

# Voorbeeldtoets Statistiek voor INF en BIT (Module 6 -201400256) – tijdsduur 2.15 uur

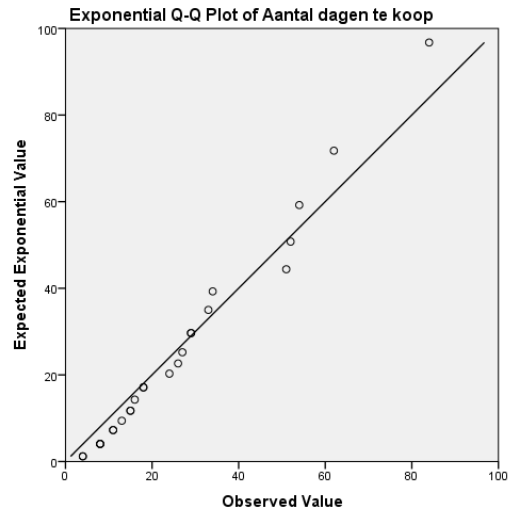
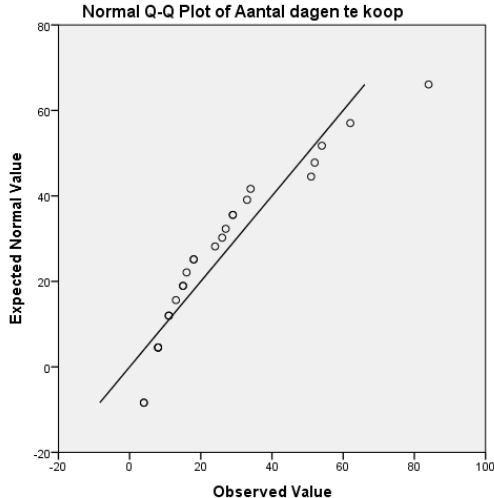
Deze toets bestaat uit 4 opgaven, een formuleblad en de  $N(0,1)$ -,  $t$ -,  $\chi^2$ - en Shapiro-Wilk-tabellen.  
Een gewone rekenmachine is toegestaan, een programmeerbare (GR) niet.

1. Een garage houdt het aantal dagen bij dat een occasion te koop staat.

Van 25 occasions zijn in onderstaande tabel de aantallen verkooptdagen en de bijbehorende numerieke samenvatting gegeven. De meetgegevens zijn al gerangschikt van klein naar groot. De verkoper vraagt zich af of de exponentiële dan wel de normale verdeling een goed model is voor het “aantal dagen in de verkoop”.

4	4	8	8	8	<b>Numerieke samenvatting:</b>		
11	11	13	15	15		Steekproefomvang	25
16	18	18	24	26		Steekproefgemiddelde	26.16
27	29	29	33	34		Steekproefstandaardafwijking	20.32
51	52	54	62	84		Steekproefvariantie	412.89
					Steekproefscheefheidcoëfficiënt	1.30	
					Steekproefkurtosis	4.38	

- Ga na of er sprake is van uitschieters volgens de  $1.5 \times IKA$ -regel.
- Vergelijk de (steekproef)mediaan met het steekproefgemiddelde: wat betekent dat voor de data?
- Wat zou, op basis van je antwoorden op onderdelen a en b en de gegeven numerieke samenvatting, jouw keuze zijn: het normale of het exponentiële model? Motiveer je antwoord kort en duidelijk.
- Wordt je keuze bij c. bevestigd door onderstaande QQ-plots? Waarom (niet)?



