Hoofdstuk 1

1.1

Vanaf nu zullen we twee of meer onafhankelijke steekproeven, met dus ook twee of meer gemiddelden, met elkaar gaan vergelijken. Om dit te kunnen doen, gebruiken we het ANOVA model (analysis of variance model). Om de twee gemiddelden te kunnen vergelijken, gebruikt dit model de variaties. De variaties worden gebruikt, omdat wanneer alle gemiddelden van de steekproeven gelijk zijn, de variaties 0 zullen zijn. De eenvoudigste vorm van het ANOVA model is de een-factor ANOVA model. In dit model wordt er gekeken naar de X (onafhankelijke variabele) en de Y (afhankelijke variabele). Wanneer een een-factor ANOVA test wordt uitgevoerd zijn we benieuwd naar het effect van de onafhankelijke variabele op de afhankelijke variabele. Wanneer je twee gemiddelden van twee onafhankelijke steekproeven wilt vergelijken, kan je de t-toets gebruiken. Maar wanneer je meer dan twee onafhankelijke steekproeven wilt vergelijken, dan moet je de meervoudige onafhankelijke t-toets (multiple independent t tests) gebruiken. Deze toets geeft de volgende hypotheses:

$$\mu 1 = \mu 2, \mu 1 = \mu 3, \mu 1 = \mu 4, \mu 2 = \mu 3, \mu 2 = \mu 4, \mu 3 = \mu 4.$$

Vanuit deze hypothese moet je meerdere onafhankelijke t-toetsen uitvoeren. Dit geeft problemen, aangezien meerdere onafhankelijke t-toetsen de kans op het maken van een Type I fout (Type I error) (kans dat de onderzoeker verwerpt de nulhypothese onjuist) vergroten. Je kan een alpha level vaststellen voor elke onafhankelijke toets. Maar een alpha-level voor een onafhankelijke toets is niet hetzelfde als een alpha-level voor alle toetsen bij elkaar. Omdat je voor elke onafhankelijke toets een risico neemt, neemt het risico toe naarmate het aantal toetsen toeneemt. De formule voor de experimentwise fout (experimentwise error):

$$\alpha\ (total) = 1 - (1 - \alpha)\hat{}C \qquad (1)$$

In deze formule,

$\alpha(total)$ = experimentwise fout.

Alpha = de kans op een Type I fout en C= de hoeveelheid of onafhankelijke toetsen.

Omdat het risico op een Type I fout vergroot wanneer je meer en meer onafhankelijke t-toetsen doet, willen we deze methode niet gebruiken.

De experimentwise fout voor C afhankelijke toetsen is moeilijker om te bepalen dus we gebruiken:

$$\alpha \leq \alpha(total) \leq C\alpha \qquad (2)$$

Om het risico op een Type I fout zo klein mogelijk te houden, moeten we het totale alpha-level controleren. Maar we moeten ook het onderscheidingsvermogen (power) (de kans om de een foute nulhypothese te verwerpen) vergroten. De omnibus F-toets kan dit totaal toetsen. Deze test wordt ook gebruikt in het ANOVA model. De een-factor ANOVA model heeft een onafhankelijke variabele of factor met twee of meer levels.

In de random-effects model, zijn in alle steekproeven de levels van de onafhankelijke variabele random genomen van de levels van de populatie. Door dit kunnen er generalisaties gemaakt worden over alle levels van de populatie.

Vanaf nu zullen we ons focussen op de fixed-effects model. Hierin worden eerst de levels van de onafhankelijke variabele geselecteerd, waarna de onderwerpen (subjects) random worden toegewezen aan de levels van de onafhankelijke variabele. In sommige situaties kan de onderzoeker deze toewijzingen controleren, maar in andere situaties kan dit niet. Daarom moet hierin een verschil worden gemaakt. De analyse verschilt niet tussen de situaties, maar de interpretatie van de resultaten verschilt wel.

Wanneer een onderzoeker de toewijzingen kan controleren, zal hij zijn resultaten meer kunnen generaliseren dan de onderzoekers die geen controle over de toewijzingen hebben. Dit is het verschil tussen een true experimental design en quasi-experiemental design. Een repeated-measures model gebruikt onderwerpen (subjects) die worden blootgesteld aan meerdere levels van een onafhankelijke variabele.

