EJERCICIO Nº 17

En la toma de datos de campo, para la realización de un inventario forestal de una masa de Pinus pinea en el valle del Tietar, se han obtenido los siguientes datos reflejados de diámetro normal (dn), altura total (h), edad , y peso de leñas gruesas, de 32 árboles muestra.

Establecer los distintos ajustes de regresión, mediante rectas, entre las variables consideradas, utilizando como variable independiente o predictora el dn, por el método de ajuste de los mínimos cuadrados. Determinar la bondad del ajuste en función de los coeficientes de correlación y determinación obtenidos del mismo.

ARBOL Nº	dn (cmts.)	h (mts.)	Edad (años)	leñas gruesas (kgs.)	ARBOL Nº	dn (cmts.)	h (mts.)	Edad (años)	leñas gruesas (kgs.)
1	20	7,5	30	50	17	29	10,5	50	50
2	20	7,5	39	40	19	31	10	49	100
3	19	7	30	40	19	32	11	47	150
4	22	8	35	60	20	29	10,5	48	50
5	20	7,5	30	50	21	37	11,5	59	250
6	22	8	30	60	22	39	12	48	300
7	20	7,5	30	50	23	35	11	47	250
8	27	11	47	60	24	35	11	47	250
9	26	7,5	44	100	25	35	11	51	200
10	23	8	39	60	26	32	11	48	150
11	24	8,5	32	40	27	36	12	67	200
12	24	8,5	32	80	28	40	12	52	300
13	27	11	49	100	29	39	12	51	300
14	27	11	47	100	30	39	12	51	300
15	24	8,5	31	60	31	40	12	52	250
16	30	11	31	200	32	42	14	54	300

RESOLUCIÓN:

En la actualidad, distintos programas estadísticos informáticos, nos permiten de manera sencilla proceder a cualquier análisis de regresión por complejo que sea el modelo de ajuste que utilicemos, en este ejercicio se realiza el ajuste por el método de los mínimos cuadrados y el método gráfico, mediante una recta ($Y = a + b \cdot X$) entre las variables seleccionadas, desarrollando todas las operaciones necesarias para ello.

1º **Relación Alturas / diámetros**

Relación h/dn (variable predictora –dn-, variable dependiente -h-)

dn(cm.) (x)	h (m.) (y)	n_i	$\sum n_i x_i$	$\sum n_i y_i$	$\sum n_i(x_i-x_m)^2$	$\sum n_i(y_i-y_m)^2$	$\sum n_i(x_i-x_m)(y_i-y_m)$
19	7	1	19	7	110,88	9,28	32,09
20	7,5	4	80	30	363,28	25,95	97,09
22	8	2	44	16	113,40	8,38	30,83
23	8	1	23	8	42,64	4,19	13,37
24	8,5	3	72	25,5	91,74	7,18	25,66
26	7,5	1	26	7,5	12,46	6,49	8,99
27	11	3	81	33	19,20	2,73	-7,23
29	10,5	2	58	21	0,56	0,41	-0,48
30	11	1	30	11	0,22	0,91	0,45
31	10	1	31	10	2,16	0,00	-0,07
32	11	2	64	22	12,20	1,82	4,71
35	11	3	105	33	89,76	2,73	15,64
36	12	1	36	12	41,86	3,81	12,64
37	11,5	1	37	11,5	55,80	2,11	10,85
39	12	3	117	36	269,04	11,44	55,49
40	12	2	80	24	219,24	7,63	40,90
42	14	1	42	14	155,50	15,63	49,30
			945	321,5	1599,97	110,68	390,21

$$\boxed{y = a + b \cdot x \Rightarrow h(m.) = a + b \cdot dn(cm.)}$$

$$\boxed{\overline{X} = dn_{medio} = \frac{945}{32} = 29{,}53 \text{ cm.}} \; ; \; \boxed{\overline{Y} = h_{media} = \frac{321{,}5}{32} = 10{,}046 \text{ m.}}$$

$$\mathbf{b} = \frac{S_{xy}}{S_x^2} = \frac{\dfrac{\sum_{i=1}^{n}(x_i-\overline{X})(y_i-\overline{Y})}{n-1}}{\dfrac{\sum_{i=1}^{n}(x_i-\overline{X})^2}{n-1}} = \frac{\dfrac{390{,}21}{31}}{\dfrac{1599{,}97}{31}} = \mathbf{0{,}2435\ m/cm.}$$

$$\overline{Y} = a + b\overline{X} \Rightarrow 10{,}046 \text{ m.} = a + 0{,}2435 \text{ m./cm.} \cdot 29{,}53 \text{ cm.}$$

$$\Rightarrow \mathbf{a} = 10{,}046 \text{ m.} - 7{,}1905 \text{ m.} = \mathbf{2{,}8555}$$

$$\boxed{y = 2{,}855 + 0{,}2435 \cdot dn \Rightarrow \mathbf{h(m.) = 2.855\ m. + 0{,}2435\ m./cm. \cdot dn(cm.)}}$$