- Om een keuze beter te onderbouwen is ook de toets van Shapiro-Wilk uitgevoerd:  $W = 0.870$ .
    - De coëfficiënt  $a_3$  in de formule van  $W$
    - het kritieke gebied voor  $\alpha = 10\%$
    - de conclusie die je daaruit kunt trekken t.a.v. de normaliteit.
  - Als de verkoper met een 95%-betrouwbaarheidsinterval voor het “gemiddelde aantal dagen in de verkoop” wil bepalen op grond van deze meetgegevens, hoe zou dan je advies luiden?
2. Een gezaghebbende verkeersdeskundige beweert dat hoogstens 10% van alle personenauto's uit een bepaalde regio niet voldoet aan de wettelijke eisen betreffende lichten, remmen en richtingaanwijzers. Gealarmeerd door deze uitspraak trekt men een steekproef door lukraak 400 personenauto's te kiezen en

te controleren,  $x = 73$  auto's bleken niet te voldoen aan de eisen. Laat  $p$  de (werkelijke) fractie personenauto's in de regio zijn die niet voldoet aan de wettelijke eisen betreffende lichten, remmen en richtingaanwijzers.

- a. Geef een zuivere schatter van  $p$  en bepaal de verwachte kwadratische fout van die schatter.
  - b. Bereken een 95%-betrouwbaarheidsinterval voor  $p$  en verwoord een interpretatie van dit interval.
  - c. Als we  $H_0: p = 0.1$  verwerpen tegen  $H_1: p > 0.1$  als  $X \geq 55$ , bereken dan:
    1. De onbetrouwbaarheid van deze toets
    2. Het onderscheidend vermogen van de toets als in werkelijkheid  $p = 0.15$ .
3. Een student moet een scriptie schrijven over de verschillen in gemeentelijke belastingtarieven. Bij veldonderzoek stuit hij op het vermoeden dat de leges voor (ver)bouwvergunningen verschillen: de grotere steden ( $> 100.000$  inwoners) zouden daar hogere tarieven voor hanteren dan plattelandsgemeenten. Hierover zijn echter geen gegevens beschikbaar. Op basis van een tweetal aselechte steekproeven van 8 steden en 8 plattelandsgemeenten tracht hij hierover een uitspraak te doen. De resultaten hebben betrekking op de leges voor bouwplannen ter waarde van € 20.000 en zijn als volgt gepresenteerd (in €):

nr.	1	2	3	4	5	6	7	8	$\bar{x}$	$s$
Stad	397	410	428	500	511	519	528	640	492	80
platteland	402	457	450	410	560	381	458	501	452	58

- a. Voer een parametrische toets uit waaruit blijkt of het vermoeden, dat de betreffende leges in steden hoger zijn dan op het platteland, juist is. Pas de toetsingsprocedure toe en bepaal daarbij het kritieke gebied met  $\alpha = 5\%$ .
  - b. Ga nu uit van niet-normale verdelingen en pas op dezelfde data set een volledige niet-parametrische toets toe, weer met  $\alpha = 5\%$ . Beslis nu met de overschrijdingskans.
4. In een opiniepeiling werd Nederlandse kiezers gevraagd welke partij ze hebben gestemd bij de laatste verkiezingen en of zij positief dan wel negatief staan ten opzichte van het beleid van het huidige VVD-PvdA-kabinet. De resultaten zijn in de volgende tabel weergegeven:

		Mening		Totaal
		positief	negatief	
Politieke voorkeur	VVD	90	130	220
	PvdA	90	100	190
	Overige partijen	120	270	390
	Totaal	300	500	800

- a. Ga met een geschikte toets na of de 3 politieke groepen kiezers verschillend denken over het huidige kabinetsbeleid. Neem  $\alpha = 1\%$  en hanteer de 8 stappen van de toetsingsprocedure.
- b. Bepaal een 95%-betrouwbaarheidsinterval voor het verschil in fracties van VVD- en van PvdA-kiezers die positief staan ten opzichte van het kabinet

**Normering:** toetscijfer =  $1 + \frac{\text{aantal punten}}{44} \times 9$

1					2			3		4		<b>Tot</b>	
a	b	c	d	e	f	a	b	c	a	b	a	b	
2	2	2	1	3	1	2	4	4	6	6	8	3	<b>44</b>