Eksamen 5. juni 2018 **IR201812 Statistikk og Simulering**
Eksamenstid 4 timar *Nynorsk*

Tillatne hjelpemiddel er kalkulator og læreboka i statistikk. Eigne notat i boka er lov.
Ta med **all mellomrekning** som trengst for å grunngje svaret.

Oppgåve 1 .... *(14%)*

I eit vekentleg minilotteri er der 10% sannsyn for å vinne 50 kr., 10% for 150 kr., og 5% for 300 kr. Loddet kostar 40 kr. (Eit lodd kan berre vinna éin av dei tre premiane, og lodd som ikkje vinn ein av desse, vinn ingenting.) Lat den stokastiske variabelen X_j vera summen ein deltakar vinn i lotteriet i veke j.

(a) Rekn ut forventingsverdien for X_j.

(b) Rekn ut standardavviket for X_j.

Jan deltek 50 veker på rad i lotteriet. Lat den stokastiske variabelen Y vera den totale summen han vinn, minus den totale summen han betaler for lodda.

(c) Skriv Y som ein funksjon av X_j.

(d) Rekn ut forventingsverdien for Y.

(e) Rekn ut standardavviket for Y.

(f) Teikn ei skisse av sannsynsfordelinga (PDF) for Y, der du òg indikerer forventingsverdien og standardavviket.

(g) Rekn ut sannsynligheten for at Jan har tjent mer enn han har betalt etter å ha delteke 50 gongar i lotteriet. (Det vil seia: Rekn ut $P(Y > 0)$).

Oppgåve 2 .... *(4%)*

Bruk sannsynstabellane i boka og finn fylgjande:

(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 fridomsgradar.

Oppgåve 3 .... *(6%)*

Ein stokastisk variabel X er binomialfordelt med n forsøk og punktsannsyn $\pi = 0{,}2$. Svar på fylgjande spørsmål, og forklar korleis du kjem fram til svaret.

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

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

Oppgåve 4 .... *(12%)*

Sjå for deg ein agentbasert modell for eit økosystem med rev og kanin (rov- og byttedyr), der agentane flytter seg over eit rutenett eller raster (*grid*), og ein *enkel* simulatorimplementasjon av denne modellen.

(a) Forklar hovudprinsippet for agentbasert simulering.

(b) Teikn ei skisse til klassediagram for ein enkel simulator for rovdyr/byttedyr-problemet.

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

(d) Kva klasser i klassediagrammet ditt er agentar?

(e) Gje to døme på at ulike agenttypar i modellen din har (eller kan ha) ulike eigenskapar.

Oppgåve 5 .... *(10%)*

Sjå no på ein agentbasert simulering av ein varemarknad, der me ynskjer å studera samanhengen mellom pris og omsetnad. Bedriftene sel svært liknande produkt, men dei kan gjera strategiske val som prissetting, reklame, osv. Kundane kan velja kor mykje dei vil kjøpa og av kven. Bedriftene er nøydde til å tena pengar, ellers går dei konkurs og forsvinn frå modellen.

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

(b) Teikn klassediagram for marknadssimulatoren, og forklar kva dei viktigaste klassene gjer.

(c) Forklar kort korleis du vil modellera og implementera kva agentane gjer kvar runde, for kvar agenttype. Modellen kan godt vera enkel.

Oppgåve 6 .... *(15%)*

Billys burger lagar burgerane for hand. Dei har fått ein del klagar på at burgerane ikkje alltid har rett storleik, og dette må dei undersøkja. Det er viktig både at vekta er rett i gjennomsnitt, og at standardavviket er so lite som råd. Me skal difor estimere både forventingsverdien og standardavviket.

Billy veg seks burgerar og finn fylgjande vekter i gram:

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

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

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

(c) Finn eit 95% konfidensinterval for forventa vekt. Presiser kva føresetnader du må gjera om sannsynsfordelinga for vekta.

(d) Finn eit (punkt)estimat for standardavviket til vekta på ein burger.

(e) Forklar korleis du kan bruka *bootstrap* for å estimera standardfeilen på estimatoren for standardavviket. (Du vel sjølv om du vil forklara vha. programkode, pseudokode, reknedøme eller anna.)

Oppgåve 7 .... *(15%)*

Me studerer samanhengen mellom reklame og omsetnad for ei bedrift. Lat X vera summen ho bruker på reklame og Y omsetnaden, begge delar i tusen kroner. Me 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 datapunkta i eit plott.

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

(c) Kva kan me seia om avhenget mellom X og Y? Forklar svaret.

(d) Bruk minste kvadratsum-metoden for å finna den beste rette lina for å beskriva datasettet.

(e) Rekn ut korrelasjonskoeffisienten mellom X og Y.

Oppgåve 8 .... *(12%)*

Ein partikkel flyttar seg på eit 2D raster. Sannsynsfordelinga etter eitt steg er gjeve som

		40%		
	20%	*start*	20%	
		20%		

(a) Rekn ut sannsynsfordelingen for posisjonen etter to steg.

(b) Rekn ut sannsynsfordelingen for posisjonen etter tre steg.

Posisjonen (X_i, Y_i) etter i steg er gjeve ved to stokastiske variablar X_i og Y_i. No skal me berre sjå på rørsla langs Y-aksa.

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

(d) Finn variansen åt Y_{100}.

Oppgåve 9 .... *(12%)*

Når me simulerer stokastiske prosessar bruker me slumptal. Ein klassisk slumptalsgenerator er den lineære kongruensgeneratoren:

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

(a) Forklar kva me meiner med perioden til generatoren. På kva måte er perioden vesentleg for bruken av generatoren i praksis?

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

(c) Slumptalsgeneratoren krev eit *frø* (*seed*). Kva meiner me med eit frø og korleis vert det brukt for å gje oss ein serie slumptal frå generatoren over?

(d) Korleis ville du ha vald frø i eit praktisk system, og kvifor?