Tentamen toegepaste statistiek voor TA (WI1275TA)

18 april 2012, 14:00-17:00

Afdeling Toegepaste Wiskunde, Faculteit Elektrotechniek, Wiskunde en Informatica

---

Toelichting: Een antwoord alleen is *niet* voldoende: er dient een berekening, toelichting en/of motivatie aanwezig te zijn. Dit alles goed leesbaar en in goed Nederlands. Rekenmachine toegestaan. Formuleblad/boek/aantekeningen niet toegestaan.

---

1. De gebeurtenissen A en B zijn onafhankelijk en bovendien is bekend dat $\mathrm{P}(A \mid B) = 1/2$ en $\mathrm{P}(B \mid A) = 1/3$. Bereken $\mathrm{P}(A \cap B)$.

2. Het is bekend dat de kans op kleurenblindheid bij jongens 1 op 20 en bij meisjes 1 op 100 is. Juffrouw Irene heeft een klas met 10 jongens en 20 meisjes.

   (a) Geef voor een willekeurige leerling uit deze klas de kans op kleurenblindheid.

   (b) Geeft de voorwaardelijke kans dat een leerling uit deze klas een jongen is, gegeven dat de leerling kleurenblind is.

3. Stel dat de stochastische variabele U een uniforme verdeling heeft op het interval $[0, 1]$. Laat zien dat $X = 1 - U$ ook uniform verdeeld is op $[0, 1]$.

4. Tijdstippen van overstromingen in een bepaald gebied vinden plaats volgens een Poisson proces. Het is bekend dat het verwachte aantal overstromingen over 50 jaar gelijk is aan 2.

   (a) Bereken de kans op precies één overstroming in 50 jaar.

   (b) Bereken de kans op tenminste twee overstromingen in 25 jaar.

5. Stel dat $X \sim \mathrm{Binom}(2, 1/3)$ (dus X is Binomiaal verdeeld) en $Y \sim \mathrm{Ber}(1/2)$ (dus Y heeft een Bernoulli verdeling). Veronderstel dat X en Y onafhankelijk zijn. Definieer $Z = XY$. Bereken de kansmassafunctie van Z.

6. We gooien een onzuivere munt achtereenvolgens totdat we voor het eerst kop gooien. We herhalen dit experiment 3 keer (met dezelfde munt). De aldus verkregen data zijn:

   - eerste experiment: eerste keer kop bij de 3de worp,
   - tweede experiment: eerste keer kop bij de 3de worp,
   - derde experiment: eerste keer kop bij de 4de worp.

   (a) Noem p de kans op kop bij één keer gooien met de munt. Toon aan dat de loglikelihood wordt gegeven door

   $$\ell(p) = 7 \log(1 - p) + 3 \log p$$

   (b) Stel dat we van het derde experiment alleen weten dat de eerste keer kop *na* de 4de worp wordt gegooid. Geef de likelihoodfunctie voor dit geval.

7. Stel dat x een realisatie van de stochastische variabele X is en dat X een exponentiële verdeling heeft met parameter λ. Dus X heeft kansdichtheid $f(x) = \lambda e^{-\lambda x}$ als $x \geq 0$. We toetsen $H_0 : \lambda = 2$ tegen $H_1 : \lambda > 2$ door H_0 te verwerpen als $x < 1/6$.
   (a) Bereken de kans op een type 1 fout.
   (b) Bereken de kans op een type 2 fout, als $\lambda = 3$.
8. Stel dat we binaire data $x_1, \ldots, x_n$ hebben. Dus iedere x_i is ofwel gelijk aan 1, ofwel gelijk aan 0. We vatten $x_1, \ldots, x_n$ op als realisaties van onafhankelijke stochasten $X_1, \ldots, X_n$ uit een Bernoulli verdeling met parameter p.
   (a) Laat zien dat de schatter $S = \bar{X}_n$ zuiver is voor p.
   (b) Definieer $T = X_1$. Bereken de relatieve efficiëntie van S ten opzichte van T.
9. In een zevendaagse studie naar het effect van ozon is een groep van 14 ratten in een ozon-vrije omgeving, en een groep van 13 ratten is een ozon-rijke omgeving geplaatst. Voor iedere rat is de gewichtstoename (in grammen) gemeten. We vragen ons af of ozon een negatieve invloed heeft op gewichtstoename. We onderzoeken dit door de hypothese $H_0 : \mu_1 = \mu_2$ te toetsen tegen de alternatieve hypothese $H_1 : \mu_1 > \mu_2$. Hierbij zijn μ_1 en μ_2 de gewichtstoename voor een rat in de ozon-vrije en ozon-rijke omgeving respectievelijk. In de ozon-vrije omgeving is de gemiddelde gewichtstoename $\bar{x}_{14} = 22.40$; in de ozon-rijke omgeving $\bar{y}_{13} = 11.01$. De gepoolde standaard-deviatie is 4.58. De niet gepoolde standaard-deviatie is 4.14.

   Veronderstel dat de data normaal verdeeld zijn, met gelijke variantie in beide groepen. Ga na of de nulhypothese verworpen wordt bij significantieniveau 0.05.

