
Eindtentamen 2011-2012

Meerkeuzevragen

1. In een enquête wordt gevraagd de vier cijfers van je postcode in te vullen. Wat voor soort data levert dit op?
 - a. intervaldata
 - b. kwantitatieve data
 - c. nominale data **JUIST**
 - d. ordinale data

2. Autofabrikant WMB beweert dat zijn auto's gemiddeld meer dan 20 kilometer kunnen afleggen per liter benzine. In een steekproef van 30 auto's van het merk WMB wordt gemiddeld 22 kilometer afgelegd per liter benzine en bedraagt de steekproefstandaarddeviatie aangaande het aantal kilometers dat een auto per liter benzine aflegt 4. Stel dat je de bewering van de fabrikant toetst met behulp van deze steekproef. Welke van de volgende beweringen is juist ervan uitgaand dat je een normale verdeling mag gebruiken?
 - a. De p -waarde van de toets ligt tussen 0,005 en 0,010; er zijn overweldigende statistische aanwijzingen dat de bewering van de fabrikant klopt. **JUIST**
 - b. De p -waarde van de toets ligt tussen 0,010 en 0,05; er zijn sterke statistische aanwijzingen dat de bewering van de fabrikant klopt.
 - c. De p -waarde van de toets ligt tussen 0,05 en 0,10; er zijn zwakke statistische aanwijzingen dat de bewering van de fabrikant klopt.
 - d. De p -waarde van de toets is groter dan 0,10; er zijn nauwelijks statistische aanwijzingen dat de bewering van de fabrikant klopt.

3. Beschouw de volgende toets aangaande een populatiegemiddelde:

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0,$$

bij een bepaald significantieniveau α . Welke van de onderstaande beweringen is juist?

- a. Als H_0 wordt verworpen, dan weet je dat H_1 waar is.

- b. Afhankelijk van α kan één specifieke waarde van het steekproefgemiddelde \bar{x} zowel leiden tot het verwerpen als het niet verwerpen van H_0 . **JUIST**
- c. Als H_0 niet wordt verworpen, dan weet je dat H_1 niet waar is.
- d. Als het steekproefgemiddelde \bar{x} gelijk is aan μ_0 , dan weet je dat H_1 niet waar is.
4. Voor Bakker Bas is het belangrijk om te weten hoe groot de spreiding in de dagelijkse vraag naar krentenbollen is. De standaardafwijking in de dagelijkse vraag was tot voor kort 10 krentenbollen, maar Bakker Bas vermoedt dat dat de laatste tijd sterk is toegenomen. Om dit te toetsen bij een significantieniveau van 0,05 noteert hij gedurende 30 dagen de dagelijkse vraag naar krentenbollen en berekent hij de bijbehorende steekproefstandaarddeviatie s . Wat is het verwerpingsgebied in termen van s bij deze toets?
- a. $s > 146,90$.
- b. $s > 42,60$.
- c. $s > 12,12$. **JUIST**
- d. $s > 6,53$.
5. Gegeven is de volgende steekproef:
7,2 8,3 6,1 5,4 6,0 4,8 6,8
Wat is de bijbehorende standaarddeviatie?
- a. 1,08.
- b. 1,17. **JUIST**
- c. 1,37.
- d. 6,37.

6. Texaan Bill gaat meedoen aan een schietwedstrijd. Hij heeft twee geweren, maar mag er slechts eentje gebruiken. Om te bepalen welk geweer de hoogste precisie heeft, voert hij een matched pairs experiment uit. Dit levert de volgende tabel op:

Afstand tot schietbord (in m)	20	30	40	50	60	70	80	90	100
Geweer A	2	3	4	6	8	10	12	7	15
Geweer B	2	2	4	5	9	10	10	10	12

Bill vermoedt dat geweer A minder precies is en toetst dit bij een significantieniveau van 10%. Welke conclusie trekt Bill?

