ECON 2130 EKSAMEN 2008 VÅR

TALLSVAR

Oppgave 1

Vi har en kortstokk bestående av 6 kort. På 2 av disse står det skrevet JA på forsiden mens det står NEI på forsiden av de 4 andre kortene. Hvis man får se kortet med baksiden vendt mot seg, er det ikke mulig å se om det står JA eller NEI på forsiden.

A. Anta kortstokken blir stokket godt og lagt ned på bordet i en bunke med baksiden opp.

(i) Hva er sannsynligheten for at det øverste kortet i bunken er et JA-kort?

(ii) Hva er sannsynligheten for at de to øverste kortene i bunken begge er JA-kort?

<<**Svar:** (i): $1/3$ (ii): $1\Big/\binom{6}{2} = 1/15$ >>

B. La generelt A og B være to begivenheter. Da gjelder

(1) $$A = (A \cap B) \cup (A \cap \overline{B})$$

(i) Forklar relasjonen (1) ved et Venn-diagram.

(ii) Anta $P(A) = \frac{4}{5}$ og $P(B \mid A) = \frac{1}{4}$. Hva blir da $P(A \cap \overline{B})$?

(iii) Anta at sannsynlighetene oppgitt i (ii) fortsatt gjelder. Hvilken verdi må $P(B)$ ha hvis A og B er uavhengige begivenheter? Hva blir i så fall $P(A \cup B)$?

<<**Svar:** (ii): $P(A \cap \overline{B}) = P(A) - P(A)P(B \mid A) = \frac{4}{5} - \frac{4}{5} \cdot \frac{1}{4} = \frac{3}{5}$.

(iii): Uavhengighet hvis $P(B) = P(B \mid A) = \frac{1}{4}$. I så fall

$P(A \cup B) = P(A) + P(B) - P(A)P(B) = \frac{4}{5} + \frac{1}{4} - \frac{4}{5} \cdot \frac{1}{4} = \frac{3}{5} + \frac{1}{4} = \frac{17}{20} = 0{,}85$ >>

C. Vi vender tilbake til situasjonen i punkt **A**. La J_1 være begivenheten at det øverste kortet i bunken er et JA-kort, og J_2 begivenheten at det nest øverste kortet er et JA-kort. Forklar intuitivt, eller ved en beregning, at $P(J_2) = P(J_1) = \frac{1}{3}$.

<< **Svar:**

$P(J_2) = P(J_1)P(J_2 \mid J_1) + P(N_1)P(J_2 \mid N_1) = \frac{1}{3} \cdot \frac{1}{5} + \frac{2}{3} \cdot \frac{2}{5} = \frac{1}{15}(1+4) = \frac{1}{3}$ >>

D. Kortstokken stokkes grundig og legges som før på bordet i en bunke med baksiden opp. Ett og ett kort åpnes deretter fra toppen av bunken og nedover inntil første JA-kort dukker opp. La X være antall kort som må snus inntil første JA-kort dukker opp. Hvis kortstokken er grundig nok stokket, vil alle mulige rekkefølger av kortene i bunken ha samme sannsynlighet, og den stokastiske variabelen X vil ha en fordeling gitt i tabell 1:

Tabell 1 *Fordelingen for X*

x	1	2	3	4	5	6
$P(X=x)$	$\frac{5}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	0

(i) Vis at $P(X=3)=\frac{3}{15}$ som gitt i tabellen.

[**Hint:** Merk at begivenheten $(X=3)$ er ekvivalent med begivenheten $N_1 \cap N_2 \cap J_3$ - dvs. der de to øverste kortene er NEI-kort og det tredje kortet ovenfra et JA-kort (der begivenheten N_i betyr at *i*-te kort fra toppen er et NEI-kort).]

(ii) Vis at $E(X)=\frac{7}{3}$ og $\text{Var}(X)=\frac{14}{9}$.