Het laatste kenmerk van het ANOVA model is de meetschaal van de onafhankelijke en de afhankelijke variabelen. Omdat het ANOVA model een toets is van gemiddelden, de meetschaal van de afhankelijke variabele is een interval of ratio level. Wanneer de afhankelijke variabele een ordinale schaal heeft, moet de Kruskal-Wallis test worden gebruikt (later besproken). Wanneer de onafhankelijke variabele kenmerken heeft van beide ordinale en interval levels, dan kunnen de ANOVA en de Kruskal-Wallis test beide worden gebruikt. ANOVA is het meeste gebruikt wanneer de onafhankelijke variabelen categoriaal zijn (nominaal of ordinaal in schaal). Maar dit model kan ook worden gebruikt met interval of ratio waarden die discreet zijn (discrete variabelen zijn variabelen die alleen een bepaalde waarde kunnen hebben en die ontstaan uit een telproces). De een-factor ANOVA model wordt ook vaak de completely randomized design genoemd. Nogmaals de kenmerken van het ANOVA model:

- De omnibus F-toets kan controleren voor de experimentwise fout.
- Er is een onafhankelijke variabele met twee of meer levels
- De onderzoeker zet de levels van de onafhankelijke variabel vast.
- Random toewijzing van de onderwerpen (subjects) aan de verschillende levels
- De onderwerpen (subjects) worden blootgesteld aan een level van de onafhankelijke variabele

Afhankelijke variabel is gemeten op een interval level. Maar de Kurskal-Wallis een-factor ANOVA kan worden gebruikt voor afhankelijke variabelen die zijn gemeten op een ordinaal level.

1.2

Zie het schema van de gegevens (tabel 1.1 op pagina 6). Elke observatie wordt Yij, genoemd waarbij j de groep kenmerkt of het level van de observatie, en i het nummer van de observatie in de groep. (Bijvoorbeeld, Y34, dit is de derde observatie van de vierde groep of level van de onafhankelijke variabele). De i heeft een gebied van i = 1,….n, en de j heeft een gebied van j=1,…,J. Dus er zijn J levels van de onafhankelijke variabele en er zijn n onderwerpen (subjects) in elke groep. Er zijn Jn=N observaties in totaal. De De ,j is de steekproef groepsgemiddelde (sample group mean), en de totale steekproef gemiddelde (overall sample mean) is

$$\bar{Y}.$$

1.3

In het ANOVA model worden de gemiddelde verschillen getest door het kijken naar de variatie in de gemiddelden. We gaan nu kijken hoe dat wordt gedaan, en we beginnen met de hypotheses die de ANOVA toets gaat testen. De hypothese voor de twee groepen situatie van de onafhankelijke t-toets zijn de nul en de alternatieve hypothese (voor een two-tailed (nondirectional toets) als volgt:

H_0: $\mu 1 = \mu 2$
H_1: $\mu 1 \neq \mu 2$

Wanneer er meer dan twee groepen zijn, gebruiken we de omnibus toets, zoals al eerder besproken. De hypotheses van de omnibus ANOVA toets zijn:

H_0: $\mu 1 = \mu 2 = \mu 3 = \cdots = \mu J$
H_1: not all the μj are equal.

De alternatieve hypothese voor de omnibus test is geschreven in een algemene vorm. Dit is gedaan om de meerdere mogelijke gemiddelde verschillen te kunnen omvatten. Deze verschillen kunnen lopen van twee verschillende gemiddelden tot verschil in alle gemiddelden. Wanneer H0 wordt verworpen kan er een multiple comparison procedure (MCP) worden toegepast. Deze MCP wordt gebruikt om te kijken welke gemiddelden of combinatie of gemiddelden aanzienlijk verschillend zijn (besproken in hoofdstuk 2).

