

Práctico 2

Psicoestadística - Facultad de Educación - UCC

Primera Parte

En las siguientes actividades se trabajará principalmente con información extraída de las siguientes bases de datos.

- Base de datos de una encuesta de evaluación de CEPRAM. CEPRAM (Centro de Promoción del Adulto Mayor) es una fundación destinada a la promoción de la salud de los adultos mayores. Una de las actividades principales desarrollada por CEPRAM es el dictado de cursos en distintas áreas de conocimiento. (Para más información ver www.cepram.org.ar) Su utilización se realiza con permiso de CEPRAM.
 - Base de datos de la Ecuesta Permanente de Hogares (EPH) realizada por el Instituto Nacional de Estadística y Censos (INDEC) correspondiente al primer trimestre de 2017. EPH es un programa nacional de producción sistemática y permanente de indicadores sociales, que permite conocer las características sociodemográficas y socioeconómicas de la población. Es una base de datos de acceso público. <http://www.indec.gob.ar/bases-de-datos.asp>.
1. De la muestra correspondiente a la EPH se seleccionaron las personas mayores a 18 años que no tuvieron ningún nivel de instrucción educativo en la region del gran Córdoba. A continuación observamos las edades para esos casos

caso	Edad	Lee
1	61	No
2	48	No
3	56	No
4	65	No
5	38	No
6	62	No
7	68	No
8	63	Si
9	63	No
10	76	No
11	63	Si
12	33	No
13	57	No
14	60	No
15	82	Si
16	62	No

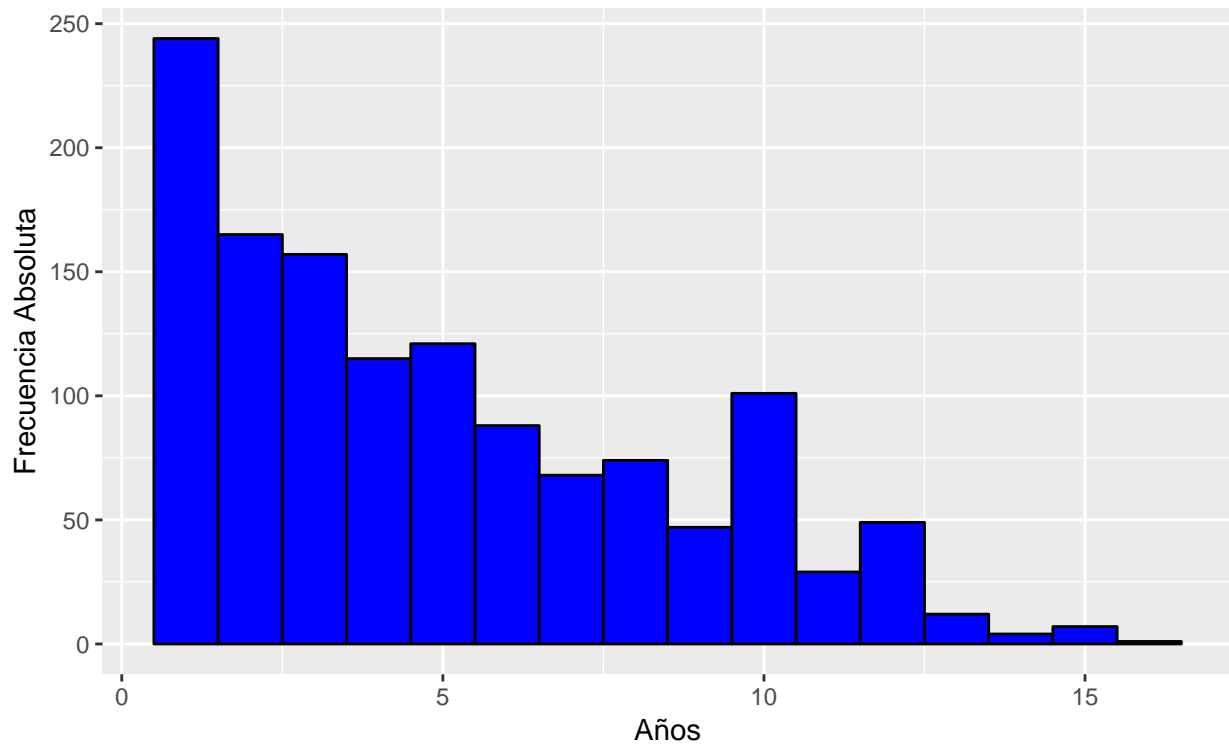
- a) Calcule la **media**, la **varianza** y la **desviación estándar** para la variable Edad.
- b) Realice una **tabla de distribución de frecuencias** para la variable **Edad**. Tome como referencia la siguiente tabla

