Tentamen Statistiek MBW/KW (deel 1, eerste kans + finale kans)

Afdeling: Propedeuse MBW/KW 2022-2023 en 2021-2022
Examinator: Dr. J.B.M. Melissen, T. Zijlstra MSc.
Datum: vrijdag 9 juni 2023 09:00 – 12:00, duur tentamen: 3 uur

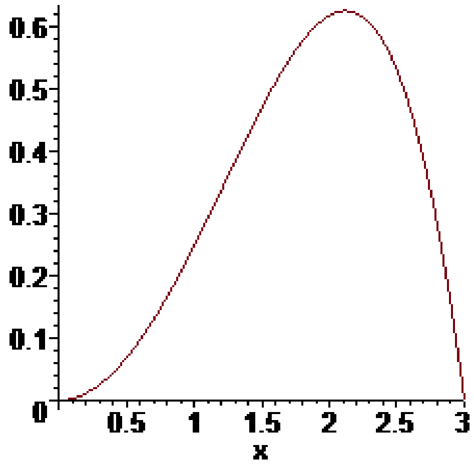
1. **Alle antwoorden moeten gemotiveerd worden**!
2. Rond eindantwoorden (kommagetallen) af op *vier* decimalen, tenzij anders vermeld.
3. Boeken, reader en aantekeningen mogen worden geraadpleegd.
4. De aanwezigheid van *communicatieapparatuur* is niet toegestaan.
5. Het gebruik van een (grafische) rekenmachine met statistische programmatuur en het raadplegen van de bijbehorende handleiding is toegestaan. Het *statistische* gebruik van deze rekenmachine is bij een aantal onderdelen ingeperkt. Let op de aanwijzingen!
6. **De opgaven dienen na afloop van het tentamen ingeleverd te worden.**

Dit tentamen bestaat uit vier opgaven van elk 25 punten. Score = Puntentotaal/10

Opgave 1 [25pt].

Voor een kansvariabele $\underline{x}$ is de volgende kansdichtheidsfunctie gegeven (zie hiernaast):

$$f(x) = \begin{cases} \frac{5}{162}x^2(9-x^2) & \text{als } 0 \le x \le 3 \\ 0 & \text{elders} \end{cases}$$



1a [5pt]. Toon aan dat $f(x)$ een goed gedefinieerde kansdichtheidsfunctie is.

Je moet aantonen dat

1. $f(x) \ge 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Er geldt

3. Tussen 0 en 3 is $x^2 \le 3^2 = 9$, dus $9 - x^2 \ge 0$, en verder is $x^2 \ge 0$, dus ook $f(x) \ge 0$. Buiten $0 \le x \le 3$ is $f(x) = 0 \ge 0$. **2pt**
1. $\int_{-\infty}^{\infty} f(x)dx = \int_0^3 \frac{5}{162}x^2(9-x^2)dx = 1.$ **3pt**

(Definities bekend max 3 pt)

1b [7pt]. Bereken de verwachtingswaarde en de standaarddeviatie van $\underline{x}$.

$\mu = E(\underline{x}) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^3 x\frac{5}{162}x^2(9-x^2)dx = \frac{15}{8} = 1{,}875$ **4pt**

$\mathrm{Var}(\underline{x}) = E(\underline{x}^2) - E(\underline{x})^2 = \int_{-\infty}^{\infty} xf(x)dx - \mu^2 = \int_0^3 x\frac{5}{162}x^2(9-x^2)dx - \mu^2 = \frac{27}{7} - \frac{225}{64} = \frac{153}{448} = 0{,}3415$ **4pt**

$\sigma = \sqrt{\mathrm{Var}(\underline{x})} = \sqrt{0{,}3415} = 0{,}5844$ **1pt**

1c [4pt]. Bereken de kans dat de waarde van $\underline{x}$ tussen 1 en 2 zit.

$$P(1 \leq \underline{x} \leq 2) = \int_1^2 f(x)dx = \int_1^2 \frac{5}{162}x^2(9-x^2)dx = \frac{37}{81} = 0{,}4568$$

1d [2pt]. De mediaan van de variabele $\underline{x}$ is de waarde m waarvoor geldt $P(\underline{x} < m) = P(\underline{x} > m)$. Toon aan dat $P(\underline{x} < m) = 0{,}5$.