We toetsen de verschillen in de gemiddelden door te kijken naar de variatie in de gemiddelden. We kijken naar de variatie, omdat het met meerdere groepen moeilijk wordt om naar de gemiddelden te kijken. Wanneer we een toets zouden doen met drie groepen, en de nulhypothese is waar, dan zijn de gemiddelden gelijk en is er dus geen variatie tussen de drie groepen. Maar wanneer de nulhypothese is niet waar, dan betekent dit dat de drie steekproeven niet afkomstig zijn van dezelfde populatie, maar van drie verschillenden populaties met verschillende gemiddelden. Dit betekent dat er dus variatie is tussen de drie verschillende groepsgemiddelden. Hiermee komt er een vraag, namelijk of het verschil tussen de steekproef gemiddelden is afkomstig van de gewone steekproef variatie die wordt verwacht in een populatie, of dat het afkomstig is van een echt verschil tussen de steekproef gemiddelden van verschillende populaties.

In-groep variatie (Within-group variabillity) is de variatie van de observaties in een groep gecombineerd over groepen. Tussen-groep variatie (between-group variatie) is de variatie tussen groepen.

Een lage in-groep variatie en tussen-groep variatie laat zien dat de prestatie consequent is. Wanneer beiden variaties laag zijn, is er een kleine kans dat de nulhypothese wordt verworpen. Wanneer de in-groep variatie groter is dan de tussen-groep variatie is er ook een kleine kans dat de nulhypothese wordt verworpen. Wanneer de tussen-groep variatie groter is dan de in-groep variatie, is de kans groot dat de nulhypothese wordt verworpen. Wanneer beide variaties hoog zijn dan kan de nulhypothese wel of niet worden verworpen.

Je kan een lage in-groep variatie zien in een grafiek omdat de verdeling erg dicht is en er is weinig spreiding. Dus een hoge in-groep variatie is te herkennen aan een grotere spreiding. Een hoge tussen-groep variatie is te zien wanneer elke verdeling losstaat van de rest, met weinig overlap. Bij een lage tussen-groep variatie is er veel overlap. (zie figuur 1.1 op pagina 8).

De partiele kwadratensom (paritioning of het sum of squares) is een nieuw onderwerp in ANOVA. De totale kwadratensom (total sum of squares) in Y wordt aangeduid met SStotal. Dit is de totale hoeveelheid variatie in Y. De volgende stap is om de totale variatie te verdelen. De verdeling between uit twee dingen, de eerste is de variatie tussen de groepen, dit wordt aangeduid met SSbetw, en de variatie in de groepen, aangeduid met SSwith.In het een-factor ANOVA model, SStotalis als volgt:

SStotal = SSbetw + SSwith
(3)

of

$$\sum_{i=1}^{n}\sum_{j=1}^{J}(Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^{n}\sum_{j=1}^{J}(\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{i=1}^{n}\sum_{j=1}^{J}(Y_{ij} - \bar{Y}_{.j})^2 \qquad \textit{(4)}$$

Waarin:

- SStotal is de totale kwadratensom door de variatie tussen alle observaties zonder rekening te houden met de verschillende groepen.
- SSbetw is de tussen-groep kwadratensom door de variatie tussen de groepen.
- SSwith is de in-groep kwadratensom door de variatie binnen de groepen over alle groepen samen.

Deze formule wordt ook wel de definitional (conceptual) formule genoemd, omdat elke term een vorm van variatie beschrijft. Deze formule wordt bijna niet gebruikt met echte gegevens, aangezien er een grote kans is voor een berekeningsfout.

Tabel 1 laat de ANOVA summary tabel zien. Dit zijn de resultaten van de analyse. De eerste column laat verschillende bronnen van de variatie zien. De tweede column laat de kwadratensom van de verschillende bronnen zien. In de derde column staan de vrijheidsgraden van de verschillende bronnen. De vrijheidsgraden zijn over het algemeen het aantal observaties die mogen variëren. Voor de tussen-groep zijn de vrijheidsgraden J-1. Waarin J het aantal groepen, categorieën of levels is van de onafhankelijke variabele. Voor de in-groep bron de vrijheidsgraden zijn N-J. Dit komt omdat er in elke groep n observaties zijn, dus in elke groep n-1 vrijheidsgraden. Er zijn J groepen, dus de vrijheidsgraden voor de in-groep is J(n-1).

