Trabajo Práctico N°4

Ejercicio 4.1:

Los datos siguientes corresponden a la determinación de Homocisteína sérica (Hcy) en ayunas en 6 mujeres con diabetes mellitus tipo 2. La Hcy se midió en mol/L por cromatografía líquida de alta performance.

11,4	11,8	13,4	12,2	13,0	12,5

en orden*: *11,4 - 11,8 - 12,2 - 12,5 - 13,0 - 13,4

Calcular la mediana, la media, el mínimo y máximo, el rango, la varianza, la desviación estándar y el coeficiente de variación.

Mediana ⇒ Med(x):= $\frac{12{,}2 + 12{,}5}{2}$= **12,35**

Media ⇒ $(\overline{x})$:= $\frac{11{,}4 + 11{,}8 + 13{,}4 + 12{,}2 + 13{,}0 + 12{,}5}{6}$≈**12,383**

Mínimo y Máximo ⇒ *Xmin*= **11,4** ; *Xmax* = **13,4**

Rango ⇒ rg(x):= *Xmax* - *Xmin* = 13,4 - 11,4 = **2**

Varianza ⇒ $sx^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_n - \overline{x})^2 := \widehat{\sigma}^2 = \widehat{Var(x)}$

$= \frac{(11{,}4 - 12{,}383)^2 + (11{,}8 - 12{,}383)^2 + (12{,}2 - 12{,}383)^2 + (12{,}5 - 12{,}383)^2 + (13{,}0 - 12{,}383)^2 + (13{,}4 - 12{,}383)^2}{(6-1)} =$

0,5537

Desviación estándar ⇒ sx = $\sqrt{0,5537}$= **0,7441**

Coeficiente de variación ⇒ CV:= $\frac{sx}{\overline{x}} = \frac{0{,}7441}{12{,}383}$= 0,0601 en porcentaje **6,01%**

Ejercicio 4.2:

En una institución se seleccionaron aleatoriamente 7 pacientes con deterioro cognitivo leve (DCL) y se les efectuó la prueba del Addenbroke cognitive examination (ACE), que tiene un valor máximo posible de 100. Se obtuvieron las siguientes mediciones:

96	93	91	96	82	88	98

en orden:* *82 - 88 - 91 - 93 - 96 - 96 - 98

Calcular la mediana, la media, el mínimo y máximo, el rango, la varianza, la desviación estándar y el coeficiente de variación.

pipiMorch

Mediana ⇒ Med(x):= **93**

Media ⇒ $(\bar{x}) := \dfrac{82+88+91+93+96+96+98}{7} \approx$ **92**

Mínimo y Máximo ⇒ *Xmin*= **82** ; *Xmax* = **98**

Rango ⇒ rg(x):= *Xmax* - *Xmin* = 98 - 82 = **16**

Varianza ⇒ $sx^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_n - \bar{x})^2 := \hat{\sigma}^2 = \widehat{Var(x)}$

$= \dfrac{(82-92)^2 + (88-92)^2 + (91-92)^2 + (93-92)^2 + (96-92)^2 \cdot 2 + (98-92)^2}{(7-1)} =$ **31**

Desviación estándar ⇒ $sx = \sqrt{31} =$ **5,5678**

Coeficiente de variación ⇒ $CV := \frac{sx}{\bar{x}} = \frac{5{,}5678}{92} =$ 0,0605 en porcentaje **6,05%**

Ejercicio 4.3:

a) Se seleccionaron aleatoriamente 10 mujeres normales y 10 mujeres con poliquistosis ovárica (PQO) y se midió la hormona folículo estimulante (FSH) en mUI/ml por el método de EQLIA. Se obtuvieron las siguientes mediciones (se encuentran ordenadas en forma ascendente para cada muestra):

Normales	6,5	6,9	7,0	7,6	7,9	8,1	8,4	8,5	8,8	9,1
PQO	5,2	5,8	6,3	6,5	6,8	6,8	6,9	7,1	7,2	7,6

Calcular, para cada uno de los dos grupos, los estadísticos descriptivos: mediana, media, mínimo y máximo, rango o amplitud, varianza, desviación estándar y coeficiente de variación.