Het totale oppervlak is 1. m verdeelt de oppervlakte in twee gelijke delen, dat moet dan 0,5 en 0,5 zijn.

$$P(\underline{x} < m) = P(\underline{x} > m) = 1 - P(\underline{x} < m)$$

dus $P(\underline{x} < m) = 0{,}5$.

Er moet gelden $\int_0^m f(x)dx = 0{,}5$.

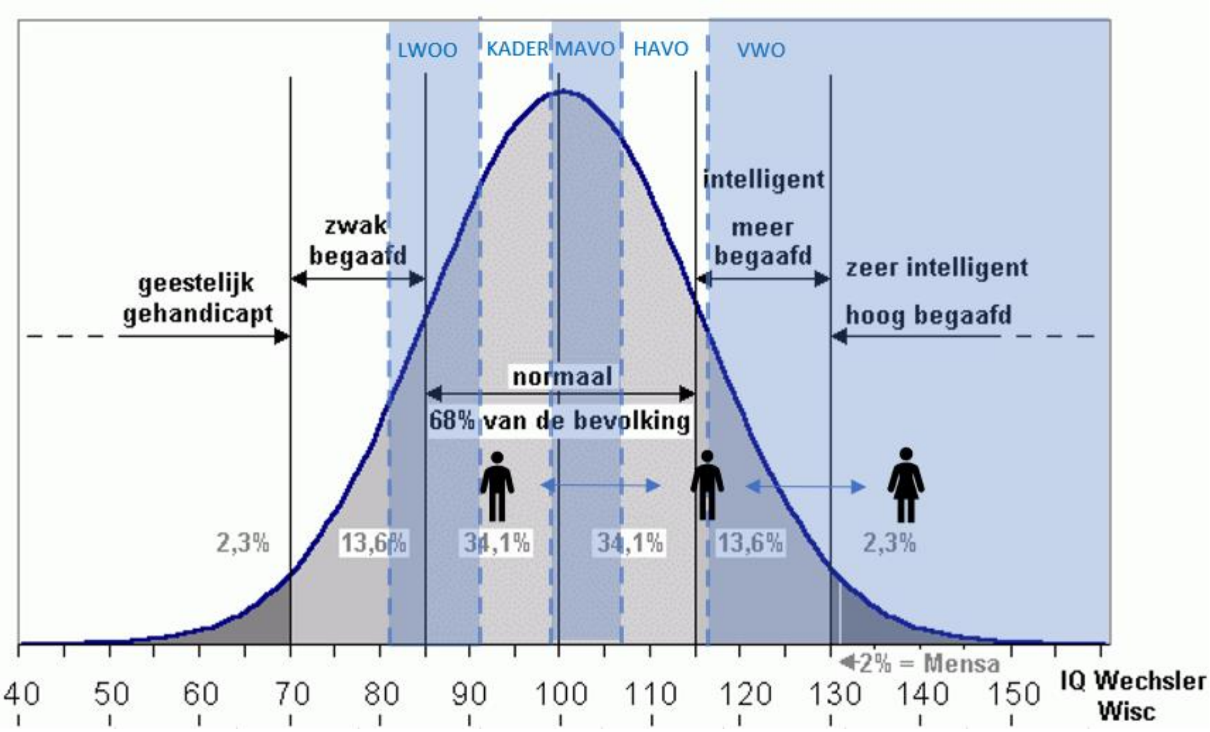
1e [4pt]. Bereken of schat de waarde van de mediaan m in twee decimalen nauwkeurig. Hint: Teken met de GR de functie $g(m) = \int_0^m f(x)dx$ (even wachten) en bereken voor welke waarde van m deze functie gelijk is aan 0,5 (intersect). Of: Bereken met de GR de waarde van $\int_0^m f(x)dx$ voor $m = 1{,}5$ en pas m net zolang aan tot de integraal in de buurt van 0,5 komt.

