

Onderzoekspracticum 2, aantekeningen college 4

Het enkelvoudig ANOVA-model

Het enkelvoudig ANOVA (Analysis Of Variance) model kan worden weergegeven als $X_{ij} = \mu_i + \varepsilon_{ij}$. Daarbij is X_{ij} de score van een persoon j in groep i , μ_i is het populatiegemiddelde van groep i en ε_{ij} is de error van een persoon j in groep i (zie ook college 3). Aangenomen wordt dat ε_{ij} normaal verdeeld is met gemiddelde 0 en standaarddeviatie σ . μ_i en σ zijn in de praktijk echter onbekend. Om het ANOVA-model te kunnen toetsen, wordt er daarom een schatter gebruikt voor deze parameters. De schatter voor μ_i is het steekproefgemiddelde van de i 'de groep:

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

De schatter voor σ wordt gegeven door de samengestelde schatter van de standaarddeviatie:

De wortel hiervan is de samengestelde standaarddeviatie. Hierbij geldt dat er I aantal groepen zijn met varianties $s_1^2, s_2^2, \dots, s_I^2$ en grootte n_1, n_2, \dots, n_I . Deze steekproeven zijn simple random samples die onafhankelijk van elkaar getrokken zijn uit populaties. De samengestelde schatter s_p mag alleen gebruikt worden als de standaarddeviaties van de populaties gelijk aan elkaar zijn. Om te bepalen of dit zo is kan een vuistregel gebruikt worden. Als de grootste standaarddeviatie minder dan twee keer de kleinste standaarddeviatie is, dan mag ANOVA gebruikt worden.

Toetsingsprocedure voor gelijkheid van meer dan twee populatiegemiddelden

In het algemeen kan een statistisch model worden weergegeven als DATA = FIT + RESIDUAL. Elk van deze drie termen kan worden gezien als bron van variatie. In SPSS kan enkelvoudige ANOVA worden uitgevoerd door Analyze → Compare Means → One-way ANOVA. De data bestaan uit één onafhankelijke variabele (de conditie, bijvoorbeeld conditie 1, 2 of 3) en scores op één afhankelijke variabele. De output in SPSS komt er als volgt uit te zien (getallen zijn fictief):

	Sum of squares	df	Mean Square	F	Sig.
Between groups	400.000	3	200.000	10.792	.000
Within groups	480.700	24	14.360		
Total	880.700	27			

Er zijn twee bronnen van variatie:

- Verschillen tussen groepen, bijvoorbeeld als gevolg van de experimentele manipulatie. Dit wordt ook wel de fit genoemd.
- Verschillen binnen groepen, bijvoorbeeld als gevolg van verschil in vermoeidheid tussen proefpersonen. Dit wordt ook wel het residu of de error genoemd. De totale spreiding is de som van de spreiding tussen groepen en de spreiding binnen groepen.

De kwadratensom (sum of squares) is een maat voor spreiding.

- SSG is de kwadratensom tussen de groepen.
- SSE is de kwadratensom binnen de groepen
- SST is de totale kwadratensom (SSG + SSE).

De formules voor het berekenen van SSG, SSE en SST staan hieronder. In deze formules is \bar{x} het gemiddelde van de proefpersonen uit alle steekproeven bij elkaar.

$$SSG = \sum_i^I n_i (\bar{x}_i - \bar{x})^2 \quad , \quad SSE = \sum_i^I (n_i - 1) s_i^2 \quad \text{en} \quad SST = \sum_i^I (x_{ij} - \bar{x})^2$$

Bij elke kwadratensom hoort een bepaald aantal vrijheidsgraden (degrees of freedom). Daarbij is I het aantal groepen en N het totaal aantal proefpersonen.

- DFG is het aantal vrijheidsgraden van de groepen ($I - 1$).
- DFE is het aantal vrijheidsgraden van de error ($N - I$).
- DFT is het totaal aantal vrijheidsgraden (DFG + DFE of $N-1$).

De gemiddelde kwadratensom (mean square) is de kwadratensom gedeeld door het bijbehorend aantal vrijheidsgraden.

- MSG is de gemiddelde kwadratensom van de groepen (SSG/DFG).
- MSE is de gemiddelde kwadratensom van de error (SSE/DFE).
- MSE is gelijk aan de samengestelde schatter van de variantie. De wortel uit MSE kan daarom gebruikt worden als schatter voor de gezamenlijke standaarddeviatie van de populatie.

