

TECHNISCHE UNIVERSITEIT EINDHOVEN
Faculteit Wiskunde en Informatica

Tentamen Mathematische Statistiek (2S990) op 27 november 2003, 9.00-12.00 uur.

U mag alleen gebruik maken van een onbeschreven Statistisch Compendium (dikt. nr. 2218) en van een zakrekenmachine. De uitwerkingen van de opgaven dienen gemotiveerd, duidelijk geformuleerd en overzichtelijk opgeschreven te worden.

Puntenwaardering

1	2	3	4	5	6	7	8	9
a	b		a	b	a	b	c	
2	2	2	3	3	2	3	2	2

(Totaal 30 punten.)

- Voor het testen van fouten in software wordt het volgende model gebruikt. Er worden onafhankelijke testen uitgevoerd die aangeven of er nog een fout in de software is. De kans om een fout in de software te vinden is $1 - \varphi^{n-i}$, waarbij n het oorspronkelijke aantal fouten in het systeem is en i het aantal reeds gevonden aantal fouten. Gevonden fouten worden op een perfecte manier verwijderd. De verdeling van het aantal tests T_i om de volgende fout te vinden na het vinden van de $(i - 1)^e$ fout is dus geometrisch verdeeld met succesparameter $1 - \varphi^{n-i+1}$. Om meer inzicht te krijgen in deze procedure, test men een stuk software waarvan n , het aantal fouten bekend is en φ onbekend.
 - Laat zien dat $\sum_{i=1}^n (n - i + 1) T_i$ een voldoende steekproefgrootte is voor φ .
 - Bepaal de Maximum-Likelihoodschatter voor φ als $n = 2$.
- Zij X_1, \dots, X_n een steekproef uit een exponentiële verdeling met parameter 1. Bepaal de asymptotische verdeling van $\sum_{i=1}^n X_i$.
- Zij $\hat{\Theta}$ een schatter voor een parameter θ gebaseerd op een steekproef X_1, \dots, X_n . Laat zien dat $\hat{\Theta}^2$ geen zuivere schatter van θ^2 kan zijn indien $\hat{\Theta}$ zuiver is en $0 < \text{Var}(\hat{\Theta}) < \infty$.
- Zij $t > 0$ een vast gekozen getal. We nemen storingsen van een machine waar tot en met tijdstip t . De tijden tussen 2 storingsen zijn onafhankelijk en exponentieel verdeeld met parameter λ , d.w.z. $f(x | \lambda) = \lambda e^{-\lambda x}$ voor $x \geq 0$. Zij K het stochastische aantal storingsen tot en met tijdstip t .
 - Laat zien dat K/t de Maximum-Likelihoodschatter is voor λ en dat dit een zuivere schatter is voor λ .
 - Beargumenteer of er een zuivere schatter is voor λ met een kleinere variantie dan K/t .

5. Construeer een 95%-betrouwbaarheidsinterval voor de parameter θ van een exponentiële verdeling, waarbij we hier de parametrizing $f(x) = \frac{1}{\theta} e^{-x/\theta}$ kiezen.
6. We willen het gewicht van vrouwelijke en mannelijke studenten vergelijken. Deze gewichten zijn normaal verdeeld met onbekende verwachting μ_1 respectievelijk μ_2 kg. Bekend is dat de standaardafwijking in gewicht zowel voor vrouwelijke als voor mannelijke studenten gelijk is aan 2,25 kg. We onderzoeken eerst het gewicht van vrouwelijke studenten en beschouwen het toetsingsprobleem

$$H_0 : \mu_1 = 68, H_1 : \mu_1 > 68$$

bij onbetrouwbaarheid $\alpha = 0,05$.

- (a) We willen dat voor $\mu_1 = 69$ kg het onderscheidingsvermogen $1 - \beta$ gelijk is aan 0,95. Hoeveel vrouwelijke studenten moeten dan gewogen worden?
 - (b) Geef bij het onder a) gevonden aantal gewogen vrouwelijke studenten een tweezijdig 95% betrouwbaarheidsinterval voor het verwachte gewicht van vrouwelijke studenten.
 - (c) We wegen nu 15 mannelijke en 10 vrouwelijke studenten. Voor de mannelijke studenten vinden we een gemiddeld gewicht van 79,2 kg en voor de vrouwelijke studenten een gemiddeld gewicht van 67,8 kg. Mogen we nu veronderstellen dat mannelijke studenten meer dan 10 kg zwaarder zijn dan vrouwelijke studenten ($\alpha = 0,05$)?
7. Zij X_1, \dots, X_n een steekproef uit een Poissonverdeling met parameter λ . Leid een Generalized-Likelihood-Ratio-toets af voor de toets $H_0 : \lambda = 1$ tegen $H_a : \lambda \neq 1$. Geef duidelijk aan wat de toetsingsgrootheid en een (benaderend) kritiek gebied zijn.
 8. Er zijn verschillende redenen waarom een computersysteem kan crashen. Een informaticus beweert dat 10% van de crashes aan software fouten ligt, 5% aan hardware fouten, 25% aan operatorfouten, 40% aan overbelasting van het systeem en de rest aan overige oorzaken. Tijdens een uitvoerig onderzoek werden 150 crashes onderzocht. Men vond 13 software fouten, 10 hardware fouten, 42 operatorfouten, 65 overbelastingen. Neem aan dat elke crash precies één oorzaak heeft. Is dit onderzoek reden om te twijfelen aan de beweringen van de informaticus?
 9. De Poissonverdeling wordt veel gebruikt om ruimtelijke effecten te beschrijven, zoals het aantal oneffenheden op glasplaat van een LCD-scherm. Een medewerker onderzoekt 32 LCD schermen en telt de volgende aantallen oneffenheden per LCD scherm: 0, 2, 3, 1, 3, 4, 2, 1, 2, 2, 1, 0, 3, 3, 3, 2, 4, 1, 2, 4, 2, 2, 1, 5, 1, 3, 2, 0, 3, 2, 2, 1. Toets of deze gegevens goed met een Poissonverdeling beschreven kunnen worden.