Antwoord: $m = 1.9294$.

1f [3pt]. Bereken $P(\underline{x} = m)$.

Omdat $\underline{x}$ een continue kansverdeling heeft is de kans op een specifieke waarde altijd gelijk aan 0, dus ook $P(\underline{x} = m) = 0$.

Opgave 2 [25pt]



Het Intelligentie Quotiënt (IQ) is een getal dat een indicatie geeft van de intelligentie van een persoon. Er wordt aangenomen dat het IQ normaal is verdeeld met een gemiddelde van 100 en een standaarddeviatie van 15. Het IQ stijgt tot de leeftijd van 18 jaar en neemt daarna niet meer toe.
Voor het vwo wordt een gemiddelde IQ waarde gehanteerd van minimaal 116, voor Havo een waarde van 107, voor VMBO-tl/Mavo een minimum IQ van 100, voor VMBO-kl een waarde van 92 en het LWOO (LeerWeg Ondersteunend Onderwijs) vanaf de waarde 82.
We nemen aan dat alle kinderen het basisonderwijs afronden en dat hun IQ zich na het basisonderwijs gedraagt volgens de normale verdeling van het IQ.

2a [4pt]. Wat is de kans dat een kind een advies VWO krijgt? Neem aan dat de IQ grenzen van hierboven strikt worden aangehouden.
De kans dat iemand een IQ heeft van minstens 116 is $\mathrm{normalcdf}(116\,,10^{10}\,,100\,,15)=$ $\mathbf{0,1431}$. (-1pt per foute grens)
13,6 + 2,3 % = 0,159 uit de grafiek klopt niet want is vanaf IQ 110, maar: 2pt

2b [5pt]. Laat zien dat de kans dat iemand die in het vervolgonderwijs (vanaf LWOO) terecht komt een VWO advies krijgt gelijk is aan 0,1617.
De kans dat iemand in het vervolgonderwijs terecht komt is gelijk aan de kans dat die persoon een IQ het van minstens 82: $\mathrm{normalcdf}(82\,,10^{10}\,,100\,,15)=0{,}8849$. De (conditionele) kans dat iemand VWO advies krijgt terwijl je weet dat hij in het vervolgonderwijs komt is dus

$$0{,}1431/0{,}8849=0{,}1617$$

2c [6pt]. Uit gegevens van het CBS blijk dat in 2020 van de 190.971 leerlingen in het middelbaar onderwijs er 42.302 op het VWO zaten. Veronderstel dat deze verdeling tot stand zou zijn gekomen puur op basis van een IQ grens, welke IQ ondergrens zou er dan gehanteerd zijn voor het VWO advies als we aannemen dat de minimale grens voor het middelbaar onderwijs op 82 blijft liggen? Geef je antwoord in één decimaal nauwkeurig.
De kans dat een leerling op het VWO zit is 42302/190971 = 0,2215. **2pt**
Dit is de conditionele kans dat iemand op het VWO zit, terwijl hij een leerling, ofwel

$$\frac{\mathrm{normalcdf}(IQ?\,,10^{10},100{,}15)}{\mathrm{normalcdf}(82,10^{10},100{,}15)}=0{,}2215$$

Ofwel:

$\mathrm{Normalcdf}(IQ?\,,10^{10},100{,}15)=0{,}2215*\mathrm{Normalcdf}(82,10^{10},100{,}15)=$ $0{,}2215*0{,}8413=0{,}1864$ **2pt**
Bij deze rechteroverschrijdingskans hoort een linkeroverschrijdingskans van 1 – 0,1864 = 0,8136, hierbij hoort een IQ waarde van $\mathrm{invNorm}(0.8136\,,100\,,15)=113{,}4$. **2pt**
Uitgerekend zonder de conditie dat IQ ≥ 82: $\mathbf{invNorm(1-0,2215\,,100\,,15)=111{,}5}$.
Max 4pt

2d [1pt]. Hoe hoog is je IQ, of hoe hoog schat je het in? Serieus antwoord graag!

