

Lösningar till prov i beskrivande statistik

8 november 2011

Uppgift 1

a) Medelvärdet blir

$$\bar{x} = \frac{0 + 5 + 0 + 7 + 9 + 0 + 3 + 4 + 33 + 0 + 5}{11} = \frac{66}{11} = 6$$

För att ta fram medianen behöver vi ordna observationerna i storleksordning

$$0 \ 0 \ 0 \ 0 \ 3 \ 4 \ 5 \ 5 \ 7 \ 9 \ 33$$

Nu får vi medianen som "mittensta" observationen, det vill säga observation 6 i storleksordning

$$\text{Md} = 4$$

b) Vi börjar med att beräkna kvadratsumman

$$\sum_{i=1}^{11} x_i^2 = 0^2 + 5^2 + 0^2 + 7^2 + 9^2 + 0^2 + 3^2 + 4^2 + 33^2 + 0^2 + 5^2 = 1294$$

Variansen blir

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1} = \frac{1294 - \frac{66^2}{11}}{10} = 89.8$$

och standardavvikelse $s = \sqrt{89.8} = 9.48$.

Vi får första kvartilen som observationen med ordningstal

$$25 \cdot \frac{n+1}{100} = 25 \cdot \frac{12}{100} = 3$$

vilket blir $q_1 = 0$ och tredje kvartilen som observationen med ordningstal

$$75 \cdot \frac{12}{100} = 9$$

vilket blir $q_3 = 7$. Kvartilavståndet blir därmed $\text{iqr} = q_3 - q_1 = 7 - 0 = 7$.

c) Den avvikande observationen 33 påverkar både medelvärde och standardavvikelse eftersom både summa och kvadratsumma blir ganska stora. Medianen och kvartilavståndet tar bara hänsyn till ordningstal och inte numeriska värden, så därför påverkar den största observationen inte så mycket.

Uppgift 2

a) Korrelationskoefficienten blir

$$r = \frac{\sum xy - (\sum x)(\sum y)/n}{\sqrt{(\sum x^2 - (\sum x)^2/n)(\sum y^2 - (\sum y)^2/n)}}$$

$$= \frac{54.20 - 1.16 \cdot 276/6}{\sqrt{(0.2430 - 1.16^2/6)(12906 - 276^2/6)}} = 0.424$$

En positiv korrelationskoefficient antyder ett positivt samband mellan kvicksilverhalt och kroppslängd. Värdet 0.42 säger dock att sambandet är inte är särskilt starkt.

b) Lutningskoefficienten b fås genom

$$b = \frac{\sum xy - (\sum x)(\sum y)/n}{\sum x^2 - (\sum x)^2/n} = \frac{54.20 - 1.16 \cdot 276/6}{0.2430 - 1.16^2/6} = 44.8$$

och a genom

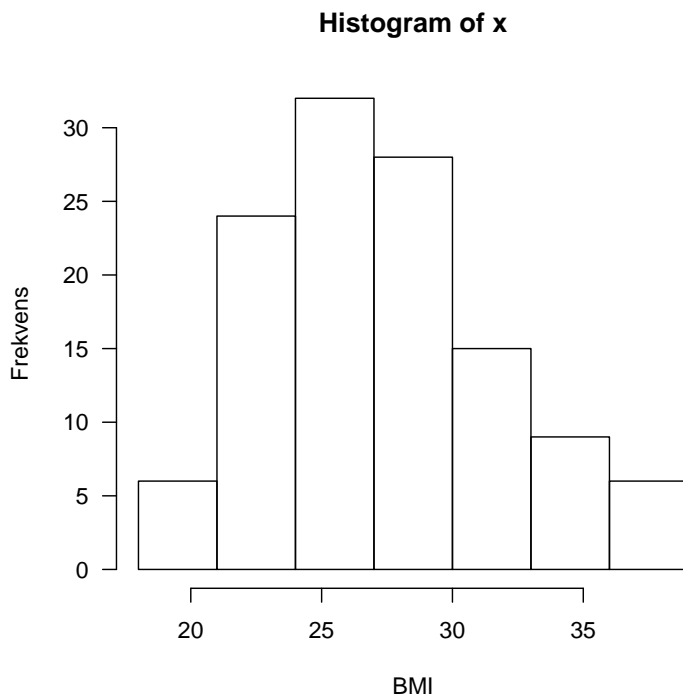
$$a = \bar{y} - b\bar{x} = \frac{\sum y}{n} - b \frac{\sum x}{n} = \frac{276}{6} - 44.8 \frac{1.16}{6} = 37.3$$