El coeficiente de determinación estimado r^2, será:

$$\mathbf{r^2} = \frac{s_{xy}^{\ 2}}{s_x^{\ 2} \cdot s_y^{\ 2}} = \frac{390{,}21^2\ m^2 \cdot cm^2}{1599{,}97 \cdot 110{,}68\ m^2 \cdot cm^2} = \mathbf{0{,}857}$$

Podemos estimar, que en el rango de datos obtenidos, la ecuación obtenida explica la evolución de las alturas de los árboles en función de la de sus diámetros normales en el 85,7% de los casos.

2º **Relación Edades / diámetros**

Relación Edad/dn (variable predictora –dn-, variable dependiente -Edad-)

dn(cm.) (x)	edad (años) (y)	n_i	$\sum n_i x_i$	$\sum n_i y_i$	$\sum n_i(x_i-x_m)^2$	$\sum n_i(y_i-y_m)^2$	$\sum n_i(x_i-x_m)(y_i-y_m)$
20	30	3	60	90	272,46	559,48	390,43
20	39	1	20	39	90,82	21,68	44,37
19	30	1	19	30	110,88	186,49	143,80
22	30	1	22	30	56,70	186,49	102,83
22	35	1	22	35	56,70	74,93	65,18
27	47	2	54	94	12,80	22,36	-16,92
26	44	1	26	44	12,46	0,12	-1,21
23	39	1	23	39	42,64	21,68	30,40
24	31	1	24	31	30,58	160,18	69,99
24	32	2	48	64	61,16	271,74	128,92
27	49	1	27	49	6,40	28,56	-13,52
30	31	1	30	31	0,22	160,18	-5,95
29	50	1	29	50	0,28	40,24	-3,36
31	49	1	31	49	2,16	28,56	7,86
32	47	1	32	47	6,10	11,18	8,26
29	48	1	29	48	0,28	18,87	-2,30
37	59	1	37	59	55,80	235,43	114,62
39	48	1	39	48	89,68	18,87	41,14
35	47	2	70	94	59,84	22,36	36,58
35	51	1	35	51	29,92	53,93	40,17
32	48	1	32	48	6,10	18,87	10,73
36	67	1	36	67	41,86	544,93	151,04
40	52	1	40	52	109,62	69,62	87,36
39	51	2	78	102	179,36	107,86	139,10
40	52	1	40	52	109,62	69,62	87,36
42	54	1	42	54	155,50	106,99	128,99
		32	945	1397	1599,97	3041,22	1785,84

$$y = a + b \cdot x \Rightarrow \text{Edad (años)} = a + b \cdot \text{dn (cm.)}$$

$$\bar{X} = \text{dn}_{\text{ medio}} = \frac{945}{32} = 29{,}53 \text{ cm.}$$

$$\bar{Y} = \text{Edad}_{\text{ media}} = \frac{1397}{32} = 43{,}65 \text{ años}$$

$$\mathbf{b} = \frac{S_{xy}}{S_x{}^2} = \frac{\dfrac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{n-1}}{\dfrac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n-1}} = \frac{\dfrac{1785{,}84}{31}}{\dfrac{1599{,}97}{31}} = \mathbf{1{,}116 \text{ años/cm.}}$$

$$\bar{X} = a + b\bar{Y} \Rightarrow 43{,}65 \text{ años} = a + 1{,}116 \text{años/cm.} \cdot 29{,}53 \text{ cm.}$$

$$\Rightarrow \mathbf{a} = 43{,}65 \text{ m. - } 32{,}956 \text{ m.} = \mathbf{10{,}694 \text{ m.}}$$

$$y = 10{,}694 + 1{,}116 \cdot \text{dn} \Rightarrow \mathbf{Edad(años) = 10{,}694 + 1{,}116 \cdot dn(cm.)}$$

El coeficiente de determinación estimado r^2, será:

$$\mathbf{r^2} = \frac{s_{xy}{}^2}{s_x{}^2 \cdot s_y{}^2} = \frac{1785{,}84^2 \text{ años}^2 \cdot \text{cm}^2}{1599{,}97 \cdot 3041{,}22 \text{ años}^2 \cdot \text{cm}^2} = \mathbf{0{,}655}$$

Podemos estimar, que en el rango de datos obtenidos, la ecuación obtenida explica la evolución de las edades de los árboles en función de la de sus diámetros normales en el 65,5% de los casos.

