UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i:	ST101 — Innføring i statistikk og sannsynlighetsregning.
Eksamensdag:	Fredag 21. november 1997.
Tid for eksamen:	09.00 – 15.00.
Oppgavesettet er på 4 sider.	
Vedlegg:	Opptrykk av tabell over kumulative normalfordelinger.
Tillatte hjelpemidler:	Formelsamling for ST 101, Rottmanns "Mathematische Formelsamlung", kalkulator.

Kontroller at oppgavesettet er komplett
før du begynner å besvare spørsmålene.

Oppgave 1.

Dersom en stokastisk variabel X følger sannsynlighetstettheten

$$f(x) = ce^{-|x|}, \qquad -\infty < x < \infty$$

sies X å være Laplace-fordelt eller dobbeleksponentielt fordelt.

a) Bestem konstanten c .

b) Beregn kumulativ fordelingsfunksjon

$$F(x) = P(X \leq x)\,.$$

c) Hva blir $P(-1 < x \leq 2)$?

d) Beregn forventning og varians. [Det oppgis at $\int_0^\infty x^2 e^{-x} dx = 2$.]

(Fortsettes side 2.)

Oppgave 2.

Bremselengden Y (målt i meter) for en bil som kjører i en hastighet $10x$ km/t antas å være normalfordelt med forventning θx^2 og standardavvik σx^2. (For en bil som kjører i 40 km/t vil dermed bremselengden være normalfordelt med forventning 16θ og standardavvik 16σ.) Parametrene θ og σ vil avhenge av en rekke forhold, f.eks. bremsenes egenskaper, dekktype, veidekke, vær- og føreforhold, målenøyaktighet. Anta i punktene a) og b) at $\theta = 0.5$ og $\sigma = 0.1$.

a) Finn sannsynligheten for at bremselengden for en bil som kjører i 40 km/t er kortere enn 6 meter. Hva er sannsynligheten for at bremselengden ligger mellom 7 og 9 meter?

b) Anta at to biler, bil I og bil II, kjører mot hverandre i 40 km/t langs den samme rette linje. Ved punktene A og B, henholdsvis, begynner de å bremse. Avstanden mellom punktene A og B antas å være 18 meter (se figur).

Hva er sannsynligheten for at de to bilene treffer hverandre? (Anta at de også under bremsing følger den samme rette linjen.)

I praksis vil θ ikke kunne avledes fra fysikk, og det vil være aktuelt å bestemme størrelsen empirisk under ulike forhold. Anta at vi har målt bremselengder $Y_1, \ldots, Y_n$ for en bil kjørt i fart $10x$ km/t. De n forsøkene ansees å være uavhengige, og fordelingen den samme som for Y i innledningen. En mulig estimator er

$$\widehat{\theta} = \frac{1}{nx^2}(Y_1 + \cdots + Y_n)$$

c) Vis at $\widehat{\theta}$ er forventningsrett.

d) Beregn $\mathrm{var}(\widehat{\theta})$ og vurder hvilken betydning den valgte hastigheten $10x$ km/t har for nøyaktigheten av estimatoren.

e) Hvilken sannsynlighetsfordeling følger $\widehat{\theta}$ under antakelsene i denne oppgaven?

Anta at vi dropper forutsetningen om at $Y_1, \ldots, Y_n$ er normalfordelt.

f) Hva kan du si om sannsynlighetsfordelingen til $\widehat{\theta}$ da?

(Fortsettes side 3.)

Oppgave 3.

Det er vanlig å bruke poisson-fordelingen til å beskrive antall tilfeller av en bestemt sykdom i en gitt befolkning. La X være antall registrerte sykdommer i løpet av ett år og anta at X følger punktsannsynligheten

$$P(X=x) = \frac{(a\lambda)^x}{x!}e^{-a\lambda}, \qquad x = 0, 1, 2, \ldots$$

der a er et *kjent* tall (befolkningens størrelse) og λ en *ukjent* sykdomsintensitet. En vanlig estimator for λ er

$$\widehat{\lambda} = \frac{X}{a}$$

a) Begrunn at $\widehat{\lambda}$ er en forventningsrett estimator for λ.

b) Vis at $\text{var}(\widehat{\lambda}) = \lambda/a^2$.

c) Gjør kort rede for at estimeringsfeilen $\widehat{\lambda} - \lambda$ vil bli forsvinnende liten når befolkningsstørrelsen $a \to \infty$.

Anta at vi også har observasjoner Y fra en annen befolkning (størrelse b) i et annet område slik at

$$P(Y=y) = \frac{(\lambda b)^y}{y!}e^{-b\lambda}, \qquad y = 0, 1, \ldots\ ,$$

der b er kjent.

d) Foreslå en forventningsrett estimator for λ basert på både X og Y.

e) Vis at estimatoren i d) er mer nøyaktig enn den første.

Oppgave 4.

Vi skal i denne oppgaven ta utgangspunkt i et eksperiment der det er tre mulige utfall A, B og C. Anta at

$$P(A) = p_a\ , \quad P(B) = p_b\ , \quad P(C) = p_c\ ,$$

der $p_a + p_b + p_c = 1$ og at A, B og C aldri inntreffer samtidig. Et eksempel på en slik situasjon er kast med en terning der $A = \{1\}$, $B = \{2, 4, 6\}$, $C = \{3, 5\}$.

a) Vis at

$$P(A \mid A \cup B) = \frac{p_a}{1-p_c}\ , \qquad P(B \mid A \cup B) = \frac{p_b}{1-p_c}$$

der $A \cup B$ markerer at A eller B (men ikke C) har forekommet.

(Fortsettes side 4.)

Anta at vi gjennomfører n *uavhengige* eksperimenter som beskrevet ovenfor. La X være antall ganger A forekommer, Y antall ganger B forekommer og Z antall ganger C forekommer. Da vil automatisk $X + Y + Z = n$.

b) Gjør kort rede for at Z er binomisk fordelt. Skriv ned punktsannsynligheten.

c) Vis at

$$E\{Z(n-Z)\} = n(n-1)p_c(1-p_c)$$

Anta at vi får oppgitt at C inntraff nøyaktig z ganger i forsøksserien slik at $Z = z$.

d) Bruk a) til å begrunne (ingen regning) at den betingete punktsannsynligheten for X gitt $Z = z$ må bli

$$P(X = x \mid z) = \binom{n-z}{x} \left(\frac{p_a}{1-p_c}\right)^x \left(1 - \frac{p_a}{1-p_c}\right)^{n-z-x}, \quad x = 0, 1, \ldots, n-z$$

e) Hva blir den betingete forventning $E(X \mid z)$ og den betingete varians $\mathrm{var}(X \mid z)$?

f) Vis at

$$E(XZ) = \sum_{z=0}^{n} z P(Z = z) \sum_{x=0}^{n-z} x P(X = x \mid z)$$

g) Bruk c), e) og f) til å vise at

$$\mathrm{cov}(X, Z) = -n\, p_a p_c\,.$$

SLUTT