- a. De toetsgroottheid ligt in het verwerpingsgebied; de nulhypothese dat er geen verschil in precisie tussen de geweren zit wordt verworpen.
 - b. De toetsgroottheid ligt in het verwerpingsgebied; de nulhypothese dat er geen verschil in precisie tussen de geweren zit wordt niet verworpen.
 - c. De toetsgroottheid ligt niet in het verwerpingsgebied; de nulhypothese dat er geen verschil in precisie tussen de geweren zit wordt verworpen.
 - d. De toetsgroottheid ligt niet in het verwerpingsgebied; de nulhypothese dat er geen verschil in precisie tussen de geweren zit wordt niet verworpen. **JUIST**
7. Beschouw de chi-kwadraat toets aangaande een samenhangstabel (*contingency table*). Wat is het effect van het verkleinen van de steekproefgrootte op het verwerpingsgebied?
- a. Het verwerpingsgebied wordt groter.
 - b. Het verwerpingsgebied blijft gelijk. **JUIST**
 - c. Het verwerpingsgebied wordt kleiner.
 - d. Dat valt niet te zeggen zonder aanvullende informatie.
8. Het is van belang dat machines die voetballen produceren zeer nauwkeurig zijn en dat ze voetballen produceren met een omtrek tussen 68 cm. en 70 cm. In een steekproef bestaande uit 25 ballen geproduceerd door één machine is de som van de omtrekken 1720 cm. en de som van de gekwadrateerde omtrekken 118366 cm.². Wat is het 95 % betrouwbaarheidsinterval voor de variantie in de populatie van alle ballen geproduceerd door deze machine?
- a. Dat loopt van 0,66 cm.² tot 3,03 cm.².
 - b. Dat loopt van 0,74 cm.² tot 2,29 cm.².
 - c. Dat loopt van 0,76 cm.² tot 2,42 cm.². **JUIST**

- d. Dat loopt van $0,82 \text{ cm.}^2$ tot $2,17 \text{ cm.}^2$.
9. Om te bepalen of de locaties van populatie A en populatie B van elkaar verschillen, wordt de Wilcoxon-rangsomtoets (*Wilcoxon Rank Sum test*) uitgevoerd. Uit elke populatie wordt een steekproef ter grootte van 5 genomen. Deze steekproeven leveren de rangsommen T_A en T_B op. Drie van de vier hieronder gegeven combinaties zijn niet mogelijk. Welke is wel mogelijk?
- $T_A = 10$ en $T_B = 45$.
 - $T_A = 20$ en $T_B = 30$
 - $T_A = 30$ en $T_B = 25$. **JUIST**
 - $T_A = 40$ en $T_B = 10$.
10. Twee onderzoekers bestuderen onafhankelijk van elkaar hetzelfde fenomeen. Daarbij dient de volgende toets te worden uitgevoerd:

$$H_0 : \mu = 355, \quad H_1 : \mu < 355, \quad \alpha = 0,10.$$

De standaarddeviatie in de populatie is gegeven: $\sigma = 60$. Beide onderzoekers trekken een steekproef ter grootte van 40. Als in werkelijkheid $\mu = 335$, wat is dan de kans dat één onderzoeker de nulhypothese verwerpt en de andere de nulhypothese niet verwerpt?

- Die kans is afgerond 0,08.
 - Die kans is afgerond 0,16.
 - Die kans is afgerond 0,24.
 - Die kans is afgerond 0,32. **JUIST**
11. Een onderzoekster voert de volgende toets uit:

$$H_0 : \mu = 40, \quad H_1 : \mu \neq 40, \quad \alpha = 0,05.$$

De onderzoekster weet dat $\sigma = 7$ en was eerst van plan een steekproef van 30 te nemen. Omdat haar budget onverwacht is toegenomen, kan ze een steekproef van 40 nemen. Wat is het effect van deze toename in de steekproefgrootte?

- De kans op een type I fout neemt af.

