
Hoorcollege 1

Statistiek: review en ANOVA

Onderzoek doen

Om onderzoek te doen moet je eerst een onderwerp kiezen. Vervolgens moet je de onderzoeksvraag formuleren en operationaliseren (hoe ga ik meten wat ik wil weten). Dan wordt de methode (hoe ga je de data verzamelen) gekozen en de hypothese opgesteld. Hierna wordt de techniek gekozen en de data geanalyseerd. Als laatste kan de conclusie getrokken worden. Een voorbeeld van deze stappen van onderzoek doen is te vinden in de PowerPoint van hoorcollege 1 van Statistiek op BlackBoard.

Het uitvoeren van een toets gaat in vijf stappen:

- Opstellen van de hypothese. De H_0 (nulhypothese) is een stelling voor geen effect, er is dan geen verschil. De H_1 (alternatieve hypothese) heeft wel effect, er is dan wel een verschil. Vervolgens wordt de significantie gekozen, dit is de waarde van α (alfa).
- Dan moeten de kritieke waarden bepaald worden. Dit wordt gedaan door de juiste waarde in de tabel op te zoeken (zoals in de t-tabel van Methode en Statistiek 1). Deze kritieke waarde bepaalt wat het kritieke gebied is. Dit zorgt ervoor dat de beslissingsregel ontstaat. Deze regel kan altijd toegepast worden en zegt dat we H_0 mogen verwerpen wanneer de getoetste waarde in het kritieke gebied valt.
- Het uitrekenen van de toetsingsgrootte. De geobserveerde waarden/verschillen worden dan gerelativeerd. Dit kan door met behulp van SPSS de p -waarde uit te rekenen. Deze geeft ons een idee over hoe extreem de waarde is die geobserveerd is. Hoe extremer de observatie is, hoe kleiner de p -waarde. De tweede beslissingsregel zegt dat we H_0 mogen verwerpen wanneer $p \leq \alpha$.
- Beslissing nemen over H_0 . Dit wordt gedaan door de toetsingsgrootte te vergelijken met de kritieke waarde en door de p -waarde te vergelijken met alfa.
- Rapporteren volgens de APA regels.

MTS 2

Er zullen verschillende onderzoeksmethoden besproken worden in MTS-2. Voorbeelden zijn het experiment, een survey, kwalitatief onderzoek en het gebruik van bestaande gegevens. Daarnaast gaan we kijken naar verschillende nieuwe onderzoeksmethoden en nieuwe statistische technieken.

ANOVA (Hoofdstuk 12)

Wanneer we 2 onafhankelijke groepen willen vergelijken, gebruiken we een t-toets. Wanneer we meer dan 2 groepen willen vergelijken kunnen we geen t-toets meer toepassen. Wanneer we dan dus 3 of meer (onafhankelijke) groepen willen vergelijken, gebruiken we de techniek die variantieanalyse of ANOVA (analysis of variance) heet. Hierbij is het belangrijk om te realiseren dat dit ongeveer hetzelfde is dan de t-toets, waar je kijkt naar het verschil. De voorwaarden die bij de t-toets gelden blijven hetzelfde, evenals de meetniveaus.

Dit wordt uitgelegd aan de hand van een voorbeeld. De onderzoeksvraag is: heeft cafeïne invloed op het concentratievermogen van studenten? Je kunt je afvragen hoe je dit het beste kunt onderzoeken. Je kan bijvoorbeeld een survey afnemen, een experiment uitvoeren of studenten interviewen. Voor dit onderzoek past het beste een experiment. In deze casus krijgt men drie verschillende experimentele groepen: espresso, gewone koffie en decafé. Dit is de operationalisering van het vaststellen van het toedienen van veel /middelmatig/geen cafeïne. Het concentratievermogen kan gemeten worden door middel van een concentratietest. Dit totale proces is de operationalisering van de onderzoeksvraag; hoe gaan we meten en wat gaan we meten. Onderzoek begint dus niet met statistiek. De manier waarop je gaat onderzoeken zou kunnen leiden tot statistiek, maar belangrijk is om bij het begin te beginnen.

Er is sprake van afhankelijke variabelen en onafhankelijke variabelen. De afhankelijke variabele (Y) heeft het minimale meetniveau interval. De onafhankelijke variabele (X) heeft als minimale meetniveau nominaal. Deze onafhankelijke variabele wordt vaak factor genoemd. De verschillende condities of waarden van deze factor worden niveaus of levels genoemd. In het voorbeeld: de Y is het concentratievermogen. Hierbij is het meetniveau interval. De X is de onafhankelijke variabele, in dit geval de hoeveelheid cafeïne, omdat het de experimentele conditie is. Dit is ordinaal, er is een bepaalde volgorde in de groepen.

