

Casus

De human resource manager van een mijnbouwbedrijf wil de samenhang onderzoeken tussen het aantal dienstjaren (X) en het aantal dagen ziekteverzuim (Y) van de arbeiders voor dat jaar. *Het aantal dienstjaren en het aantal dagen ziekteverzuim zijn normaal verdeeld, wat blijkt uit jarenlange waarneming.* De resultaten van 10 willekeurige arbeiders zijn in onderstaande tabel verwerkt:

Arbeider	A	B	C	D	E	F	G	H	I	J
X = Aantal dienstjaren	3	4	5	4	7	7	6	3	1	10
Y = Aantal dagen ziekteverzuim	12	15	18	20	24	25	22	19	6	26

Regressie= model: $y^c = \alpha + \beta x + \varepsilon$

Opgave 1

Bereken $\hat{\beta} = b$ en $\hat{\alpha} = a$ (2 dec.) (10 pnt)

Opgave 2

Toets b met een betrouwbaarheid van 95 %. Formuleer hierbij eerst de nul- en alternatieve hypothese. (25 pnt)

Opgave 3

Bereken de storingsterm (2 dec.) (10 pnt)

Opgave 4

Hoeveel procent van de variantie van het aantal dagen ziekteverzuim wordt verklaard door de regressie? M.a.w. Bereken de determinatie-coëfficiënt R^2 (2 dec.) (15 pnt)

Opgave 5

Bereken een 95 % betrouwbaarheidsinterval van het aantal dagen ziekteverzuim, voor een arbeider met 6 dienstjaren. (op gehelen afronden) (30 pnt)

Formules:

$$y_i^c - t_{0.5\alpha} S_F < \hat{y} < y_i^c + t_{0.5\alpha} S_F$$

$$S_F = S_\varepsilon \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x - \bar{x})^2}}$$

$$S_\varepsilon = \sqrt{\frac{\sum (y - y^c)^2}{n - 2}}$$

$$Cijer = \frac{Scoe + 10}{10}$$

Correctie-model:

Opgave 1 (10 pnt)

x	y	$(x - \bar{x}) \cdot (y - \bar{y})$	$(x - \bar{x})^2$	$(y - y^c)^2$	$(y - \bar{y})^2$
3	12	13,4	4	5,76	44,89
4	15	3,7	1	2,4025	13,69
5	18	0	0	0,49	0,49
4	20	-1,3	1	11,9025	1,69
7	24	10,6	4	1	28,09
7	25	12,6	4	4	39,69
6	22	3,3	1	1,3225	10,89
3	19	-0,6	4	21,16	0,09
1	6	50,8	16	16,81	161,29
10	26	36,5	25	11,9025	53,29
$\bar{x} = 5$	$\bar{y} = 18,7$	$\Sigma = 129$	$\Sigma = 60$	$\Sigma = 76,75$	$\Sigma = 354,1$

$$b = \frac{129}{60} = 2.15 \text{ en } a = 18.7 - (2.15 * 5) = 7.95$$

Opgave 2 (25 pnt)

$$H_0: \beta = 0 \wedge H_1: \beta \neq 0$$

$$t_b = \frac{b}{s_b} = \frac{2.15}{0.4} = 5.38 \quad t_{0.025 [10-2=8]} = 2.306$$

Dus b verschilt significant van nul, nulhypothese wordt verworpen.

Opgave 3 (10 pnt)

$$S_\epsilon = \sqrt{\frac{76.75}{10-2}} \approx 3.10$$

Opgave 4 (15 pnt)

$$R^2 = 1 - \frac{76.75}{354.1} \approx .78$$

Opgave 5 (30 pnt)

$$y^c(6) = 7.95 + (6 * 2.15) = 20.85$$

$$S_F = 3.10 * \sqrt{1 + \frac{1}{10} + \frac{(6-5)^2}{60}} = 3.28$$

$$20.85 - (2.306 * 3.28) < \hat{y} < 20.85 + (2.306 * 3.28)$$

$$13 < \hat{y} < 28 \text{ verzuimen per jaar}$$