
Instructiecollege 4

Hoofdstuk 16

Opdracht 2

De opdracht is om $\hat{Y} = 3X - 2$ berekenen met de getallen 0, 1, 3 en -2.

$$\hat{Y} = (3 * 0) - 2 = -2$$

$$\hat{Y} = (3 * 1) - 2 = 1$$

$$\hat{Y} = (3 * 3) - 2 = 7$$

$$\hat{Y} = (3 * -2) - 2 = -8$$

Opdracht 4

De vergelijking met de kleinste kwadraten (least squares regression).

Opdracht 6

Hoe kleiner de standaardfout, hoe groter de correlatie.

Of: hoe groter de standaardfout, hoe kleiner de correlatie.

Extra opdrachten

Opdracht 1

Een aantal gegevens zijn al bekend:

Voor de SAT: $M = 580$ en $SS = 22400$.

Voor de GPA: $M = 3,10$; $SS = 1,26$ en $SP = 84$.

- $\hat{Y} = 0,00375X + 0,925 = (0,00375 * 620) + 0,925 = 3,25$.
- $\text{Residu} = Y - \hat{Y} = 2,9 - 3,25 = -0,35$.
- Lager, want bij opgave a hebben we het gemiddelde geschat op 3,25 en 2,9 ligt daaronder. Maar ook is het te zien aan de uitkomst bij vraag b. Deze uitkomst is namelijk negatief en daardoor is het ook te zien dat het lager is.
- $r^2 = SP / \sqrt{(SS_X * SS_Y)} = 84 / \sqrt{(22400 * 1,26)} = 0,50$.
 $r^2 = (0,50)^2 = 0,25$. Als we dit vergelijken met de richtlijnen van Cohen dan is het een groot effect.
- $F = MS_{\text{regression}} / MS_{\text{residual}} = 0,315 / 0,0727 = 4,33$.

$$MS_{\text{regression}} = SS_{\text{regression}} / 1 = 0,315 / 1 = 0,315.$$

$$SS_{\text{regression}} = r^2 * SS_Y = 0,25 * 1,26 = 0,315.$$

$$MS_{\text{residual}} = SS_{\text{residual}} / n - 2 = 0,945 / (15 - 2) = 0,0727.$$

$$SS_{\text{residual}} = (1 - r^2) * SS_Y = (1 - 0,25) * 1,26 = 0,945.$$

- $df = n - 2 = 15 - 2 = 13$. Het aantal vrijheidsgraden wordt (1,13). Die 1 is het aantal richtingscoëfficiënten - 1. We hebben hier 2 richtingscoëfficiënten dus wordt de uitkomst 1.
- De kritieke waarde (uit tabel B.4) is 4,67. Het is dus niet significant want $F = 4,33$ is kleiner dan F_{kritiek} dus H_0 kan niet verworpen worden. Er is dus geen significante relatie tussen SAT en GPA.

Opdracht 2

- In de grafiek is dan af te lezen dat $X = 17$ en $Y = 14$.
- Omdat dit punt onder de lineaire lijn ligt is het residu negatief.
- Onderaan is te zien dat $r^2 = 0,446$. Dit betekent dat 44,6% van Y verklaard kan worden door de relatie met X .
- Hierbij wordt de ANOVA-tabel gebruikt. In de tabel staat een sig. van $< .001$. Dit wordt vergeleken met de $\alpha = .05$. Het is dus significant want de p waarde is kleiner dan α .
- $r^2 = SS_{\text{regressie}} / SS_Y = 690,022 / 1546,875 = 0,446$. De gegevens om deze formule in te vullen zijn af te lezen in de ANOVA-tabel.
- $df_{\text{total}} = n - 1 = 31$. Er zijn dus 32 studenten. Ook deze gegevens zijn af te lezen in de tabel bij df_{total} .
- $\hat{Y} = bX + a = 1,122X + 0,570$.
Om deze formule in te kunnen vullen wordt de coëfficiënts tabel gebruikt en wordt in kolom B gekeken. a wordt dan constant en de waarde die hierbij hoort is $.570$, b is $1,122$.
- H_0 : de regressiecoëfficiënt is gelijk aan 0.
 H_1 : de regressiecoëfficiënt is niet gelijk aan 0.
Het verschilt significant van 0, $b = 1,12$, $t(30) = 4,92$, $p < .001$.