Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:

Odd Kolbjørnsen	73 59 35 27
Håkon Tjelmeland	73 59 35 38
Arvid Næss	73 59 70 53

SIF5060/SIF5505 Statistikk

Torsdag 29.november 2001

Tid: 09:00–14:00

Hjelpemidler: Bestemt enkel kalkulator
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.

Sensuren faller: 13. januar 2002.

Oppgave 1 Lottotipping

I denne oppgaven skal vi analysere to forskjellige aspekter ved lottotipping.

I lotto spiller en deltager en enkeltrekke ved å velge 7 av 34 tall. Det er også tillatt å spille system. Når en deltager spiller system velger deltageren ut m av 34 tall, hvor m betegner antall tall i systemet. Når en deltager spiller et system av størrelse m, vil antall enkeltrekker som spilles være lik antall mulige kombinasjoner av 7 tall blant de m tallene i systemet.

a) Hvor mange enkeltrekker spilles i et system som inneholder 8 tall?

Hvor mange enkeltrekker spilles i et system som inneholder m tall?

Når en leverer inn en lottokupong, må en betale kr 3,- per enkeltrekke som spilles. Hvor mye koster det å levere inn et system med 12 tall? (Dette er det største systemet som er tillatt av Norsk Tipping.)

Det finnes totalt 5 379 616 mulige enkeltrekker i lotto. Sannsynligheten for at en tilfeldig valgt rekke skal sammenfalle med vinnerrekka er dermed $p = (5\,379\,616)^{-1} \approx 1.86 \cdot 10^{-7}$. I trekningen den 17. november, ble det spilt $n = 21\,481\,335$ rekker. I denne oppgaven skal vi regne som om alle de n rekkene som ble spilt, var trukket tilfeldig og uavhengig av hverandre blant mengden av mulige enkeltrekker. La X betegne antall enkeltrekker som sammenfaller med vinnerrekka, dvs. antall vinnere av toppgevinsten. Under antagelsene over kan X regnes som binomisk fordelt, $b(x; n, p)$.

b) Under hvilke betingelser kan en binomisk fordeling tilnærmes med en poissonfordeling?

Beregn poissontilnærmelsen for fordelingen til X.

Beregn punktsannsynlighetene for hendelsene $X = 0$ og $X = 1$, både for den eksakte binomiske fordelingen og for poissontilnærmelsen. Er poissontilnærmelsen rimelig i dette tilfellet ?

Oppgave 2 Første mål vinner ?

Mange mennesker har i dag en lidenskapelig interesse for elitefotball og (såkalte) eksperter har ofte klare meninger om spillet. I denne oppgaven skal vi konsentrere oss om kamper mellom to spesielle lag, som vi benevner henholdsvis R og L. En ekspertkommentator på fjernsyn kom med følgende påstand om kamper mellom R og L: "som oftest vil det laget som får det første målet også vinne kampen". I denne oppgaven skal vi regne litt med utgangspunkt i denne påstanden.

For en fotballkamp mellom lagene R og L, la følgende hendelser være definert:
R: Lag R vinner kampen.
F: Lag R får mål før lag L.
I: Kampen ender målløs, dvs. 0-0.

a) I dette punktet skal du anta at $\mathrm{P}(R) = 0.4$, $\mathrm{P}(F) = 0.5$, $\mathrm{P}(R \cap F) = 0.3$ og $\mathrm{P}(I) = 0.05$.

Tegn hendelsene R, F og I i et venndiagram.

Bestem sannsynligheten for at lag R vinner gitt at lag R får mål før lag L, dvs. $\mathrm{P}(R \mid F)$.

Bestem sannsynligheten for at lag R vinner kampen gitt at kampen ikke ender målløs, dvs. $\mathrm{P}(R \mid I')$, hvor I' betegner komplementærhendelsen til I.

Vi skal videre kun analysere de kampene mellom R og L som ikke endte målløse. La p benevne sannsynligheten for at det laget som får det første målet også vinner kampen. Vi forutsetter at denne sannsynligheten ikke avhenger av om det er R eller L som har hjemmekamp. Vi skal estimere p ut fra resultatene i de siste n seriekampene mellom R og L (kun kamper med minst ett mål blir tatt med). La X benevne antall av de n kampene hvor laget som fikk det første målet også vant kampen. Vi antar at X er binomisk fordelt med parametre n og p og bruker estimatoren

$$\widehat{p} = \frac{X}{n}.$$

b) Hva er de generelle forutsetninger for en binomisk fordeling? Er det ut fra dette rimelig å anta at X er binomisk fordelt? (begrunn svaret)

Redegjør kort for det generelle resultatet i sentralgrenseteoremet.

Vis hvordan sentralgrenseteoremet gir at

$$\frac{\widehat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

er tilnærmet standard-normalfordelt, dersom n er stor.