- b. De kracht (*power*) van de toets neemt af.
 - c. De kans op een type I fout neemt toe.
 - d. De kracht (*power*) van de toets neemt toe. **JUIST**
12. In welke van de onderstaande situaties is het 95% betrouwbaarheidsinterval aangaande het populatiegemiddelde het smalst?
- a. Als $\sigma = 2$ en $n = 50$.
 - b. Als $\sigma = 2$ en $n = 70$.
 - c. Als $s = 2$ en $n = 70$.
 - d. Als $s = 2$ en $n = 90$. **JUIST**
13. Er zijn verkiezingen in Verweggistan. Bij de vorige verkiezingen haalde Leefbaar Verweggistan 25% van de stemmen. Uit een exit poll blijkt dat deze keer 21% van de stemmers op Leefbaar Verweggistan heeft gestemd. Deze exit poll is simpelweg een steekproef bestaande uit het stemgedrag van 500 personen. Men toetst aan de hand van de exit poll of de populariteit van Leefbaar Verweggistan is veranderd bij een significantieniveau van $\alpha = 0,025$. Welke conclusie wordt getrokken?
- a. De p -waarde is groter dan 0,025 en daarom wordt de nulhypothese dat de populariteit niet is veranderd niet verworpen. **JUIST**
 - b. De p -waarde is groter dan 0,025 en daarom wordt de nulhypothese dat de populariteit niet is veranderd wel verworpen.
 - c. De p -waarde is kleiner dan 0,025 en daarom wordt de nulhypothese dat de populariteit niet is veranderd niet verworpen.
 - d. De p -waarde is kleiner dan 0,025 en daarom wordt de nulhypothese dat de populariteit niet is veranderd wel verworpen.
14. De kans dat het op een februaridag in Groningen gaat regenen als het de vorige dag regende is 0,8 en de kans dat het op zo'n dag droog blijft als het de vorige dag niet regende is 0,9. Stel het blijft vandaag droog. Wat is de kans dat we aanstaand weekend minstens één regenachtige dag hebben?
- a. Die kans is afgerond 0,20.
 - b. Die kans is afgerond 0,25. **JUIST**

- c. Die kans is afgerond 0,30.
 - d. Die kans is afgerond 0,35.
15. Op 2 januari 2011 had Albert Heijn een marktaandeel van 40%, Super de Boer had een marktaandeel van 25%, Jumbo had 20% en Aldi had 15%. Op 2 januari 2012 heeft men 500 willekeurig gekozen Nederlanders gevraagd bij welke supermarkt ze hun boodschappen doen. Het resultaat: 222 mensen gingen naar Albert Heijn, 136 naar Super de Boer, 64 naar Jumbo en 78 naar Aldi. Toets bij een significantieniveau van 1% of de marktaandelen van de vier supermarktketens zijn veranderd in de loop van 2011. Welke conclusie trek je?
- a. De nulhypothese dient verworpen te worden als $\chi^2 > 13,3$ en we verwerpen de nulhypothese niet.
 - b. De nulhypothese dient verworpen te worden als $\chi^2 > 13,3$ en we verwerpen de nulhypothese wel.
 - c. De nulhypothese dient verworpen te worden als $\chi^2 > 11,3$ en we verwerpen de nulhypothese niet.
 - d. De nulhypothese dient verworpen te worden als $\chi^2 > 11,3$ en we verwerpen de nulhypothese wel. **JUIST**

Open vragen

1. Onze Bakker Bas heeft een nieuw type brood ontwikkeld. Zeven klanten wordt gevraagd dit nieuwe type te beoordelen op een vijfpuntsschaal (1: heel vies, 2: beetje vies, 3: neutrale smaak, 4: lekker, 5; zeer lekker). Een andere groep van acht klanten wordt gevraagd hun oordeel te vellen (op dezelfde vijfpuntsschaal) over het populairste type brood dat Bakker Bas al in het assortiment heeft. De resultaten zijn:

