

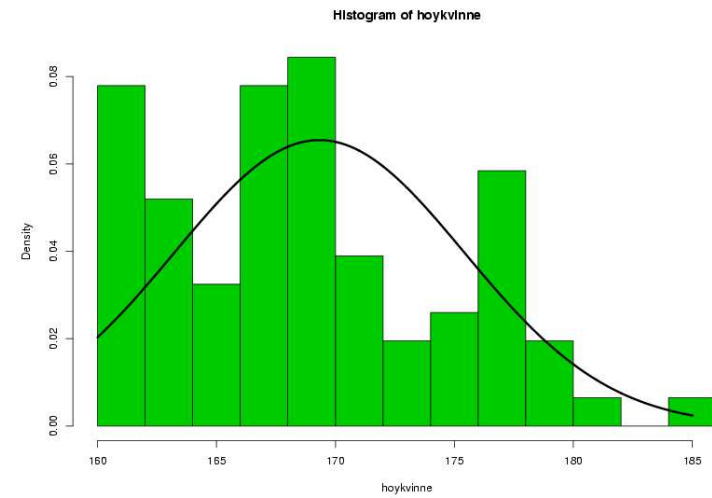
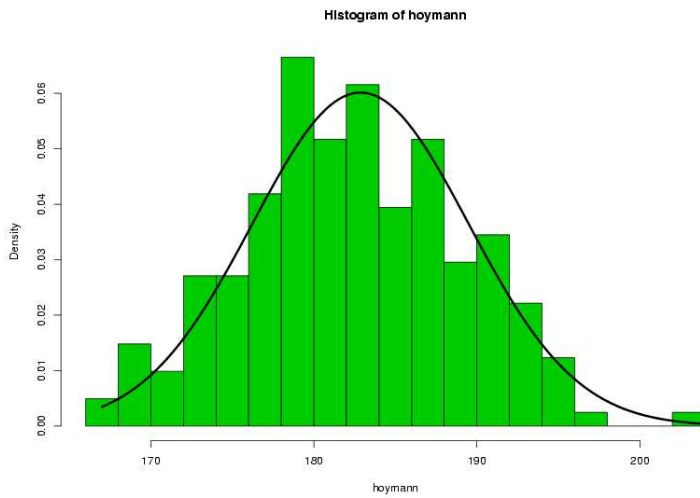
Kapittel 11: Enkel lineær regresjon og korrelasjon

