Ingen hjelpemiddel er tillatne.
Ta med **all mellomrekning** som trengst for å grunngje svaret.

Oppgåve 1 .... *(10%)*

Lat X og Y vera uavhengige stokastiske variablar trekt uniformt frå $\{0, 1\}$. Lat $Z = X \cdot Y \mod 2$. (Dersom me tolkar 0 som usann og 1 som sann, er Z logisk «X og Y».)

(a) Kva meiner me med at X og Y er uavhengige?

Solution: $P(X \cap Y) = P(X) \cdot P(Y)$ Det finnes flere likeverdige definisjoner; se bok.

(b) Kva er utfallsrommet til Z?

Solution: Z kan anta verdiene $0 \cdot 0 = 0$, $0 \cdot 1 = 0$, $1 \cdot 0 = 0$ og $1 \cdot 1 = 1$. Utfallsrommet er derfor $\{0, 1\}$.

(c) Set opp sannsynsfordelinga til Z i ein tabell.

Solution:

Z	P(Z)
0	0,75
1	0,25

(d) Kva er forventingsverdien $E(Z)$?

Solution: $E(Z) = \sum Z \cdot P(Z) = 0{,}75 \cdot 0 + 0{,}25 \cdot 1 = 0{,}25$

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

Solution: $\mathsf{var}(Z) = \sum (Z - E(Z))^2 \cdot P(Z) = 0{,}75 \cdot (0 - 0{,}25)^2 + 0{,}25 \cdot (1 - 0{,}25)^2 = 0{,}1875$

Oppgåve 2 .... *(6%)*

Me kastar ein rettferdig mynt to gongar. Lat den stokastiske variabelen X vera talet på kron i løpet av desse to kasta. Rekn ut forventingsverdien $\mu = E(X)$ på to måtar, der du viser utrekninga i detalj:

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

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

Solution: $\mu = \sum_{utfallsrom} x \cdot p(x) = 0 \cdot p(0) + 1 \cdot p(1) + 2 \cdot p(2) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$

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

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

Solution: $\mu = \frac{1}{n} \sum_{populasjon} x = \frac{1}{4}(0 + 2 \cdot 1 + 2) = 1$

Me gjentek eksperimentet fire gongar og får utvalet $\mathcal{U} = \{2, 0, 1, 1\}$.

(c) Rekn ut utvalsmiddelverdien $\overline{x}$.

Solution: $\overline{x} = \frac{1}{n} \sum\limits_{utvalg} x = \frac{1}{4}(2 + 0 + 1 + 1) = 1$

Oppgåve 3 . . . . *(4%)*

Bruk sannsynstabellane i boka og finn fylgjande:

1. $P(Z < -0{,}5)$ der $Z \sim N(0, 1)$ (standard normalfordeling)

Solution: $P(Z < -0{,}5) = 0{,}3085$

2. Den kritiske verdien t slik at $P(T > t) = 0{,}05$ når T har Students t-fordeling med fem fridomsgradar.

Solution:

$$t = 2{,}015$$

Oppgåve 4 . . . . *(7%)*

Ein partikkel finst i eit firkanta, todimensjonalt raster med uendeleg utstrekking. For kvart tidssteg går partikkelen eitt steg til ein av dei fire naboposisjonane med fylgjande sannsynsfordeling:

	40%	
30%	*start*	30%
	0%	

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

Solution:

		16%		
	24%		24%	
9%		18%		9%

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

Solution:

			6,4%			
		14,4%		14,4%		
	10,8%		21,6%		10,8%	
2,7%		8,1%		8,1%		2,7%

Oppgåve 5 .... *(8%)*

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

(a) Rekn ut sannsynet $P(X \leq 2)$.

> **Solution:** $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \binom{10}{0} \cdot 0{,}2^0 \cdot 0{,}8^{10} + \binom{10}{1} \cdot 0{,}2^1 \cdot 0{,}8^9 + \binom{10}{2} \cdot 0{,}2^2 \cdot 0{,}8^8 = 0{,}678$

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

(b) Rekn ut forventingsverdien $E(Y)$.

> **Solution:** $E(Y) = n \cdot \pi = 10000 \cdot 0{,}2 = 2000$

(c) Rekn ut variansen σ^2 for Y.