Nieuw type: 3 5 3 2 2 4 4

Populairste type: 2 4 2 3 2 4 2 1

- a. Bakker Bas besluit het nieuwe type brood in het assortiment op te nemen als de waarderingen aangaande dit nieuwe brood over het algemeen hoger liggen dan de waarderingen aangaande z'n populairste type brood. Hierbij hanteert hij een significantieniveau van

0,05. Formuleer de hypotheses die horen bij de toets die Bakker Bas dient uit te voeren en voer die toets uit. Neemt Bakker Bas het nieuwe type op in het assortiment? (14 punten)

De *Wilcoxon rank sum test* dient hier uitgevoerd te worden (3 pt.).
 Bakker Bas neemt alleen het nieuwe brood op in het assortiment als de bijbehorende waarderingen over het algemeen hoger zijn, de alternatieve hypothese is dus eenzijdig. De hypotheses zijn: H_0 : De locaties van de twee populaties verschillen niet, H_1 : De locatie van de waarderingen voor het nieuwe brood (populatie 1) liggen rechts van die van het populairste brood (populatie 2). (2pt.)
 Berekening T_1 : $T_1 = 9 + 15 + 9 + 4\frac{1}{2} + 4\frac{1}{2} + 12\frac{1}{2} + 12\frac{1}{2} = 67$. (3 pt.)
 Omdat $n_1 = 7$, $n_2 = 8$ en $\alpha = 0,05$, verwerp je de nulhypothese als $T_1 > 71$. (3 pt.)
 Bakker Bas verwerpt de nulhypothese dus niet (2 pt.) en het nieuwe type wordt niet in het assortiment opgenomen. (1 pt.)

- b. De zeven personen die hun mening gaven over het nieuwe type werd ook gevraagd wat ze een geschikte prijs voor dit brood vonden. Deze vraag leverde een gemiddelde prijs van 1,59 euro op met een standaardafwijking van 0,25 euro. Bakker Bas had zelf een prijs van 1,75 euro in gedachte. Om na te gaan of 1,75 euro niet teveel is, toetst hij bij een significantieniveau van 0,05 of de door klanten gemiddeld geschikt geachte prijs lager is dan 1,75 euro. Voor deze toets uit door de p -waarde te benaderen. Je mag ervan uitgaan dat de geschikt geachte prijzen normaal verdeeld zijn. (11 punten)

De vraag leverde een standaardafwijking op, dit betekent dat alleen een steekproefstandaardafwijking bekend is ($s = 0,25$) en er een t -toets met $7-1=6$ vrijheidsgraden uitgevoerd dient te worden. (3 pt.)
 De vraag is of 1,75 euro te duur is: de alternatieve hypothese is eenzijdig. De hypotheses zijn: $H_0 : \mu = 1,75$, $H_1 : \mu < 1,75$. (2 pt.)
 De toetsgrootte: $t = \frac{\bar{x}-1,75}{0,25/\sqrt{7}} \approx -1,693$. (2 pt.)
 De waarde van de toetsgrootte ligt tussen $-t_{0,050;6}$ en $-t_{0,100;6}$, de p -waarde ligt dus tussen 0,05 en 0,10. (2 pt.)
 Aangezien de p -waarde groter is dan $0,05 = \alpha$, wordt de nulhypothese niet verworpen. (2 pt.)

2. In Texas zijn twee grote ranches, die van Hank en die van Butch. Beide

boeren hebben momenteel 200 koeien in de aanbieding. Schietliefhebber Bill handelt in Texaanse koeien. Tot voor kort kocht hij alle koeien van boer Hank en alle koeien van boer Butch. Door de crisis heeft hij niet meer voldoende geld om van beide boeren alle koeien te kopen. Hij besluit daarom eerst te bepalen welke boer de beste aanbieding heeft, dat wil zeggen wie de vetste koeien aanbiedt. Een bevriende veearts stelt de volgende steekproefgegevens beschikbaar:

$$\begin{aligned} n_1 &= 50, & \bar{x}_1 &= 700, & s_1^2 &= 30, \\ n_2 &= 40, & \bar{x}_2 &= 690, & s_2^2 &= 40, \end{aligned}$$

waarbij steekproef 1 slaat op de gewichten van de koeien van boer Hank en steekproef 2 op die van Butch.