TMA4245 Statistikk V2005

11.1-11.4 foreleses uke 17, 2005



Høyde studenter

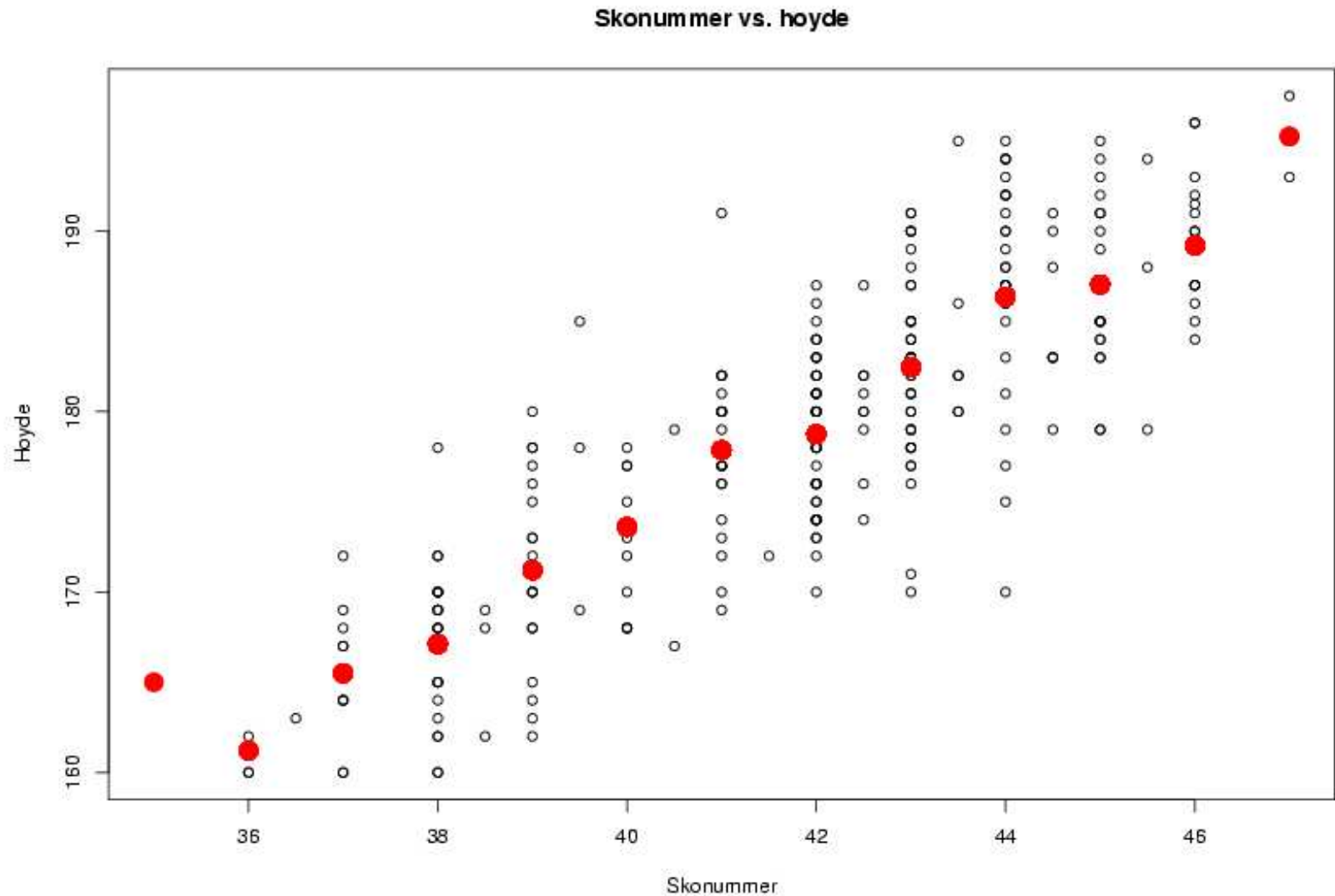


- Ønsker å gi et anslag med usikkerhet på høyden til en ny student basert fra spørreundersøkelse TMA4245V2005, 201 menn og 77 kvinner.
- Beste estimat for høyden til en mannlig student er 183 cm, 95% prediksjonsintervall: [170,196]
- Beste estimat for høyden til en kvinnlig student er 169 cm, 95% prediksjonsintervall: [157,182]