De F-toets

Bij de F-toets voor gelijkheid van meer dan twee populatiegemiddelden geldt: $F = MSG/MSE$. De F-toets toetst of er tussen groepen een grotere spreiding is dan binnen groepen. F heeft onder de nulhypothese H_0 een F-verdeling met DFG en DFE vrijheidsgraden. Als populatiegemiddelden gelijk zijn, dan zal F ongeveer een waarde van 1 hebben. Dit is het geval bij H_0 . Als F veel groter is dan 1 verwerp je H_0 . Let op: je toetst hierbij altijd eenzijdig. De waarde van F is terug te vinden in de tabel van SPSS. In het voorbeeld heeft F een waarde van 10.792. In de kolom daarnaast is te zien dat de kans dat je bij geen populatieverschillen in gemiddelden een F-waarde vindt van 10.792 of hoger, kleiner dan 5% is. De nulhypothese wordt in dat geval verworpen: er zijn wel populatieverschillen in gemiddelden.

De effectgrootte kan berekend worden door R^2 . R^2 geeft aan hoeveel procent van de totale variantie verklaard kan worden door de onafhankelijke variabele. $R^2 = SSG / SST$. Vanaf 0.010 spreken we van een klein effect. Vanaf 0.059 is er een gemiddeld effect en vanaf 0.138 is er sprake van een groot effect.

Toetsingsprocedure voor gelijkheid van twee varianties

De F-toets voor gelijkheid van twee varianties toetst of varianties van twee steekproeven gelijk zijn aan elkaar. Er moet namelijk sprake zijn gelijke varianties om de samengestelde t-toets te kunnen gebruiken. De hypothesen die horen bij de F-toets voor gelijkheid van varianties zijn:

- H_0 : $\sigma_1 = \sigma_2$ (de populaties hebben gelijke standaarddeviaties)
- H_a : $\sigma_1 \neq \sigma_2$ (de populaties hebben geen gelijke standaarddeviaties)

Onder H_0 heeft de statistische maat $F = s_1^2/s_2^2$ een F-verdeling met n_1-1 en n_2-1 vrijheidsgraden. s_1 en s_2 zijn de standaarddeviaties van de steekproeven, n_1 en n_2 zijn de steekproefgroottes. Omdat de F-verdeling niet symmetrisch is maar scheef naar rechts, is de linker kritieke grens niet gelijk aan de rechter kritieke grens. In Tabel E (*Introduction to the Practice of Statistics* van Moore, McCabe en Craig, 2011) staan alleen de rechter kritieke grenzen van de F-verdeling. Daarom bereken je F altijd door de grootste variantie te delen door de kleinste variantie. De teller moet groter zijn dan de noemer. De p-waarde vind je door de F-waarde op te zoeken in tabel E. Om dit te doen kijk je in de tabel horizontaal naar het aantal vrijheidsgraden dat hoort bij de standaarddeviatie uit de teller. Verticaal kijk je naar het aantal vrijheidsgraden dat hoort bij de standaarddeviatie uit de noemer. Let op: je toetst hierbij altijd tweezijdig, dus je moet de gevonden p-waarde nog vermenigvuldigen met twee.

Belangrijke punten

Het is belangrijk om er goed op te letten of je een- of tweezijdig toetst. Bij de F-toets bij ANOVA (voor gelijkheid van meer dan twee populatiegemiddelden) toets je altijd eenzijdig en vermenigvuldig je p nooit met 2. Onder de alternatieve hypothese is de spreiding tussen groepsgemiddelden namelijk altijd groter dan de spreiding binnen groepen. Bij de F-toets voor gelijkheid van twee varianties toets je juist altijd tweezijdig

en vermenigvuldig je p standaard met twee. Onder de alternatieve hypothese kan namelijk de variantie van zowel de ene als de andere populatie groter zijn.

Als de rechter kritieke grens van de F-toets voor gelijkheid van varianties bijvoorbeeld 1.5 is, dan is de linker kritieke grens $1 / 1.5$.

De F-toets voor gelijkheid van varianties kun je niet in SPSS uitrekenen, maar SPSS maakt wel gebruik van Levene's test voor gelijkheid van varianties. Deze toetst hetzelfde als de F-toets.

De F-toets voor gelijkheid van varianties heeft een nadeel. De toets is namelijk extreem gevoelig voor schending van normaliteit en kan daarom alleen gebruikt worden als de data normaal verdeeld zijn. Wanneer er geen normaalverdeling is of als er niets bekend is wat betreft de verdeling van data, dan gebruik je de vuistregel. Als de grootste standaarddeviatie minder dan twee keer de kleinste standaarddeviatie is, dan mag de F-toets gebruikt worden.

Bij t-toetsen voor gemiddelden ben je vaak geïnteresseerd in de alternatieve hypothese: gemiddelden zijn dan ongelijk. Bij de F-toets voor gelijkheid van varianties ben je juist geïnteresseerd in de nulhypothese. Om de samengestelde t-toets te kunnen gebruiken, die exact een t-verdeling heeft, moeten de varianties namelijk gelijk zijn. Je wilt bij de F-toets dus juist géén significant effect vinden, zodat je de nulhypothese niet hoeft te verwerpen.