

Toets Statistiek voor TI en BIT (Module 6 -201500118) – 25 januari 2016, 13.45-16.00

Deze toets bestaat uit 5 opgaven. Een formuleblad en tabellen zijn separaat toegevoegd.
Een gewone rekenmachine is toegestaan, een programmeerbare (GR) niet.

1. Voor een project hebben IT-studenten de mogelijke verbeteringen van een energieverbruik-monitor voor huishoudens onderzocht: is het mogelijk om deze meer gebruiksvriendelijk te maken? Als eerste hebben de studenten het gebruik van de huidige energie-monitor onderzocht: een belangrijk aspect is hoeveel tijd het kost om het elektriciteitsverbruik van, bijvoorbeeld, afgelopen maand uit te lezen. 16 van de huidige gebruikers werd gevraagd dit verbruik uit te lezen. De tijden, die het de gebruikers kostte, werden geregistreerd:

2.5 5.5 4.1 4.4 5.8 6.2 2.7 6.0 3.8 3.2 6.0 5.0 1.8 6.1 3.9 7.0

SPSS gaf voor deze gegevens de volgende output (de Kurtosis is de in het dictaat gedefinieerde waarde):

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Dev.	Variance	Skewness	Kurtosis
Completion time old design	16	1.8	7.0	4.625	1.552	2.409	-0.310	1.955

Test of Normality			
	Shapiro-Wilk		
	Statistic	df	Sig.
Completion time old design	0.947	16	0.450

- Bepaal de z-score van de hoogste gemeten tijd: zou je deze waarde extreem noemen? Waarom (niet)?
 - Wat is de betekenis van de waarde voor de scheefheid (*Skewness*) en voor de *Kurtosis* met betrekking tot de eventuele veronderstelling dat de normale verdeling voor de gemeten tijden geldt?
 - Laat zien hoe Shapiro-Wilk's W berekend is, d.w.z.:
 - Bepaal de waarde van de noemer en
 - Geef de eerste en laatste term van de sommatie in de teller: $a_1 X_{(1)} = \dots$ en $a_{16} X_{(16)} = \dots$
 - Kunnen we normaliteit veronderstellen volgens Shapiro-Wilk's toets? Beantwoord deze vraag in 3 stappen:
 - Geef de hypothesen,
 - Bepaal het kritieke gebied voor $\alpha = 10\%$ en
 - Trek je conclusie
 - Als we aannemen dat de 16 tijden gebaseerd zijn op een aselechte steekproef uit alle huidige gebruikers, bepaal dan een 95%-betrouwbaarheidsinterval voor de verwachte tijd. Geef ook een correcte interpretatie van dit interval in woorden.
2. In deze opgave borduren we voort op de in opgave 1 beschreven situatie: de studenten kregen de opdracht een nieuw, gebruiksvriendelijker interface te ontwerpen zodat gebruikers sneller relevante informatie met de monitor kunnen achterhalen, zoals het maandelijks elektriciteit-verbruik. Een nieuwe aselechte steekproef van 16 gebruikers kreeg het nieuwe ontwerp 2 maanden in bruikleen en daarna werden de tijden voor het achterhalen van het elektriciteit-verbruik gemeten. In onderstaande tabel zijn de resultaten van zowel de eerste (opgave 1) als de tweede steekproef opgenomen.

Oud	2.5	5.5	4.1	4.4	5.8	6.2	2.7	6.0	3.8	3.2	6.0	5.0	1.8	6.1	3.9	7.0	$\bar{x}_1 = 4.625, s_1 = 1.552$
Nieuw	2.0	3.2	5.4	3.0	2.4	4.3	2.2	2.3	4.5	4.0	2.9	3.1	0.9	4.3	3.9	4.3	$\bar{x}_2 = \dots, s_2 = \dots$

- Bepaal de niet-vermelde waarden van het steekproefgemiddelde \bar{x}_2 en van de standaardafwijking s_2 .
- Kunnen we aannemen dat de varianties gelijk zijn? Bepaal daartoe voor de geschikte (parametrische) toets met $\alpha = 5\%$:
 - De hypothesen,
 - de waarde van de toetsingsgrootte,
 - het kritieke gebied en
 - de conclusie aangaande de gelijkheid van varianties.

