

Kansrekening en statistiek
wi2105IN—deel 2
11 april 2011, 14.00–16.00 uur

Bij dit examen is het gebruik van een (evt. grafische) rekenmachine toegestaan. Tevens krijgt u een formuleblad uitgereikt—na afloop inleveren alstublieft.

Meerkeuzevragen

Toelichting: In het algemeen zijn niet altijd vijf van de zes alternatieven 100% fout, het juiste antwoord is het meest volledige antwoord. Maak op het bijgeleverde antwoordformulier het hokje behorende bij het door u gekozen alternatief zwart of blauw. Doorstrepen van een fout antwoord heeft geen zin: u moet het òf uitgummen, òf verwijderen met correctievloeistof òf een nieuw formulier invullen. Vergeet niet uw studienummer in te vullen en aan te strepen.

1. Laat gegeven zijn dat $X_1, X_2, \dots, X_{60}$ onafhankelijke en uniform verdeelde stochasten op het interval $[0, 3]$ zijn. Er geldt $\text{Var}(X_1) = \frac{3}{4}$. De centrale limietstelling geeft dat $P(X_1 + X_2 + \dots + X_{60} \geq 80)$ bij benadering gelijk is aan
- a. 0.9319 b. 0.7967 c. 0.8159 d. 0.8289 e. 0.9236 f. 0.9871

2. We beschikken over een dataset van omvang 20 die we opvatten als realisatie van onafhankelijke stochasten $X_1, X_2, \dots, X_{20}$. Gegeven is dat elke X_i een continue verdeling heeft met kansdichtheid $f(x) = \frac{1}{2\theta}$, als $-\theta \leq x \leq \theta$, en $f(x) = 0$ voor alle andere waarden van x . Verder is $\theta > 0$ onbekend. Definieer de volgende schatter voor θ :

$$T = a(|X_1| + |X_2| + \dots + |X_{20}|).$$

Voor welke a is T een zuivere schatter voor θ ?

- a. $\frac{1}{40}$ b. $\frac{1}{20}$ c. $\frac{3}{20}$ d. $\frac{1}{10}$ e. $\frac{5}{40}$ f. $\frac{3}{40}$
3. Gegeven is de volgende dataset

.0018	.1571	.1687	.2006	.2093
.2691	.2774	.3233	.3279	.3922
.4417	.4931	.5704	.5998	.6013
.6221	.6719	.9470	.9531	.9870

Voor welk getal in de dataset is de waarde van de empirische verdelingsfunctie van deze dataset gelijk aan 0.6?

- a. 0.2691 b. 0.5704 c. 0.4931 d. 0.9470 e. 0.4417 f. 0.4991
4. Gegeven zijn twee steekproefgrootheden (Engels: sample statistics) T en S . We willen een parameter θ schatten, die uit het interval $(1, \infty)$ komt. Bekend is dat $E[T] = 2\theta$, en dat $E[S] = \sqrt{\theta}$. Welke van de volgende schatters is een zuivere schatter voor θ ?
- a. $2T$ b. T^2 c. $T - 2$
d. $T + S$ e. $\sqrt{S}$ f. i.h.a. géén van deze vijf
5. Gegeven is een dataverzameling die een realisatie is van een steekproef van omvang 16 uit een $N(\mu, \sigma^2)$ verdeling, met μ en $\sigma^2 > 0$ onbekend. Verder is gegeven dat het gemiddelde van de data 15 is, en de steekproefvariantie gelijk is aan 4. Het 90% betrouwbaarheidsinterval voor μ wordt dan gegeven door:
- a. (13.934, 16.066) b. (13.940, 16.006) c. (14.020, 15.980)
d. (14.124, 15.877) e. (14.127, 15.873) f. (14.177, 15.822)

6. Gegeven data $x_1, x_2, \dots, x_{15}$ die we opvatten als een steekproef uit een verdeling met kansdichtheid

$$f(x) = \begin{cases} 0 & x < \theta \\ e^{-(x-\theta)} & x \geq \theta \end{cases}.$$

We toetsen de nulhypothese $\theta = 0$ tegen de alternatieve hypothese $\theta > 0$. Als toetsingsgrootheid nemen we $T = \min\{X_1, X_2, \dots, X_{15}\}$. Grote waarden van T duiden op de alternatieve hypothese. Stel dat de geobserveerde waarde van T gelijk is aan $t = 0.1$. De p -waarde bij deze toets (afgerond op 2 decimalen) is gelijk aan:

- a. 0.14 b. 0.90 c. 0.22 d. 0.06 e. 0.12 f. 0.43

Hint: Als $X_1, X_2, \dots, X_n$ een steekproef uit een $\text{Exp}(\lambda)$ verdeling is, dan heeft

$$\min\{X_1, X_2, \dots, X_n\}$$

een $\text{Exp}(n\lambda)$ verdeling.