2e [5pt]. Met een IQ van minstens 130 wordt iemand hoogbegaafd genoemd. Acht je het waarschijnlijk dat er onder 80 eerstejaars MBW/KW studenten (60 cadetten en 20 adelborsten) wel een hoogbegaafde adelborst is, maar geen hoogbegaafde cadet? Ondersteun je mening met een berekening.

We nemen aan dat cadetten en adelborsten niveau VWO hebben. De kans dat een VWO-er hoogbegaafd is is (conditionele kans) **2pt**

$$\frac{\text{Normalcdf}(130, 10^{10}, 100{,}15)}{\text{Normalcdf}(116, 10^{10}, 100{,}15)} = \frac{0{,}0228}{0{,}1431} = 0{,}1593$$

Onder 80 studenten zijn er gemiddeld 80*0,1593 = 12,7 hoogbegaafde studenten.

De kans dat er onder 20 adelborsten (minstens) een hoogbegaafde is is: **1pt**

$$1 - \text{binomcdf}(n = 20, p = 0.1593, k = 0) = 0{,}9689$$

De kans dat er onder 60 cadetten **geen** hoogbegaafde is is: **1pt**

$$\text{binomcdf}(n = 60, p = 0.1593, k = 0) = 0{,}00003009$$

De gevraagde kans is dus $0{,}9689 * 0{,}00003009 = 0{,}000029 = 0$ (afgerond) **1pt**

VWO niet meegenomen: max 3pt

Gerekend met 2,3%: max 2pt.

2f [4pt]. Met een IQ onder de 70 is een persoon geestelijk gehandicapt. Leg zonder een berekening te maken uit waarom het aantal geestelijk gehandicapten naar verwachting gelijk is aan het aantal hoog begaafden.

De geestelijk gehandicapten liggen qua IQ twee standaarddeviaties (2*15 = 30) of meer onder het gemiddelde, terwijl hoogbegaafden met IQ twee standaarddeviaties of meer boven het gemiddelde scoren. De normale verdeling is symmetrisch, dus dit zijn er ongeveer evenveel.

Alleen op grond van de tweemaal 2,3% in de grafiek: 2pt.

Opgave 3 [25pt]

Begin 2020 had Nederland 8,7 miljoen auto's op 8 miljoen huishoudens. We nemen aan dat het aantal auto's in een gezin verdeeld is volgens een Poissonverdeling.

3a [4pt]. Toon aan dat de gemiddelde waarde en de standaarddeviatie van het aantal auto's in een huishouden gelijk zijn aan respectievelijk 1,0875 en 1,0428 auto's per huishouden.

Gemiddeld zijn er per huishouden $\mu = \frac{8{,}7 \cdot 10^6}{8 \cdot 10^6} = 1{,}0875$ auto's. **3pt**

De standaarddeviatie in de Poissonverdeling is $\sigma = \sqrt{\mu} = 1{,}0428$ auto's per huishouden. **1pt**

3b [6pt]. De Poissonverdeling is discreet, maar wel oneindig, d.w.z., er is een kans dat een huishouden 100 auto's bezit of zelfs 1000, maar die kans is wel astronomisch klein. Bereken het kleinste aantal auto's waarvoor geldt dat de kans dat een huishouden zoveel auto's bezit kleiner is dan 0,0001.

Er geldt (let op: pdf!):

$$\text{poissonpdf}(1{,}0875\,, 7) = 0{,}000120 > 0{,}0001$$
$$\text{poissonpdf}(1{,}0875\,, 8) = 0{,}000016 < 0{,}0001$$

Het kleinste aantal auto's is dus 8.

Berekening met cdf: max 4pt

3c [8pt]. Bereken de kans dat 100 huishoudens bij elkaar meer dan 110 auto's hebben. Maak hierbij gebruik van de centrale limietstelling. Hint: Denk aan de continuïteitscorrectie.

