Skrevne og trykte hjelpemidler samt kalkulator er tillatt.
Ta med **all mellomregning** som er nødvendig for å grunngi svaret.

Oppgave 1 ... *(15%)*

La den stokastiske variabelen X være antallet ganger kron i fire kast med en rettferdig mynt.

(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) Gi den kumulative sannsynlighetsfordelingen for X.

Vi gjentar eksperimentet fem ganger og får som resultat: 2; 1; 4; 2; 1.

(f) Regn ut utvalgsmiddelverdien.

(g) Regn ut utvalgsvariansen.

Oppgave 2 ... *(4%)*

Bruk sannsynlighetstabellene i boken og finn følgende:

(a) $P(T > 1{,}5)$ der T har Students t-fordeling med fem frihetsgrader.

(b) Den kritiske verdien z slik at $P(Z < z) = 0{,}95$ der $Z \sim N(0, 1)$ (standard normalfordeling).

Oppgave 3 ... *(6%)*

En stokastisk variabel X er binomialfordelt med n forsøk og punktsannsynlighet $\pi = 0{,}3$. 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 $P(X \leq 60)$.

Oppgave 4 ... *(9%)*

Mange dataprogrammer trenger tilfeldige tall.

(a) Gi tre eksempler på praktiske problemer som krever programmering med slumptall.

En vanlig måte å generere slike slumptall er den lineære kongruensgeneratoren

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

(b) 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?

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

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

Det er vanlig å teste kvaliteten på en slumptallsgenerator ved hjelp av statistisk hypotesetesting. La $[x_1, x_2, \ldots, x_n]$ være en heltallsfølge fra slumptalsgeneratoren. For å teste at de tre siste bitsene er uniformt fordelt kan vi bruke χ^2-statistikken (kji-kvadrat-statistikken)

$$G = \sum_{i=0}^{7} \frac{(F_y - E_y)^2}{E_y}$$

der $y_i = x_i \mod 8$, og F_y er frekvensen av hver mulige verdi av y.

(a) Hva er nullhypotesen for testen?

(b) Hvilken fordeling har G under nullhypotesen? Forklar svaret. Husk å spesifisere antall frihetsgrader.

(c) Hva er den kritiske verdien for testen ved 5% signifikansnivå?

(d) Hva er forventet frekvens E_y i denne testen?

(e) Forklar hvordan vi kan bruke χ^2-testen til å teste om påfølgende verdier i slumptallsfølgen er uavhengig fordelt.

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

I mange former for simulering, blir landskapet modellert som et rutenett eller raster (*grid*). En spesiell situasjon oppstår ved yttergrensen i landskapet. Dette kan man håndtere på forskjellige måter. Forklar kort (gjerne supplert med en tegning) hva vi mener med hver av de følgende randbetingelsene:

(a) periodisk

(b) reflekterende

(c) Hvilke randbetingelser er best egnet for simulering av svømmeoppførsel for fisk i et akvarium? Forklar svaret.

Oppgave 7 .... *(5%)*

Av 15000 ganger vi kjører en rovdyr-byttedyrsimulering, dør byttedyrpopulasjonen ut 10 ganger. Har vi statistisk belegg for å si at sannsynligheten for at byttedyrpopulasjonen dør ut er mindre enn 0,1 %? Gjennomfør hypotesetesten på 5 % signifikansnivå.

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

En *random walker* beveger seg på et 1D raster. La den stokastiske variabelen X være forflytningen til partikkelen i løpet av et tidssteg. Sannsynlighetsfordelingen for X er:

$$P(X = -2) = P(X = 2) = 0,1 \tag{5}$$

$$P(X = -1) = P(X = 1) = 0,4. \tag{6}$$

$$\tag{7}$$

(a) Regn ut populasjonsstandardavviket for posisjonen etter ett tidssteg.

(b) Regn ut populasjonsstandardavviket for posisjonen etter 100 tidssteg.

(c) Vi er interessert i den eksakte sannsynlighetsfordelingen for posisjonen til partikkelen etter 100 tidssteg. Forklar hvordan vi kan finne denne sannsynlighetsfordelingen ved hjelp av en datamaskin.

(d) Forklar hvordan vi kan generere et utvalg på størrelse 10 av posisjonen etter 100 tidssteg.

Oppgave 9 .... *(15%)*

Et maskinlæringsprogram får en stor datamengde som inndata. Målet er å studere avhengigheter mellom forskjellige variabler. Et tilfeldig utvalg av datasettet for variablene X og Y er gitt i tabellen under.

X	Y
2	6
3	2
5	8
4	7
0,5	0
1	3
0,5	2
1	2

For datasettet gjelder:

$$\sum x = 17$$
$$\sum y = 30$$
$$\sum x^2 = 56,5$$
$$\sum y^2 = 170$$
$$\sum xy = 92$$

(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 10 .... *(18%)*

Tenk deg at vi vil simulere et feltslag mellom to hærer vha. en agentbasert modell. Kvar hær har 2-3 ulike typar soldatar (t.d. riddar, fotsoldat med helebard, bogeskyttar) med ulike eigenskapar. Slagmarken kan vi modellere som et rutenett (raster, *grid*).

(a) Tegn en skisse til klassediagram for en *enkel* simulator.

(b) Forklar kort funksjon og formål for hver av klassene i skissen din.

(c) Drøft: hvilke klasser er spesielle for feltslagmodellen, og hvilke kan gjenbrukes på andre problemer?

(d) Hvilke klasser i klassediagrammet ditt er agenter?

(e) Hvilke egenskaber (tilstand og oppførsel) er felles for alle agentene i systemet?

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