**Tentamen Statistiek KW/MBW (deel 1, finale herkansing)**

Afdeling: Propedeuse KW/MBW 2020-2021

Examinator: Dr. J.B.M. Melissen

Datum: 29 april 2022, duur tentamen: 2 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. **Lever de antwoorden in op het geprinte antwoordformulier (zet je naam erop), de berekeningen en uitleg op gelinieerd papier.**

7. **De opgaven dienen na afloop van het tentamen ingeleverd te worden.**

Dit tentamen bestaat uit vier opgaven (30, 25, 25, 20 punten). Score = Puntentotaal/10

**Opgave 1 (Totaal 30 punten)**

Om de overlast als gevolg van een militaire vliegbasis in kaart te brengen is onder 7328 burgers in de omgeving van de vliegbasis een enquête gehouden. De geënquêteerden werd gevraagd naar de overlast die ze gedurende het afgelopen jaar hadden ervaren door het vliegverkeer op de basis. Antwoorden konden worden gegeven op een schaal van 0 – 4 (0 = geen overlast, 1 = nauwelijks overlast, 2 = geregeld overlast, 3 = behoorlijk veel overlast, 4 = zeer veel overlast. De respons op deze vraag wordt beschreven door een kansvariabele . Verder is ook de afstand van de woning van de ondervraagden in kaart gebracht. Deze aantal wordt beschreven door een kansvariabele met : afstand minder dan 3 km, : afstand tussen 3 en 6 km, en : afstand groter dan 6 km. De gezamenlijke kansverdeling is hieronder gegeven:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  | 0,0000 | 0,0120 | 0,0820 | 0,0590 | 0,0990 | 0,2520 |
|  | 0,0480 | 0,1340 | 0,0910 | 0,1230 | 0,0450 | 0,4410 |
|  | 0,2010 | 0,0590 | 0,0240 | 0,0110 | 0,0120 | 0,3070 |
|  | 0,2490 | 0,2050 | 0,1970 | 0,1930 | 0,1560 | 1,0000 |

**1a [4pt].** Bereken de waarden van de kansverdelingen en die in de tabel ontbreken.

**1b [8pt].** Bereken de verwachtingswaarde, de variantie en de standaarddeviatie van . Geef tussenresultaten van de berekeningen en maak geen gebruik van het statistisch menu van de grafische rekenmachine (hooguit ter controle).

.  **2pt**

**2pt**

**2pt**

. **2pt**

**1c [4pt].** Hoeveel geënquêteerden ervaren behoorlijk veel of zeer veel overlast van de vliegbasis?

**2pt**

Het bijbehorend aantal geënquêteerden is dus (afgerond) **2pt**

**1d [4pt].** Bereken de kans dat er onder 100 willekeurige ondervraagden hoogstens 40 personen zijn die behoorlijk veel of zeer veel overlast van de vliegbasis ervoeren (Gebruik geen benadering).

Kans op behoorlijk veel of zeer veel overlast is

De kans dat er onder 100 ondervraagden hoogstens 40 personen zijn die behoorlijk veel of zeer veel overlast van de vliegbasis ervoeren is . **4pt**

**1e [10pt].** Leg uit wat de volgende vijf kansen betekenen in termen van ervaren overlast en afstand tot de vliegbasis en bereken ze

**1pt**  
Dit is de kans dat een geënquêteerde niet of nauwelijks overlast heeft. **1pt**

**1pt**  
Dit is de kans dat een geënquêteerde binnen 6 km afstand woont. **1pt**

(zie tabel). **1pt**

Dit is de kans dat een geënquêteerde niet of nauwelijks overlast heeft en ook binnen 6 km afstand woont.  **1pt**

**1pt**  
Dit is de kans dat een geënquêteerde die binnen 6 km afstand woont niet of nauwelijks overlast heeft.  **1pt**

**1pt**

Dit is de kans dat een geënquêteerde die niet of nauwelijks overlast heeft binnen 6 km afstand woont.  **1pt**

**Opgave 2 (Totaal 25 punten)** De reistijd van de sergeant Buldermans van woonplaats naar de kazerne (enkele reis) kan worden beschouwd als een normaal verdeelde kansvariabele met = 25 minuten en = 4 minuten.

**2a. [5pt]** Bereken de kans dat hij op een willekeurige ochtend minder dan 20 minuten reistijd nodig heeft om op de kazerne te komen.