Edad	f	F	f'	F'
[30, 40]				
(40, 50]				
(50, 60]				
(60, 70]				
(70, 80]				

- c) Identifique los intervalos que contienen a la **mediana** y la **moda** en la tabla anterior. Realice alguna descripción de lo observado. Por ejemplo: La mediana se encuentra entre los valores ... Esto quiere decir que el 50 % de los casos se encuentra aproximadamente por debajo de ...
- d) Realice un **tabla de distribución de frecuencias** para la variable **Lee**. En este caso, no es necesario que calcule las frecuencias acumuladas (F y F').
2. La siguiente es una tabla de distribución de frecuencias para la variable nivel educativo para personas mayores de 18 años de la EPH:

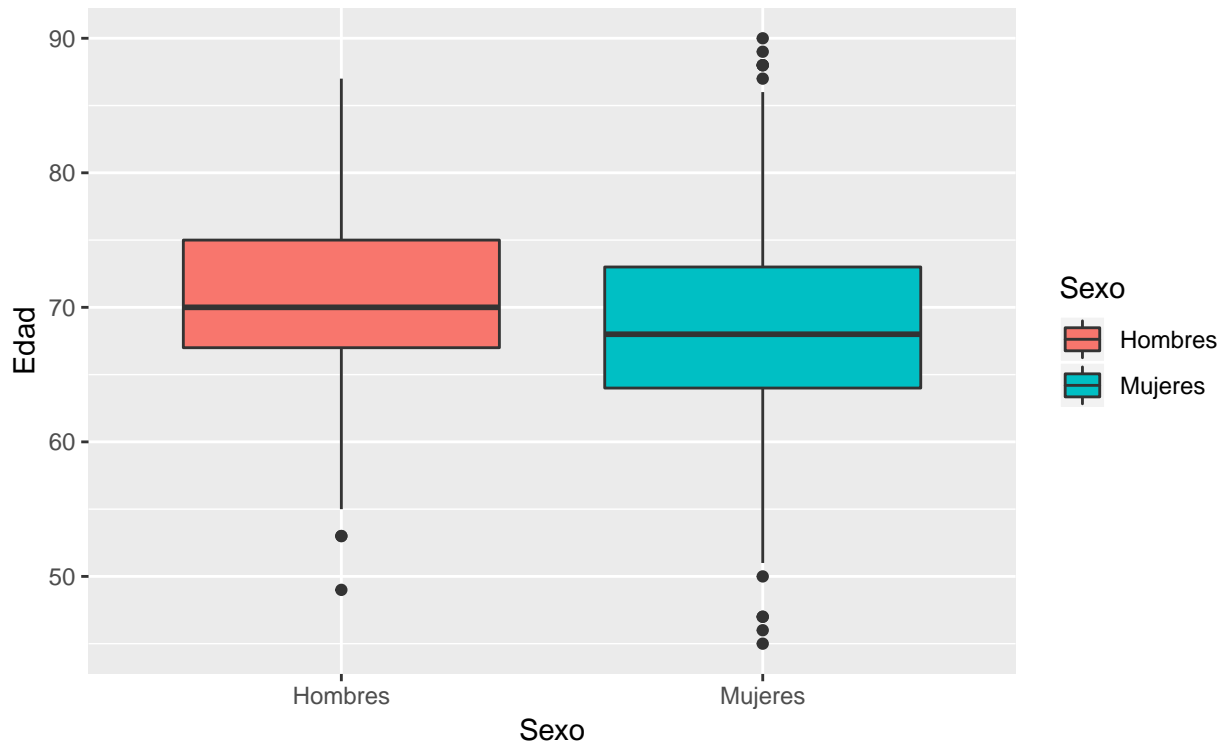
Nivel	f	Fa	Fr	Fra
Sin instrucción	369	369	0.01	0.01
Primaria Incompleta	2578	2947	0.06	0.07
Primaria Completa	7960	10907	0.19	0.26
Secundaria Incompleta	6734	17641	0.16	0.43
Secundaria Completa	10832	28473	0.26	0.69
Superior Universitaria Incompleta	6060	34533	0.15	0.84
Superior Universitaria Completa	6780	41313	0.16	1.00

- a) ¿Qué nivel de medición tiene la variable? Identifique el **modo**, la **mediana**.
- b) ¿Qué porcentaje de la muestra tiene un nivel educativo inferior o igual al Primario Completo? ¿Qué porcentaje tiene un nivel Universitario Incompleto o Superior?
- c) ¿Cual es la suma de frecuencias de personas con formación universitaria completa e incompleta?
3. La siguiente imagen es un histograma para los años que llevan asistiendo a la institución los participantes de la encuesta de CEPAM.