3º Relación Cantidad de leñas / diámetros

Relación Leñas/dn (variable predictora –dn-, variable dependiente -leñas-)

$$y = a + b \cdot x \Rightarrow \text{Leñas (Kg.)} = a + b \cdot \text{dn(cm.)}$$

$$\bar{X} = \text{dn}_{\text{medio}} = \frac{945}{32} = 29{,}53 \text{ cm.}$$

$$\bar{Y} = \text{Leñas}_{\text{media}} = \frac{4550}{32} = 142{,}19 \text{ Kg.}$$

$$\mathbf{b} = \frac{S_{xy}}{S_x^2} = \frac{\dfrac{\sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})}{n-1}}{\dfrac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{n-1}} = \frac{\dfrac{20402{,}81}{31}}{\dfrac{1599{,}97}{31}} = \mathbf{12{,}752\ Kg./cm.}$$

$$\overline{X} = a + b\overline{Y} \Rightarrow 142{,}1875\ \text{Kg.} = a + 12{,}753\ \text{kg./cm.} \cdot 29{,}53\ \text{cm.}$$

$$\Rightarrow \mathbf{a} = 142{,}1875\ \text{Kg.} - 376{,}56\ \text{Kg.} = \mathbf{-234{,}38\ Kg.}$$

$$\boxed{y = 234{,}38 + 12{,}752 \cdot x \Rightarrow \mathbf{Leñas(Kg.) = -234{,}38 + 12{,}752 \cdot dn(cm.)}}$$

dn(cm.) (x)	Leñas (Kg.) (y)	n_i	$\sum n_i x_i$	$\sum n_i y_i$	$\sum n_i(x_i-x_m)^2$	$\sum n_i(y_i-y_m)^2$	$\sum n_i(x_i-x_m)(y_i-y_m)$
20	50	3	60	150	272,53	25495,61	2635,99
20	40	1	20	40	90,84	10442,29	973,97
19	40	1	19	40	110,91	10442,29	1076,16
22	60	2	44	120	113,44	13509,57	1237,95
27	60	1	27	60	6,41	6754,79	208,04
26	100	1	26	100	12,47	1779,79	148,97
23	60	1	23	60	42,66	6754,79	536,79
24	40	1	24	40	30,59	10442,29	565,22
24	80	1	24	80	30,59	3867,29	343,97
24	60	1	24	60	30,59	6754,79	454,60
27	100	2	54	200	12,81	3559,57	213,57
30	200	1	30	200	0,22	3342,29	27,10
29	50	2	58	100	0,56	16997,07	97,95
31	100	1	31	100	2,16	1779,79	-61,96
32	150	2	64	300	12,19	122,07	38,57
37	250	1	37	250	55,78	11623,54	805,22
39	300	3	117	900	268,97	74714,36	4482,86
35	200	1	35	200	29,91	3342,29	316,16
35	250	2	70	500	59,81	23247,07	1179,20
36	200	1	36	200	41,84	3342,29	373,97
40	250	1	40	250	109,59	11623,54	1128,66
40	300	1	40	300	109,59	24904,79	1652,10
42	300	1	42	300	155,47	24904,79	1967,72
		32	945	4550	1599,97	299746,88	20402,81

El coeficiente de determinación estimado r^2, será:

$$r^2 = \frac{s_{xy}{}^2}{s_x{}^2 \cdot s_y{}^2} = \frac{20402{,}81^2 \ Kg^2 \cdot cm^2}{1599{,}97 \cdot 299746{,}88 \ Kg^2 \cdot cm^2} = \mathbf{0{,}868}$$

Podemos estimar, que en el rango de datos obtenidos, la ecuación obtenida explica la evolución de la cantidad de leñas gruesas obtenida de cada árbol en función de la de sus diámetros normáles en el 86,8% de los casos.

