Onderzoekspracticum 2, aantekeningen college 3

Onafhankelijke steekproeven: samengestelde t-toets

In college 2 is de t-toets voor twee onafhankelijke steekproeven aan de orde gekomen. De bijbehorende t-verdeling is echter slechts een benadering van de werkelijke steekproevenverdeling, omdat het precieze aantal vrijheidsgraden niet bekend is. Om dit probleem op te lossen, kun je een andere t-toets gebruiken: de samengestelde t-toets. Deze toets heeft exact een t-verdeling, maar ook een extra aanname. De samengestelde t-toets is namelijk alleen te gebruiken wanneer de varianties van twee normaal verdeelde populaties gelijk zijn aan elkaar.

Bij gelijke populatievarianties is de som van de twee aparte varianties gelijk aan de variantie van het gemiddelde verschil $(\bar{x}_1 - \bar{x}_2)$. Dit kan als volgt worden weergegeven:

$$\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

Deze formule kan worden ingevuld in de z-toets voor twee gemiddelden. De resulterende z-toets wordt:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Voor de samengestelde t-toets wordt er gebruik gemaakt van een schatter voor de gezamenlijke populatievariantie σ^2. Deze schatter wordt gegeven door de volgende formule:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Door σ in de z-toets te vervangen door s_p ontstaat de samengestelde t-toets. Deze t-toets heeft een t-verdeling met $n_1 + n_2 - 2$ vrijheidsgraden. De samengestelde t-toets wordt dus:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Onafhankelijke steekproeven: betrouwbaarheidsinterval samengestelde t-toets

Als twee (onafhankelijke) populaties dezelfde variantie hebben, dan wordt het C%-betrouwbaarheidsinterval van μ_1-μ_2 op de volgende manier berekend:

$$(\bar{x}_1 - \bar{x}_2) \pm t^* s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

De conclusie die hierbij hoort, is: in de populatie ligt het gemiddelde verschil tussen … (ondergrens) en … (bovengrens) met …(C%) zekerheid.

Onafhankelijke steekproeven: effectgrootte samengestelde t-toets

Het berekenen van de effectgrootte door middel van Cohen's d werkt bij de samengestelde t-toets op dezelfde manier als bij de niet-samengestelde t-toets voor onafhankelijke steekproeven (zie college 2).

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p}$$

SPSS
In SPSS kan de samengestelde t-toets worden berekend door middel van Analyze → Compare Means → Independent Samples T-Test. De uitvoer geeft verschillende toetsen: één t-toets waarbij de varianties van de populaties niet gelijk zijn en één t-toets waarbij de varianties wel gelijk zijn.

Toetsingsschema samengestelde t-toets
1. Formuleer je onderzoeksvraag.
2. Ga de aannamen af en ga na of eraan is voldaan.
   - De afhankelijke variabele moet normaal verdeeld zijn. Bij schending zijn de resultaten nog wel betrouwbaar zolang de totale steekproef groot genoeg is (≥ 40).
   - Proefpersonen moeten aselect getrokken zijn. Als dit niet het geval is, dan zijn de resultaten beperkt generaliseerbaar.
   - Proefpersonen moeten onafhankelijk van elkaar zijn getrokken. Als dit niet het geval is, dan gebruik je de gepaarde t-toets (zie college 2).
   - Er moeten gelijke populatievarianties zijn. Als hier niet aan voldaan is, zijn de resultaten nog wel betrouwbaar bij (vrijwel) gelijke steekproefgroottes. Bij (sterk) ongelijke steekproefgroottes moet je de niet-samengestelde t-toets gebruiken.
3. Stel je hypothesen op.
4. Kies de juiste toets, in dit geval de samengestelde t-toets, en bepaal alpha α (vaak 0.05)
5. Voer je berekening uit.
6. Zoek de bijbehorende p-waarde op.
7. Neem een beslissing: vergelijk p met α. Als $p < \alpha$, dan verwerp je de nulhypothese.
8. Bereken de effectgrootte door middel van Cohen's d.
9. Trek je conclusie en beschrijf deze inhoudelijk. Geef de richting aan van het effect dat je gevonden hebt en plaats kanttekeningen als je (één van) de aannamen geschonden hebt.

Om te bepalen of populatievarianties gelijk zijn, kun je gebruik maken van een vuistregel. Als de grootste standaarddeviatie niet meer dan twee keer de kleinste standaarddeviatie is, dan kun je de varianties als gelijk beschouwen.