Stap 1

De eerste stap is de hypothesen bepalen en significantieniveau kiezen. Hierbij is de nulhypothese de gemiddelde concentratievermogen scores die in de drie populaties gelijk zijn. Dit kan ook op een statistische manier geformuleerd worden.

$H_0: \mu_1 = \mu_2 = \mu_3$.

De H_1 hypothese is als de populatiegemiddelden niet allemaal gelijk zijn. Anders geformuleerd: er is ten minste één populatie gemiddelde anders dan de anderen. De alternatieve hypothese kan niet statistisch opgeschreven worden omdat er een enkele uitzondering kan zijn, er staat geen 'allemaal' in de zin. Daarom moet deze altijd in woorden.

Bij het kiezen van een significantieniveau moet je goed nadenken over wat de gevolgen zijn voor het onderzoek. Wat zijn de consequenties als de toets fout blijkt te zijn? Hierop moet je je significantieniveau aanpassen. In het voorbeeld wordt een significantieniveau van 5% gebruikt.

Stap 2

De tweede stap is de kritieke waarden bepalen. Deze stap wordt vaak uitgevoerd aan de hand van SPSS-output, die later gevonden wordt. Daarom is deze stap niet altijd relevant. Meer uitleg hierover staat onder stap 4.

Stap 3

De derde stap is de toetsingsgrootte uitrekenen. Van iedere steekproef wordt het steekproefgemiddelde uitgerekend. Ook wordt er gekeken naar de spreiding binnen elke steekproef. Het is niet het kijken naar het verschil tussen twee gemiddelden in onze casus, maar

naar hoeveel het groepsgemiddelde verschilt van het totale gemiddelde. Als de nulhypothese waar is zijn de verschillen nul of zeer klein. Dit kan pas gezegd worden na het kijken naar het verschil in de spreiding binnen de groepen met de spreiding tussen groepen, omdat je dan pas ziet of het gemiddelde relatief afwijkt of niet. Dit is de variantieanalyse. Het geobserveerde variantie tussen gemiddelden gedeeld door het verwachte gemiddelde variantie zonder effect levert de F op (dit is hetzelfde als bij de t-toets, maar daar is het woord variantie vervangen door verschil).

Er zijn twee maten van spreiding. De spreiding binnen de groepen wordt de binnenvariantie (Mean Square within of Mean Square error) genoemd. De spreiding tussen de groepen wordt de tussenvariantie (Mean Square between) genoemd.

De tussenvariantie (MS_{between}) wordt uitgerekend door de SS_{between} te delen door de df_{between} (dit is $k-1$). De k is het aantal groepen die vergeleken worden.

Dus: $MS_{\text{between}} = SS_{\text{between}} / k - 1$.

De binnenvariantie (MS_{within}) wordt uitgerekend door de SS_{within} te delen door de df_{within} (dit is $N-k$). N is de totale steekproefomvang dus $n_1 + n_2 + n_3$ enzovoort. De k staat voor het aantal groepen.

Dus: $MS_{\text{within}} = SS_{\text{within}} / N - k$.

Om de verschillen tussen de groepen te relativiseren, wordt de tussenvariantie door de binnenvariantie gedeeld. Deze breuk wordt gebruikt als de toetsingsgrootte, zodat meerdere gemiddelden vergeleken kunnen worden. Het symbool is F .

Dus de formule is $F = MS_{\text{between}} / MS_{\text{within}}$.

Stap 4

Bij de vierde stap wordt de beslissing genomen over de H_0 hypothese. We verwerpen H_0 als de p -waarde kleiner of gelijk is aan α . Deze beslissing gebeurt dus aan de hand van de p -waarde. De H_0 wordt ook verworpen als de toetsingsgrootte in het kritieke gebied ligt.

Als H_0 waar is, dan zijn er geen verschillen tussen de groepen, er is geen treatment effect. Als H_0 gelijk is zijn de verschillende varianties van de binnenspreiding en de tussenspreiding ongeveer gelijk. De F is dan ongeveer gelijk aan 1.

Als H_0 niet waar is wordt de tussenspreidingsmaat groter ten opzichte van de binnenspreiding. De F -ratio wordt dan veel groter dan 1.

Maar wanneer is iets veel groter dan 1? Om te bepalen of iets veel groter is dan 1 is de grootte van de steekproef van belang en het aantal groepen wat vergeleken wordt. Dan kan er naar de F -verdeling gekeken worden. De F -verdeling is een scheve verdeling naar rechts. Aan de hand van de F -verdeling kan een kritieke waarde of p -waarde opgezocht worden. Die F -verdeling begint bij nul en wordt dan groter. In het kritieke gebied vinden we de F -ratio zo groot dat de H_0 verworpen wordt.

