

Harjoitus 10

1. $X \sim \text{Exp}(a)$; tiheysfunktio $f(x) = ae^{-ax}$, kun $x \geq 0$

(x_1, x_2, \dots, x_n) satunnaisotos X :stä

Likelihood-funktio

$$L(a) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n) = ae^{-ax_1} \cdot \dots \cdot ae^{-ax_n} = a^n \cdot e^{-a \sum_{i=1}^n x_i}, \text{ kun } x_i \geq 0.$$

Etsitään $L(a)$:lle maksimikohta: L :llä ja $\ln L$:llä on sama maksimikohta.

$$l(a) = \ln L(a) = n \ln a - a \sum_{i=1}^n x_i.$$

Maksimikohta löytyy derivaatan nollakohdasta. Ylläolevan funktion derivaatan nollakohta on parametrin a estimaattori.

$$l'(a) = \frac{n}{a} - \sum_{i=1}^n x_i = 0$$

$$\Rightarrow \hat{a} = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}}.$$

Havainnoista saadaan $\bar{x} = 3,41 \Rightarrow \hat{a} = \frac{1}{3,41} \approx 0,293$.

2. $p = P(\text{bitti 1 vaihtuu bitiksi 0})$

Likelihood-funktio

$$L(p) = p^{90} \cdot (1-p)^{9910}$$

Log-Likelihood-funktio $l(p) = \ln L(p) = 90 \ln p + 9910 \ln(1-p)$

Ylläolevan funktion derivaatan nollakohta on parametrin p estimaatti.

$$l'(p) = \frac{90}{p} - \frac{9910}{1-p} = \frac{90 - 90p - 9910p}{p(1-p)} = 0$$

$$\Rightarrow 90 - 10000p = 0 \quad \Rightarrow \hat{p} = \frac{90}{10000} = 0,009.$$

3.

Merkitään X = pituus; Y = paino.

a) Korrelaatiokerroin $r = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}} = 0,98845$

Selitysaste $r^2 = 0,9770$.

Jos et saa laskimestasi suoraan r :ää, niin ota laskimestasi \bar{x}, \bar{y}, s_x ja s_y (saa tavallisesta funktiolaskimestakin!) ja laske

$$S_{xx} = \Sigma(x_i - \bar{x})^2 = (n - 1)s_x^2 = 293,57875;$$

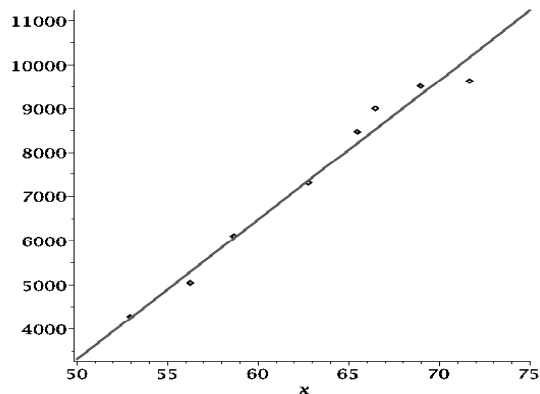
$$S_{yy} = \Sigma(y_i - \bar{y})^2 = (n - 1)s_y^2 = 30148896,875;$$

$$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - n\bar{x}\bar{y} = 92993,9375.$$

b) Vakiot $a = \frac{S_{xy}}{S_{xx}} = 316,75977$ ja $b = \bar{y} - a\bar{x} = -12529,19$

Regressiosuora $y = 316,8 \cdot x - 12529$.

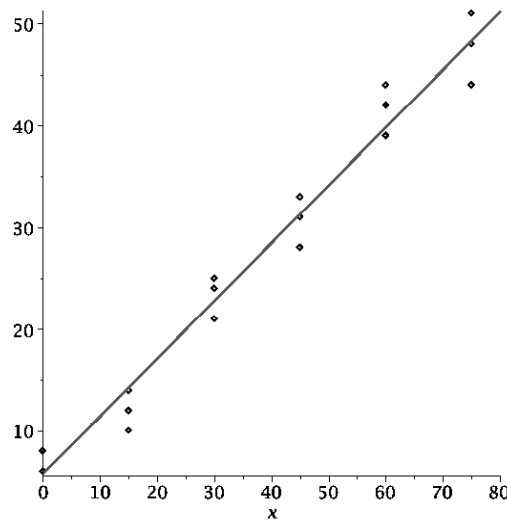
Havaintopisteet ja suora:



c) Painoennuste $y = 316,8 \cdot 180 - 12529 = 44495[\text{g}] \approx 44,5 \text{ kg}$.

Painoennuste on huono, miksi?

4. a) Sirontakuvio:



b) r ja r^2 suoraan laskimesta tai seuraavista arvoista: $\bar{x} = 37,5$ $s_x = 26,360$

$n = 18$ $\bar{y} = 27,11$ $s_y = 15,169$ $S_{xx} = (n-1)s_x^2 = 11812,5$ $S_{yy} = (n-1)s_y^2 = 3911,78$

$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = 6705$.

Korrelaatiokerroin $r = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}} = 0,986372$

Selitysaste $r^2 = 0,9729$.

c) Regressiosuoran kertoimet suoraan laskimesta tai laskemalla edellä olleista arvoista: vakiot

$a = \frac{S_{xy}}{S_{xx}} = 0,5676$ ja $b = \bar{y} - a\bar{x} = 5,8254$, joten regressiosuora on

$\hat{y} = 0,5676x + 5,8254$.

Pisteittäiset residuaalit (pisteparit $(x_i, y_i - \hat{y}_i)$, missä $\hat{y}_i = 0,5676x_i + 5,8254$):