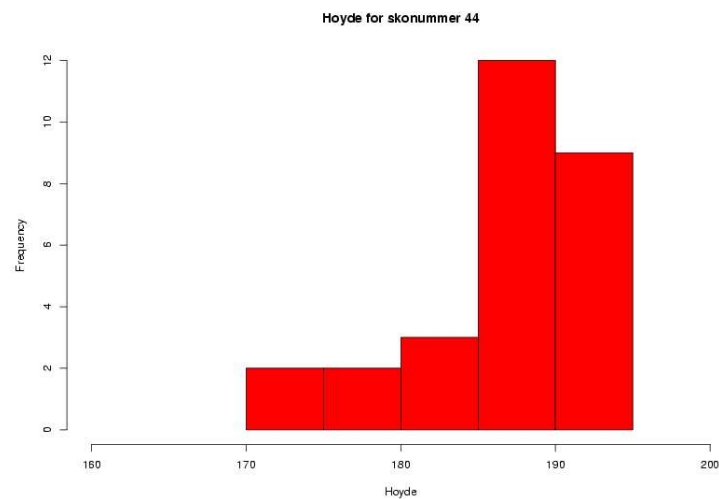
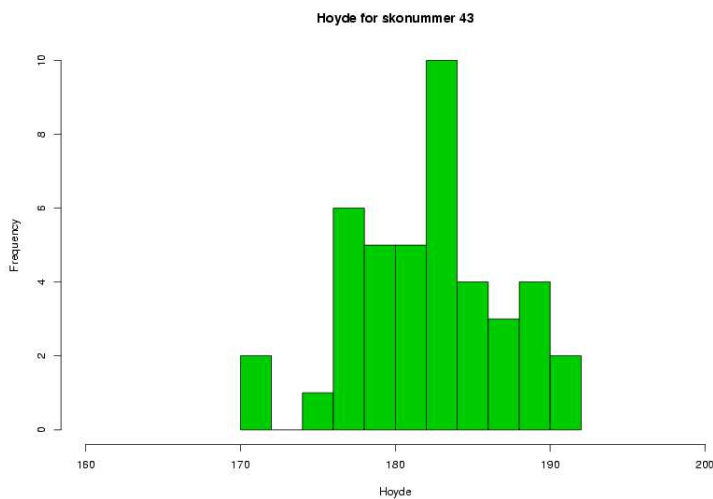
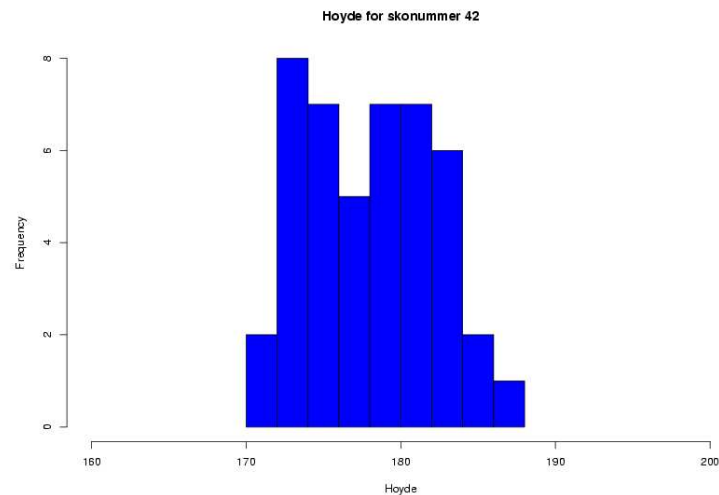
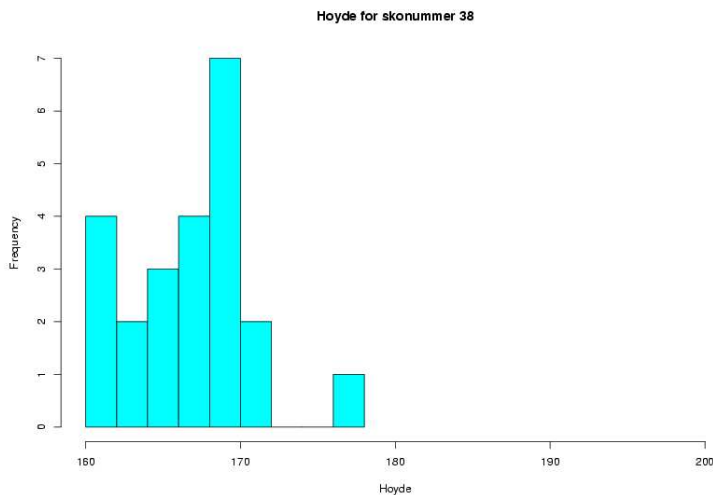
$$\bar{x} - t_{\frac{\alpha}{2}, (n-1)} s \sqrt{1 + \frac{1}{n}} < x_0 < \bar{x} + t_{\frac{\alpha}{2}, (n-1)} s \sqrt{1 + \frac{1}{n}}$$

- Men, hva hvis vi har andre opplysninger om andre mål? Hvilke?

Høyde vs. skonommer for studenter



Høyde av studenter, gitt skonummer



Estimatorer for α og β



Mellomregninger

$$\begin{aligned}\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= \sum_{i=1}^n (x_i - \bar{x})y_i - \sum_{i=1}^n (x_i - \bar{x})\bar{y} \\ &= \sum_{i=1}^n (x_i - \bar{x})y_i - \bar{y} \left(\sum_{i=1}^n x_i - n\bar{x} \right) = \sum_{i=1}^n (x_i - \bar{x})y_i \\ &= \sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^n (x_i - \bar{x})^2 &= \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + n\bar{x}^2 \\ &= \sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2\end{aligned}$$

