

Institutt for matematiske fag

Eksamensoppgave i **ST0103 Brukerkurs i statistikk**

Faglig kontakt under eksamen: Øyvind Bakke

Tlf: 73 59 81 26, 990 41 673

Eksamensdato: 19. desember 2015

Eksamenstid (fra–til): 9.00–13.00

Hjelpemiddelkode/Tillatte hjelpemidler: Gult A4-ark med egne håndskrevne notater, bestemt kalkulator (Casio fx-82ES Plus, Citizen SR-270X, Citizen SR-270X College eller HP 30s), *Tabeller og formler i statistikk* (Tapir forlag), *Matematisk formelsamling* (K. Rottmann)

Annen informasjon:

I vurderingen teller hvert av de ti bokstavpunktene likt.

Alle svar skal begrunnes (f.eks. ved at mellomregning tas med eller ved henvisning til teori eller eksempler fra pensum).

Målform/språk: bokmål

Antall sider: 2

Antall sider vedlegg: 0

Kontrollert av:

Dato

Sign

Oppgave 1

I forbindelse med undersøkelse av en mineralforekomst blir det boret n hull. For hvert hull er det sannsynlighet p for at mineralet blir funnet. Anta at antall hull, X , der mineralet blir funnet, er binomisk fordelt med parametre n og p .

- a) Hvilke forutsetninger må være oppfylt for at X skal være binomisk fordelt? Finn $P(3 \leq X \leq 4)$ når $n = 20$ og $p = 0,3$.
- b) Finn en tilnærmet verdi for $P(50 \leq X \leq 75)$ når $n = 200$ og $p = 0,3$.

Anta nå at p er ukjent. En ny serie på m hull blir boret. Antall hull, Y , med forekomst av mineralet i den nye serien antas å være uavhengig av X og binomisk fordelt med suksessannsynlighet p . To estimatorer for p blir foreslått, $\hat{p} = (X + Y)/(n + m)$ og $\tilde{p} = \frac{1}{2}(X/n + Y/m)$.

- c) Hvilken av de to estimatorene vil du foretrekke? Det er tilstrekkelig at du undersøker egenskapene til \hat{p} og \tilde{p} for $n = 20$ og $m = 10$.

Oppgave 2

Et bestemt mønster av nukleotider opptrer på en DNA-sekvens som en poissonprosess med intensitet (rate) $\lambda = 0,03$, slik at antall ganger mønsteret forekommer på en DNA-sekvens av lengde l (målt i antall tusen basepar) er poissonfordelt med parameter λl .

- a) Finn sannsynligheten for at mønsteret forekommer 2 eller flere ganger på en DNA-sekvens av lengde 100. Finn den betingede sannsynligheten for at mønsteret forekommer nøyaktig 2 ganger gitt at det forekommer 2 eller flere ganger.
- b) Begrunn kort hvorfor avstanden fra et mønster til neste sted mønsteret forekommer er eksponentielt fordelt med parameter λ (dvs. med forventningsverdi $1/\lambda$). Hva er sannsynligheten for at denne avstanden er større enn 50?

Vi måler avstanden fra første sted mønsteret forekommer til andre sted, avstanden fra andre til tredje sted, og avstanden fra tredje til fjerde sted mønsteret forekommer. Fra egenskapene til en poissonprosess følger det at disse tre avstandene er uavhengige (du skal ikke vise dette).

- c) Finn sannsynligheten for at den minste av de tre avstandene er større enn $1/\lambda$.

Oppgave 3

Konsentrasjonen (målt i mM) av et stoff som dannes i en reaksjon er normalfordelt med forventningsverdi 9,0 og standardavvik 2,0. Reaksjonen gjennomføres to ganger. Konsentrasjonene blir X første gang og Y andre gang, og X og Y er uavhengige.

- a) Finn $P(X > 10,0)$ og $P\left(\frac{1}{2}(X + Y) > 10,0\right)$.

En analysemetode er avhengig av at konsentrasjonen etter første reaksjon er større enn 10,0, og av at konsentrasjonen etter andre reaksjon er større enn konsentrasjonen etter første reaksjon.

- b) Finn sannsynligheten for dette – det vil si $P(10,0 < X < Y)$. (Vink: Det kan være til hjelp å skissere hendelsen i xy -planet.)

Oppgave 4

John James Strutt (Lord Rayleigh) publiserte i 1894 *On an anomaly encountered in determinations of the density of nitrogen gas* (*Proc. R. Soc. Lond.* 1894 **55**, s. 340–344). I artikkelen beskriver han hvordan han i en serie på 8 forsøk fra november 1893 til januar 1894 produserte nitrogengass fra nitrogenholdige forbindelser og veide innholdet av denne gassen i en kolbe med kjent volum. De 8 massene ble målt til 2,30143, 2,29890, 2,29816, 2,30182, 2,29869, 2,29940, 2,29849 og 2,29889 g.

Anta at målingene er uavhengige og kommer fra samme normalfordeling. Det oppgis at gjennomsnittet av de 8 målingene er 2,29947 g, og at utvalgsstandardavviket er 0,00138 g.

- a) Finn et 95 %-konfidensintervall for standardavviket i normalfordelingen.

I en annen serie på 7 forsøk fra desember 1893 til februar 1894 veide Strutt nitrogengass utvunnet fra luft (i motsetning til i de 8 nevnte forsøkene, der gassen ble produsert fra nitrogenholdige forbindelser). Massene ble målt til 2,31017, 2,30986, 2,31010, 2,31001, 2,31024, 2,31010 og 2,31028 g.

Anta at også disse målingene er uavhengige (også fra forrige forsøksserie) og kommer fra en normalfordeling. Det oppgis at gjennomsnittet av de 7 målingene er 2,31011 g, og at utvalgsstandardavviket er 0,00014 g.

- b) Test nullhypotesen som sier at forventningsverdien for målt masse ved de to metodene er like, mot alternativet som sier at forventningsverdiene er ulike. Bruk signifikansnivå 0,001. Hvilke forutsetninger må gjelde for at testmetoden skal være riktig? Diskuter kort om forutsetningene ser ut til å holde.