Om de verschillende groepen te vergelijken kun je niet de kwadratensommen van de verschillende groepen gebruiken. Dit zou geen goede resultaten opleveren, aangezien dit het aantal observaties de kwadratensom kan beïnvloeden. Daarom wordt de mean square term (MS) gebruikt. Deze nummers staan in de vierde column van de tabel en zijn als volgt berekend:

MSbetw = SSbetw/dfbetw
(5)

MSwith = SSwith/dfwith.
(6)

De mean square terms zijn ook een schatting van de variatie omdat ze de kwadratensom afwijkingen van het gemiddelde gedeeld door de vrijheidsgraden voorstellen. In de laatste column is de F-waarde weergegeven. Dit is de samenvattende toets van de table. De F-waarde is berekend als volgt:

F = MSbetw/MSwith (7)

Source	SS	Df	MS	F
Between groups	SSbetw	J-1	MSbetw	MSbetw/MSwith
Within groups	SSwith	N-J	MSwith	
Total	SStotal	N-1		

Tabel 1

De F-waarde verteld of laat zien of er meer tussen-groep variatie is dan in-groep variatie. Dit is nodig omdat we hieruit kunnen bepalen of de nulhypothese verwerpen of niet. De F waarde is groter dan 1 wanneer er meer tussen-groep variatie is dan in-groep variatie. Wanneer de beide variaties ongeveer gelijk is zal de F-waarde ongeveer 1 zijn. Om de nulhypothese te kunnen verwerpen moeten we grote F-waarden vinden. Om te beslissen of de nulhypothese wordt verworpen wordt de F-waarde uit de tabel vergeleken met de kritische F-waar uit tabel A.4. Voor de kritische waar, kijk je naar F (J-1, N-J). Het ligt er dus ook aan welke alpha-waarde je kiest, voor de verschillende alpha-waardes zijn er verschillende tabellen. De dfbetw= J-1 is de teller van de F-waarde en dfwith = N-J is de noemer van de F-waarde. Deze toets is een one-tailed toets. De nulhypothese wordt verworpen wanneer de F-waarde groter is dan de kritische F-waarde. Deze omnibus F-toets bewijst of er tenminste een statistisch significant gemiddelde verschil is tussen de groepen.

Ook hierbij is het niet duidelijk waar het verschil tussen de gemiddelden is, wanneer de nulhypothese wordt verworpen omdat er meer dan twee groepen zijn. Dan ook hier kan een MCP worden gebruikt (hoofdstuk 2). Bij een twee-groepen toets is het duidelijk dat het verschil van de gemiddelden tussen de twee groepen ligt. Voor de twee-groepen situatie, de F en t-toets hebben de regel: F=t2 voor een niet directionele alternatieve hypothese in the onafhankelijke t-toets.

1.4

We zullen nu kijken naar het lineaire model van ANOVA, de schatting van de parameters van het model, de effectgroottes, betrouwbaarheidsintervallen, onderscheidingsvermogen (power) en de verwachte mean squares.

De een-factor ANOVA fixed-effects model kan worden geschreven met populatie parameters als volgt:

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \qquad (8)$$

Waarin:

- Y is de geziene score van de afhankelijke variabele voor een individu i in groep j.

- μ is het populatiegemiddelde.
- α is het groep effect voor groep j.
- ε is random residual error voor het individu i in groep j.

De residual error kan afkomstig zijn van verschillende dingen, zoals verschillen tussen individuen, een fout in de meting, of andere factoren. De populatie groep effect en de residual error worden berekend als volgt:

$$\alpha_j = \mu_{.j} - \mu \qquad (9)$$

$$\varepsilon_{ij} = Y_{ij} - \mu_{.j} \qquad (10)$$

μ is het populatiegemiddelde voor groep j. Het groep effect (αj) kan je ook zien als het gemiddelde effect van een individu van een bepaalde groep. Een positief groep effect laat zien dat het groepsgemiddelde groter is dan het algemene gemiddelde. Een negatief groep effect laat het tegenovergestelde zien. In een een-factor fixed-effect model zijn het populatie groep effecten samen 0. De residual (rest) in het ANOVA model laat het deel van Y zien dat niet verantwoord is door X.

