NTNU
Kunnskap for en bedre verden

Avdeling for ingeniør- og realfag, NTNU i Ålesund

# Eksamensoppgave i IR201812 Statistikk og Simulering

**Faglig kontakt under eksamen:** **Siebe van Albada**

**Eksamensdato:** **31.05.16**
**Eksamenstid:** **09:00 – 13:00**
**Hjelpemiddelkode/Tillatte hjelpemidler:**
**Kalkulator uten muligheter for kommunikasjon**
**Statistikkbok**
**Annen informasjon:** **Det er lov med håndskrevne notater i statistikkboka.**

**Målform/språk:** **Bokmål**
**Antall sider (uten forside):** **3**
**Antall sider vedlegg:** **-**

| **Informasjon om trykking av eksamensoppgave** | |
|---|---|
| **Originalen er:** | |
| ~~**1-sidig**~~ X | **2-sidig** ☒ |
| **sort/hvit** ☐ | ~~**farger**~~ X |

**Kontrollert av:**

Dato Sign

Merk! Studenter finner sensur i Studentweb. Har du spørsmål om din sensur må du kontakte instituttet ditt. Eksamenskontoret vil ikke kunne svare på slike spørsmål.

**Eksamen 31. mai 2016** **IR201812 Statistikk og Simulering**

*Tillate hjelpemidler: Kalkulator og statistikkbok med notater skrevet i boka.*

**Oppgave 1**
Vi kaster en rettferdig firesidet terning to ganger. La den stokastiske variabelen X være summen av antall øyne i de to kastene.

a) Hva er utfallsrommet til X?

b) Hva er sannsynlighetsfordelingen til X?

c) Regn ut forventningsverdien for X.

d) Regn ut populasjonsvariansen for X.

e) Hva er den kumulative sannsynlighetsfordelingen for X?

Vi gjentar eksperimentet fire ganger og får som resultat: 6; 3; 5; 5.

f) Regn ut utvalgsmiddelverdien.

g) Regn ut utvalgsvariansen.

**Oppgave 2**
Bruk sannsynlighetstabellene i boka og finn følgende:

a) $P(0.5 < Z < 1.5)$ der Z er standard normalfordelt.

b) $P(0.5 < X < 1.5)$ der X er normalfordelt med forventningsverdi $\mu = 1$ og standardavvik $\sigma = 0.5$.

c) $P(1025 < Y < 1500)$ der Y er binomialfordelt med $\pi = 0.2$ og $n = 5000$. (Bruk sentralgrensesetningen).

**Oppgave 3**
På et lager til en butikk står 25 datamaskiner. Fem av disse er ødelagt uten at det viser.

a) Hvor stor er sannsynligheten for at minst to av de første 10 datamaskinene som blir solgt, er ødelagt?

Butikken er en del av en internasjonal kjede. Det selges i gjennomsnitt 3.3 datamaskiner per minutt totalt. Anta at antall solgte datamaskiner per tidsenhet er poissonfordelt.

b) Hvor stor er sannsynligheten for at kjeden selger færre enn to datamaskiner på et bestemt minutt?
c) Hvor stor er sannsynligheten for at kjeden selger færre enn to datamaskiner på et bestemt minutt når vi vet at antallet er lavere enn tre?

**Oppgave 4**
Av 1000 ganger vi kjørte en Monte-Carlo-simulering, konvergerte den 980 ganger.

Har vi statistisk belegg for å kunne si at andelen av simuleringene som ikke konvergerer ligger over 1 %? Gjennomfør hypotesetesten på 5 % signifikansnivå.

**Oppgave 5**
Anta at du har tilgang til en (pseudo-)slumptallgenerator som sampler den uniforme fordelingen i intervallet [0, 1>. Hvordan kan en sample de følgende sannsynlighetsfordelingene? Beskriv algoritmene detaljert, for eksempel med pseudokode.

a) Den diskrete sannsynlighetsfordelingen som er definert ved P(X=A)=0.2, P(X=B)=0.3 og P(X=C)=0.5.
b) Den binomiske fordelingen med n=10 og $\pi = 0.1$.
c) Den eksponentielle fordelingen $f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

**Oppgave 6**
La oss si at vi har implementert Box-Muller transformasjonen for å sample standardnormalfordelingen. Hvordan kan vi anvende funksjonen for å sample en normalfordeling med middelverdi $\mu = 10$ og standardavvik $\sigma = 2$?

