Statistiek – Werkcollege 4 / Hoofdstuk 5

Vraag 1

A) Dan gaan we maar de Dixon Q-test uitvoeren.



Qtab,n=9,α=0.05 = 0.492. Qcal > Qtab dus we nemen aan dat dit een uitbijter is.

B) Aangezien we grenzeloos vertrouwen in de analist hebben (wat ik wat dubious vind) gaan we de Wilcoxon-signed-rank doen. H0: waarnemingen komen uit verzameling met 0.0997M als mediaan. H1: dit is niet het geval.

Uitvoeren van test:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Referentiewaarde: | 0,0997 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| Waarden: | 0,1013 | 0,1005 | 0,1011 | 0,1 | 0,1 | ~~0,0912~~ | 0,1008 | 0,0968 | 0,1017 |
| Verschil ref: | 0,0016 | 0,0008 | 0,0014 | 0,0003 | 0,0003 |  | 0,0011 | -0,0029 | 0,002 |
| Gesorteerd: | 0,0003 | 0,0003 | 0,0008 | 0,0011 | 0,0014 | 0,0016 | 0,002 | -0,0029 |  |
| Rang: | 1,5 | 1,5 | 3 | 4 | 5 | 6 | 7 | -8 |  |
| T- | 8 |  |  |  |  |  |  |  |  |
| T+ | 28 |  |  |  |  |  |  |  |  |
| Tcal | 8 |  |  |  |  |  |  |  |  |

**Ttab is 3 (tabel A.9), dus Tcal > Ttab, en de nulhypothese wordt geaccepteerd (*tegenovergesteld aan wat normaal gebeurt als Tcal > Ttab!)*. Waarschijnlijk zit de zaal dus op dezelfde lijn als de analist.**

C) Aangezien we twee reeksen waarnemingen willen vergelijken pakken we de Mann-Whitney U-test (‘tegenhanger van de two sample t-toets). H0: medianen verschillen significant, H1: medianen niet significant verschillend. α = 0.05. Toetsingsgrootheid:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Waarden: | 0,1013 | 0,1005 | 0,1011 | 0,1 | 0,1 | ~~0,0912~~ | 0,1008 | 0,0968 | 0,1017 | 0,1101 | 0,1201 | 0,1142 | 0,0929 | 0,1098 |
| Gesorteerd: | 0,0929 | 0,0968 | 0,1 | 0,1 | 0,1005 | 0,1008 | 0,1011 | 0,1013 | 0,1017 | 0,1098 | 0,1101 | 0,1142 | 0,1201 |  |
| Rang: | 1 | 2 | 3,5 | 3,5 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |  |
| ni | 8 | 5 |  |  |  |  |  |  |  |  |  |  |  |  |
| Ri | 44 | 47 |  |  |  |  |  |  |  |  |  |  |  |  |

Ucal = 8. Utab,n1=8,n2=5,α=0.05 = 6. Ucal > Utab dus de nulhypothese wordt geaccepteerd, en de zalen zitten (verbazend genoeg) waarschijnlijk op dezelfde lijn.

*Overigens is die waarde 0.0929 van de tweede reeks geen uitbijter, maar dat had je wel kunnen testen.*

D) Analoog vorige werkcollege:

Dus ‘gewone’ 2-sample t-toets, maar voor kleine steekproeven (n < 30).

Eerst

s12 = 0,0000023, s22 = 0.000102607

tν = 11,α = 0.05 = 2.201, tcal = 2,575, tcal > ttab, dus hier zou je concluderen dat de twee zalen waarschijnlijk *niet* op eenzelfde lijn zitten (μA = μB met α = 0.05).

Vraag 2

(Tentamen 7/11/7, p152, opgaven 1b-1d.)

B) Ik heb de keuze uit de volgende niet-parametrische toetsen uit hoofdstuk 5:

* *Wilcoxon signed-rank test:* Tegenhanger van de one-sample t-toets. Kijkt of de gemeten waardes uit een verdeling komen met een bepaalde mediaan.
* *Mann-Whitney U-test:* Tegenhanger van de two-sample t-toets. Kijkt of de gemeten waardes uit een verdeling komen met dezelfde mediaan.