De parameter in het model,

$$\mu, \alpha_j \; and \; \varepsilon_{ij},$$

,moet worden geschat. De steekproef schattingen worden aj and eij genoemd. De laatste twee worden als volgt berekend:

represented by $\bar{Y}$, a_j and e_{ij}. The last two are computed as:

$$a_j = \bar{Y}_{.j} - \bar{Y}_{..} \qquad (11)$$

$$e_{ij} = Y_{ij} - \bar{Y}_{.j} \qquad (12)$$

Y is het totale steekproef gemiddelde. De dubbele punt hierbij laat zien dat er is gemiddeld tussen beide i en j. Y.j staat voor het steekproef gemiddelde van groep j. Hierbij betekent de punt dat het gemiddelde is genomen van alle i individuen in groep j.

Er zijn verschillende manieren om de effectgroottes te meten. Deze effectgroottes laten de sterkte van de associatie tussen X en Y zien, dus de relatieve sterkte van het groep effect. Er zijn drie verschillende manieren om dit te meten:

De correlatie ratio (correlation ratio) (een generalisatie van R2). Dit laat de hoeveelheid van variatie in Y zien dat is toegelicht in de verschillen in groepsgemiddelden in X. Deze waarde ligt tussen 0 en 1. Wanneer de waarde is 0, dan is niets van de totale variatie van de afhankelijke variabele gekomen door verschillen tussen de groepsgemiddelden. Wanneer de waarde 1 is, is alle variatie van de afhankelijke variabele gekomen door verschillen in de groepsgemiddelden. Deze test is positief bevooroordeeld, en dit vooroordeel is het meest duidelijk bij n<30.

$$\eta^2 = \frac{SS_{betw}}{SS_{total}} \qquad (13)$$

ω2 wordt op dezelfde manier geïnterpreteerd als de eerste manier. Alleen deze manier heeft minder vooroordelen dan de eerste.

$$\omega^2 = \frac{SS_{betw} - (J-1)MS_{with}}{SS_{total} + MS_{with}} \quad (14)$$

f is de laatste manier om effectgroottes te meten. Deze methode kan waardes aannemen van 0 (wanneer de gemiddelden gelijk zijn) tot oneindig hoge positieve waardes. Bij deze manier kan het worden geïnterpreteerd als een correlatie index, maar ook als een standaarddeviatie van de gestandaardiseerde gemiddelden (standardized means).

$$f = \sqrt{\frac{\eta^2}{1-\eta^2}} \quad (15)$$

f kan worden gebruikt om de effectgrootte, d te berekenen. De formules om d te berekenen met f verschillen. Dit hangt namelijk af van de grootte van het effect, of er een minimale, gemiddelde of maximale variatie is tussen de gemiddelde van de groepen.

De verschillende effectgroottes moeten als volgt geïnterpreteerd worden:

- small effect: f= .1, η^2 or $\omega^2 = .01$
- medium effect: f=.25, η^2 or $\omega^2 = .06$
- large effect: f= .40, η^2 or $\omega^2 = .14$

De betrouwbaarheidsintervallen worden voornamelijk gebruikt om de intervallen van de geschatte populatie parameters te laten zien. Hiermee kunnen we namelijk de nauwkeurigheid van de steekproef schatting bepalen.

Het onderscheidingsvermogen (power) kan worden verdeeld in geplande power (a priori) of geobserveerde power (post hoc). We weten dat het onderscheidingsvermogen afhangt van de grootte van de steekproef, alpha-level en de effectgrootte. Geplande power wordt gebruikt om geschikte steekproefgroottes te bepalen. Deze vorm van power laat de geschikte grootte van de steekproef zien om een gewenst level van het onderscheidingsvermogen te hebben.

(Een voorbeeld van alle theorie staat op pagina 15 en 16).

De verwachte mean square van een bepaalde bron laat de gemiddelde mean square waarde zien voor de bron die zou zijn verkregen wanneer de studie een oneindig keer zou worden herhaald. De verwachte mean square waarde wordt genoteerd als E(MSbetw).

Op dit moment kunnen er zich weer twee situaties voordoen. De eerste is dat de nulhypothese waar is, en de andere dat de nulhypothese niet waar is. In de eerste situatie is er eigenlijk geen verschil tussen de populatie groepsgemiddelden, dan zijn de verwachte mean squares als volgt:

$$E(MS_{betw}) = \sigma_\varepsilon^2$$
$$E(MS_{with}) = \sigma_\varepsilon^2 \quad (16)$$

Waarin de sigma de variatie van de residual errors van de populatie is.