Realizado con calculadora [mode(SD)]

Grupo de mujeres normales:

Mediana: ***Med(x)***$=\frac{7{,}9+8{,}1}{2}$ = **8**

Media $(\bar{x})$= **7,88**

Mínimo y Máximo= *Xmin*= **6,5** *Xmax*= **9,1**

Rango o amplitud= *rg(x):= Xmax - Xmin* = 9,1 - 6,5 = **2,6**

Varianza= $\widehat{Var(x)}$ = **0,7507**

Desviación estándar (sx)= **0,8664**

Coeficiente de variación= CV:= $\frac{sx}{\bar{x}} = \frac{0,8664}{7,88} =$ 0,1099 en porcentaje **10,99%**

Grupo de mujeres con poliquistosis ovárica (PQO):
Mediana: ***Med(x)*= 6,8** (se repite)

Media $(\bar{x})$**= 6,62**

Mínimo y Máximo= *Xmin=* **5,2** *Xmax=* **7,6**

Rango o amplitud= *rg(x):= Xmax - Xmin* = 7,6 - 5,2 = **2,4**

Varianza= $\widehat{Var(x)}$ = **0,4973**

Desviación estándar (sx)= 0,7052

Coeficiente de variación= CV:= $\frac{sx}{\bar{x}} = \frac{0,7052}{6,62} =$ 0,1065 en porcentaje **10,65%**

b) Si las muestras hubieran sido las de la tabla siguiente, analizar las dos situaciones en vista a la comparación del nivel de hormona entre los dos grupos:

Normales	4,9	5,1	6,0	7,2	7,9	8,1	8,4	8,5	10,2	11,9
PQO	4,1	4,5	4,6	5,8	6,8	6,8	7,5	7,9	8,6	10,0

Para hacerlo de una manera más simplificada, colocamos los datos en una tabla /similar al infostat):

Estadístico	Mediana	Media	Min	Max	Rango	Var(x)	DS (sx)	CV
Normales	8	7,82	4,9	11,9	7	4,7573	2,181	27,89%
PQO	6,8	6,66	4,1	10,0	5,9	3,7115	1,926	28,93%

Ejercicio 4.4

Para estudiar si una dieta disminuye el contenido medio de colesterol en la sangre, a 7 pacientes se les midió el contenido de colesterol (mg/100mL) antes y después de la dieta. Los datos obtenidos fueron:

Pacientes	1	2	3	4	5	6	7
X_1: contenido de colesterol antes de la dieta	270	410	350	360	350	430	268

X_2: contenido de colesterol después de la dieta	175	308	248	231	196	190	154
Disminución (X_1-X_2).	**95.**	**102.**	**102.**	**129.**	**154.**	**240.**	**114.**

a) Hallar la disminución media del contenido de colesterol de esa muestra.
b) Calcular la desviación estándar muestral de la disminución del contenido de colesterol.
c) ¿Los investigadores pueden concluir que con la dieta la disminución media del contenido de colesterol en mayor que 130mg/100mL?

a) Para el cálculo de la disminución media, utilizamos disminución X_1- X_2 ={95; 102; 102; 114; 129; 154; 240}
Cómo ingresar los datos en calculadora f_x-82MS
https://youtu.be/qguhqq0xvM0

Media= 133,71 mg/100mL

b) **Desviación estándar= 51,04 mg/100mL**

c) No se puede concluir que con la dieta la disminución media del colesterol es mayor que 130mg/100mL, ya que este fue hecho en un grupo de 7 pacientes a los cuales no se obtuvo un seguimiento previo al estudio (de esta manera no se puede saber si la disminución es debida a la dieta o a otro factor).