- c. Één van de studenten meent dat de waargenomen tijden aantonen dat de tijden bij het oude ontwerp gemiddeld groter zijn dan die bij het nieuw ontwerp.
Voer een volledige toets uit, uitgaande van normaliteit en gelijke varianties, waaruit blijkt of de conclusie van de student gerechtvaardigd is bij een onbetrouwbaarheid van 1%.
- d. Omdat een aantal studenten in de projectgroep twijfelen aan de juistheid van de aanname van de normale verdeling werd ook een parameter-vrij alternatief voor de toets in c. uitgevoerd: SPSS bepaalde de waarde van Wilcoxon's rangsom voor de tijden bij het oude ontwerp op $W = 326$. Geef voor deze toets de benaderende verdeling van W , bereken de overschrijdingskans bij de waargenomen waarde en beslis weer met een onbetrouwbaarheid van 1%.

3. Behandeling van chronische pijn aan de Achillespees

In het *British Journal of Sports Medicine* (1 februari, 2004) werd een studie naar chronische pijn aan de Achillespees gepubliceerd: elk van de 16 patiënten met chronische pijn aan de Achillespees werd aan een intensieve nieuwe training van 12 weken onderworpen. De dikte van de pees (bij ontstekingen wordt de pees dikker) werd vóór en na de training bleken als volgt te zijn:

Patiënt	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Vóór	11.0	4.0	6.3	12.0	18.2	9.2	7.5	7.1	7.2	6.7	14.2	7.3	9.7	9.5	5.6	8.7
Na	11.5	6.4	6.1	10.0	14.7	3.7	6.1	6.4	5.7	6.5	13.2	7.5	7.4	7.2	6.3	6.0

Gebruik een niet-parametrische toets om na te gaan of de pijnscores zijn afgenomen na de training. Gebruik de 8 stappen van de toetsingsprocedure met $\alpha = 5\%$.

4. Museum Management

In het blad *Museum Management and Curatorship* (Juni 2010) werd een wereldwijd onderzoek onder 30 bekende musea voor modern kunst gepubliceerd. Elke museumdirecteur is gevraagd welke prestatie maat het belangrijkste is bij interne evaluaties. De resultaten zijn samengevat in nevenstaande tabel.

Prestatiemaat	Aantal musea
Totaal aantal bezoekers	8
Aantal betalende bezoekers	5
Aantal grote evenementen	6
Omvang donaties	7
Aantal leden	4

Is één van de prestatie maten meer in gebruik dan de andere?

Beantwoord deze vraag door een geschikte toets uit te voeren met een significantieniveau van 10%.

5. X_1, X_2, \dots, X_{10} is een aselechte steekproef uit een $N(\mu, \mu^2)$ -verdeling met onbekende verwachting $\mu = \sigma$ (μ en σ zijn gelijk!). Beschouw de volgende twee schatters van μ : $T_1 = \frac{X_1 + X_2}{2}$ en $T_2 = \frac{X_1 + X_2 + \dots + X_{10}}{11}$.
- Ga voor beide schatters na of deze zuiver zijn.
 - Bepaal voor beide schatters de verwachte kwadratische fout (uitgedrukt in μ).
 - Welke van de twee schatters is de beste?

Cijfer = $1 + \frac{\text{aantal punten}}{44} \cdot 9$,

afgerond op 1 decimaal.

1					2				3	4	5			
a	b	c	d	e	a	b	c	d			a	b	c	Tot
2	1	2	3	3	2	4	6	3	6	6	2	3	1	44