> **Solution:** $\sigma^2 = n \cdot \pi \cdot (1 - \pi) = 10000 \cdot 0{,}2 \cdot 0{,}8 = 1600$

(d) Finn sannsynet $P(Y \leq 2000)$. Forklar korleis du finn svaret.

> **Solution:** Vi bruker sentralgrensesetningen fordi n er stor. For en binomisk variabel er $\mu = n \cdot \pi = 10000 * 0{,}2 = 2000$ og $\sigma = \sqrt{n \cdot \pi \cdot (1 - \pi)} = \sqrt{10000 \cdot 0{,}2 \cdot 0{,}8} = 40$. Vi finner: $P(Y \leq 2000) = P\left(\frac{Y-\mu}{\sigma} \leq \frac{2000-\mu}{\sigma}\right) = P\left(Z \leq \frac{2000-2000}{40}\right) = P(Z \leq 0) = 0{,}5$

Oppgåve 6 .... *(10%)*

Eit dataprogram skal simulera ein rettferdig firesida terning.

(a) Rekn ut forventingsverdien μ til terningen.

> **Solution:** $\mu = \frac{1}{n} \sum_i x_i = \frac{1}{4}(1 + 2 + 3 + 4) = 2{,}5$

Me testar dataprogrammet ved å køyra det 10000 gongar og finn:

Tal på augo	Frekvens
1	2411
2	2340
3	2654
4	2595

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

(b) Finn eit 99% konfidensintervall for forventingsverdien μ for talet på augo per terningkast i dette dataprogrammet.

> **Solution:** Konfidensintervall for μ når σ er ukjent:
> $\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$
> Vi setter inn verdiene: $\bar{x} = 2{,}5433$, $s = 1{,}2494$, $n = 10000$ og $t_{\alpha/2} = 2{,}576$ og finner:
> $2{,}5433 - 2{,}576\frac{1{,}2494}{\sqrt{10000}} < \mu < 2{,}5433 + 2{,}576\frac{1{,}2494}{\sqrt{10000}}$
> $2{,}511 < \mu < 2{,}575$

(c) Verkar denne virtuelle terningen rettferdig? Bruk svara i oppgåve a) og b) for å grunngje svaret ditt.

Solution: For en rettferdig firesidet terning er $\mu = 2{,}5$. Denne verdien ligger utenfor 99% konfidensintervallet. Terningen virker derfor ikke rettferdig. Likevel beviser dette ikke at terningen ikke er rettferdig.

Oppgåve 7 .... *(10%)*

I denne oppgåva ser me på agent-basert simulering.

(a) Forklar hovudprinsippa ved agent-basert simulering. Kva kjenneteiknar ein agent?

Solution: Agent-basert simulering legg vekt på å simulera einskilde individ eller komponentar mest mogleg uavhengig av kvarandre. Kvart individ (kvar komponent) vert modellert som ein *agent*. Agent-modellen definerer kva agenten gjer, som ein funksjon av det agenten oppfattar av miljøet rundt seg. Agenten treng ikkje informasjon om heile det kringliggande systemet; berre om det lokale miljøet som direkte påverkar agenten. Dette opnar for å modellera agentane i detalj, utan å analysera systemsamanhengane på førehand.

(Kjernepunktet her er at ein agent er ein modell av eit individ (t.d. i eit økosystem) eller ein komponent (t.d. i eit mekanisk system), som kan modellerast i detalj. Slik vert systemet modellert nedanfrå og opp.)

Sjå for deg ein økonomisk modell der agentane er bedriftar som kan kjøpa og selja varer og aksjar. Ved implementasjon er det naturleg, i alle fall, å ha ein *simulator*-klasse og ein *agent*-klasse, gjerne med underklassar.

(b) Forklar kva ansvar og oppgåver *agent*-klassen har.

Solution: Agent-klassa har ansvar for oppførselen til agenten, og interagera med miljøet som del av denne oppførselen. Klassa har ein metode (*act()*) som vert kalt for kvart tidssteg, og denne metoden definerer kva agenten gjer i løpet av intervallet.

(c) Forklar kva ansvar og oppgåver *simulator*-klassen har.

Solution: Simulator-klassa må halda greie på klokka, og kalla ein *act()*-rutine i kvar agent for kvart tidssteg.

Ved oppstart kan simulator-klassa ha ansvar for å instantiera heile modellen med alle agentane.