**Oppgave 7**
Vi stiller inn temperaturen på en termostat. Deretter måler vi den faktiske temperaturen i midten av rommet. Resultatene er:

| **Innstilt temperatur (°$C$)** | **Målt temperatur (°$C$)** |
|---|---|
| 10 | 10,2 |
| 15 | 14,9 |
| 20 | 19,5 |
| 25 | 23,8 |
| 30 | 27,1 |
| 35 | 32,0 |

For datasettet gjelder:
$\sum x = 135$; $\sum y = 127{,}5$; $\sum x^2 = 3475$; $\sum y^2 = 3031{,}15$ og $\sum xy = 3243{,}5$

a) Bruk minste kvadratsums metode for å finne den beste rette linja som beskriver datasettet.
b) Regn ut korrelasjonskoeffisienten mellom målt temperatur og innstilt temperatur.

**Oppgave 8**
Vi ønsker å modellere / simulere de kjemiske reaksjonene

**Reaksjon 1:** A + B → C (rate constant: $k_f$)

**Reaksjon 2:** C → A + B (rate constant: $k_b$)

Hvilken simuleringsmetode er best egnet om:

a) det er få partikler i systemet (ca. 100), og systemet er homogent (det vil si uniform konsentrasjon over rommet)?

b) det er få partikler i systemet (ca. 100), og systemet er inhomogent (det vil si at konsentrasjonen varierer over rommet)?

Beskriv metodene detaljert.

**Oppgave 9**
De følgende reaksjonene blir simulert med Gillespie-algoritmen:

**Reaksjon 1:** $A + B \rightarrow C + D$ med rate $20 \times N_A \times N_B$ per sekund og

**Reaksjon 2:** $C + D \rightarrow A + B$ med rate $10 \times N_C \times N_D$ per sekund.

Her er $N_A$, $N_B$ og $N_C$ og $N_D$ antallene partikler av hhv. type A, B og C i systemet. Vi starter simuleringen med 200 partikler av type A, og 100 partikler av type B, C og D.

a) Regn ut ratene for reaksjon 1 og 2 ved starten av simuleringen.
b) Hva er sannsynligheten for at den første reaksjonen som skjer er reaksjon 1?
c) Hvilke bevaringslover gjelder for partiklene av type A, B, C og D? Dvs, om vi vet antallet A-partikler, hvordan kan vi finne antallet B-, C- og D-partikler?

Systemet er i likevekt når ratene til reaksjon 1 og 2 er like store.

d) For hvilke antall partikler er systemet i likevekt?

**Oppgave 10**
En partikkel befinner seg i et endimensjonalt raster med uendelig utstrekning. For hvert tidssteg beveger partikkelen seg ett steg til en av de to naboposisjonene med følgende sannsynlighetsfordeling:

| | | | | |
|---|---|---|---|---|
| | *40 %*<br>*Y = -1* | *Start*<br>*Y = 0* | *60 %*<br>*Y = 1* | |

Hvert steg er uavhengig av de forrige. Vi definerer den stokastiske variabelen Y som posisjonen til partikkelen etter n steg.

a) Regn ut sannsynlighetsfordelingen for posisjonen P(Y) etter to steg.

b) Regn ut sannsynlighetsfordelingen for posisjonen P(Y) etter tre steg.

c) Forklar at sannsynlighetsfordelingen for posisjonen etter *n* steg er binomialfordelt: $P(Y = 2x - n) = \binom{n}{x} 0.6^x 0.4^{n-x}$, om vi definerer den stokastiske variabelen X som antallet ganger at partikkelen har beveget seg mot høyre.

d) Regn ut forventningsverdien for posisjonen Y etter *n* steg.