Puntenverdeling bij vragen:

Opgave:	1	2a	2b	3	4a	4b	5	6a	6b	7a	7b	8a	8b	9
Punten:	1	2	2	2	1	2	2	2	2	2	2	1	2	2

Antwoorden:

1 Door onafhankelijkheid geldt $\mathrm{P}(A \mid B) = \mathrm{P}(A) = \frac{1}{2}$ en $\mathrm{P}(B \mid A) = \mathrm{P}(B) = \frac{1}{3}$. Dus $\mathrm{P}(A \cap B) = \mathrm{P}(A \mid B)\,\mathrm{P}(B) = \frac{1}{2}\frac{1}{3} = \frac{1}{6}$.

2 Als we willekeurig een leerling uit de klas kiezen, dan geldtt $\mathrm{P}(J) = 1/3$ en $\mathrm{P}(M) = 2/3$. Laat K de gebeurtenis zijn dat een leerling kleurenblind is. Gegeven is dat $\mathrm{P}(K \mid J) = 1/20$ en $\mathrm{P}(K \mid M) = 1/100$.

a. Gevraagd is $\mathrm{P}(K)$. Er geldt

$$\mathrm{P}(K) = \mathrm{P}(K \mid J)\,\mathrm{P}(J) + \mathrm{P}(K \mid M)\,\mathrm{P}(M) = \frac{1}{20}\frac{1}{3} + \frac{1}{100}\frac{2}{3} = \frac{7}{300} \approx 0.023$$

b. Gevraagd is $\mathrm{P}(J \mid K)$. Er geldt

$$\mathrm{P}(J \mid K) = \frac{\mathrm{P}(K \mid J)\,\mathrm{P}(J)}{\mathrm{P}(K)} = \frac{\frac{1}{20}\frac{1}{3}}{\frac{7}{300}} = \frac{5}{7} \approx 0.71$$

3 Merk op dat X waarden aanneemt tussen $(0,1)$. Het volstaat om te laten zien dat de verdelingsfunctie van X overeenkomt met die van een uniforme verdeling. Dit is het geval, aangezien voor $x \in (0,1)$ geldt

$$\mathrm{P}(X \leq x) = \mathrm{P}(1 - U \leq x) = \mathrm{P}(-U \leq x - 1) = \mathrm{P}(U \geq 1 - x) = x.$$

4 a. Het aantal overstromingen in 50 jaar volgt een Poisson verdeling met parameter 2, daarom is de gevraagde kans

$$2 \cdot e^{-2} \approx 0.27.$$

b. Het aantal overstromingen in 25 jaar volgt een Poisson verdeling met parameter $25 \times (2/50) = 1$. Als $X \sim \mathrm{Pois}(1)$, dan is de gevraagde kans dus

$$\mathrm{P}(X \geq 2) = 1 - \mathrm{P}(X = 0) - \mathrm{P}(X = 1) = 1 - e^{-1} - e^{-1} = 1 - 2 \cdot e^{-1} \approx 0.26.$$

5 Merk op, Z neemt de waarden $0, 1, 2$ aan. Nu geldt

$$P(Z = 2) = P(X = 2, Y = 1) = \frac{1}{9}\frac{1}{2} = \frac{1}{18},$$

$$P(Z = 1) = P(X = 1, Y = 1) = \frac{4}{9}\frac{1}{2} = \frac{4}{18}.$$

Dus $\mathrm{P}(Z = 0) = 1 - \frac{1}{18} - \frac{4}{18} = \frac{13}{18}$.

6 a. Zie MIPS opgave 21.2 De likelihood is

$$L(p) = [(1-p)^2 p][(1-p)^2 p][(1-p)^3 p] = (1-p)^7 p^3.$$

Neem de logaritme

$$\ell(p) = 7 \log(1-p) + 3 \log p$$

b. Aangezien de kans op eerst 4 keer munt gelijk is aan $(1-p)^4$ geldt

$$L(p) = [(1-p)^2 p][(1-p)^2 p][(1-p)^4] = (1-p)^8 p^2.$$

7 a. $\mathrm{P}(X < 1/6 \mid \lambda = 2) = 1 - e^{-1/3} \approx 0.28$.

b. $\mathrm{P}(X \geq 1/6 \mid \lambda = 3) = e{-}1/2 \approx 0.61$.

8 a. $\mathrm{E}\left[\bar{X}_n\right] = \mathrm{E}\left[X_1\right] = p$.

b.

$$\frac{\mathrm{Var}(S)}{\mathrm{Var}(T)} = \frac{p(1-p)/n}{p(1-p)} = \frac{1}{n}.$$

9 Dit is een variatie op MIPS opgave 28.3(a). Aangezien de varianties gelijk verondersteld worden, gebruiken we als toetsingsgrootheid

$$S = \frac{\bar{X}_{23} - \bar{Y}_{22}}{S_p},$$

waarbij S_p de gepoolde standaard-deviatie is. We verwerpen de nulhypothese voor grote waarden van S. Onder H_0 heeft S een t-verdeling met $14 + 13 - 2 = 25$ vrijheidsgraden. De geobserveerde waarde van de toetsingsgrootheid bedraagt $(22.4 - 11.01)/4.58 \approx 2.487$. De kritieke waarde bij de toets wordt gegeven door $t_{25;0.05} \approx 1.708$. We verwerpen de nulhypothese dus.