

Afdeling Wiskunde	Volledig tentamen Algemene Statistiek voor BWI
	Deeltentamen II Algemene Statistiek voor BWI
Vrije Universiteit	27 mei 2010

Gebruik van een niet-programmeerbare rekenmachine is toegestaan.
 U krijgt de beschikking over een tabellenboekje.
 Na de correctie liggen de tentamens ter inzage bij het onderwijsbureau.
 De normering staat bij de vragen aangegeven.
 Cijfer voor volledig tentamen = $\frac{\text{totaal}+4}{4}$. (opgaven 1,2,3,4)
 Cijfer voor deeltentamen II = $\frac{\text{totaal}+3}{3}$. (opgaven 3,4,5)

SUCCES!

Opgave 1 [11 punten] Alleen voor gehele tentamen

Zij $X_1, \dots, X_n$ een steekproef uit de exponentiële verdeling met parameter $1/\theta$ met dichtheid $p_\theta(x) = (1/\theta)e^{-x/\theta}$ en $Y_1, \dots, Y_m$ een steekproef uit de exponentiële verdeling met parameter $2/\theta$. We nemen aan dat de twee steekproeven onafhankelijk zijn.

- [2 punten] Behoren de twee genoemde exponentiële verdelingen tot dezelfde lokatie-schaalfamilie? Beargumenteer uw antwoord met een korte berekening.
- [5 punten] Laat zien dat de momentenschatter voor θ gebaseerd op $Y_1, \dots, Y_m$ (op basis van het eerste moment) gelijk is aan de maximum likelihood-schatter voor θ gebaseerd op $Y_1, \dots, Y_m$.
- [4 punten] Zij $T_c = c\bar{X} + 2(1-c)\bar{Y}$ met $c \in [0, 1]$ een schatter voor θ . Voor welke waarde van c is deze schatter optimaal?

Opgave 2 [5 punten] Alleen voor gehele tentamen

Zij $X_1, \dots, X_n$ een steekproef uit de verdeling met kansdichtheid

$$p_\theta(x) = \frac{2\theta^2}{x^3} 1_{[\theta, \infty)}(x)$$

met $\theta > 0$ een onbekende parameter.

- [3 punten] Stel de likelihood-functie op en schets deze functie. Geef in de schets aan wat er op de x -as en y -as uit staat. Bepaal de maximum likelihood-schatter voor θ .
- [2 punten] Laat zien dat de momentenschatter voor θ in dit geval niet gelijk is aan de maximum likelihood-schatter voor θ .

Opgave 3 [12 punten]

Zij $X_1, \dots, X_n$ een steekproef uit de alternatieve verdeling met parameter $\theta \in [0, 1]$. De kansmassafunctie van de verdeling is te schrijven als

$$p_\theta(x) = \theta^x (1 - \theta)^{1-x}, \quad x \in \{0, 1\}.$$

De maximum likelihood-schatter voor θ is gelijk aan $\hat{\theta} = \bar{X}$.

- [3 punten] Bepaal de Fisher-informatie i_θ .

- b. [3 punten] Wat is de asymptotische verdeling van $\sqrt{n}(\hat{\theta} - \theta)$? Bepaal een benaderend $(1 - \alpha)$ -betrouwbaarheidsinterval voor θ op basis van $\hat{\theta}$ en de plug-in schatter voor i_θ .
- c. [3 punten] Bepaal de likelihood-ratiostatistiek $\lambda(X)$ gebaseerd op $X = (X_1, \dots, X_n)$ voor het toetsen van $H_0 : \theta = \theta_0$.
- d. [3 punten] Bepaal een benaderend $(1 - \alpha)$ -betrouwbaarheidsinterval voor θ op basis van de likelihood-ratiostatistiek.

Opgave 4 [8 punten]

Veronderstel dat $X_1, \dots, X_n$ een steekproef is uit de Poisson-verdeling met parameter θ en definieer $X = \sum_{i=1}^n X_i$. Dan volgt X de Poisson-verdeling met parameter $n\theta$. We willen de nulhypothese $H_0 : \theta = \theta_0$ toetsen.

- a. [2 punten] Bepaal een kritiek gebied K_X voor de toetsingsgrootheid X om H_0 te toetsen bij onbetrouwbaarheidsdrempel α . Er zijn meerdere antwoorden mogelijk. (Bedenk zelf een notatie voor de kwantielen van een Poisson-verdeling.)

Net als bij de binomiale toets, kan ook hier een normale benadering worden toegepast mits n groot genoeg is. De continuïteitscorrectie wordt hier achterwege gelaten.

- b. [1 punt] Welke normale verdeling benadert de verdeling van X onder de nulhypothese voor grote waarden van n ?
- c. [3 punten] Bepaal een kritiek gebied K_T voor een geschikte toetsingsgrootheid T om H_0 te toetsen bij onbetrouwbaarheidsdrempel α met behulp van de normale benadering. Baseer de toetsingsgrootheid T op X .

Met behulp van de relatie tussen toetsen en betrouwbaarheidsgebieden is het mogelijk om een benaderend $(1 - \alpha)$ -betrouwbaarheidsgebied voor θ op te stellen op basis van de toets in onderdeel c die bij benadering van niveau α is.

- d. [2 punten] Bepaal een benaderend $(1 - \alpha)$ -betrouwbaarheidsgebied voor θ met behulp van de toets uit onderdeel c.

Opgave 5 [7 punten] Alleen voor deeltentamen II

Om de juli-temperaturen in De Bilt en Eelde te vergelijken heeft het KNMI de volgende gemiddelde juli-temperaturen verzameld:

jaar	1965	1966	1967	1968	1969	1970	1971	1972	1973	gem	var
De Bilt	13.7	14.2	15.3	14.1	18.1	15.7	17.2	17.2	17.1	15.84	2.625
Eelde	13.4	14.1	14.9	14.1	17.5	15.1	17.0	17.1	16.7	15.54	2.385
verschil	0.3	0.1	0.4	0.0	0.6	0.6	0.2	0.1	0.4	0.3	0.0475

De onderste rij geeft de verschillen weer. De waarde in de kolom *gem* is het gemiddelde van de rij, de waarde in de kolom *var* is de steekproefvariantie (S_X^2) van de rij. We nemen aan dat de waarnemingen over de verschillende jaren onafhankelijk zijn.

- a. [4 punten] Toets de nulhypothese dat er geen verschil in verwachte temperatuur ($\mu - \nu$) bestaat bij een onbetrouwbaarheidsdrempel van 5%. Vermeld een geschikt statistisch model, de nulhypothese, de toetsingsgrootheid, de verdeling van de toetsingsgrootheid onder de nulhypothese, de overschrijdingskans of het kritieke gebied en de conclusie. Vermeld duidelijk welke aannames u maakt.
- b. [3 punten] Stel een numeriek 95%-betrouwbaarheidsinterval op voor $\mu - \nu$ op basis van een geschikte pivot.