Oppgave 1 Oppgave 10.35 fra læreboka.

Oppgave 2 Oppgave 10.45 fra læreboka.

Oppgave 3 Oppgave 10.55 fra læreboka.

Oppgave 4 Oppgave 10.65 fra læreboka.

Oppgave 5 Oppgave 10.71 fra læreboka.

Oppgave 6 Flypassasjerer — Eksamen juni 1994, oppgave 3 av 4

Ved innsjekking på rutefly veies normalt ikke passasjerer med håndbagasje. Vekten av de enkelte passasjerer med håndbagasje, X_1, $X_2, ...$, antas uavhengige og normalfordelte (Gaussisk fordelte), hvor forventningsverdi μ og varians σ^2 normalt er kjent fra internasjonale statistikker. En bestemt flyrute betjenes av en flytype med 16 passasjerplasser.

a) La $X_1, ..., X_{16}$ være vekten av 16 passasjerer med håndbagasje, og la $Y = X_1 + ... + X_{16}$. Anta at $\mu = 80$ (kg) og $\sigma^2 = 18^2$ (kg^2). Beregn følgende:

i) $P(X_1 > 90)$ ii) $E(Y)$ iii) $\text{Var}(Y)$ iv) $P(Y > 16 \cdot 90)$

Ville noen av de fire beregningene du har gjort vært korrekte uten å anta uavhengighet men med antagelse om normalfordeling? Angi i så fall hvilke.

Ville noen av de fire beregningene du har gjort vært korrekte uten å anta normalfordeling men med antagelse om uavhengighet? Angi i så fall hvilke.

I resten av oppgaven antas μ og σ^2 å være ukjente. For å skaffe informasjon om vekten av passasjerer med håndbagasje på denne flyruten, veies 20 tilfeldig trukne passasjerer med håndbagasje, med følgende resultat: $x_i, \;\; i = 1, 2, ..., 20$:

76, 97, 82, 73, 65, 74, 104, 55, 97, 69, 78, 86, 73, 90, 69, 86, 76, 66, 107, 68.

Disse gir: $\sum_{i=1}^{20} x_i = 1591, \;\; \sum_{i=1}^{20} x_i^2 = 130221$.

b) Beregn estimat for μ og σ^2 basert på dette datasettet. Finn et 90% konfidensintervall for μ.

c) Flyselskapet vil undersøke om $\sigma^2 < 15^2$ på denne flyruten. I så fall kan det eventuelt tillates økt vekt av frakt. Selskapet setter opp hypotesen

$$H_0: \quad \sigma^2 \geq 15^2 \qquad\qquad H_1: \quad \sigma^2 < 15^2.$$

Hva blir konklusjonen av en test med 1% signifikansnivå, ut fra datasettet ovenfor?

Oppgave 7 Første mål vinner? — Eksamen november 2001, oppgave 2 av 3

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. $\mathrm{P}(R \,|\, F)$.

Bestem sannsynligheten for at lag R vinner kampen gitt at kampen ikke ender målløs, dvs. $\mathrm{P}(R \,|\, I')$, 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 8 Atle, du lyver!

Utspørring av deltakere i humor- og realityprogram på TV har den siste tiden blitt svært populært. Spørsmålene som stilles kan vi dele inn i tre typer, og vi definerer følgende disjunkte hendelser:

A_1="det stilles et spørsmål av ikke-sensitiv natur, f.eks. hva heter du?",
A_2="det stilles et spørsmål av delvis sensitiv natur, f.eks. hvor gammel er du?",
A_3="det stilles et spørsmål av sensitiv natur, f.eks. har du vært utro?".

I tillegg definerer vi hendelsen:
L="deltakeren lyver"

Følgende sannsynligheter er oppgitt:
$P(A_1) = 0.1$, $P(A_2) = 0.4$, $P(A_3) = 0.5$, $P(L|A_1) = 0.05$, $P(L|A_2) = 0.2$, $P(L|A_3) = 0.6$.

a) Vis de fire hendelsene i et venndiagram.

Gitt at en deltaker blir spurt et spørsmål av type A_2, hva er sannsynligheten for at deltakeren ikke lyver, $P(L'|A_2)$?

Hva er sannsynligheten for at en tilfeldig valgt deltaker lyver, $P(L)$?

Et av spørsmålene som regnes å være av delvis sensitiv natur er "hvor gammel er du?". En gruppe på n personer ble stilt dette spørsmålet, deretter ble svarene registrert og sammenlignet med informasjon i offentlige registre. La X være en stokastisk variabel som angir antall personer som lyver blant n personer, og la p være sannsynligheten for at en person lyver.

b) Under hvilke antagelser vil X være binomisk fordelt?

Vi antar at $p = 0.2$ og at vi spør $n = 20$ personer. Hva er $P(X = 4)$?

Hva er $P[(X \leq 2) \cup (X > 5)]$?

Vi antar nå at p er ukjent. For å estimere p er det foreslått to estimatorer,

$$\hat{p} = \frac{X}{n} \quad \text{og} \quad p^* = \frac{X}{n-1}.$$

c) Finn forventningsverdi og varians til hver av estimatorene $\hat{p}$ og p^*.

Hvilke to egenskaper kjennetegner en god estimator?

Hvilken av estimatorene $\hat{p}$ og p^* vil du foretrekke? Begrunn svaret.

Vi forutsetter at n er stor og velger å benytte estimatoren $\hat{p}$ videre. Da er

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

tilnærmet standard normalfordelt.

Basert på tidligere undersøkelser antar vi at $p = 0.2$. Vi ønsker å undersøke om de innsamlede data gir oss grunn til å tro at p er større enn 0.2.

d) Formulér dette som en hypotesetest ved å definere nullhypotese og alternativ hypotese.

Bruk at Z er tilnærmet standard normalfordelt til å bestemme et forkastningsområde for nullhypotesen når vi velger signifikansnivå $\alpha = 0.01$.

Hva blir konklusjonen på testen når $n = 200$ personer ble spurt og $x = 55$ personer løy?

Vil p-verdien til testen være mindre eller større enn 0.01? Begrunn svaret. (Det kreves ikke at du regner ut p-verdien.)

Fasit

6. **a**) i) 0.289, ii) 1280, iii) 72^2, iv) 0.013 **b**) 79.55, 13.87^2, (74.2,84.9) **c**) Ikke forkast H_0

7. **a**) $\mathrm{P}(R|F) = 0.6$, $\mathrm{P}(R|I') = 0.421$ **c**) Forkaster ikke H_0 **d**) 156

8. **a**) 0.8, 0.385 **b**) 0.218, 0.402