es significativa ya que el formulario ha sido distribuido principalmente entre estudiantes universitarios.

Por último la marca que más relación tiene con el tiempo de uso es HP. Esto se puede llegar a deber a que dicha marca cuente con una mayor autonomía o unas baterías más potentes.