Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Nynorsk

Fagleg 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

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

Sensuren fell: 13. januar 2002.

Oppgave 1 Lottotipping

I denne oppgåva skal vi analysere to ulike aspekt ved lottotipping.

I lotto spelar ein deltakar ei enkeltrekkje ved å velgje 7 av 34 tal. Det er også tillete å spele system. Når ein deltakar spelar system velgjer deltakaren ut m av 34 tal, kor m angjev antal tal i systemet. Når ein deltakar spelar eit system av storleik m, vil antal enkeltrekkjer som vert spela vere lik antal moglege kombinasjonar av 7 tal blant dei m tala i systemet.

a) Kor mange enkeltrekkjer vert spela i eit system som inneheld 8 tal?

Kor mange enkeltrekkjer vert spela i eit system som inneheld m tal?

Når ein leverer inn ein lottokupong, må ein betale kr 3,- per enkeltrekkje som vert spela. Kor mykje kostar det å levere inn eit system med 12 tal? (Dette er det største systemet som er tillete av Norsk Tipping.)

Det finst totalt 5 379 616 moglege enkeltrekkjer i lotto. Sannsynet for at ei tilfeldig valt rekkje skal felle saman med vinnarrekkja er dermed $p = (5\,379\,616)^{-1} \approx 1.86 \cdot 10^{-7}$. I trekninga den 17. november, vart det spela $n = 21\,481\,335$ rekkjer. I denne oppgåva skal vi rekne som om alle dei n rekkjene som vart spela, var trekt tilfeldig og uavhengig av kvarandre blant mengda av moglege enkeltrekkjer. La X angje antal enkeltrekkjer som fell saman med vinnarrekkja, dvs. antal vinnarar av toppgevinsten. Under føresetnadane over kan ein rekne X som binomisk fordelt, $b(x; n, p)$.

b) Under kva vilkår kan ei binomisk fordeling verte tilnærma med ei poissonfordeling?

Berekn poissontilnærminga for fordelinga til X.

Berekn punktsannsyna for hendingane $X = 0$ og $X = 1$, både for den eksakte binomiske fordelinga og for poissontilnærminga. Er poissontilnærminga rimeleg i dette tilfellet ?

Oppgave 2 Første mål vinn ?

Mange menneske har i dag ei lidenskapeleg interesse for elitefotball og (såkalla) ekspertar har ofte klare meninger om spelet. I denne oppgåva skal vi konsentrere oss om kampar mellom to spesielle lag, som vi namngjev henhaldsvis R og L. Ein ekspertkommentator på fjernsyn kom med følgjande påstand om kampar mellom R og L: "som oftast vil det laget som får det første målet også vinne kampen". I denne oppgåva skal vi rekne litt med utgangspunkt i denne påstanden.

For ein fotballkamp mellom laga R og L, la følgjande hendingar vere definert:

R: Lag R vinn kampen.
F: Lag R får mål før lag L.
I: Kampen ender mållaus, dvs. 0-0.

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

Teikn hendingane R, F og I i eit venndiagram.

Bestem sannsynet for at lag R vinn gjeve at lag R får mål før lag L, dvs. $\mathrm{P}(R \mid F)$.

Bestem sannsynet for at lag R vinn kampen gjeve at kampen ikkje endar mållaus, dvs. $\mathrm{P}(R \mid I')$, kor I' angjev komplementærhendinga til I.

Vi skal vidare kun analysere dei kampane mellom R og L som ikkje enda mållause. La p vere sannsynet for at det laget som får det første målet også vinn kampen. Vi føreset at dette sannsynet ikkje avhenger av om det er R eller L som har heimekamp. Vi skal estimere p ut frå resultata i dei siste n seriekampane mellom R og L (kun kampar med minst eitt mål blir tekne med). La X angje antal av dei n kampane kor laget som fikk det første målet også vann kampen. Vi føreset at X er binomisk fordelt med parametrane n og p og bruker estimatoren

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

b) Kva er dei generelle føresetnadane for ei binomisk fordeling? Er det ut frå dette rimeleg å føresetje at X er binomisk fordelt? (grunngje svaret)

Grei kort ut om det generelle resultatet i sentralgrenseteoremet.