De reistijd is normaal verdeeld:

**2b. [5pt]** Bereken hiermee de kans dat hij in een maand (20 werkdagen) minstens één keer binnen 19 minuten reistijd naar zijn kazerne nodig heeft.

Elke ochtend geldt dat de reistijd naar de kazerne minder dan 19 minuten is met (succes)kans

Volgens de binomiale verdeling is de kans dat dit minstens één keer per 20 ochtenden gebeurt:

**2c. [5pt]**  Hoe groot is de kans dat gedurende een week (5 dagen) de gemiddelde **reistijd naar de kazerne** minder dan 20 minuten is? Neem aan dat de reistijden onafhankelijk zijn en dat de verdeling elke dag hetzelfde is.

De gemiddelde reistijd over vijf dagen is normaal verdeeld met = 25 minuten en = 4/ minuten (zie formuleblad). De kans dat het gemiddelde kleiner is dan 20 minuten is

**2d. [5pt]** Het vertrektijdstip van de sergeant is normaal verdeeld met gemiddeld 6:28 en minuten. Hoe groot is de kans dat hij na 7:00 aankomt?

Aankomst = vertrek + reistijd is normaal verdeeld. De kans dat hij later dan 7:00 aankomt is

De laatste regel is gerekend vanaf 6:00 in minuten (dat is willekeurig), dus 7:00 is dan 60 minuten en 6:51 is 51 min. Dit is gedaan omdat je met de tijden niet kunt rekenen.

**2e. [5pt]** Hoeveel minuten moet de sergeant ’s morgens gemiddeld eerder vertrekken om om 7:00 binnen te zijn met een kans die hoogstens 0,02 is?

Aankomst = vertrek + reistijd is normaal verdeeld. De kans dat hij te laat is, is

Oplossen met solver:

Of: , dus hij moet 2,3263\*5 = 10,2687 minuten te vroeg zijn. Hij is normaal al gemiddeld 7 minuten te vroeg, dus hij moet 10,2687 – 7 = 3,2687 minuten eerder vertrekken.

**Opgave 3 (Totaal 25 punten)** In het kader van verscherpte grensbewaking in COVID-19 tijden voert de Koninklijke Marechaussee op lokale wegen in de grensregio’s 100%-controles uit. Voor een bepaalde dag wordt een inzet gepland waarbij gedurende drie uur alle voertuigen op een weg in Limburg gecontroleerd moeten worden. Gebaseerd op gegevens van de RDW is de verwachting dat zich gedurende de controle gemiddeld 26,5 voertuigen per uur aandienen en er wordt aangenomen dat dit volgens een Poissonverdeling zal gebeuren. Uit eerdere inzetten is gebleken dat de benodigde tijd per voertuig uniform is verdeeld tussen 3 en 15 minuten. Elke controle wordt uitgevoerd door een team van twee marechaussees.

**3a. [4pt]** Bereken hoeveel teams er gemiddeld nodig zijn om deze controles uit te voeren.

3a. Gemiddeld 26,5 voertuigen per uur, gemiddelde controletijd is 9 minuten per voertuig, dus er is gemiddeld 238,5 minuten controletijd per uur nodig. 238,5/60 = 3,975 teams nodig. Er kunnen alleen gehele aantallen teams aanwezig zijn, dus afronden naar boven: 4 (onder de aanname dat de teams continu bezet zijn en dat er zich geen wachtrijen vormen door een te lage capaciteit).

**3b. [6pt]** Bereken de kans dat zich gedurende de controletijd van 3 uur meer dan 100 voertuigen aandienen.

3b. Gedurende 3 uur is het verwachte aantal voertuigen 26,5\*3 = 79,5. Volgens de Poissonverdeling (kies als periode 3 uur) is

**3c. [6pt]** Neem aan dat elk team direct van start kan gaan en vervolgens continu bezig is met controles. Bereken hoeveel tijd een team nodig heeft om met 97% zekerheid 20 controles te kunnen uitvoeren. Maak hiervoor gebruik van een geschikte benadering op grond van de centrale limietstelling en de parameters van de uniforme verdeling.

3c. De tijd die nodig is wordt bepaald door de som van 20 tijden die elk getrokken worden uit de uniforme verdeling tussen 3 en 15 minuten. Daarvan is de gemiddelde waarde (3+15)/2 = 9 minuten en de standaarddeviatie is minuten per controle (zie formuleblad).

