Eksamen 11. mai 2015 **IR201812 Statistikk og Simulering**
Eksamenstid 4 timer *Bokmål*

Ingen hjelpemidler er tillatt.
Ta med **all mellomregning** som er nødvendig for å grunngi svaret.

Oppgave 1 .... *(10%)*

La X og Y være uavhengige stokastiske variabler trukket uniformt fra $\{0,1\}$. La $Z = X \cdot Y \mod 2$. (Dersom vi tolker 0 som usann og 1 som sann, så er Z logisk «X og Y».)

(a) Hva betyr det at X og Y er uavhengige?

(b) Hva er utfallsrommet til Z?

(c) Sett opp sannsynlighetsfordelingen til Z i en tabell.

(d) Hva er forventningsverdien $E(Z)$?

(e) Hva er variansen $\mathsf{var}(Z)$?

Oppgave 2 .... *(6%)*

Vi kaster en rettferdig mynt to ganger. La den stokastiske variabelen X være antallet kron i løpet av disse to kastene. Regn ut forventningsverdien $\mu = E(X)$ på to måter, der du viser utregningen i detalj:

(a) ved hjelp av utfallsrommet $\mathcal{S}$ og formelen

$$\mu = \sum_{x \in \mathcal{S}} x \cdot p(x)$$

(b) ved hjelp av populasjonen $\mathcal{P}$ og formelen

$$\mu = \frac{1}{n} \sum_{x \in \mathcal{P}} x$$

Vi gjentar eksperimentet fire ganger og får utvalget $\mathcal{U} = \{2,0,1,1\}$.

(c) Regn ut utvalgsmiddelverdien $\overline{x}$.

Oppgave 3 .... *(4%)*

Bruk sannsynlighetstabellene i boken og finn følgende:

1. $P(Z < -0{,}5)$ der $Z \sim N(0,1)$ (standard normalfordeling)
2. Den kritiske verdien t slik at $P(T > t) = 0{,}05$ når T har Students t-fordeling med fem frihetsgrader.

Oppgave 4 .... *(7%)*

En partikkel befinner seg i et firkantet, todimensjonalt raster med uendelig utstrekning. For hvert tidssteg beveger partikkelen seg ett steg til en av de fire naboposisjonene med følgende sannsynlighetsfordeling:

		40%		
	30%	*start*	30%	
		0%		

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

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

Oppgave 5 .... *(8%)*

En stokastisk variabel X er binomialfordelt med $n = 10$ forsøk og punktsannsynlighet $\pi = 0{,}2$.

(a) Regn ut sannsynligheten $P(X \leq 2)$.

En annen stokastisk variabel Y er binomialfordelt med $n = 10000$ forsøk og punktsannsynlighet $\pi = 0{,}2$.

(b) Regn ut forventningsverdien $E(Y)$.

(c) Regn ut variansen σ^2 for Y.

(d) Finn sannsynligheten $P(Y \leq 2000)$. Forklar hvordan du finner svaret.

Oppgave 6 .... *(10%)*

Et dataprogram skal simulere en rettferdig firesidet terning.

(a) Regn ut forventningsverdien μ til terningen.

Vi tester dataprogrammet ved å kjøre det 10000 ganger og finner:

Antall øyner	Frekvens
1	2411
2	2340
3	2654
4	2595

Middelverdien over utvalget er 2,5433 og standardavviket er 1,2494.

(b) Finn et 99% konfidensintervall for forventningsverdien μ for antall øyne per terningkast i dette dataprogrammet.

(c) Virker denne virtuelle terningen rettferdig? Bruk svarene i oppgave a) og b) for å grunngi svaret ditt.

Oppgave 7 .... *(10%)*

I denne oppgaven ser vi på agent-basert simulering.

(a) Forklar hovedprinsippene ved agent-basert simulering. Hva kjennetegner en agent?

Se for deg en økonomisk modell der agentene er bedrifter som kan kjøpe og selge varer og aksjer. Ved implementasjon er det naturlig, i alle fall, å ha en *simulator*-klasse og en *agent*-klasse, gjerne med underklasser.

(b) Forklar hvilket ansvar og hvilke oppgaver *agent*-klassen har.

(c) Forklar hvilket ansvar og hvilke oppgaver *simulator*-klassen har.

Oppgave 8 .... *(9%)*

På en vei med fartsgrense på 80 km/h måler vi hastigheten på politibiler. Fire politibiler kjører forbi i løpet av en time, og farten blir målt til 90, 88, 80 og 94 km/h. Anta at hastighetene for politibiler på denne veien er normalfordelt. Har vi statistisk belegg for å si at gjennomsnittlig hastighet for politibiler ligger over fartsgrensen på denne veien? Formuler nødvendige hypoteser og test på 5% signifikansnivå.

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

I mange former for simulering, blir landskapet modellert som et rutenett eller raster (*grid*). På yttergrensene av rutenettet må man ta spesielle hensyn, og der finnes forskjellige randbetingelser som kan brukes.

Tenk på en agent-basert simulering, og se på det som skjer når en agent kommer til ytterkanten av rutenettet. Gi tre eksempler på forskjellige randbetingelser, og forklar hvordan du vil programmere agentens oppførsel i hvert tilfelle.

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

Vi simulerer en vei og teller bilene som passerer et kryss per tidsenhet.

Tid (min)	Antall biler
3	102
6	193
9	292
12	399

For datasettet gjelder:

$$\sum x = 30$$
$$\sum y = 986$$
$$\sum x^2 = 270$$
$$\sum y^2 = 292118$$
$$\sum xy = 8880.$$

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

(b) Regn ut korrelasjonskoeffisienten mellom tid og antall biler.

Oppgave 11 .... *(14%)*

En partikkel beveger seg i et én-dimensjonalt raster. Sannsynlighetsfordelingen for bevegelsen i hvert tidssteg er som følger:

Forflytning	Sannsynlighet
I ro	30%
Ett steg til venstre	25%
Ett steg til høyre	25%
To steg til venstre	10%
To steg til høyre	10%

La X være bevegelsen i løpet av t tidssteg, der negativ verdi viser bevegelse mot venstre og positiv mot høyre.

(a) Vis at $\mu = E(X) = 0$ for $t = 1$.

(b) Vis at $\mu = E(X) = 0$ for vilkårlig t.

(c) Regn ut diffusjonskoeffisienten til partikkelen. Formelen er $E(X^2) = 2dDt$ der $d = 1$ er antall dimensjoner, D er diffusjonskoeffisienten og t er tiden.

(d) Forklar hvorfor $E(X^2)$ (forventningsverdi for kvadratisk forflytning) etter ett tidssteg er lik variansen $\mathsf{var}(X)$.

Partikkelen begynner på tid $t = 0$ og posisjon $p_0 = 500$. La den stokastiske variabelen Y være posisjonen til partikkelen etter $t = 1000$ tidssteg.

(e) Regn ut forventningsverdien $\mu = E(Y)$.

(f) Regn ut variansen $\sigma_Y^2 = \mathsf{var}(Y)$ til Y.

(g) Hvilken fordeling har Y? Grunngi svaret.