Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:
Odd Kolbjørnsen 73 59 35 20

SIF5060/SIF5062/SIF5505/SIF5506 Statistikk

Onsdag 4. august 2001
Tid: 09:00–14:00

Hjelpemidler: Godkjent lommekalkulator med tomt minne.
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.
Ordliste utdelt på forelesning.

Oppgave 1

La A og B være to hendelser hvor $\mathrm{P}(A \cap B) > 0$. La hendelsene C og D være gitt ved at $C = A \cap B$ og $D = A \cap B'$, der B' er komplementærhendelsen til B.

Tegn hendelsene A og B inn i et venndiagram og skraver i dette venndiagrammet hendelsene C og D.

Er hendelsene C og D disjunkte? (Begrunn svaret)

Er hendelsene C og D uavhengige? (Begrunn svaret)

Oppgave 2

La X og Y være to uavhengige normalfordelte stokastiske variable, der $\mathrm{E}(X) = \mathrm{E}(Y) = 1$, $\mathrm{Var}(X) = 1$ og $\mathrm{Var}(Y) = 4$.

Bestem følgende sannsynligheter

$$\mathrm{P}(X \leq 2) \quad , \quad \mathrm{P}(X \leq 2 \cap Y \leq 1) \quad \text{og} \quad \mathrm{P}(X + 2Y > 2).$$

Oppgave 3 Politiske meningsmålinger

Det er i dag svært vanlig at man utfører meningsmålinger for å skaffe seg informasjon om de politiske partiers velgeroppslutning. Dette utføres ved at et utvalg av de stemmeberettigede blir spurt hvilket parti de ville ha stemt på dersom det hadde vært valg den dagen. I denne oppgaven skal vi regne litt på denne situasjonen og vi skal fokusere på oppslutningen til ett bestemt parti, som vi benevner P. For å forenkle situasjonen noe skal vi se bort fra muligheten for at noen ikke vil stemme eller at noen ikke vil svare eller svarer usant når de blir spurt om sitt partivalg.

La N betegne antall stemmeberettigede og la p være andelen av disse N som vil stemme på partiet P. Anta at n personer blir spurt i meningsmålingen og la X betegne antall av disse n som svarer at de ville ha stemt på P. Anta til slutt at de n personene som blir spurt er trukket tilfeldig uten tilbakelegging fra de N stemmeberettigede.

a) Forklar hvorfor X er hypergeometrisk fordelt.

Forklar hvorfor X i denne situasjonen er tilnærmet binomisk fordelt med n forsøk og sannsynlighet p.

Som estimator for p benyttes $\widehat{p} = X/n$.

b) Benytt at X er tilnærmet binomisk fordelt til å bestemme forventning og varians for $\widehat{p}$.

Hvis partiet P har en oppslutning på $p = 0.079$ blant de stemmeberettigede, hvor mange personer, n, må man minst spørre for at standardavviket til $\widehat{p}$ ikke skal overstige 0.010.

Som kjent kan en binomisk fordeling tilnærmes med en normalfordeling dersom np og $n(1-p)$ begge er tilstrekkelig store. I resten av oppgaven kan du anta at dette er oppfylt slik at

$$\frac{X - np}{\sqrt{np(1-p)}}$$

er tilnærmet standard normalfordelt.

La $p_0 = 0.079$ betegne oppslutningen til partiet P ved forrige valg. En ønsker nå å benytte resultatet av meningsmålingen til å undersøke om oppslutningen om P har gått ned siden den gang.

c) Formuler dette som et hypotesetestingsproblem. Velg testobservator og lag en hypotesetest med signifikansnivå $\alpha = 0.05$.

Hva blir konklusjonen på testen dersom man har spurt 1000 personer og 52 av disse svarte at de ville ha stemt på partiet P.

Oppgave 4 Absorpsjon av lys

Når lys av vilkårlig retning treffer en kule så absorberes en viss andel av lyset. Vi betegner denne andelen med X. Det kan vises at det er rimelig å oppfatte X som en stokastisk variabel med sannsynlighetstetthet

$$f(x) = \begin{cases} \frac{1}{\theta} & \text{for } 0 \leq x \leq \theta, \\ 0 & \text{ellers,} \end{cases}$$

der θ er en parameter som avhenger av kulas radius og type overflate.

a) Bestem kumulativ fordelingsfunksjon for X, F(x).

Skisser $f(x)$ og $F(x)$.

For $\theta = 2.0$, finn sannsynligheten $\mathrm{P}(X \leq 0.4)$.

Anta at θ for en bestemt kule er ukjent. For å skaffe informasjon om θ for denne kula gjøres n målinger $X_1, X_2, \ldots, X_n$, der $X_1, X_2, \ldots, X_n$ er uavhengige stokastiske variable, alle med sannsynlighetsfordeling som gitt over.

b) Vis at sannsynlighetsmaksimeringsestimatoren (SME), $\widehat{\theta}$, for θ er

$$\widehat{\theta} = \max(X_1, X_2, \ldots, X_n).$$

Videre i oppgaven innfører vi notasjonen $Y = \max(X_1, X_2, \ldots, X_n)$ slik at $\widehat{\theta} = Y$.

c) Bestem kumulativ fordelingsfunksjon for Y, $G(y) = \mathrm{P}(Y \leq y)$.

Benytt så dette til å vise at sannsynlighetstettheten for Y er

$$g(y) = \begin{cases} \frac{ny^{n-1}}{\theta^n} & \text{for } 0 \leq y \leq \theta, \\ 0 & \text{ellers.} \end{cases}$$

d) Benytt sannsynlighetstettheten for Y gitt i punkt **c**) til å vise at estimatoren $\widehat{\theta} = Y$ **ikke** er forventningsrett.

Bestem k slik at $\widetilde{\theta} = kY$ blir en forventningsrett estimator for θ.

e) Finn et 95% konfidensintervall for θ basert på $\widetilde{\theta}$. (Hint: Bestem fordelingen til $Z = \widetilde{\theta}/(k\theta)$ og ta utgangspunkt i Z for å bestemme konfidensintervallet.)

Beregn også intervallet numerisk når $n = 10$ og den største målte verdien er 0.46.