Hieruit volgt dat de ratio van de verwachte mean squares:

$$E(MS_{betw}) / E(MS_{with}) = 1 \tag{17}$$

Hiervan kunnen we de verwachte waar de van F bepalen:

$$E(F) = df_{with}/(df_{with}-2) \tag{18}$$

In de tweede situatie is de nulhypothese niet waar. In dit geval zijn er dus wel verschillen tussen de populatie groepsgemiddelden, dan zijn de verwachte mean squares als volgt:

$$\begin{aligned} E(MS_{betw}) &= \sigma_\varepsilon^2 + \left(n \textstyle\sum_{j=1}^{J} \alpha_j^2\right) / (J-1) \\ E(MS_{with}) &= \sigma_\varepsilon^2 \end{aligned} \tag{19}$$

Dus hier is de ratio van de verwachte mean squares:

$$E(MS_{betw})/E(MS_{with}) > 1 \tag{20}$$

De verwachte waarde van F is:

$$E(F) > df_{with}/(df_{with}-2) \tag{21}$$

Dus er zijn verschillen tussen de verwachte mean squares in de verschillende situaties. Er is een sommatie teken bij gekomen, deze staat voor de kwadratensom van de groepseffecten.

Hieruit kunnen we concluderen dat de F-ratio is:

$$F = (\text{systematic variability} + \text{error variability}) / (\text{error variability}) \tag{22}$$

De systematische variatie (systematic variability) is de variatie tussen de groepen. De error variatie (error variabillity) is de variatie in de groepen.

1.5

Er zijn drie aannames die worden gemaakt in het ANOVA model:

1. Onafhankelijkheid (independence)

De observaties moeten onafhankelijk van elkaar zijn, in de steekproeven en tussen de verschillende steekproeven. Om deze aanname waar te maken moeten (a) de toewijzing van de individuen aan de groepen apart zijn, dit kan door het ontwerp van het experiment, en (b) de individuen moeten apart van elkaar worden gehouden om deze aanname waar te maken. Random toewijzing (random sampling) is nodig in het ANOVA model. Als dit niet het geval is, zal het risico op een type I of II fout worden vergroot. Daarnaast kan schending van de aanname de standaardfout (standard error) van de steekproef gemiddelden beïnvloeden. De makkelijkste manier om te kijken of deze aanname wordt waargemaakt, is door de residuals in een grafiek te zetten. De aanname is waargemaakt wanneer de punten random over de grafiek zijn verdeeld voor elke groep.

Daarnaast kan de Dubin-Watson toets gebruikt worden om te testen voor autocorrelatie. Schending van deze aanname gebeurt meestal in een van deze drie situaties:

- De observaties zijn verzameld na verloop van tijd
- De observaties zijn gemaakt binnen blokken
- Wanneer de observaties vaker gedaan worden.

2. Homogeniteit van de variatie (homogeneity of variance)

Wanneer de variatie van elke populatie gelijk is, is er homogeniteit van de variatie, of homoscedasticiteit. Wanneer deze aanname niet waar wordt gemaakt zal er een vooroordeel in de SSwith term zijn. Dit kan het risico op een Type I of Type II fout vergroten. Er zitten grote gevolgen wanneer deze aanname niet wordt waargemaakt. Er zijn verschillende manieren om hiermee om te gaan:

- Het gebruik van andere methoden (Welch, Brown-Forysthe, and James methoden)
- Verkleinen van alpha
- Transformatie van Y

3.Normaliteit (Normality)

Elke populatie moet een normale verdeling volgen. De F-toets is redelijk sterk wanneer er een middelmatige schending van deze aanname is. De effecten van het niet waarmaken van de normaliteit aanname zijn minimaal, behalve wanneer n klein is, of wanneer n ongeluk is (unqeual n's) of/en bij extreme niet-normaliteit. Uitbijters (oultliers) kunnen ervoor zorgen dat de aanname niet wordt waargemaakt. Er zijn verschillende manieren om te bepalen of de aanname waargemaakt is of niet:

- Kijken naar een boxplot, histogram etc.
- Een Q-Q plot (Quantile-Quantile plot)
- Een plot met de groepsgemiddelde tegen de groep variaties
- Scheefheid (skewness) en platheid (kurtosis) van de residuen.

