

Formuleblad tussentoets Statistiek

Populatie en steekproef

Gegeven een populatie met waarden $X_1, \dots, X_N$ en gemiddelde μ , zijn variantie σ^2 en standaardafwijking σ van de populatie gedefinieerd als

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} \quad \sigma = \sqrt{\sigma^2}.$$

Gegeven een steekproef met waarden $X_1, \dots, X_n$ en gemiddelde $\bar{X}$ zijn variantie s^2 en standaardafwijking s van de steekproef gedefinieerd als

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad s = \sqrt{s^2}.$$

Correlatie en regressie

Gegeven twee stochasten X en Y , waarbij (X_i, Y_i) de waarde van element i is in een steekproef ter grootte n , is de *Pearson correlatiecoëfficiënt*:

$$r = \frac{\sum z_X z_Y}{n-1} = \frac{\sum_{i=1}^n z_{X_i} z_{Y_i}}{n-1}.$$

Waarbij z_{X_i}, z_{Y_i} de standaarscores van X_i, Y_i t.o.v. de steekproef zijn:

$$z_{X_i} = \frac{X_i - \bar{X}}{s_{\bar{X}}} \quad z_{Y_i} = \frac{Y_i - \bar{Y}}{s_{\bar{Y}}}.$$

Gebruikmakend van de notatie $x = X - \bar{X}$:

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}}.$$

Ook geldt

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}}.$$

In de regressielijn $\hat{Y} = bX + a$ is de regressiecoëfficiënt b is gedefinieerd als:

$$b = r \frac{s_Y}{s_X}.$$

waarbij s_X en s_Y de standaardafwijking van respectievelijk X en Y zijn. De regressieconstante a is gedefinieerd als:

$$a = \bar{Y} - b\bar{X}.$$

Wanneer $z_{\hat{Y}}$ de standaarscores van $\hat{Y}$ en z_X van X aanduiden, dan geldt dat

$$z_{\hat{Y}} = r z_X.$$

De standaardfout van de benadering, $s_{Y \cdot X}$, is gedefiniëerd als de standaardafwijking van de stochast $(Y - \hat{Y})$ bij $n - 1$ scores. D.w.z., voor $Z = (Y - \hat{Y})$:

$$s_{Y \cdot X} = \sqrt{\frac{\sum (Z - \bar{Z})^2}{n - 2}} = s_Y \sqrt{1 - r^2} \sqrt{\frac{(n - 1)}{(n - 2)}}.$$

Integreren

In het volgende is c een constante ongelijk 0, n een geheel getal ongelijk 0 en ongelijk -1, en x een variabele.

functie	primitieve
c	0
x^n	$\frac{x^{n+1}}{n+1}$
$\frac{c}{x}$	$c \ln(x)$
e^{cx}	$\frac{e^{cx}}{c}$

Verder geldt dat als $F(x)$ en $G(x)$ de primitieven van respectievelijk $f(x)$ en $g(x)$ zijn, dan is $cF(x)$ de primitieve van $cf(x)$ en $F(x) + G(x)$ is de primitieve van $f(x) + g(x)$.