7. Gegeven is een dataverzameling die een realisatie is van een steekproef van omvang 16 uit een kansverdeling met verwachting μ . Het gemiddelde van de data is 10 en de steekproefvariantie is 4. Men voert een bootstrapsimulatie uit voor het gestudentiseerde gemiddelde met 1000 herhalingen. Van de 1000 geordende bootstrapwaarden is het volgende bekend:

25-ste	50-ste	100-ste	901-ste	951-ste	976-ste
-3.42	-2.94	-2.01	1.11	1.39	1.62

Het 90% bootstrap betrouwbaarheidsinterval voor μ wordt gegeven door:

- a. (9.305, 11.470) b. (8.290, 10.825) c. (8.530, 10.695)
d. (8.995, 10.555) e. (9.175, 11.710) f. (9.445, 11.005)

8. Stel dat $X_1, X_2, \dots, X_n$ een steekproef is uit een verdeling met verwachting μ en variantie σ^2 . Beschouw de beweringen

- A Het steekproefgemiddelde is een zuivere schatter voor μ .
B De steekproefvariantie S_n^2 is een zuivere schatter voor σ^2 .
C De steekproefstandaarddeviatie S_n is een zuivere schatter voor σ .

Welke van deze beweringen zijn waar?

- a. geen b. B en C c. A
d. A, B en C e. A en C f. A en B

9. Stel dat $X_1, X_2, \dots, X_n$ een steekproef is uit een normale verdeling met verwachting μ en variantie σ^2 . Beschouw de volgende beweringen:

- A Het 95%-betrouwbaarheidsinterval voor μ wordt smaller naarmate n groter wordt.
B Het 95%-betrouwbaarheidsinterval voor μ is smaller dan het 99%-betrouwbaarheidsinterval voor μ

Welke van deze beweringen zijn waar?

- a. geen b. A c. B d. A en B

10. Voor een zekere statistische toets wordt de toestingsgrootheid gegeven door T . Het significantie-niveau is gelijk aan 0.1. Het kritieke gebied behorende bij deze toets noteren we met K . De kans dat we de nulhypothese verwerpen als deze niet waar is, is gelijk aan 0.6. De kans op een fout van de tweede soort is gelijk aan:

- a. 0.1 b. 0.9 c. 0.4
d. 0.6 e. 0.5 f. 0.7

Open vragen

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.

1. Men beschikt over een dataset van 13 elementen die men opvat als een realisatie van een steekproef uit een $Pois(\mu)$ verdeling. Door omstandigheden zijn de 0-en en 1-en op een hoop terecht gekomen. Zodoende is alleen het volgende bekend:

waarde	0 of 1	2	3
aantal	10	2	1

- a. Beargumenteer dat de likelihood wordt gegeven door

$$L(\mu) = C \cdot \frac{e^{-13\mu} \mu^7 (\mu + 1)^{10}}{24},$$

waarbij de constante C het aantal volgorden is om 10 keer een 0 of een 1, 2 keer een 2 en 1 keer een 3 te hebben.

- b. Bepaal de loglikelihood en bereken de maximum likelihood schatting voor μ .

2. Stel dat X een uniforme verdeling heeft op het interval $[0, \theta]$, waarbij $\theta > 0$ een onbekende parameter is. We toetsen $H_0 : \theta = 3$ versus $H_1 : \theta > 3$. Als toetsingsgrootheid nemen we X zelf.

- a. Bereken het kritieke gebied bij significantieniveau 0.1.
b. Stel dat we $x = 1$ observeren. Bereken de bijbehorende p -waarde.

Normering:

Voor iedere open vraag is 1 punt te behalen.

Voor de open vragen geldt:

- 1a: 1 punt
1b: 2 punten
2a: 1 punt
2b: 1 punt

In totaal kun je dus 15 punten halen.

Het eindcijfer wordt gegeven door $1 + \text{aantal behaald punten} * \frac{9}{15}$.

