Statistiek
Hoorcollege 5, 07-12-2015

Population regression model: y = $\boldsymbol{\beta_0 + \beta_1 x + \epsilon}$

a. Bepaal de afhankelijke variabele Y en de onafhankelijke variabele X.
b. Maak een plot van de punten
c. Schat de regressielijn: $\hat{y} = \beta_0 + \beta_1 x$
d. Test of de relatie statistically significant is
e. Interpreteer de resultaten

Redisials: $e_i = y_1 - \hat{y}$

Estimating Regression Coefficients

$\hat{y} = \beta_0 + \beta_1$x

$\beta_1 = \dfrac{S_{xy}}{s_x^2} = r\dfrac{S_y}{S_x}$

$\beta_0 = \bar{y} - \beta_1\bar{x}$

Voor de formules voor Sample variance, Population variance en Sample covariance: zie slide 12 en 13.

Sum of Squares Error (SSE)

SSE = $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

$S_\varepsilon^2 = \dfrac{SSE}{n-2}$

Als s_ε klein is verklaart het model groot deel van de variatie van de onafhankelijke variabele. Dit is dus goed.

Sampling distribution of the slope coefficient $\boldsymbol{\beta_1}$

B_1 is een estimator van de populatie parameter β_1

Estimated variance van b_1 is (gebaseerd op steekproef): $s_{b1}^2 = \frac{s_\varepsilon^2}{(n-1)s_x^2}$

Variantie verminderd als:

- Punten liggen dicht bij de regressie lijn
- Grote range in de onafhankelijke variabele x
- Grootte steekproef n

Assessing the significance of the model

1. $H_0 : B_1 = 0, B_1 \neq 0$ (tweezijdige toets)
2. Significantie = 0.05
3. De test statistiek is t = $\frac{b1-\beta 1}{S_{b1}}$ = b1 / S_{b1}
   De geschatte standard error van b1 is: $S_{bf} = \frac{s_\epsilon}{\sqrt{(n-1)s_x^2}}$
4. De test statistic is (student's) t gedistribeerd met (n-2) graden vrijheid onder de nul hypothese
5. De kritieke waarden zijn: $-t_{a/2,\, n-2}$ en $t_{a/2,\, n-2}$

Betrouwbaarheids interval voor $\boldsymbol{\beta_1}$

$b_1 \pm s_{b1} \cdot t_{\alpha/2,n-2} = (b_1 - s_{b1} \cdot t_{\alpha/2,n-2},\ b_1 + s_{b1} \cdot t_{\alpha/2,n-2})$

R^2

Variatie in y (SST) = SSE + SSR
R^2 = 1- (SSR/SST) = (SSR/SST)

Assessing the power of a test

De kans op het plegen van een type I error wanneer we de nul hypothese verwerpen is minder dan of gelijk aan a. Wanneer we de nul hypothese niet kunnen verwerpen weten we dat:

- De nul hypothese klopt, of:
- Het verwerpen van de nul hypothese mislukt wanneer het alternatief wel waar is (type II error)

We beschouwen een type twee error voor specifieke parameters die zitten in de alternatieve hypothese μ_1^*.

Voorbeelden: slide 31 t/m 36