Eksamen 5. juni 2018 **IR201812 Statistikk og Simulering**
Eksamenstid 4 timer *Bokmål*

Tillatne hjelpemiddel er kalkulator og læreboka i statistikk. Eigne notat i boka er lov. Ta med **all mellomregning** som er nødvendig for å grunngi svaret.

Oppgave 1 .... *(14%)*

I et ukentlig minilotteri er der 10% sannsynlighet for å vinne 50 kr., 10% for 150 kr., og 5% for 300 kr. Loddet koster 40 kr. (Et lodd kan bare vinne én av de tre premiene, og lodd som ikke vinner en av disse, vinner ingenting.) La den stokastiske variabelen X_j være beløpet en deltager vinner i lotteriet i uke j.

(a) Regn ut forventningsverdien for X_j.

(b) Regn ut standardavviket for X_j.

Jan deltar 50 uker på rad i lotteriet. La den stokastiske variabelen Y være det totale beløpet han vinner, minus det totale beløpet han betaler for loddene.

(c) Skriv Y som en funksjon av X_j.

(d) Regn ut forventningsverdien for Y.

(e) Regn ut standardavviket for Y.

(f) Tegn en skisse av sannsynlighetsfordelingen (PDF) for Y, der du også indikerer forventningsverdien og standardavviket.

(g) Regn ut sannsynligheten for at Jan har tjent mer enn han har betalt efter å ha deltatt 50 ganger i lotteriet. (Det vil si: Regn ut $P(Y > 0)$).

Oppgave 2 .... *(4%)*

Bruk sannsynlighetstabellene i boken og finn følgende:

(a) $P(Z > 2{,}0)$ der Z er standardnormalfordelt.

(b) Den kritiske verdien t slik at $P(T < t) = 0{,}05$ der T har Students t-fordeling med fem frihetsgrader.

Oppgave 3 .... *(6%)*

En stokastisk variabel X er binomialfordelt med n forsøk og punktsannsynlighet $\pi = 0{,}2$. Svar på følgende spørsmål, og forklar hvordan du kommer frem til svaret.

(a) La $n = 6$, og finn $P(X \leq 2)$.

(b) La $n = 180$, og finn $P(X \leq 40)$.

Oppgave 4 .... *(12%)*

Se for deg en agentbasert modell for et økosystem med rev og kanin (rov- og byttedyr), hvor agentene beveger seg over et rutenett eller raster (*grid*), og en *enkel* simulatorimplementasjon av denne modellen.

(a) Forklar hovudprinsippet for agentbasert simulering.

(b) Tegn en skisse til klassediagram for en enkel simulator for rovdyr/byttedyr-problemet.

(c) Forklar kort funksjon og formål for hver av klassene.

(d) Hvilke klasser i klassediagrammet ditt er agenter?

(e) Gi to eksempler på at ulike agenttyper i modellen din har (eller kan ha) forskjellige egenskaber.

Oppgave 5 .... *(10%)*

Se nu på en agentbasert simulering av et varemarked, der vi ønsker å studere sammenhengen mellom pris og omsetning. Bedriftene selger svært lignende produkter, men de kan gjøre strategiske valg som prissetting, reklame, osv. Kundene kan velge hvor meget de vil kjøpe og av hvem. Bedriftene er nødt til å tjene penger, ellers går de konkurs og forsvinner fra modellen.

(a) Sammenlign problemet med rovdyr/byttedyr-modellen i forrige oppgave. Kva element kan du gjenbruka, og kva må skrivast om?

(b) Tegn klassediagram for markedssimulatoren, og forklar hva de viktigste klassene gjør.

(c) Forklar kort hvordan du vil modellere og implementere hva agentene gjør hver runde, for hver agenttype. Modellen kan godt være enkel.

Oppgave 6 .... *(15%)*

Billys burger laver burgerne for hånd. De har fått en del klager på at burgerne ikke alltid har rett størrelse, og dette må de undersøke. Det er viktig både at vekten er rett i gjennomsnitt, og at standardavviket er så lite som mulig. Vi skal derfor estimere både forventningsverdien og standardavviket.

Billy veier seks burgere og finner følgende vekter i gram:

$$160, 170, 180, 180, 185, 190$$

(a) Finn et (punkt)estimat for forventa vekt (forventingsverdien).

(b) Regn ut (eit estimat for) standardfeilen for estimatoren som du bruker.

(c) Finn et 95% konfidensinterval for forventet vekt. Presiser hvilke forutsetninger du må gjøre om sannsynlighetsfordelingen for vekten.

(d) Finn et (punkt)estimat for standardavviket til vekten på en burger.

(e) Forklar hvordan du kan bruke *bootstrap* for å estimere standardfeilen på estimatoren for standardavviket. (Du velger selv om du vil forklare vha. programkode, pseudokode, regneeksempler eller annet.)

Oppgave 7 .... *(15%)*

Vi studerer sammenhengen mellom reklame og omsetning for en bedrift. La X være beløpet de bruker på reklame og Y omsetningen, begge deler i tusen kroner. Vi observerer følgende:

X	Y
10	68
45	210
60	250
72	305
90	400
100	440

$$\sum_i x_i = 377$$
$$\sum_i y_i = 1673$$
$$\sum_i x_i^2 = 29\,009$$
$$\sum_i y_i^2 = 557\,849$$
$$\sum_i x_i y_i = 127\,090$$

(a) Skisser datapunktene på et plott.

(b) Finnes der korrelasjon mellom X og Y? Forklar svaret.

(c) Hva kan vi si om avhengigheten mellom X og Y? Forklar svaret.

(d) Bruk minste kvadratsums metode for å finne den beste rette linjen for å beskrive datasettet.

(e) Regn ut korrelasjonskoeffisienten mellom X og Y.

Oppgave 8 .... *(12%)*

En partikkel beveger seg på et 2D raster. Sannsynlighetsfordelingen etter ett steg er gitt som

		40%		
	20%	*start*	20%	
		20%		

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

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

Posisjonen (X_i, Y_i) etter i steg er gitt ved to stokastiske variabler X_i og Y_i. Nu skal vi bare se på bevegelsen langs Y-aksen.

(c) Finn forventningsverdien $E(Y_{100})$.

(d) Finn variansen for Y_{100}.

Oppgave 9 .... *(12%)*

Når vi simulerer stokastiske prosesser bruker vi slumptall. En klassisk slumptallsgenerator er den lineære kongruensgeneratoren:

$$x_i = a \cdot x_{i-1} + c \pmod{m}.$$

(a) Forklar hva vi mener med perioden til generatoren. På hvilken måte er perioden vesentlig for bruken av generatoren i praksis?

(b) Finn den minste perioden til generatoren over når $a = 2$, $c = 3$ og $m = 18$.

(c) Slumptalsgeneratoren krever et *frø* (*seed*). Hva mener vi med et frø og hvordan blir det brukt for å gi oss en serie med slumptall fra generatoren over?

(d) Hvordan ville du ha valgt frø i et praktisk system, og hvorfor?