Volgens de centrale limietstelling geldt voor de totale tijd (som) van 20 controles een normale verdeling met minuten en minuten (formuleblad).

De tijd waarin de controles met 97% zekerheid kunnen worden gedaan is de waarvoor

Dit kun je met de solver uitrekenen of met-waarde , waarin . De gevraagde tijd is 180 + 1,8808 × 15,4919 = 209,1372 minuten = 3,3856 uur.

**3d. [4pt]** Uit de voorgaande berekeningen volgt dat een team met 97% zekerheid 20 voertuigen kan controleren binnen 3,5 uur, en dat met 98,9% zekerheid niet meer dan 100 voertuigen zijn te controleren.

Bereken met behulp van deze gegevens de kans dat vijf teams voldoende zijn om binnen 3,5 uur de controles uit te voeren, dus (voor het gemak): de kans dat er in 3 uur niet meer dan 100 voertuigen komen én dat elk van de teams daarvan 20 auto’s binnen 3,5 uur kan controleren. Dit is een worst-case inschatting, in werkelijkheid is de kans groter.

3d. 0,989 \* 0,975 = 0,8493.

Dit is worst-case want er is vanuit gegaan dat elk team 20 auto’s afhandelt. In de praktijk kan het ook voorkomen dat bijvoorbeeld één team er 19 doet en een ander team 21, waardoor het totaal toch wordt gehaald.

**3e. [5pt]** Hoe groot is de kans dat het, na het aanhouden van een voertuig, minimaal 6 minuten duurt voordat het volgende voertuig arriveert? Maak gebruik van de negatief exponentiële verdeling.

3e. De tijd tussen twee voertuigen wordt beschreven door exp(), waarbij de tijd in minuten is en het aantal verwachte voertuigen per minuut, dus . De kansfunctie is

Of met de kansdichtheidsfunctie

**Opgave 4 (totaal 20 punten)**

**9600 tests afgenomen sinds openen corona-nummer GGD**

**Via het landelijke nummer voor een coronatest, waren kort na het uitbreken van corona al bijna 16.000 afspraken gemaakt. Intussen werden bijna 9600 tests afgenomen en 94 mensen positief bevonden, maar nog niet alle uitslagen waren bekend. (Bron: RTL Nieuws 3 juni 2020)**

**4a. [2pt]** Bereken op grond van deze gegevens een schatting voor de fractie Nederlanders die op dat moment besmet waren met het coronavirus.

4a. 94/9600 = 0,009792.

**4b. [2pt]** Bereken op grond van deze fractie hoeveel van de 17,3 miljoen Nederlanders naar schatting besmet waren met het Coronavirus.

4b. 94/9600 \* 17300000 = 169.399.

**4c. [5pt]** Is de waarde uit 4b. waarschijnlijk een ondergrens, een bovengrens of is dat niet duidelijk? Geef hiervoor argumenten.

4c. Niet duidelijk. Nog niet alle uitslagen zijn bekend, dus het percentage kan in werkelijkheid hoger zijn. Het gaat om mensen die zelf om een toets vragen, dus ook mensen met klachten, dus het percentage in deze groep kan juist hoger zijn dan in de rest van Nederland, maar als er veel gezonde mensen bij zijn, die alleen maar duidelijkheid willen kan het juist weer lager zijn dan in de rest van Nederland. De ernstige gevallen zitten hier niet bij, die liggen al in het ziekenhuis. Dan gaat het ook nog om een steekproef, waardoor het werkelijke percentage naar boven of naar beneden kan afwijken.

# Is echt maar 5,5 procent van de Nederlanders besmet geweest met corona?

**Zo'n 5,5 procent van de Nederlandse bloeddonoren heeft antistoffen aangemaakt tegen het coronavirus. Dat werd vanochtend duidelijk. (Bron: RTL Nieuws 3 juni 2020)**

**4d. [5pt]** Bereken op grond van deze 5,5% hoeveel van 90 cadetten en adelborsten besmet zijn geweest met het coronavirus. Bereken ook de standaarddeviatie in deze waarde.

4d. 0,055\*90 = 4,95.

Standaarddeviatie (binomiale verdeling, zie formuleblad) =

**4e. [6pt]** Bereken de kans dat van de 90 studenten er 3, 4, 5, 6 of 7 besmet zijn geweest.

4d.

**=== EINDE TENTAMEN ===**