Identifique el intervalo que contiene al **modo**. Especifique si la distribución es **simétrica**, **asimétrica a la izquierda** o **asimétrica a la derecha**.

4. El siguiente gráfico es un boxplot (también llamado gráfico de cajas o bigotes) para la Edad de los participantes del CEPRAM, particionado por sexo.



Describa el gráfico. Además, indique qué representan los extremos de las cajas, los extremos de los bigotes, la línea horizontal al medio de la caja y los puntos extremos.

Segunda parte

1. Ingrese a <http://aprenderdatos.educacion.gob.ar/binarg/RpWebEngine.exe/Portal?lang=esp>. Acceda a los datos de 5to o 6to año del nivel secundario (Diríjase a la pestaña Aprender -> Nivel Secundario -> 5to o 6to Año del Nivel Secundario). Responda las siguientes preguntas:
 - a. Vaya a “Característica de los estudiantes -> Características del estudiante y trayectoria educativa” y solicite un análisis por Sexo/Género ¿Qué cantidad de personas realizaron la prueba Aprender en Córdoba? ¿Cuál es el porcentaje de hombres y cuál el porcentaje de mujeres? Copie y pegue la tabla de resultados.
 - b. Vaya a “Característica de los estudiantes -> Acceso a TICS”. Analice la variable “Tenencia de conexión a Internet” para la ciudad de Córdoba. ¿Qué porcentaje de alumnos evaluados posee y qué porcentaje no posee conexión a Internet? Considere la variable “Edad cuando empezó a usar celular/computadora/Tablet/netbook/notebook” ¿Qué porcentaje de alumnos comenzó a utilizar la computadora antes de los 10 años? Copie y pegue los gráficos de resultados.

Imagine que está realizando un estudio en el que desea explorar variables que puedan estar correlacionadas con el rendimiento en Matemática. Para realizar el análisis considere toda la base de datos.

- c. Realice una tabla de contingencia para el rendimiento en Matemática (desagregado en 4 categorías) en función de “Edad cuando empezó a usar celular/computadora/Tablet/netbook/notebook”. Calcule los porcentajes marginales por fila. Interprete la salida. ¿Que sugieren los resultados? ¿El acceso temprano a los dispositivos electrónicos está relacionado con el rendimiento en Matemática? Justifique. Copie y pegue la tabla de resultados.
 - d. Realice una tabla de contingencia para el rendimiento en Matemática (desagregado en 4 categorías) en función de “Autoconcepto en Matemática”. Calcule los porcentajes marginales por fila. Interprete la salida. ¿Que sugieren los resultados? ¿El autoconcepto en Matemática está relacionado con el rendimiento en Matemática? Justifique. Copie y pegue la tabla de resultados.
2. La Encuesta Permanente de Hogares (EPH) es un programa nacional de producción sistemática y permanente de indicadores sociales de propósitos múltiples. Permite conocer características sociodemográficas y socioeconómicas de la población. Actualmente se realiza con una periodicidad de 3 meses y abarca 31 conglomerados urbanos. Descargue la base de datos de la EPH correspondiente al 2do Trimestre de 2017. (<https://www.indec.gob.ar/bases-de-datos.asp>). Abra la base usu_individual.

- a. Ralice una tabla de distribución de frecuencias y un gráfico de barras para la variable CH12 (“¿Cuál es el nivel más alto que cursa o cursó?”). Indique el nivel de medición de la variable.

Imagine que está interesado en estudiar la inserción laboral y la capacidad adquisitiva de personas que recibieron educación especial.

- b. Seleccione en la base los casos que recibieron Educación Especial ($CH12 = 9$) y son mayores a 18 años ($CH06 \geq 18$). Construya y guarde una nueva tabla con estos datos. De aquí en adelante trabajará con esa nueva tabla.
- c. Construya un histograma para la Edad (CH06).
- d. Elabore un boxplot para el ingreso total individual (P47T) particionado por sexo (CH04). Interprete los resultados. Indique los cuartiles, la mediana, los valores mínimos y máximos y los resultados extremos.
- e. Realice una tabla de distribución de frecuencias para la variable ESTADO y CAT_INAC. Describa los resultados.