Vis korleis sentralgrenseteoremet gjev at

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

er tilnærma standard-normalfordelt, dersom n er stor.

Då ekspertkommentatoren som vart nevnt i byrjinga av oppgåva vart beden om å konkretisere påstanden sin om at i kampar mellom R og L er det som oftast laget som får det første målet som vinn kampen, sa han at sannsynet p er minst lik 0.80. Vi ønskjer no å undersøkje om vår observerte verdi for X gjev grunnlag for å seie at eksperten si utsegn er feil.

c) Formuler dette som eit hypotesetestingsproblem. Velg signifikansnivå 5% og bestem ein regel for når ein skal forkaste H_0.

Kva vert konklusjonen på testen når $n = 24$ og $x = 17$?(Dette er resultat frå kamper mellom Rosenborg og Lillestrøm i perioden 1990-2001. Ingen av desse kampane enda forøvrig mållause.)

d) Føreset at forkastningsregelen frå **c)** vert nytta, men at p i røynda er 0.7. Kor mange kampobservasjoner må ein då ha for at sannsynet for å oppdage at eksperten si utsegn er feil skal være minst 0.9?

Oppgave 3 Parkeringsbota

Ein bileigar som ikkje betalar parkeringsavgift, vert ilagt ei bot pålydande kr 300,- dersom ei parkeringsvakt oppdagar lovbrotet. Dersom fleire parkeringsvakter oppdagar det same lovbrotet er storleiken til bota framleis den same. I denne oppgåva skal vi analysere *statistiske* aspekt ved denne situasjonen.

Vi føreset at parkeringsvaktene kjem til ein bestemt parkeringsplass i følgje ein poissonprosess med parameter λ. La T angje tida frå bilen vert parkert til den første parkeringsvakta kjem. Under føresetnaden om ein poissonprosess vil som kjent T vere eksponensialfordelt, slik at sannsynstettleiken 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 oppgåva utan bevis nytte 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) Føreset i dette punktet at $\lambda = 1/5$ pr time.

Katrine gløymer å betale parkeringsavgift og ho er vekke frå bilen i 2 timar. Kva er sannsynet for at ho har fått bot når ho kjem attende?

Føreset no at Katrine gløymer å betale parkeringsavgift og ho er vekke frå bilen i t timar. Dersom ei parkeringsvakt oppdagar at parkeringsavgifta ikkje er betalt, vert Katrine ilagt ei bot pålydande kr 300,-. Vis at forventa kostnad for parkeringsoppholdet til Katrine er kr

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

Ordinær parkeringsavgift er på kr 30,- pr time. Vil det løne seg for Katrine å ikkje betale parkeringsavgift dersom ho står parkert i 8 timar ?

Eigarane av budbilfirmaet Snusk og Snask får høyre at det kan løne seg å ikkje betale parkeringsavgift. Dei ser dermed ei moglegheit til å spare penger. For å vurdere ulike strategiar, ønskjer dei å estimere λ. Dei set derfor ein nytilsett til å halde vakt over ein parkeringsplass. Den nytilsette rapporterer tidsintervalla, $T_1, T_2, ..., T_n$ mellom kvar gong ei parkeringsvakt kjem til parkeringsplassen. Under føresetnaden om ein poissonprosess, vil desse tidene vere uavhengige og eksponensialfordelte.

Det vert observert følgjande $n = 20$ 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 vert oppgjeve at $\sum_{i=1}^{20} t_i = 42.51$.

b) Bestem sannsynsmaksimeringsestimatoren (SME) for λ.

Er estimatoren forventningsrett? Viss ikkje føreslå ein ny forventningsrett estimator ved å ta utgangspunkt i sannsynsmaksimeringsestimatoren.

For den forvenningsrette estimatoren, kva blir estimatet når dataene er som gjeve over?

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

d) Bruk resultatet i **c)** til å utleie eit 95% konfidensintervall for λ.

I **a)** vart forventa kostnad ved eit parkeringsopphald av lengde t berekna til å vere $\gamma = 300(1 - \exp\{-\lambda t\})$, dersom parkeringsavgift ikkje var betalt. Bruk intervallet du laga for λ til å lage eit 95% konfidensintervall for γ.

Berekn konfidensintervallet for γ numerisk for $t = 8$ timar når dataene er som gjeve over.