Da ekspertkommentatoren som ble nevnt i begynnelsen av oppgaven ble bedt om å konkretisere sin påstand om at i kamper mellom R og L er det som oftest laget som får det første målet som vinner kampen, sa han at sannsynligheten p er minst lik 0.80. Vi ønsker nå å undersøke om vår observerte verdi for X gir grunnlag for si at ekspertens uttalelse er feil.

c) Formuler dette som et hypotesetestingsproblem. Velg signifikansnivå 5% og bestem en regel for når H_0 skal forkastes.

Hva blir konklusjonen på testen når $n = 24$ og $x = 17$?(Dette er resultater fra kamper mellom Rosenborg og Lillestrøm i perioden 1990-2001. Ingen av disse kampene endte forøvrig målløse.)

d) Anta at forkastningsregelen fra **c)** benyttes, men at p i virkeligheten er 0.7. Hvor mange kampobservasjoner må man da ha for at sannsynligheten for å oppdage at ekspertens uttalelse er feil skal være minst 0.9.

Oppgave 3 Parkeringsboten

En bileier som ikke betaler parkeringsavgift, blir ilagt en bot pålydende kr 300,- dersom en parkeringsvakt oppdager forseelsen. Dersom flere parkeringsvakter oppdager den samme forseelsen er botens størrelse fortsatt den samme. I denne oppgaven skal vi analysere *statistiske* aspekter ved denne situasjonen.

Vi antar at parkeringsvaktene ankommer en bestemt parkeringsplass i følge en poissonprosess med parameter λ. La T betegne tiden fra bilen parkeres til den første parkeringsvakten ankommer. Under antagelsen om en poissonprosess vil som kjent T være eksponensialfordelt, slik at sannsynlighetstettheten til T er

$$f(t;\lambda) = \begin{cases} \lambda \exp\{-\lambda t\}, & \text{hvis } t > 0, \\ 0 & \text{ellers.} \end{cases} \tag{1}$$

Du kan i denne oppgaven uten bevis benytte at dersom $T_1, T_2, ..., T_n$ er uavhengige og eksponensialfordelte, så er

$$\mathrm{E}\left\{\frac{1}{\sum_{i=1}^{n} T_i}\right\} = \frac{\lambda}{n-1}.$$

a) Anta i dette punktet at $\lambda = 1/5$ pr time.

Katrine glemmer å betale parkeringsavgift og hun er borte fra bilen i 2 timer. Hva er sannsynligheten for at hun har fått bot når hun kommer tilbake ?

Anta nå at Katrine glemmer å betale parkeringsavgift og hun er borte fra bilen i t timer. Dersom en parkeringsvakt oppdager at parkeringsavgiften ikke er betalt, ilegges Katrine en bot pålydende kr 300,-. Vis at forventet kostnad for Katrines parkeringsopphold er kr

$$300(1 - \exp\{-\lambda t\}).$$

Ordinær parkeringsavgift er på kr 30,- pr time. Vil det lønne seg for Katrine å ikke betale parkeringsavgift dersom hun står parkert i 8 timer ?

Eierne av budbilfirmaet Snusk og Snask får høre at det kan lønne seg å ikke betale parkeringsavgift. De ser dermed en mulighet til å spare penger. For å vurdere ulike strategier, ønsker de å estimere λ. De setter derfor en nyansatt til å holde vakt over en parkeringsplass. Den nyansatte rapporterer tidsintervallene, $T_1, T_2, ..., T_n$ mellom hver gang en parkeringsvakt ankommer parkeringsplassen. Under antagelsen om en poissonprosess, vil disse tidene være uavhengige og eksponensialfordelt.

Det blir observert følgende $n = 20$ tider

0.56	2.79	2.62	5.95	0.92	3.23	0.04	3.27	5.13	1.82
1.79	1.01	3.01	2.27	0.70	0.05	3.84	0.35	0.81	2.35

Det oppgis at $\sum_{i=1}^{20} t_i = 42.51$.

b) Bestem sannsynlighetsmaksimeringsestimatoren (SME) for λ.

Er estimatoren forventningsrett? Hvis ikke foreslå en ny forventningsrett estimator ved å ta utgangspunkt i sannsynlighetsmaksimeringsestimatoren.

For den forventningsrette estimatoren, hva blir estimatet når dataene er som gitt over?

c) Vis ved bruk av momentgenererende funksjoner at $V = 2\lambda \sum_{i=1}^{n} T_i$ er χ^2-fordelt med $2n$ frihetsgrader.

d) Bruk resultatet i **c)** til å utlede et 95% konfidensintervall for λ.

I **a)** ble forventet kostnad ved et parkeringsopphold av lengde t beregnet til å være $\gamma = 300(1 - \exp\{-\lambda t\})$, dersom parkeringsavgift ikke var betalt. Bruk intervallet du lagde for λ til å lage et 95% konfidensintervall for γ.

Beregn konfidensintervallet for γ numerisk for $t = 8$ timer når dataene er som gitt over.