- a. Toets of de variantie aangaande de gewichten in de twee populaties koeien hetzelfde is. Gebruik hierbij een significantieniveau van 0,05. (11 punten)

De hypothesen van de toets zijn: $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$, $H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$. (2 pt.)
 De waarde van de toetsgrootte F is $\frac{s_1^2}{s_2^2} = 0,75$. (1 pt.)
 Onder de nulhypothese is deze toetsgrootte $F_{49,39}$ -verdeeld. (2 pt.)
 Je verworpt de nulhypothese als $F < F_{0,975;49;39} = \frac{1}{F_{0,025;39;49}} \approx \frac{1}{F_{0,025;40;50}} = \frac{1}{1,80} \approx 0,56$ en als $F > F_{0,025;49;39} \approx F_{0,025;50;40} = 1,83$. (4 pt.)
 Aangezien $F = 0,75$ niet in het verworpsgebied ligt, wordt de nulhypothese niet verworpen. (2 pt.)

- b. Toets of het gemiddelde gewicht van Hanks koeien significant afwijkt van het gemiddelde gewicht van Butchs koeien. Gebruik weer een significantieniveau van 0,05. (10 punten)

De hypothesen zijn: $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$. (2 pt.)
 Beide populatievarianties zijn onbekend, maar uit a. blijkt dat we mogen aannemen dat $\sigma_1 = \sigma_2$. De gelijke-variantiestoetsgrootte aangaande $\mu_1 - \mu_2$ mag dus gebruikt worden. Onder de nulhypothese volgt dat een t -verdeling met 88 vrijheidsgraden. (3 pt.)
 De gepoolde steekproefvariantie is $s_p^2 = \frac{49 \times 30 + 39 \times 40}{88} = 34,432$. De waarde van de toetsgrootte t is dus $\frac{700 - 690 - 0}{\sqrt{34,432(\frac{1}{50} + \frac{1}{40})}} = 8,034$. (2 pt.)
 De nulhypothese wordt verworpen als $t < -t_{0,025;88} \approx -t_{0,025;90} = -1,987$ of als $t > t_{0,025;88} \approx 1,987$. (2 pt.)
 Omdat $t = 8,034$ binnen het verwerpingsgebied valt, kunnen we concluderen dat de twee populatiegemiddelden significant van elkaar verschillen. (1 pt.)

- c. Door een merkwaardig incident op de schietbaan hebben Bill en boer Hank ruzie gekregen en is Bill minder geneigd om de koeien van Hank te kopen. Aan de andere kant, zaken zijn zaken. Hij besluit alleen de koeien van Hank te kopen als de nulhypothese van de volgende toets verworpen kan worden:

$$H_0 : \mu_1 = \mu_2 + 10, \quad H_1 : \mu_1 > \mu_2 + 10, \quad \alpha = 0,05.$$

Als hij de nulhypothese niet kan verwerpen, koopt hij de koeien van Butch. Van wie koopt Bill koeien? (9 punten)

Onder de nulhypothese volgt $t = \frac{\bar{x}_1 - \bar{x}_2 - 10}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}$ een t -verdeling met 88 vrijheidsgraden. Dit volgt uit het feit dat 10 optellen bij alle waarnemingen in populatie 2 geen invloed heeft op de variantie σ^2 en ook niet op s_p^2 . (4 pt.)
 De waarde van de toetsgrootte t is $\frac{700 - 690 - 10}{\sqrt{34,432(\frac{1}{50} + \frac{1}{40})}} = 0$. (1 pt.)
 De nulhypothese wordt verworpen als $t > t_{0,05;88} \approx t_{0,05;90} = 1,662$. (2 pt.)
 De nulhypothese wordt niet verworpen, Bill koopt daarom de koeien van Butch. (2 pt.)