Ejercicio 4.5

Los siguientes datos corresponden a 12 pacientes con esclerosis múltiple (EM) y a 12 controles; todos diestros. Se les tomó una prueba de destreza manual y se registró el puntaje:

Controles	15,2	16,4	17,0	17,5	17,5	17,6	19,4	19,5	20,0	20,0	20,1	20,3
EM	15,7	16,4	16,6	18,3	18,5	18,6	18,8	19,4	19,9	21,2	21,3	22,4

Calcular los estadísticos descriptivos: media, desviación estándar, mediana y rango o recorrido.
*Utilizando la calculadora***:**

Estadísticos	Media	Desv. Estand.	Mediana	Mínimo	Máximo	Rango
Controles	18,375	1,7131	18,5	17,6	19,4	5,1
EM	18,925	2,0658	18,7	18,6	18,8	6,7

Ejercicio 4.6

La siguiente tabla muestra la distribución de salarios de los 100 trabajadores de un gran laboratorio, que incluye desde el personal de mantenimiento hasta los directores generales:

Salario ($)	1500	3000	4500	7800	80000	195000

Frecuencia	51	37	8	2	1	1

a) Calcular la media, la mediana, la desviación estándar y la amplitud.
b) Discutir los estadísticos de tendencia central y de dispersión más adecuados para describir la situación.

a) Media $\Rightarrow \overline{X} = \frac{1}{n}\sum_{k=1}^{n} xk \Rightarrow \frac{(1500 \cdot 51) + (3000 \cdot 37) + (4500 \cdot 8) + (7800 \cdot 2) + 80000 + 195000}{100} = \frac{514100}{100} \approx$ **5141**

Mediana $\Rightarrow$ **1500** *como es un número par (100) promediamos el centro (50).*

Desvío estándar $\Rightarrow sx \sqrt{\left(\frac{1}{n-1}\right) \sum_{k=1}^{n} (x_n - \overline{x})^2}$

$\Rightarrow$

$$\sqrt{\frac{(1500-5141)^2 \cdot 51 + (3000-5141)^2 \cdot 37 + (4500-5141)^2 \cdot 8 + (7800-5141)^2 \cdot 2 + (80000-5141)^2 + (195000-5141)^2}{(100-1)}}$$

$\Rightarrow$DS *(sx)*= **20722,66**

Amplitud o rango $\Rightarrow$*rg(x)= Xmax - Xmin*= 195000 - 1500 = **193500**

Ejercicio 4.7

Un laboratorio envasa ampollas en envases de 3 unidades. Sea X = número de ampollas defectuosas en cada envase. Por estudios realizados se conoce la función de probabilidad de X:

X_i	0	1	2	3
$p(X_i)$	0,729	0,243	0,027	0,001

a) ¿Qué distribución sigue X y cuáles son sus parámetros?
b) Hallar E(X) y Var(X).
c) Al tomar una muestra aleatoria de 100 envases se obtuvo:

X_i	0	1	2	3
f_i	71	26	3	0

Hallar la media y la varianza muestrales. Interpretar estos estadísticos como estimadores de los parámetros calculados en b)

a)

X = número de ampollas defectuosas en cada envase
X ~Bi(3; p) p= P(ampollas defectuosas) sería el Éxito

Como desconocemos *p*, la vamos a estimar. Sabiendo que E(x)= m·p entonces, despejamos de la siguiente manera: p= $\frac{E(x)}{m}$

$\widehat{E(X)} = \overline{x} = \frac{0 \cdot 0{,}729 + 1 \cdot 0{,}243 + 2 \cdot 0{,}027 + 3 \cdot 0{,}001}{1} = 0{,}3$

$\hat{p} = \frac{\widehat{E(X)}}{m} = \frac{\overline{X}}{3} = \frac{0{,}03}{3} = \mathbf{0{,}1}$

Nos queda ***X~Bi(3; 0,1)***

b) La esperanza es **E(x)= 0,3** ***[E(x)= m·p]***
Calculamos la varianza→ Var(x)= m·p·q *siendo (1 - p)=q*
Var(x) = 3·0,1·(1 - 0,1) = **0,27**

c) Hallamos la media ($\overline{x}$) y varianza muestral (s^2):

$\overline{x} \Rightarrow \frac{71 \cdot 0 + 26 \cdot 1 + 3 \cdot 2 + 0 \cdot 3}{100} = \mathbf{0{,}32}$