De Poissonverdeling is een discrete verdeling met $\mu = 1{,}0875$ en $\sigma = 1{,}0428$. Volgens de centrale limietstelling is het aantal auto's in 100 huishoudens dan verdeeld als een normale verdeling met $\mu = 1{,}0875 * 100 = 108{,}75$ **3pt**

en $\sigma = 1{,}0428 * \sqrt{100} = 10{,}428.$ (is ook $\sqrt{\mu}$) **2pt**

De kans dat er minstens 111 auto's zijn is dus (met continuïteitscorrectie):

$$\mathrm{normalcdf}(111 - 0.5\,,\ 10^{10}\,, 108{,}75\,, 10.428) = 0{,}4334$$

2+1(continuïteitscorrectie)pt

3d [7pt]. Bereken de kans dat 100 huishoudens bij elkaar meer dan 110 auto's hebben. Maak hierbij gebruik van de Poissonverdeling.

Het aantal auto's van 100 huishoudens gedraagt zich als een Poissonverdeling met $\mu = 108{,}75.$ **2pt**

De gevraagde kans is dus

$$P(\underline{k} > 110) = 1 - P(\underline{k} \leq 110) = 1 - \mathrm{poissoncdf}(108.75\,, 110) = 0{,}4273$$

5pt

Opgave 4 [25pt]

In een militair depot in oorlogstijd worden materialen aangevoerd door middel van 12-tonners in 20ft zeecontainers. Trucks keren terug met een lege container en materiaal uit de volle containers wordt gesorteerd en opgeslagen in het depot. Tegelijkertijd arriveren er lege viertonners die lading komen halen om naar locaties in de buurt van operaties te vervoeren. Voor het gemak nemen we aan dat 12-tonners 12 ton lading aanvoeren en dat viertonners 4 ton lading afvoeren.

De 12-tonners arriveren met een gemiddelde van μ_{12} voertuigen per uur, van de viertonners arriveren er gemiddeld μ_4 per uur. De aankomsten van viertonners en van 12-tonners kunnen elke 24 uur met een Poissonverdeling worden beschreven. Per 24 uur zijn deze waarden vast, maar elke dag kunnen de waarden veranderen.

4a [7pt]. Op een bepaalde dag (24 uur) arriveren er zeven 12-tonners en er arriveren gemiddeld $\mu_4 = 0{,}9$ viertonners per uur. Wat is de kans dat de geleverde lading van deze 24 uur door de arriverende viertonners kan worden afgevoerd? Houd hierbij geen rekening met transactietijd in het depot (uitladen, opslaan, orderpicken, inladen).

In deze 24 uur arriveert er 7×12 = 84 ton lading, hiervoor zijn 7×12/4 = 21 viertonners nodig. **2pt**

De viertonners arriveren met een gemiddeld aantal van 0,9×24 = 21,6 viertonners per 24 uur. **2pt**

De kans dat er minstens 21 viertonners arriveren is **3pt**

$$P(\underline{k} \geq 21) = 1 - P(\underline{k} \leq 20) = 1 - \mathrm{poissoncdf}(21.6\,, 20) = 0{,}5802$$

4b [5pt]. Er is net een viertonner binnengekomen. Bereken de kans dat de volgende viertonner binnen een uur binnenkomt. **5pt, onjuiste μ: -1pt**

Deze kans is $F(1) = 1 - e^{-0{,}9 \cdot 1} = 0{,}5934.$

Ofwel, voor één uur geldt: $P(\underline{k} \geq 1) = 1 - \text{Poissoncdf}(0.9\,, 0) = 0{,}5934$

Niet helemaal juiste aanpak is: $\mathbf{poissonpdf(0.9\,,1) = 0,365}$ **max 3pt**

4c [8pt]. De commandant van het depot verwacht voor de volgende 24 uur gemiddeld $\mu_{12} = 0{,}35$ 12-tonners per uur. Hij wil met 99% zekerheid weten wat het maximale aantal 12-tonners is dat die dag kan arriveren. Bereken dat aantal.

