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: **09.09.16**

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 X	farger □

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 9. september 2016 **IR201812 Statistikk og Simulering**

Tillatte hjelpemidler: Kalkulator og statistikkbok med notater skrevet i boka.

Oppgave 1 (14 %; per deloppgave: 3, 2, 2, 3, 4)

En kontinuerlig stokastisk variabel X har sannsynlighetstetthet

$$f(x) = \begin{cases} kx, \ 0 \le x < 1 \\ k, \ 1 \le x < 2 \\ 3k - kx, \ 2 \le x < 3 \\ 0, \ for\ øvrig \end{cases}$$

a) Vis at k=$\frac{1}{2}$.
b) Tegn grafen til sannsynlighetstettheten.
c) Regn ut forventningsverdien for X.
d) Regn ut populasjonsvariansen for X.
e) Finn den kumulative fordelingsfunksjonen for X.

Oppgave 2 (11 %; per deloppgave: 3, 4, 4)

To kuler blir valgt tilfeldig med tilbakelegging fra en sekk som inneholder fem hvite og tre svarte kuler. Finn sannsynligheten at:

a) Begge kulene har samme farge.
b) Minst én av kulene er hvit.
c) Gjenta oppgave b) om vi velger kulene uten tilbakelegging.

Oppgave 3 (9%; per deloppgave: 2, 3, 4)

Bruk sannsynlighetstabellene i boka og finn følgende:

a) P ($-1 < T < 1$) der T er t-fordelt med to frihetsgrader.

b) P ($-4 < X < 0$) der X er normalfordelt med forventningsverdi $\mu = -2$ og standardavvik $\sigma = 2$.

c) P ($4950 \le Y < 5050$) der Y er binomialfordelt med $\pi = 0.5$ og n = 10000. (Bruk sentralgrensesetningen).

Oppgave 4 ***(15%, per deloppgave: 3, 4, 4, 4)***

I en simulering beveger en partikkel seg over en rett linje. Partikkelen begynner på posisjon x = 0 på tidspunkt t = 0. Hvert tidssteg beveger den seg med 30% sannsynlighet over en avstand 1 mot høyre (forflytning $\Delta x=1$), med 10% sannsynlighet over en avstand 1 mot venstre ($\Delta x= -1$), og med 60% sannsynlighet blir den på samme plass ($\Delta x= 0$).

a) Hva er sannsynlighetsfordelingen for posisjonen etter to tidssteg?

b) Vis at sannsynlighetsfordelingen til posisjonen etter ett tidssteg har middelverdi $\mu = 0.2$ og standardavvik $\sigma = 0.6$.

c) La den stokastiske variabelen Y antyde posisjonen til partikkelen etter 100 tidssteg. Gi verdiene for μ_Y og σ_Y.

d) Hva er sannsynligheten for at posisjonen til partikkelen etter hundre tidssteg ligger mellom x=14 og x=26?

Oppgave 5 (14 %; per deloppgave: 3, 4, 3, 4)

En bedrift jobber med 10 uavhengige prosjekter. For hvert prosjekt er det 75% sannsynlighet for at det avsluttes med overskudd.

a) Hva er sannsynligheten for at bedriften avslutter eksakt ni av prosjektene med overskudd?
b) Hva er sannsynligheten for at ni eller flere prosjekter ikke avsluttes med overskudd?

Bedriften har 20 medarbeidere: 5 administrativt ansatte og 15 ingeniører. Bedriftslederen velger 6 medarbeidere vilkårlig for en medarbeidersamtale.

c) Hva er sannsynligheten for at 2 eller flere administrativt ansatte blir valgt?
d) Hva er sannsynligheten for at 2 eller flere administrativt ansatte blir valgt, gitt at færre enn 4 administrativt ansatte blir valgt?

Oppgave 6 (10%; 5 per deloppgave)

I en bestemt simulering forventer vi at gjennomsnittlig posisjon y til en partikkelsverm varierer lineært med tiden x. Vi skriver ut y på sju påfølgende tidspunkt. For datasettet gjelder:

$\Sigma x = 140$; $\Sigma y = 122{,}1$; $\Sigma x^2 = 3500$; $\Sigma y^2 = 2577{,}01$ og $\Sigma xy = 3000{,}5$ og $n = 7$.

a) Regn ut korrelasjonskoeffisienten mellom gjennomsnittlig posisjon y og tid x.

b) Bruk minstekvadratsums metode for å finne den beste rette linja gjennom målingene.

Oppgave 7 (5%)

I en simulering måler vi den såkalte frie energien til et fysisk system i fem uavhengige eksperimenter. Resultatene er (i kcal/mol): 28.5, 28.5, 28.3, 28.6 og 28.8. Har vi statistisk belegg for å kunne si at middelverdien til den frie energien ligger under 28.8 kcal/mol? Anta at resultatene kommer fra en normalfordeling, og gjennomfør hypotesetesten på 5 % signifikansnivå.

Oppgave 8 (5%)

Hvordan kan en bruke acceptance-rejection sampling for å estimere volumet til en kule av radius 3? Beskriv fremgangsmåten detaljert.

Oppgave 9 (13%; per deloppgave 4, 4, 5)

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

a) En Bernoulliprosess med π=0.6.
b) Den binomiske fordelingen med n=5 og π = 0.6.
c) Den eksponentielle fordelingen $f(x) = \begin{cases} 3e^{-3x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

Oppgave 10 (4%)
Vi har implementert Box-Muller transformasjonen i en funksjon for å sample standardnormalfordelingen. Hvordan kan vi anvende denne funksjonen for å simulere en partikkel som beveger seg på en rett linje med steg av normalfordelt størrelse, når lengden på stegene har middelverdi μ = 1 og standardavvik σ = 10?