Oppgåve 8 .... *(9%)*

På ein veg med fartsgrense på 80 km/h måler me farten på politibilar. Fire politibilar køyrer forbi i løpet av ein time, og farten vert målt til 90, 88, 80 og 94 km/h. Gå ut frå at farten på politibilar på denne vegen er normalfordelt. Har me statistisk grunnlag for å seia at gjennomsnittleg fart for politibilar ligg over fartsgrensa på denne vegen? Formuler dei hypotesane som du treng, og test på 5% signifikansnivå.

Solution: Dette er en énsidig hypotesetest for μ når σ er ukjent. Vi finner: $\bar{x} = 88$ og $s = 5{,}89$.
Vi setter opp hypotesene:
$H_0 : \mu = 80$
$H_1 : \mu > 80$
Den observerte t-verdien er: $t_{obs} = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{88-80}{5{,}89/\sqrt{4}} = 2{,}7$ p-verdi $= P(T \geq 2{,}7) = 0{,}037$
Siden p-verdi $= 0{,}037 < 0{,}05 = \alpha$, forkaster vi nullhypotesen og aksepterer den alternative hypotesen: Gjennomsnittlig hastighet til politibiler på denne veien ligger over $80 km/h$.

Oppgåve 9 .... *(12%)*

I mange formar for simulering, vert landskapet modellert som eit rutenett eller raster (*grid*). På yttergrensene av rutenettet må ein ta særskilde omsyn, og der finst ulike randvilkår som kan brukast.

Tenk på ei agent-basert simulering, og sjå på kva som skjer når ein agent kjem til ytterkanten av rutenettet. Gje tre døme på ulike randvilkår, og forklar korleis du vil programmera oppførselen åt agenten i kvart tilfelle.

Solution: Randbetingelsene kan være reflekterende, periodiske eller absorberende.
Reflekterende: Hvis en agent etter i et tidsteg har valgt å bevege seg utenfor rasteret, plasserer vi agenten på den posisjonen, speilet i kanten av brettet. Dette kan vi gjøre uavhengig for hver hovedakse. La oss si at x-koordinaten går fra 0 til n, og agenten ønsker å bevege seg fra posisjon $x_{old} = n$ til posisjon $x_{new,try} = n + 1$. Vi kan da speile x-koordinaten i posisjon $x_{boundary} = n + 0{,}5$, slik at x-koordinaten blir $x_{new} = x_{boundary} - max(0, x_{new,try} - x_{boundary}) = n{+}0{,}5{-}max(0, n{+}1{-}(n{+}0{,}5)) = n{+}0{,}5{-}max(0, 0{,}5) = n$. Et alternativ er å velge $x_{boundary} = n$.
Ved undergrensen $x = 0$ gjelder også: $x_{new} = x_{boundary} - min(0, x_{new,try} - x_{boundary})$.
Periodisk: Betrakt x-koordinaten (w.l.o.g.). Da gjelder: $x_{new} = x_{new,try} \% n$.
Absorberende: Agenten annihileres når den overskrider en av grensene.

Oppgåve 10 .... *(10%)*

Me simulerer ein veg og tel bilane som passerer ein vegeskil per tidseining.

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

For datasettet gjeld:

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

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

Solution: Vi finner:
$n = 4$
$\sum xy = 3 \cdot 102 + 6 \cdot 193 + 9 \cdot 292 + 12 \cdot 399 = 8880$
$\sum x^2 = 3^2 + 6^2 + 9^2 + 12^2 = 270$
$\sum y^2 = 102^2 + 193^2 + 292^2 + 399^2 = 292118$
$\sum x = 3 + 6 + 9 + 12 = 30$
$\sum y = 102 + 193 + 292 + 399 = 986$
$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \left(\sum x\right)^2} = \frac{4 \cdot 8880 - 30 \cdot 986}{4 \cdot 270 - 30^2} = 33$
$a = \bar{y} - b\bar{x} = \frac{\sum y}{n} - b\frac{\sum x}{n} = \frac{986}{4} - 33 \cdot \frac{30}{4} = -1{,}0$
Dermed blir den estimerte regresjonslinja: $\hat{y} = a + bx = 33x - 1$

(b) Rekn ut korrelasjonskoeffisienten mellom tid og talet på bilar.