ANOVA (Analysis of variance)
Als je in een onderzoek gebruik maakt van meer dan twee steekproeven, kun je gebruik maken van variantieanalyse (ANOVA). In het geval van één onafhankelijke variabele spreek je van enkelvoudige variantieanalyse. De F-toets wordt hierbij gebruikt om de t-toets te generaliseren naar meerdere gemiddelden.

Om de relatie tussen de t-toets en de F-toets duidelijk te maken zal nu eerst de F-toets worden besproken voor twee gemiddelden. Let op: deze formule gebruiken wij verder niet!
Er geldt: $F = s_1^2/s_2^2$. De F-toets kan enkel positieve waarden aannemen en is daarom niet symmetrisch, maar scheef naar rechts verdeeld. Door de formule voor de samengestelde t-toets te herleiden tot t^2, kan er ook een andere formule voor F worden afgeleid. Bij ANOVA geldt dan dat t^2 gelijk is aan F.

$$F = \frac{\frac{n}{2}(\bar{x}_1 - \bar{x}_2)^2}{s_p^2}$$

De teller (boven) is een maat voor de spreiding tussen twee groepen. De noemer (onder) is een maat voor de spreiding tussen individuen. Wat de F-toets doet, is toetsen of de spreiding *tussen* twee groepen significant groter is dan de spreiding *binnen* deze groepen.

Hypothesen en aannamen ANOVA
Bij ANOVA is er sprake van drie of meerdere groepen. Dit betekent dat de nulhypothese en de alternatieve hypothese anders moeten worden weergeven dan bij twee groepen. Stel dat er een experiment wordt uitgevoerd met drie verschillende groepen. De nulhypothese en alternatieve hypothese geef je dan als volgt weer:
H_0: $\mu_1 = \mu_2 = \mu_3$
H_a: niet alle μ_i's zijn gelijk (je geeft geen richting van het effect aan).

Kiezen of je eenzijdig of tweezijdig toetst is bij de F-toets niet van toepassing. Bij het toetsen van twee gemiddelden zal de uitkomst van de F-toets namelijk gelijk zijn aan die van de tweezijdige samengestelde t-toets. In dat geval kan de F-toets dus als tweezijdig worden gezien, of als 'veelzijdig' bij meer dan twee gemiddelden. Bij het toetsen van spreiding kun je de F-toets echter als eenzijdig zien. De F-toets wordt namelijk gebruikt om te bepalen of de spreiding tussen groepen *groter* is dan de spreiding binnen groepen. Daarom gebruik je de F-toets bij variantieanalyse altijd eenzijdig en vermenigvuldig je p nooit met 2.

Bij ANOVA horen twee aannamen:
- De populaties moeten normaal verdeeld zijn.
- De populaties moeten dezelfde varianties hebben. Om te bepalen of hieraan voldaan is, gebruik je de eerder besproken vuistregel.

Statistische modellen voor steekproeven
Observaties in een onderzoek kunnen we aangeven met X_j. Hierbij staat *j* voor de *j*'de persoon in de steekproef. Als persoon 6 bijvoorbeeld een cijfer 7 heeft gescoord op een test, kunnen we dit weergeven als $X_6 = 7$. Als alle proefpersonen uit de steekproef precies hetzelfde cijfer halen, dan geldt: $X_j = \mu$. Dit is echter nooit het geval, want er zijn altijd individuele afwijkingen van het populatiegemiddelde. Dit noemen we de error (van persoon *j*) en dit wordt weergegeven met ε_j. Het statistische model dat hieruit volgt, is: $X_j = \mu + \varepsilon_j$. Wanneer X_j een normaalverdeling heeft met gemiddelde μ en standaarddeviatie σ, dan heeft ε_j een normaalverdeling met gemiddelde 0 en standaarddeviatie σ.

Als er meerdere onderzoeksgroepen zijn, dan wordt het aantal groepen aangegeven met *I* en het totaal aantal personen met *J*. Een observatie van een persoon *j* in groep *i* wordt dan aangegeven met X_{ij}. Als alle personen per groep hetzelfde cijfer hebben, geldt: $X_{ij} = \mu_i$. Dit is echter nooit het geval, omdat er altijd variantie is tussen personen. De error van een persoon *j* in groep *i* wordt weergegeven met ε_{ij}. Hieruit volgt het volgende model: $X_{ij} = \mu_i + \varepsilon_{ij}$. Dit wordt het enkelvoudig ANOVA-model genoemd. ε_{ij} heeft een normaalverdeling met gemiddelde 0 en standaarddeviatie σ.