Estimatorer for α og β

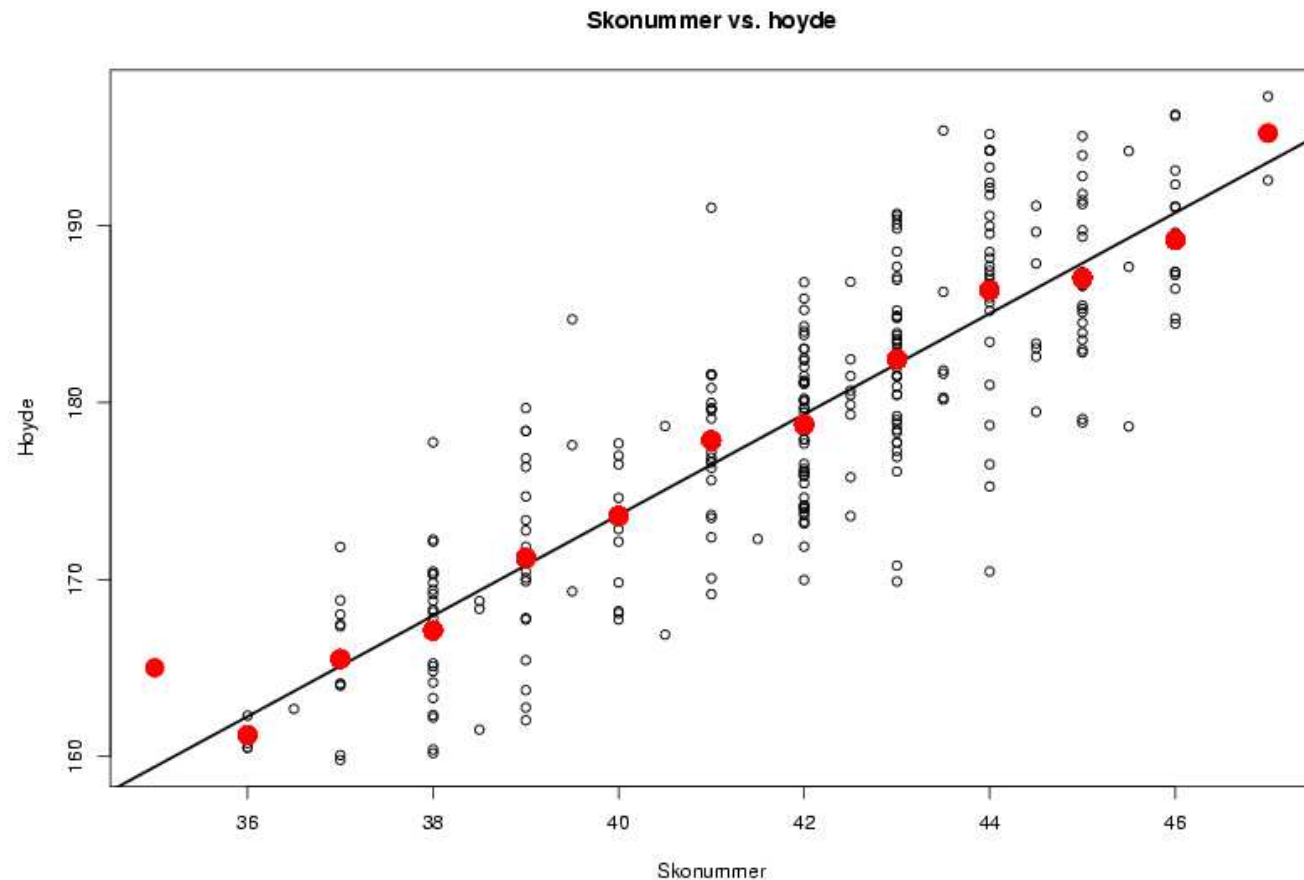
- Gitt utvalget $\{(x_i, y_i); i = 1, \dots, n\}$, så er minste kvadratsum estimatorene A og B for koeffisientene α og β gitt som

$$B = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$A = \bar{Y} - B\bar{X} = \frac{\sum_{i=1}^n Y_i - B \sum_{i=1}^n x_i}{n}$$

- Dette er også sannsynlighetsmaksimeringsestimatorene.

Høyde vs. skonommer for studenter



● $\sum_{i=1}^{278} (x_i - \bar{x})y_i = 5285.4, \sum_{i=1}^{278} (x_i - \bar{x})^2 = 1854.8$

● $\bar{x} = 41.9$ og $\bar{y} = 179.1$

Høyde og skonummer

Call:

```
hoyde = alpha + beta*skonummer
```

Residuals:

Min	1Q	Median	3Q	Max
-15.0416	-3.4354	0.1310	3.1088	14.5071

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
alpha	59.6612	4.6417	12.85	<2e-16 ***
beta	2.8496	0.1106	25.77	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.762 on 276 degrees of freedom

Multiple R-Squared: 0.7064, Adjusted R-squared: 0.7053

F-statistic: 664.1 on 1 and 276 DF, p-value: < 2.2e-16