Daarnaast wanneer in een residuen plot de punten random verdeeld zijn, laat dit ook normaliteit zien.

1.6

Tot nu toe hebben we verondersteld dat er gelijke n’s (equal n’s, balanced case) zijn, maar dit is niet altijd het geval. Ongelijke n’s (unequal n’s or unbalanced case) kunnen ook voorkomen. Wanneer dit zo is, is de interpretation van de analyse, de aannames enzovoort gelijk aan de situatie wanneer er gelijke n’s zijn.

1.7

Er zijn verschillende alternatieven voor het een-factor fixed-effects ANOVA model:

Kruskal-Wallis test

In deze test wordt er geen aanname gemaakt over de normaliteit van de populatie verdeling, maar er wordt wel een aanname gemaakt over gelijke variaties tussen de groepen. Wanneer de normaliteit aanname wordt waargemaakt, kan je beter het ANOVA model gebruiken, omdat er minder risico is op een Type II fout. In andere gevallen kan je beter de Kurskal-Wallis toets gebruiken.

De Kruskal-Wallis toets werkt als volgt. De observaties van de afhankelijke variabele worden in volgorde gezet van hoog naar laag, zonder te kijken tot welke groep ze behoren. Hiermee kan worden getest of de gemiddelden verschillend zijn tussen groepen. Hiermee kan worden vastgesteldof ze een random steekproef van de populatie zijn. Dus H0 is dat het gemiddelde hetzelfde is voor elke groep en H1is dat deze gemiddelden niet hetzelfde zijn tussen de groepen. De toetsingsgrootheid wordt H genoemd, en wordt vergeleken met de kritische waarde (table A.3). Wanneer H> kritische waarde, dan wordt je nulhypothese verworpen.

Er zijn twee situaties. De eerste is dat chi-kwadraat kritische waarden alleen handig zijn wanneer er op zijn minst drie groepen zijn en op zijn minst vijf observaties per groep. Ten tweede is dat wanneer de rankingen dicht bij elkaar zin dat de steekproef verdeling H kan worden beïnvloed.

De Welch toets, Brown, en Forsythe procedure en de James first and second-order procedures.

Deze testen worden gebruikt om te toetsen voor de heteroscedasticiteit aanname. Voor deze testen is homogeniteit niet nodig. Onderzoek laat zien dat (a) wanneer er homogeniteit is, je beter de F-toets kan gebruiken dan een van deze methoden en (b) wanneer er geen homogeniteit is, je beter een van deze toetsen kan gebruiken.

1.8

Om een one-way ANOVA toets uit te voeren in SPSS, moet je een dataset hebben dat op zijn minst twee variabelen of columns heeft. Eentje laat de levels of categorieën zien van de onafhankelijke variabele, de andere is de afhankelijke variabele. Dit zijn de stappen om de ANOVA toets in SPSS uit te voeren:

1. Ga naar “Analyze” en selecteer “General Linear Model” en selecteer “Univariate”.
2. Sleep de onafhankelijke variable in de box “Dependent variable” en sleep de afhankelijke variabele in de box “Fixed Factors”.
3. Klik op “Options” , hier kan je verschillende opties selecteren.
4. Ga terug naar “Univariate”, en klik op “Plots”. De onafhankelijke variable moet geplaatst worden op de “Horizontal Axis”. Klik op “add” en sleep de variabele in “plots”.
5. Vanaf de “Univariate” box klik op “save” en selecteer de dingen die je wilt opslaan. Vanaf de “Univariate” box klik op “Ok” om de resultaten te genereren.

Nu hebben we de resultaten, en deze moeten geïnterpreteerd worden (kijk naar pagina 28 en 29). De tabel die “Between Subjects Factors” genoemd wordt, laat de steekproefgroottes zijn voor elke categorie van de onafhankelijke variabele. De tabel “descriptive statistics” laat alle basis statistieken zien, zoals de gemiddelden en standaarddeviaties.