Solution: $S_{xx} = \sum x^2 - n \cdot \bar{x}^2 = \sum x^2 - n \cdot \left(\frac{\sum x}{n}\right)^2 = 270 - 4 \cdot \left(\frac{30}{4}\right)^2 = 45$
$S_{yy} = \sum y^2 - n * \bar{y}^2 = \sum y^2 - n \cdot \left(\frac{\sum y}{n}\right)^2 = 292118 - 4 \cdot \left(\frac{986}{4}\right)^2 = 49069$
$S_{xy} = \sum xy - n\bar{x}\bar{y} = \sum xy - n \cdot \frac{\sum x}{n} \frac{\sum y}{n} = 8880 - 4 \cdot \frac{30}{4} \frac{986}{4} = 1485$
Utvalgskorrelasjonskoeffisienten er da: $r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{1485}{\sqrt{45 \cdot 49069}} = 0{,}9993$

Oppgåve 11 .... *(14%)*

Ein partikkel rører seg i eit éin-dimensjonalt raster. Sannsynsfordelinga for rørsla i kvart tidssteg er som fylgjer:

Rørsle	Sannsyn
I ro	30%
Eitt steg til venstre	25%
Eitt steg til høgre	25%
To steg til venstre	10%
To steg til høgre	10%

Lat X vera rørsla i løpet av t tidssteg, der negativ verdi viser rørsle mot venstre og positiv mot høgre.

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

> **Solution:** $\mu = E(X) = \sum x \cdot p(x) = 0 \cdot 0{,}30 + (-1) \cdot 0{,}25 + 1 \cdot 0{,}25 + (-2) \cdot 0{,}10 + 2 \cdot 0{,}10 = 0$

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

> **Solution:** Forflytningen etter t tidssteg er summen av t uavhengige og identisk fordelte variable, alle likt fordelt med X. Forventningsverdien for forflytning etter t tidssteg blir da $t \cdot \mu_X = t \cdot 0 = 0$.

(c) Rekn ut diffusjonskoeffisienten til partikkelen. Formelen er $E(X^2) = 2dDt$ der $d = 1$ er talet på dimensjonar, D er diffusjonskoeffisienten og t er tida.

> **Solution:** Vi betrakter partikkelen etter ett tidssteg. Forventningsverdien for kvadratisk forflytning er da:
> $E(X^2) = \sum(x^2 \cdot p(x)) = 0^2 \cdot 0{,}30 + (-1)^2 \cdot 0{,}25 + 1^2 \cdot 0{,}25 + (-2)^2 \cdot 0{,}10 + 2^2 \cdot 0{,}10 = 1{,}3$
> $E(X^2) = 2dDt$ gir:
> $1{,}3 = 2 \cdot 1 \cdot D \cdot 1$, derfor
> $D = \frac{1{,}3}{2} = 0{,}65$

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

> **Solution:** $\sigma_x^2 = E((X - \mu_x)^2) = E((X - 0)^2) = E(X^2)$
> N.B. Variansen i posisjonen $Z = X + p_0$ etter ett tidssteg (hvor p_0 er startposisjonen) er også lik $E(X^2)$. Siden $E(X) = \mu_x = 0$ gjelder
> $\sigma_z^2 = E(Z - \mu_z)^2) = E(((X + p_0) - E(X + p_0))^2) = E(((X + p_0) - E(X) - p_0))^2) = E(((X + p_0) - 0 - p_0))^2) = E(X^2)$

Partikkelen byrjar på tid $t = 0$ og posisjon $p_0 = 500$. Lat den stokastiske variabelen Y vera posisjonen til partikkelen etter $t = 1000$ tidssteg.

(e) Rekn ut forventingsverdien $\mu = E(Y)$.

> **Solution:** Vi ser: $Y = X + p_0 = X + 500$
> Derfor: $E(Y) = E(X + 500) = E(X) + 500 = 0 + 500 = 500$

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

> **Solution:** $\sigma_Y^2 = \sigma_{X+500}^2 = \sigma_X^2 = 1000 \cdot \sigma_{X_1}^2 = 1000 \cdot 1{,}3 = 1300$ hvor vi har brukt notasjonen X_1 for forflytningen etter ett tidssteg.

(g) Kva fordeling har Y? Grunngje svaret.

> **Solution:** Variabelen Y er ein sum av mange uavhengige variablar og me kan difor bruka sentralgrensesatsen, som seier at Y er tilnærma normalfordelt.