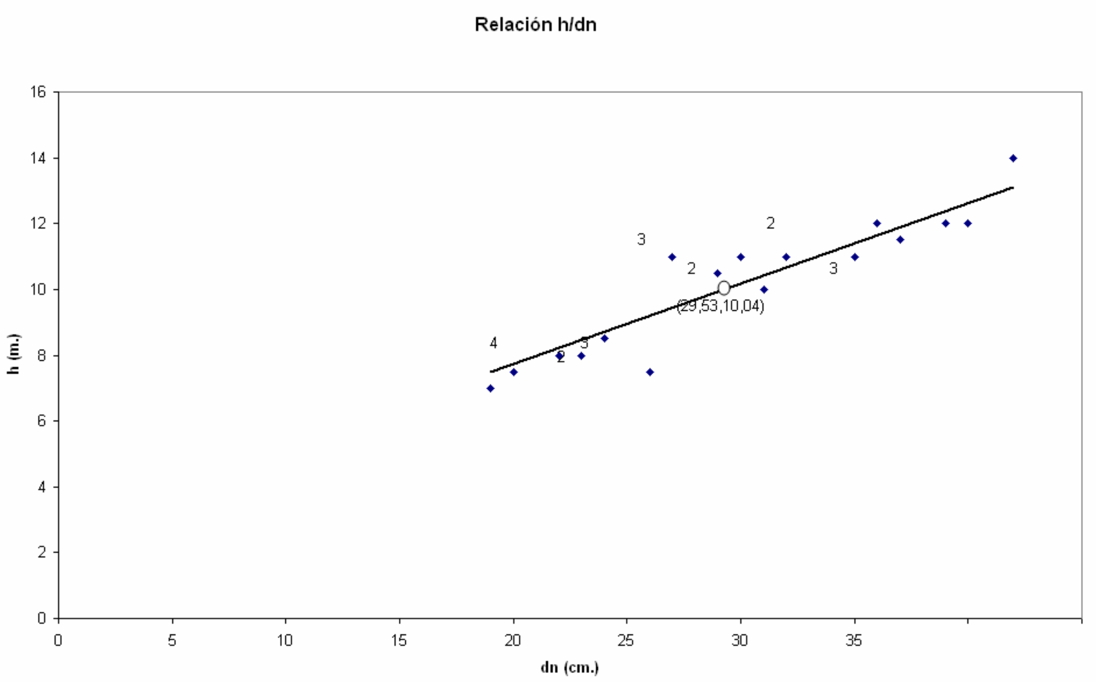
Ajuste lineal por el método gráfico:

El método gráfico, ha sido el método de ajuste utilizado antiguamente y que hasta fechas recientes ha tenido su utilidad.

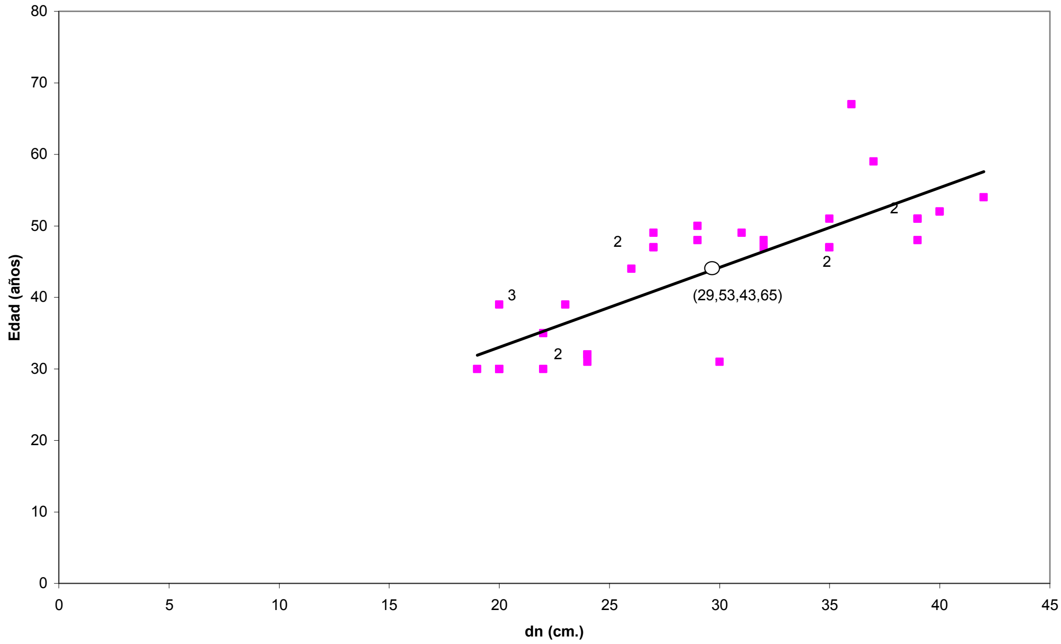
Está basado en la representación gráfica sobre un eje de coordenadas de las variables consideradas interdependientes, en el eje "Y", la variable dependiente y en el eje "X", la variable predictora o independiente.

En la nube de puntos obtenida, a partir de su centro de gravedad $(\overline{X}, \overline{Y})$ debemos trazar una línea recta que deje, (aproximadamente), el mismo número de puntos a un lado y al otro de la misma.

Relación Alturas / diámetros normales



Relación Edades / diámetros normales



Relación cantidad leñas gruesas / diámetros normales