Desarrollo Ensayo 2º Control

1. a) i) Población: Sector Norte de Antofagasta
   ii) Elementos: Niños de 1 mes
   iii) Variable: Peso
   iv) Tipo de variable: Cuantitativa Continua

b) R = 4.35 – 3.63 = 0.72 $\Rightarrow A = \frac{0.67}{8} = 0.09$ es la amplitud de los intervalos

Título: Distribución de niños de un mes según peso. Sector norte de Antofagasta. Junio de 2010

i	Categorías	C_i	n_i	f_i	N_i	F_i
1	[3.63,3.72]	3.675	5	0.05102	5	0.05102
2	(3.72,3.81]	3.765	8	0.08163	13	0.1327
3	(3.81,3.90]	3.855	13	0.1327	26	0.2653
4	(3.90,3.99]	3.945	25	0.2551	51	0.5204
5	(3.99,4.08]	4.035	23	0.2347	74	0.7551
6	(4.08,4.17]	4.125	17	0.1735	91	0.9286
7	(4.17,4.26]	4.215	5	0.05102	96	0.9796
8	(4.26,4.35]	4.305	2	0.02041	98	1.0000
			98	1		

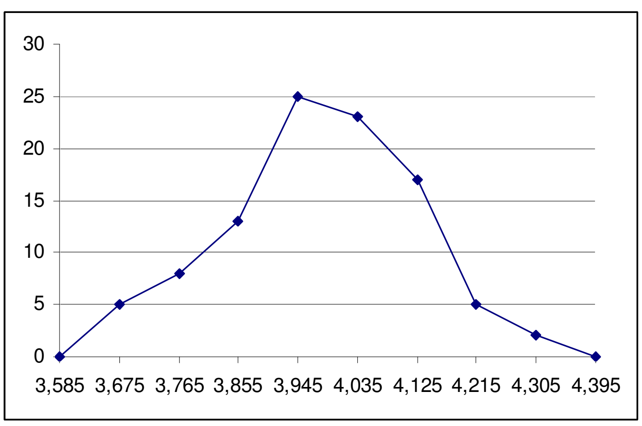
i) $N_3 = 26$

ii) $n - N_2 = 98 - 13 = 85$

iii) $N_7 - N_1 = 96 - 5 = 91$

iv) $(1 - F_4)*100 = (1 - 0.5204)*100 = 47.96\%$

c)



d) $\overline{X} = 3.9781 \quad S_X = 0.1422 \Rightarrow \overline{X} + S_X = 4.1203$

17 niños de la muestra tiene un peso mayor a 4.1203 kilos y corresponde al 17.35%, luego el 17.35% de los niños tiene riesgo de obesidad

2. a) Distribución condicional de Y si $X \in (8,12]$

i	Y/ $X \in (8,12]$	n_i	N_i
1	[30,40]	1	1
2	(40,50]	6	7
3	(50,60]	11	18
4	(60,70]	13	31
		31	

i) $M_e = 50 + \left(\frac{15.5 - 7}{11}\right)*10 = 57.7272$; ii) $M_o = 60 + \left(\frac{13 - 11}{(13-11)+(13-0)}\right)*10 = 61.3333$

Interpretación: i) El 50% de los estudiantes con una edad mayor que 8 y menor o igual a 12 años tiene un puntaje menor o igual a 57.7272 puntos
ii) El puntaje más frecuente de los estudiantes con una edad mayor a 8 y menor o igual a 12 años es 61.3333 puntos.

b) Distribución marginal de X

i	X	n_i	N_i
1	[6,8]	32	32
2	(8,10]	11	43
3	(10,12]	20	63
4	(12,14]	27	90
		90	

Se pide el percentil 80, luego $\alpha = \frac{90*80}{100} = 72 < 90 \Rightarrow P_{80} \in (12,14]$

Luego $P_{80} = 12 + \left(\frac{72-67}{27}\right)*2 = 12.6667$

Respuesta: La edad mínima del 20% de los estudiantes mayores es 12.6667 años

c) Distribución condicional de Y si $X \in (6,10]$

i	Y/ $X \in (6,10]$	n_i	N_i
1	[30,40]	20	20
2	(40,50]	14	34
3	(50,60]	3	37
4	(60,70]	6	43
		43	

$$42 = 40 + \left(\frac{\alpha - 20}{14}\right)*10 \Rightarrow \alpha = \left(\frac{42-40}{10}\right)*14 + 20 = 22.8 = \frac{43*i}{100}$$

$$\Rightarrow i = \frac{22.8*100}{43} = 53.02 \Rightarrow 100 - i = 46.97\%$$

Respuesta: El 46.97% de los niños con una edad mayor a 6 y menor o igual a 10 años tiene un puntaje en el examen mayor o igual a 42 puntos.

d) Distribución marginal de X

i	X	C_i	n_i
1	[6,8]	7	32
2	(8,10]	9	11
3	(10,12]	11	20
4	(12,14]	13	27
			90

$$CV(X) = \left(\frac{2.4980}{9.9333}\right)*100 = 25.15\%$$

Distribución marginal de Y

j	Y	C_j	n_j
1	[30,40]	35	21
2	(40,50]	45	17
3	(50,60]	55	19
4	(60,70]	65	33
			90

$$CV(Y) = \left(\frac{11.8551}{52.1111}\right)*100 = 22.75\%$$

Respuesta: El puntaje es más homogéneo que la edad por que tiene menor dispersión relativa o menor coeficiente de variación.

3. a) i) Directamente con la calculadora se obtiene r(D,V) = 0.5355

ii) Linealizando el modelo $V = AD^B \Rightarrow \log V = B(\log D) + \log A$

e ingresando en la calculadora los datos transformados $(\log D_i, \log V_i)$ se obtiene r(log D, log V) = 0.5490

Respuesta: Se elige el modelo ii) por tener mayor coeficiente de correlación lineal en valor absoluto.

b) Estimación de parámetros:

Con la calculadora se obtiene

$$\hat{k} = 1.6782 = \log \widehat{A} \Rightarrow \widehat{A} = 10^{1.6782} = 47.6683$$

$$\hat{m} = 1.8090 = \widehat{B} \Rightarrow \widehat{B} = 1.8090$$

Luego el modelo ajustado es:

$$\boxed{\widehat{V} = f(D) = 47.6683 * D^{1.8090}}$$

c) Estimación del error estándar

D_i	V_i	$\widehat{V}_i = f(D_i)$	$e_i^2 = (V_i - f(D_i))^2$
0.45	0.75	11.2432	110.1072
0.55	43.5	16.1639	747.2624
0.60	47.99	18.9193	845.1056
0.70	52.1	25.0041	734.1877
0.85	55.8	35.5262	411.02696
0.88	56.8	37.8267	359.9861
1.00	59.1	47.6683	130.6838
2.00	60.2	167.0294	11412.5207

$$\sum_{i=1}^{8} e_i^2 = 14750.7554$$

Luego el error estándar del modelo ajustado es: $\widehat{S}_e = \sqrt{\dfrac{14750.7554}{8}} = 42.94$

La viscosidad estimada para una densidad de 2.5 es $f(2.5) = 250.0939$

Un intervalo de amplitud 3 errores estándar para esta estimación es:

$$[250.0939 - 64.41, 250.0939 + 64.41] = [185.6839, 314.5039]$$

d) Como r(D;V) = 0.5355 > 0 entonces la relación lineal entre la viscosidad y la densidad es directa o creciente, luego a mayor viscosidad aumenta la densidad, por lo tanto la afirmación del investigador es errónea,