

Naam:

19 januari 2017

Studentennummer:

Statistiek en gegevensverwerking

Mathieu Labare – Examen eerste zit 2016-2017

- 8u30 – 13u
- Voor de **theorie** is er enkel schrijfgerief toegestaan.
- Leg alles grondig uit. Je mag eventueel een tekening maken om de situatie beter uit te leggen. Niet-verklaarde of –relevante dingen leveren minpunten op.
- Voor de **oefeningen** zijn de cursus, de slides, grafische rekenmachine ... toegelaten. Je kan enkel je cursus erbij nemen als het theoriedeel is afgegeven.
- Denk aan je tijdsindeling!
- Veel succes

1. THEORIE

- I. Gegeven is de waarschijnlijkheidsverdeling $P(x; a_1, a_2)$. Er kan een schatter van a_1 en a_2 gemaakt worden.
 - a. Geef de 3 voornaamste schattingsmethodes en geef in het bijzonder voor elke methode \hat{a}_1 en \hat{a}_2 .
 - b. Bespreek de voor- en nadelen van elke methode. Wanneer zou je welke methode gebruiken?

- II. De conditionele kans $P(a|b)$ is de waarschijnlijkheid van a op voorwaarde dat theorie b waar is.
 - a. Pas de definitie van het Theorema van Bayes toe op $P(a|b)$ en geef een mening en interpretatie van elke term. Waarom is een Bayesiaanse aanpak onvermijdelijk in de realiteit, zelfs met de frequentistische definitie van waarschijnlijkheid.
 - b. $f_{XY}(x,y)$. Geef de conditionele en marginale waarschijnlijkheidsverdeling. Wat is hun verband, zowel voor het algemene geval als voor het geval van onafhankelijke variabelen)
 - c. Leg het verband uit tussen het begrip conventionele waarschijnlijkheidsverdeling en het opbouwen en de interpretatie van confidentieintervallen. Leg het praktische gebruik uit van de Bayesiaan voor afleiden van confidentieniveaus
 - d. Beschrijf de conditionele kans in het kader van beslissingstheorie. Geef de optimale test met twee eenvoudige hypothesen.

2. OEFENINGEN

- I. We willen het loodgehalte bepalen in een bodemtype. Een eerste methode leerde ons dat dit 86 deeltjes per miljoen zijn (ppm). We willen nu een nieuwe methode testen. Voor 20 meetresultaten krijgen we een gemiddelde waarde van 83 ppm, met een standaardafwijking van 10 ppm.
- Kunnen we met een significantie van 0,1 aantonen dat de nieuwe methode gemiddeld minder extraheert dan de eerste methode? Vermeld je eventuele veronderstelling(en) die je maakt.
 - Welk type fout kan je hier mogelijks maken? Leg uit waarom.
 - Een kritische onderzoeker zegt dat het verschil afhangt van variaties in de bodem. Leg uit hoe we het experiment kunnen aanpassen om dit probleem op te lossen.

- II. Er was eens een beroemde wiskundige die de groei van populaties bestudeerde. Hij noemde zijn verdeling de logistische verdeling. Deze werd gegeven als volgt:

$$f(x; \mu, s) = C \frac{e^{-z}}{(1 + e^{-z})^2}, \text{ met } z = \frac{x - \mu}{s}$$

- Bepaal de constante C zodat f werkelijk een kansdichtheid is.
 - Bereken de cumulatieve verdeling van deze kansdichtheid en leid op een simpele manier af om een reeks van toevallige getallen te genereren die f(x) volgen.
 - Gegeven zijn N metingen X_1, \dots, X_N . Deze metingen zijn verdeeld volgens de logistische verdeling met $\mu = 0$ en $s = 1$.
Geef de verdeling van $M_n = \text{Max}\{X_1, \dots, X_N\}$
 - Stel nu de nieuwe verdeling $Y_n = M_n + \ln(n)$. Wat is de cumulatieve verdeling van Y_n ? Bewijs dat Y_n convergeert naar de Gumbelverdeling $e^{-e^{-t}}$ met $n \rightarrow \infty$.
- III. Een student uit de bachelor wil de valversnelling $g = 9,81\text{m/s}^2$ berekenen. Hij heeft hiervoor een opstelling gemaakt waarin hij een knikker in een buis van 10 meter laat vallen naar beneden. Op het moment dat het luikje op 10m opengaat start hij de tijd. De tijd wordt ook gemeten op telkens exact 2 m. Bij zijn eerste poging verkreeg hij volgend resultaat:

Hoogte (h)	8,0	6,0	4,0	2,0	0,0
Tijd (t)	0,8	1,05	1,26	1,43	1,59

- Gegeven is dat de hoogte van t afhangt volgens $h(t) = h_0 - g \frac{t^2}{2}$. Komen de gegevens overeen met de verwachte waarde van g? Geef een volledige kwantitatieve rechtvaardiging.
- Stel dat de h_0 exact gekend is, bepaal dan g (+ de fout erop). Wat is je conclusie?
- Verklaar het vorige resultaat en stel een oplossing voor dit probleem.