Waar de kritieke grens ligt, hangt af van het aantal groepen en hoe groot de steekproeven zijn. De F-verdeling wordt gedefinieerd door 2 paar vrijheidsgraden, namelijk boven de streep ($df_{\text{between}} = k-1$) en onder de streep ($df_{\text{within}} = N-k$). Aan de hand van de df en de kritieke waarde vinden we kritieke waarden in de tabel. Als de vrijheidsgraden er niet in staan, altijd het dichtstbijzijnde, lagere getal nemen.

De waarden in de tabel die niet dikgedrukt staan, horen bij een significantieniveau van 0.05. De dikgedrukte waarden in de tabel horen bij een significantieniveau van 0.01.

In het boek staan geen definitieformules voor SS. Deze zijn als volgt:

- Totale spreiding: $SS_{\text{total}} = \text{alle } x (X - M_{\text{total}})^2$
- Tussenspreiding: $SS_{\text{between}} = \text{alle } x (M_i - M_{\text{total}})^2$. Deze meet verschillen tussen de groepen.
- Binnenspreiding: $SS_{\text{within}} = \text{groep 1 } (X - M_1)^2 + \text{groep 2 } (X - M_2)^2 + \dots$
oftewel $SS_{\text{within}} = SS_1 + SS_2 + SS_3 + \dots + SS_k$. Deze meet verschillen binnen de groepen.

Rekenformules in het boek geven de G (grand total), T (total) en SS. Deze mogen overgeslagen worden omdat ze niets laten zien over wat je meet en wat je met de data doet. Je zult in het algemeen geen SS zelf hoeven te berekenen voor een ANOVA. Wel moet je *met* SS kunnen rekenen. Zie dia 50 van de Powerpoint van Statistiek 1.

Om alle verschillende *sources* helder te houden, wordt de ANOVA-tabel gebruikt. Waarin overzichtelijk alle formules weergegeven worden. Verticaal wordt er opgeteld en horizontaal wordt er gedeeld. Op dia 52 van het hoorcollege 1 van Statistiek staat deze tabel afgebeeld. Er zijn een aantal handige formules voor deze tabel:

$$SS_{\text{total}} = SS_{\text{between}} + SS_{\text{within}}$$

$$MS = SS / df$$

Door de MS_{between} te delen door de MS_{within} wordt de F uitgerekend.

Stap 5

Voor de vijfde stap moet de effectgrootte berekend worden. Hiervoor wordt het eta-kwadraat gebruikt. Dit eta-kwadraat geeft het percentage van de verklaarde variantie en is een indicatie van de sterkte van het effect van de factor op de afhankelijke variabele. Kleine waarden geven een klein effect en grote waarden geven een groot effect. De hoeveelheid van de totale spreiding wordt verklaard door de verschillen tussengroepen. Dit levert een percentage op wat gerapporteerd moet worden. De interpretatie hiervan kan aan de hand van de richtlijnen van Cohen. De formule om het eta-kwadraat is als volgt: $\eta^2 = SS_{\text{between}} / SS_{\text{total}}$.

Richtlijnen van Cohen: blz. 299 van Gravetter

$\eta^2 = 0,01$ klein effect

$\eta^2 = 0,09$ medium effect

$\eta^2 = 0,25$ groot effect

Bij ANOVA gebeurt de rapportage aan de hand van de tabel, je zet dus geen gemiddelden in de tekst omdat het dan een grote rommel met allemaal getallen wordt. De tabel laat de gemiddelden zien. Anders mag het via de APA-regel.

Hier volgen nog enkele praktische handreikingen. ANOVA mag gebruikt worden bij 2 of meer dan 2 groepen. Bij 2 groepen mag er gekozen worden welke toets je wilt gebruiken, namelijk een t-toets voor 2 onafhankelijke groepen of een ANOVA. De p-waarde van de twee toetsen zal identiek zijn.

De voorkeur gaat uit volgens de manier van Tukey, die de betrouwbaarheidsintervallen samen gaat vergelijken. De post hoc toets zorgt dat het totale betrouwbaarheidsniveau blijft waar het moet zijn. Dit hoeft alleen via SPSS.

Met post hoc toetsen kun je paarsgewijze vergelijkingen maken. Er zijn verschillende technieken: Bonferroni, Tukey en Scheffé. Voor elk paar van vergelijkingen gebruiken we het betrouwbaarheidsinterval. Alle technieken passen het betrouwbaarheidsniveau van het individuele interval aan, zodat het totale betrouwbaarheidsniveau op 95% blijft. Dit houdt in dat als we de procedure meerdere keren zouden uitvoeren, 95% van de resultaten binnen het interval ligt.