$s^2 \Rightarrow \frac{(0 - 0{,}32)^2 \cdot 71 + (1 - 0{,}32)^2 \cdot 26 + (2 - 0{,}32)^2 \cdot 3 + (3 - 0{,}32)^2 \cdot 0}{(100 - 1)} = \frac{7{,}2704 + 12{,}0224 + 8{,}4672 + 0}{99} =$
0,28

Ejercicio 4.8

Las ampollas de un preparado biológico para vacunas pierden su efectividad al cabo de cierto tiempo. Después de 6 meses de su preparación se tomaron 40 cajas con 10 ampollas cada una y se observó cuántas ampollas resultaban eficaces. Se obtuvieron los datos del archivo Ej.4.8.
a) Definir la variable aleatoria, indicar su recorrido y la distribución que puede seguir.
b) Estimar la probabilidad de que una ampolla no pierda su efectividad después de 6 meses de su preparación.
c) Estimar la cantidad de cajas que podrán esperarse con por lo menos 6 ampollas eficaces.

Infostat

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana
CANTIDAD DE AMPOLLAS EFICA..	40	7,65	1,37	1,87	17,90	5,00	10,00	8,00

a)

X = número de ampollas eficientes por caja, con un total de 10 cada una
X~Bi(10; p)

b) Estimamos la probabilidad de que una ampolla no pierda la efectividad después de 6 meses

$\widehat{E(X)} = \overline{x} = 7{,}65$

$\hat{p} = \frac{\widehat{E(X)}}{m} = \frac{\overline{X}}{10} \Rightarrow \frac{7{,}65}{10} = \mathbf{0{,}765}$

Ejercicio 4.9

Un virólogo diluye partículas de virus en una solución. Luego toma de dicha solución 150 muestras de 1 cc y encuentra, en cada muestra, el número de partículas que se detalla en el archivo Ej.4.9. Suponiendo que se ha comprobado que se necesitan por lo menos tres partículas

de virus para infectar un embrión de pollo estimar, usando la distribución de Poisson, la probabilidad de que una muestra de 1 cc infecte un embrión de pollo.

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana
CANTIDAD DE PARTÍCULAS	150	2,11	1,50	2,24	71,10	0,00	7,00	2,00

X= número de partículas de virus diluidas en solución de vol 1cc

$$X \sim P(\lambda) \Rightarrow \hat{\lambda} = \bar{x} = 2{,}11$$

Estimar la probabilidad de que una muestra de 1cc infecte un embrión de pollo, siendo necesarios por lo menos 3 partículas del virus:

$\hat{P}(X \geq 3) = 1 - \hat{P}(X \leq 2) = 1 - 0{,}6496 =$ **0,3504**

Tabla 4: Función de Distribución de Poisson

λ

k	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9	3
0	0,1225	0,1108	0,1003	0,0907	0,0821	0,0743	0,0672	0,0608	0,0550	0,0498
1	0,3796	0,3546	0,3309	0,3084	0,2873	0,2674	0,2487	0,2311	0,2146	0,1991
2	0,6496	0,6227	0,5960	0,5697	0,5438	0,5184	0,4936	0,4695	0,4460	0,4232
3	0,8386	0,8194	0,7993	0,7787	0,7576	0,7360	0,7141	0,6919	0,6696	0,6472

Ejercicio 4.10

Las observaciones que se detallan en el archivo Ej.4.10 corresponden al peso, en gramos, de ambos riñones en 50 hombres presumiblemente normales cuyas edades están comprendidas entre 40 y 49 años. Suponiendo que el peso de los riñones se distribuye en forma normal, estimar la probabilidad de que dicho peso supere los 350 g.

