
46. Point-biserial correlatie

The Point-biserial correlatie wordt gebruikt om de sterkte en richting van de associatie te meten die bestaat tussen een continue variabele en een dichotome variabele. Een dichotome variabele is een categorische variabele met twee categorieën (bijvoorbeeld ja/nee). De uitkomst zal net als bij andere correlatiecoëfficiënten altijd binnen -1 en $+1$ vallen. Een voorbeeld is het meten van de relatie tussen het salaris (continue variabele) en geslacht (dichotome variabele; man vs. vrouw). Een ander voorbeeld is het meten tussen het cholesterolgehalte (continue variabele) en de rokersstatus (of iemand rookt ja of nee).

Assumpties

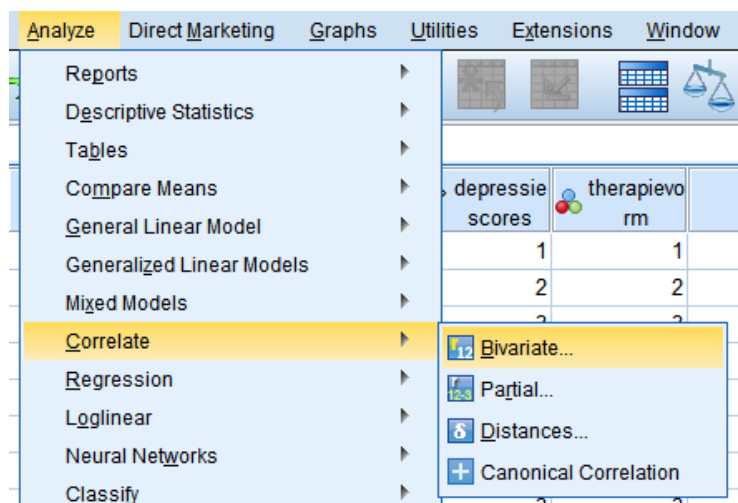
Bij een Point-biserial correlatie gelden de volgende assumpties:

- Eén van de twee variabelen moet zijn gemeten op een continue schaal.
- Eén van de twee variabelen moet dichotoom zijn.
- Er mag geen sprake zijn van uitbijters voor de continue variabele voor elke categorie van de dichotome variabele. Je kunt uitbijters controleren door middel van een Boxplot.
- De continue variabele moet nagenoeg normaal verdeeld zijn voor elke categorie van de dichotome variabele. Je kunt dit testen door middel van de Shapiro-Wilk test van normaliteit.
- De continue variabele moet gelijke varianties hebben voor elke categorie van de dichotome variabele. Dit kun je testen door middel van Levene's test of equality of variances.

Procedure

Nu volgt de procedure voor het berekenen van een Point-biserial correlatie.

1. Kies **Analyze**, selecteer **Correlatie** en kies vervolgens **Bivariate**.
2. Selecteer de variabelen naar keuze (in dit geval geslacht en depressiescore) en sleep deze naar de **Variables** sectie door te klikken op de pijltjestoets (wijzend naar rechts).
3. Zorg ervoor dat **Pearson** is aangevinkt onder **Correlation Coefficients** (de Point-biserial correlatie is namelijk eigenlijk een soort Pearson correlatie, maar dan met een eigen naam).
4. Klik op **Options**. Als je beschrijvende statistiek wil laten zien kun je dit aanvinken onder **Statistics**.
5. Klik op **Continue** en op **OK** (of op **Paste** als je de analyse wil opslaan in de Syntax Editor).



Interpretatie van de output

Correlations

		geslacht	depressiescores
geslacht	Pearson Correlation	1	-,394 [*]
	Sig. (2-tailed)		,031
	N	30	30
depressiescores	Pearson Correlation	-,394 [*]	1
	Sig. (2-tailed)	,031	
	N	30	30

*. Correlation is significant at the 0.05 level (2-tailed).

Hierboven is in de tabel de output te zien van de correlatieanalyse. De correlatie tussen geslacht en depressiescores is $-,394$. Zoals is te zien is deze correlatie significant. Dit betekent dat er sprake is van een verband tussen geslacht en depressiescores. De correlatie is negatief. In de eerste rij, in de kolom geslacht, staat een 1. Deze correspondeert met vrouw (0 correspondeert immers met man). Dit betekent dat voor vrouwen geldt dat de depressiescores lager zijn dan voor mannen. Was de correlatie in dit voorbeeld positief geweest ($+,394$) dan konden we stellen dat depressiescores hogere zijn voor vrouwen dan voor mannen.