Den sjunde abborren borde ha kroppslängden

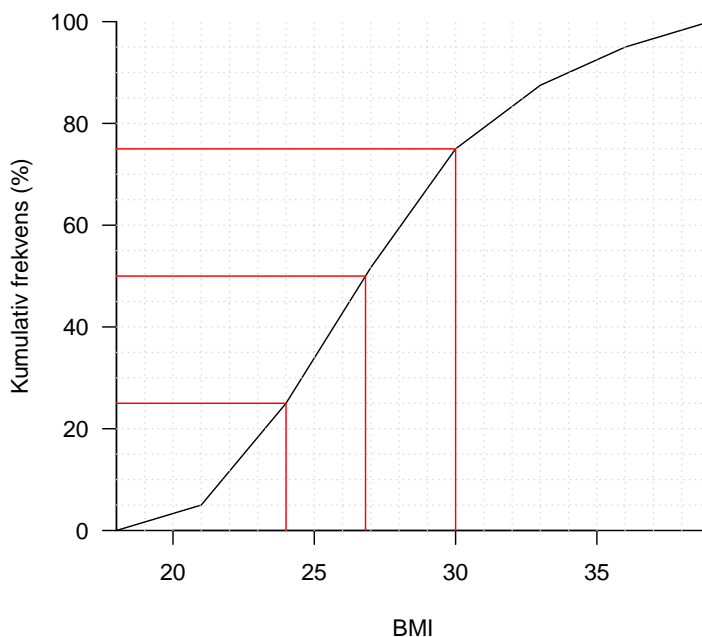
$$y = 37.3 + 44.8 \cdot 0.20 = 46 \text{ cm}$$

Uppgift 3

a) Histogrammet blir



b) Summakurvan blir



Kvartilerna fås ur figuren som $q_1 = 24.0$, $q_2 = 26.8$ och $q_3 = 30.0$.

c) Medelvärdet blir

$$\begin{aligned}\bar{x} &\approx \frac{\sum f \cdot x}{n} = \frac{6 \cdot 19.5 + 24 \cdot 22.5 + 32 \cdot 25.5 + 28 \cdot 28.5 + 15 \cdot 31.5 + 9 \cdot 34.5 + 6 \cdot 37.5}{120} \\ &= \frac{3279}{120} = 27.3\end{aligned}$$

Kvadratsumman blir

$$\begin{aligned}\sum f \cdot x^2 &= 6 \cdot 19.5^2 + 24 \cdot 22.5^2 + 32 \cdot 25.5^2 + 28 \cdot 28.5^2 + 15 \cdot 31.5^2 + 9 \cdot 34.5^2 + 6 \cdot 37.5^2 \\ &= 92016\end{aligned}$$

vilket ger variansen

$$s^2 = \frac{\sum f \cdot x^2 - \frac{(\sum f \cdot x)^2}{n}}{n - 1} = \frac{92016 - \frac{3279^2}{120}}{119} = 20.3$$

och standardavvikelsen $s = \sqrt{20.3} = 4.5$.

d) Medelfelet ges av

$$\frac{s}{\sqrt{n}} = \frac{4.5}{\sqrt{120}} = 0.41$$

vilket är en uppskattning av skattningsfelet i medelvärdet. Med andra ord ligger stickprovsmedelvärdet cirka 0.4 enheter från populationsmedelvärdet.