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana
PESO DEL RIÑÓN	50	336,47	29,34	860,65	8,72	269,50	419,40	330,70

X = peso (gramos) en riñones en adultos presumiblemente normales

$X \sim N(\mu; \sigma)$

siendo los estimadores de μ y σ; sabiendo que $E(x)=\mu$ y $Var(x)=\sigma^2$

$$\hat{\mu} = \bar{x}$$

$$\widehat{\sigma^2} = s^2$$

Estimamos la probabilidad de que X supere los 350 g

$\hat{P}(X>350) = \hat{P}\left(\frac{x-\hat{\mu}}{\hat{\sigma}} > \frac{350 - 336{,}47}{29{,}34}\right) = \hat{P}(Z > 0{,}46) = 1 - F_X(0{,}46) = 1 - 0{,}6772 =$ **0,3228**

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
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

Ejercicio 4.12

Utilizando la base de Datos Ej.4.12.

a) Calcular la media, la mediana, la varianza, la desviación estándar, el máximo, el mínimo, el rango y el coeficiente de variación para las variables edad, peso, altura y número de hijos.

b) Calcular los porcentajes de fumadores y de hipertensos.

Medidas resumen

a)

Variable	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana	Rango
EDAD	42,96	16,34	267,12	38,05	15,00	90,00	40,00	75,00
PESO	80,86	12,41	153,93	15,34	47,00	140,00	80,00	93,00
ALTURA	1,76	0,08	0,01	4,53	1,51	2,01	1,75	0,50
Nª DE HIJOS	1,56	1,57	2,46	100,51	0,00	10,00	1,00	10,00

b) Para el cálculo del porcentaje de fumadores y de hipertensos (en un total de 540 casos)

Tablas de frecuencias

Variable	Clase	Categorías	FA	FR	
FUMA	1	No	352	0,65	→ 65,00%
FUMA	2	Si	188	0,35	→ 35,00%

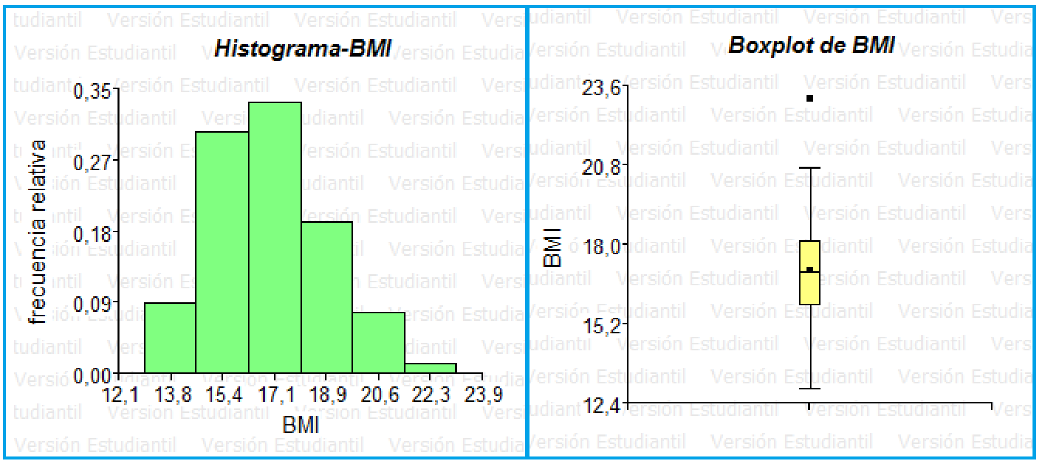
Variable	Clase	Categorías	FA	FR	
HIPERTENSO	1	No	401	0,74	→ 74,00%
HIPERTENSO	2	Si	139	0,26	→ 26,00%

Ejercicio 4.13

En el año 2005 se evaluaron niños de jardines de infantes de Capital Federal en el marco de un estudio nutricional. La base de datos que se encuentra en el archivo Ej. 4.13 corresponde a 80 niños entre 2 y 5 años de edad y comprende, entre otras variables, la medición del índice de masa corporal (Body Mass Index: BMI = peso(kg)/[estatura(m)]2). Efectuar una estadística descriptiva de los datos de BMI y confeccionar un histograma y un box-plot para visualizar la distribución de la variable.

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana
BMI	80	17,08	1,86	3,45	10,87	12,90	23,10	17,00



pipiMorch