Hier moet even bij worden opgemerkt dat je ook de *Wilcoxon signed-rank test* kunt gebruiken om gepaarde waarnemingen te vergelijken. Dat doe je door de waardes van elkaar af te trekken en “0” te gebruiken als referentiewaarde.

Dat laatste gaan we dan dus natuurlijk ook doen. (NB: Dat je hier al te maken hebt met gemiddelden maakt verder niets uit.)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Vit. D | 4,25 | 4,15 | 4,16 | 4,17 | 4,68 | 6,59 | 6,77 | 6,04 | 6,34 | 6,21 |
| Control | 4,16 | 4,12 | 4,14 | 4,13 | 4,54 | 4,37 | 5,5 | 4,95 | 4,51 | 4,22 |
| Versch. | 0,09 | 0,03 | 0,02 | 0,04 | 0,14 | 2,22 | 1,27 | 1,09 | 1,83 | 1,99 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Data | 0,02 | 0,03 | 0,04 | 0,09 | 0,14 | 1,09 | 1,27 | 1,83 | 1,99 | 2,22 |
| Rang | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Tcal,+ = 55; Tcal,- = 0. Dan kijken we in tabel A.9: je zou ervan uit kunnen gaan dat dit eenzijdig is, omdat je alleen een positief verschil verwacht (en ziet), maar ik vind dat toch een aanname die je niet snel moet maken. In elk geval is in beide gevallen bij n = 10 de getabelleerde waarde (10 en 8, respectievelijk eenzijdig en tweezijdig) groter dan onze toetsingswaarde, die namelijk 0 is. De nulhypothese (medianen reeksen waarnemingen niet sign. verschillend) moet dus worden verworpen, en de H1 wordt geaccepteerd (de medianen zijn sign. verschillend). Er is dus waarschijnlijk sprake van een effect.

C) Als parametrisch alternatief moet je natuurlijk ook een gepaarde waarnemingen toets doen. Om te testen of je deze mag doen moet je testen of de *verschillen* *tussen* dewaarnemingen normaal verdeeld zijn. De volgende testen testen naar verdelingen:

- *χ2-test:* Indelen in hokjes en testen tegen verwachting.

- *Kolmogorov-Smirnov test:* Doet in feite hetzelfde maar dan op continue schaal (plotten met verwachting en onderling verschil testen). Kan echter alleen met volledig gespecificeerde verdelingen (bv. testen of uit N(6.00,0.10)).

- *Shapiro-Wilks:* Test alleen of iets normaal verdeeld is. Test voor afwijking van ideale rechte lijn met kleinste kwadraten methoden, gebeurd gecomputeriseerd.

Deze laatste test is dus het meest ideaal, maar kan ik nu niet uitvoeren. Ook Kolmogorov-Smirnov test kan ik niet uitvoeren omdat je dan gemiddelde en standaarddeviatie moet weten. Ik zou daarom de χ2-test ook kunnen doen. Dan ga je uit van een normale verdeling. Je schat dan het gemiddelde en de standaarddeviatie (van de waarnemingen op de verschillende tijdstippen), en maakt hier een theoretische verdeling mee (die voldoet aan de eisen op pagina 71). Aan de hand daarvan deel je de data in in “hokjes” en bereken je de toetsingsgrootheid.

Het is wel héél moeilijk op het oog te zien of het normaal verdeeld is of niet, vooral omdat het ook nog een gepaarde test is. Als je een grafiekje van plot van de verschillen (zie onder, links) zou je verwachten dat de meeste waarnemingen rond een centrale waarde liggen (zie onder, rechts, voor voorbeeldje). Dit is niet het geval (er zijn in plaats daarvan 5 kleine en 5 grote verschillen), en ik verwacht dus niet dat hier sprake is van een normale verdeling.