Hij wil de grootste k weten zodanig dat $P(\underline{k} \leq k) \geq 0{,}99$. Dat betekent dat

$$\text{Poissoncdf}(0.35 * 24\,, k) \geq 0{,}99$$

De grootst mogelijke k is dan 16, want

$$\text{Poissoncdf}(0.35 * 24\,, 15) = 0{,}9875$$
$$\text{Poissoncdf}(0.35 * 24\,, 16) = 0{,}9941$$

4d [5pt]. De commandant wil vervolgens weten wat de minimale μ_4 voor die 24 uur moet zijn zodat hij met 99% zekerheid de aangevoerde lading ook weer kan afvoeren in het geval van de situatie in 4c. Bereken deze μ_4. Als je niet zeker bent van je antwoord in 4c, ga dan uit van 17 12-tonners.

Voor het afvoeren van de lading van 16 12 tonners zijn 48 viertonners nodig. De kleinste μ_4 waarvoor de kans minstens 99% is d3at er minstens 48 viertonner zullen zijn voldoet aan

$$1 - \text{Poissoncdf}(24\mu_4\,, 47) = 0{,}99$$

De solver levert:

$$\mu_4 = 2{,}7321$$

4e. [Bonusvraag 10 extra punten] De commandant laat zijn berekeningen aan zijn PLV zien en zegt dat hij nu met de in 4d berekende waarde 99% zekerheid heeft dat de aangeleverde lading kan worden afgevoerd. Na wat nadenken zegt de PLV dat hij het daar niet mee eens is. Volgens hem kan het in 4c mis gaan als er teveel 12-tonners komen, daar is 1 % kans op, maar het kan ook in 4d mis gaan als er te weinig viertonners komen, daar is ook 1 % kans op. Er is dus in totaal 2% kans dat het mis gaat, dus de betrouwbaarheid is volgens hem maar 100 – 2 = 98%. Wie heeft er gelijk? Wat is uiteindelijk de betrouwbaarheid als de waarde uit 4d wordt gehanteerd? Hint: Dit kost je minstens een kwartier. Bekijk alle mogelijkheden voor elk inkomend aantal 12-tonners en reken in minstens 5 cijfers nauwkeurig.

De betrouwbaarheid hangt af van hoeveel 12-tonners er arriveren:

Bij 16 12-tonners (kans $\text{Poissonpdf}(0.35 * 24\,, 16) = 0{,}0066$) moeten er minstens 48 viertonners zijn (kans $1 - \text{Poissoncdf}(24 * 2{,}7321\,, 47) = 0{,}9900$

Alle mogelijkheden staan in de volgende tabel:

Aantal 12-ton	Kans	Minim 4-ton	Kans	Kans totaal (product)
0	0,000225	0	1	0,000225
1	0,001889	3	1	0,001889
2	0,007933	6	1	0,007933
3	0,022213	9	1	0,022213
4	0,046648	12	1	0,046648
5	0,078369	15	1	0,078369
6	0,109716	18	1	0,109716
7	0,131659	21	1	0,131659
8	0,138242	24	1	0,138242

9	0,129026	27	1	0,129026
10	0,108382	30	1	0,108382
11	0,082764	33	0,999997	0,082764
12	0,057935	36	0,999974	0,057933
13	0,037435	39	0,999842	0,037429
14	0,022461	42	0,999231	0,022444
15	0,012578	45	0,99695	0,01254
16	0,006604	48	0,989999	0,006537
17	0,003263	51	0,972491	0,003173
18	0,001523	54	0,935612	0,001425
19	0,000673	57	0,870012	0,000586
20	0,000283	60	0,770575	0,000218
21	0,000113	63	0,641103	7,25E-05
22	4,32E-05	66	0,495242	2,14E-05
21	1,58E-05	69	0,352117	5,55E-06
Totaal				0,99945

De betrouwbaarheid is 0,99945, dus eigenlijk hebben ze geen van beiden gelijk.