TECHNISCHE UNIVERSITEIT EINDHOVEN
Faculteit Wiskunde en Informatica

Tentamen Mathematische Statistiek (2WS05 en 2S990 (uitloop)) op 23 maart 2007, 14.00-17.00 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 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | a | b | c | d | e | f | a | b | c | d | a | b |
| 1 | 2 | 2 | 2 | 3 | 1 | 3 | 2 | 2 | 2 | 3 | 3 | 2 | 2 |

(Totaal 30 punten.)

---

**NB: Studenten kiezen óf voor de versie 2WS05 (nieuwe versie Mathematische Statistiek sinds studiejaar 2006-2007) óf voor de versie 2S990 (uitloop oude versie Mathematische Statistiek).**

1. Twee soorten motorolie worden getest op hun octaangehalte. Bekend is dat de variantie van de eerste soort octaangehalte gelijk is aan $1,5$ en de variantie van de tweede soort octaangehalte gelijk is aan $1,2$. Er worden uit beide soorten onafhankelijk aselecte steekproeven met een omvang 15 uit de eerste soort en 20 uit de tweede soort. Het gemiddeld gemeten octaangehalte voor de eerste soort is gelijk aan $89,6$ en voor de tweede soort $92,5$. We veronderstellen dat de waarnemingen normaal verdeeld zijn.

   (a) Bereken een 95%-betrouwbaarheidsinterval voor het verschil van het gemiddelde octaangehalte van de beide soorten motorolie.

   (b) Als de tweede soort motorolie een hoger octaangehalte heeft, wenst de fabrikant dat vast te stellen. Formuleer een hierbij passende nulhypothese en alternatieve hypothese. Voer de toets uit met $\alpha = 0,05$.

2. Zij $X_1, \dots, X_n$ een steekproef uit een exponentiële verdeling met dichtheid $\frac{1}{\theta}\, e^{-x/\theta}$ voor $x > 0$. Indien nodig mag U gebruiken dat de klasse van exponentiële verdelingen volledig is.

   (a) Laat zien dat $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ een voldoende steekproefgrootheid is voor $\theta$.

   (b) Bepaal de Maximum-Likelihoodschatter voor $\theta$.

   (c) Laat zien dat $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ een zuivere schatter met minimale variantie (UMVU schatter) is voor $\theta$.

   (d) Bepaal de Maximum-Likelihoodschatter voor $1/\theta$.

(e) Laat zien dat $(n-1)/(n\,\overline{X})$ een zuivere schatter met minimale variantie (UMVU schatter) is voor $1/\theta$.

(f) Is er op statistische gronden een reden om de exponentiële verdeling te parametriseren met $\theta$ zoals in de inleiding van deze opgave of met $\lambda = 1/\theta$. Betrek bij Uw antwoord de resultaten van de voorgaande onderdelen van deze opgave en een in deze opgave niet genoemd kwaliteitscriterium voor schatters.

3. Zij $X_1, \ldots, X_{10}$ een steekproef uit een Poissonverdeling met verwachting $\lambda$. We toetsen $H_0 : \lambda = 0,1$ tegen $H_1 : \lambda = 0,5$ en gebruiken als kritiek gebied $\left\{(x_1, \ldots, x_{10}) \,\middle|\, \sum_{i=1}^{10} x_i \geq 3\right\}$.

(a) Bepaal het significantieniveau van deze toets (de kans op een type I fout).

(b) Bepaal het onderscheidingsvermogen van deze toets als $\lambda = 0,5$.

(c) Laat zien dat deze toets het beste onderscheidingsvermogen (MP toets) heeft van alle toetsen met hetzelfde significantieniveau. N.B. Deze vraag kunt U ook oplossen als U a) niet heeft kunnen beantwoorden.

(d) Bepaal een toets met optimaal onderscheidingsvermogen voor $H_0 : \lambda = 0,1$ tegen $H_1 : \lambda > 0,1$.

4. (**alleen voor 2S990**) Zij $X_1, \ldots, X_n$ een steekproef uit een verdeling met dichtheid

$$f_\theta(x) = \frac{1}{\pi(1 + (x-\theta)^2)} \qquad -\infty < x < \infty, -\infty < \theta < \infty.$$

(a) Laat zien dat de variantie van zuivere schatters voor $\theta$ minimaal $2/n$ is.

(b) Wat is de asymptotische verdeling van de Maximum-Likelihoodschatter voor $\theta$?

4. (**alleen voor 2WS05**) Beschouw het volgende toetsingsprobleem: objecten worden geclassificeerd volgens $k$ kenmerken. Elk object krijgt precies één kenmerk bij de classificatie. Er wordt nu onderzocht of de kenmerken even vaak voorkomen in een aselecte steekproef van $n$ objecten. De toetsingsgrootheid voor dit probleem wordt gegeven door de volgende `R` code (N.B. `rep` genereert een rijtje van steeds dezelfde waarde):

```
statistic <- function(cellcounts){
k <- length(cellcounts)
n <- sum(cellcounts)
proportions <- rep(n/k,times=k)
sum((cellcounts-proportions)^2/proportions)
}
```

(a) Als $X_1, \ldots, X_k$ de aantallen objecten per kenmerk zijn, geef dan de toetsingsgrootheid in gewone formulevorm i.p.v. als `R` code.

(b) Neem aan dat de toetsingsgrootheid (bij benadering) verdeeld is als een $\chi^2$-stochast met $n-k-1$ vrijheidsgraden. Gebruik dit om het kritiek gebied van deze toets af te leiden. N.B. U hoeft alleen de vorm te beredeneren.