De F-toets (en de p-waarde) van de Levene's Test for Equality of Error Variances kan worden bekeken om te zien of je kan aannemen dat er gelijke variaties zijn. De df1 zijn de vrijheidsgraden van de teller, dus J-1. De df2 zijn de vrijheidsgraden van de noemer, dus N-J.

Je kan kijken of er een normale verdeling is van de residuen (een van de aannames van ANOVA model) door in SPSS te klikken op "explore" en een histogram te maken van de residuen. Door te kijken naar de scheefheid (skewness) en platheid (kurtosis) kunnen we concluderen of het een normale verdeling is of niet.

De S-W toets kan dit ook testen. En zoals al eerder genoemd kan je ook kijken naar Q-Q plots of naar een boxplot.

We moeten in SPSS ook testen voor de aanname van onafhankelijkheid. In sommige gevallen is random toewijzing (random assignment) mogelijk, maar dat is niet zo in alle gevallen. Door de residuen in een grafiek te zetten met de onafhankelijke variabele kan worden bekekenof de aanname is waargemaakt of niet. Dit kan je in SPSS doen door te klikken op "Simple Scatterplot" en de residuen in de box "Y axis" te plaatsen, en de onafhankelijke variabele in de box "X axis" te plaatsen. Wanneer de punten random in de grafiek en onder de referentie lijn vallen is er bewijs voor onafhankelijkheid.

De Kruskal-Wallis toets in SPSS heeft de volgende stappen:

1. Ga naar "Analyze" en selecteer "Nonparametric Tests", selecteer "Legacy Dialogs" en selecteer "K Indepdent Samples"

2. Vanaf de "Test for Several Independent Samples" box sleep de afhankelijke variabele in de "Test Variable List" box. Klik daarna de groepering variabele in de "Grouping Variable" box. In deze box moet je aangeven welke categorieën je wilt meerekenen. Dit doe je door op "Define Range" te klikken. Vink daarna de "Kruskall-Wallis H" aan in het schermpje "Test Type".

De nulhypothese van de Kruskal-Wallis toets is dat de gemiddelde van de groepen van de onafhankelijke variabele niet significant verschillend zijn. Door te kijken naar de p-waarde kan worden besloten om de nulhypothese wel of niet te verwerpen.

De Welch and Brown-Forsythe toets heeft in SPSS de volgende stappen:

1. Ga naar "Analyze" en selecteer "Compare Means". Selecteer "One-way ANOVA".

2. Sleep de afhankelijke variabele in de "Dependent list", en sleep de onafhankelijke variabele in de "Factor" box.

3. Klik op "Options". Hier moet je Brown-Forsythe aanvinken.

4. Genereer de resultaten.

De p-waarde die wordt gegenereerd, kan worden gebruikt om te besluiten of de nulhypothese moet worden verworpen of niet. Wanneer de nulhypothese is verworpen, bestaater een significant verschil tussen het gemiddelde van de verschillende groepen.

Om de post hoc power te vinden voor de one-way ANOVA methode gebruiken we G*Power. Hiervoor selecteren we "Tests", daarna selecteer "Means" and daarna "Many groups: ANOVA: One-way (one independent variable)". Selecteer "Type of Power Analysis". Hierin selecteren wij "Post hoc: Compute achieved power-given , sample size and effect size". (de standaard instellingen voor de "Test Family" is de "t-test". De standaard instelling voor "Statistical Test" is "Correlation: Point biserial model". Deze veranderen vanzelf wanneer je de stappen volgt).

De “Input Parameter” moeten worden gespecificeerd met:

1. effectgrootte f
2. Alpha level
3. Totale steekproef grootte
4. Aantal groepen van de onafhankelijke variabele

Wanneer de input parameters zijn bepaald, klik dan op “Calculate” om de power te berekenen. De “output parameters” geven de waarden gegeven de “input parameters”. Uiteindelijk krijg je een post hoc power van de toets:de kans op het verwerpen van de nulhypothese wanneer deze eigenlijk niet waar is. Een power > 0.80 (80%) is voldoende.

Voor de priori power kunnen we de totale steekproefgrootte die we nodig hebben bepalen wanneer we de volgende dingen weten: de geschatte effectgrootte, het alpha niveau, de power die we willen en het aantal groepen van de onafhankelijke variabele.