<< **Svar:** (i) $P(X=3)=P(N_1)P(N_2 \mid N_1)P(J_3 \mid N_1 \cap N_2)=\frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}=\frac{1}{5}=\frac{3}{15}$

(ii)

$$E(X)=\frac{1\cdot 5+2\cdot 4+\cdots}{15}=\frac{35}{15}=\frac{7}{3}, \qquad E(X^2)=\frac{1^2\cdot 5+2^2\cdot 4+\cdots}{15}=\frac{105}{15}=7$$

og $\text{Var}(X)=7-\left(\frac{7}{3}\right)^2=\frac{63-49}{9}=\frac{14}{9}$ >>

E. Å stokke en kortstokk godt er vanskeligere enn folk flest tror. Jens mener at han vanligvis stokker kortene godt nok, men er villig til å gjennomføre følgende test. Eksperimentet beskrevet i punkt **D** gjentas $n = 40$ ganger. I hvert forsøk stokker Jens kortene så godt han mener er tilstrekkelig, legger deretter kortene på bordet i en bunke med baksiden opp, åpner ett og ett kort fra toppen og nedover i bunken inntil første JA-kort dukker opp og registrerer til slutt hvor mange kort som må snus før JA-kortet viser seg. La for forsøk i X_i betegne antall kort som må snus inntil det første JA-kortet kommer ($i=1,2,\ldots,40$).

Vi antar at $X_1, X_2, \ldots, X_{40}$ er uavhengige og identisk fordelt, med felles forventning, $E(X_i)=\mu$ og varians, $\text{Var}(X_i)=\sigma^2$. Hvis stokkingen er god nok, er den felles fordelingen kjent og gitt i tabell 1. Dette utgjør vår null-hypotese. Hvis stokkingen ikke er god nok, vil den felles fordelingen være ukjent med ukjent forventning og varians.

Gjennomsnittlig X-verdi for de 40 stokke-forsøkene til Jens ble 1,98. Sett opp og gjennomfør (basert på dette resultatet) en test med signifikansnivå 5% for hypotesen:

$$H_0: \quad \mu = \frac{7}{3} = 2{,}33 \quad \text{mot} \quad H_1: \quad \mu \neq 2{,}33$$

Beregn p-verdien for testen og formuler en konklusjon.

[**Hint:** Bruk sentralgrenseteoremet. Merk at både $\mu = \frac{7}{3}$ og $\sigma^2 = \frac{14}{9}$ er kjente under H_0, som i dette tilfellet er det eneste vi trenger å vite for å kunne beregne kritiske verdier og p-verdi.]

<<**Svar:** To-sidig test. Alternativer:

Testobservator	**Observert verdi**	**Nedre kritisk verdi**	**Øvre kritisk verdi**
$\bar{X}$	1,98	$2{,}33-(1{,}96)\sqrt{\frac{14}{9\cdot 40}}=1{,}95$	$2{,}33+(1{,}96)\sqrt{\frac{14}{9\cdot 40}}=2{,}72$
$Z=\frac{\bar{X}-7/3}{\sqrt{14/9}}\sqrt{40}$	$-1{,}77$	-1,96	1,96

P-verdi = $P_{H_0}(|Z|>|z_{obs}|)=P_{H_0}(|Z|>1{,}77)=2\cdot P(Z<-1{,}77)=2\cdot 0{,}0384=0{,}0768$

H_0 forkastes ikke på 5% >>

Oppgave 2

Et problem ved spørreundersøkelser er å få pålitelige svar på sensitive spørsmål. For å sikre anonymiteten benytter en forsker følgende intervju-metode. Anta spørsmålet er, "Har du kjøpt smuglersprit i år?". Forskeren viser respondenten (dvs. personen som blir intervjuet) en kortstokk av samme type som i oppgave 1 der JA står på 2 av kortene og NEI på 4. Respondenten får opplyst hvor mange av kortene som har JA og hvor mange som har NEI. Respondenten blir bedt om ikke å svare direkte på spørsmålet, men i stedet å trekke et kort og si om det som står på kortet stemmer eller ikke. For eksempel hvis respondenten faktisk ikke har kjøpt smuglersprit og trekker et kort med NEI på, så svarer vedkommende at det som står på kortet stemmer. Deretter blir kortet lagt tilbake i kortstokken uten at forskeren får se hva som står på det.

For enkelthets skyld vil vi i denne oppgaven se bort ifra de tilfellene der respondenten nekter å svare eller lyver, og antar i stedet at alle gir et av to mulige svar, "det stemmer", eller "det stemmer ikke", og at alle snakker sant.

A. Metoden brukes på et rent tilfeldig utvalg av voksne personer fra befolkningen. Gjør rede for at antall personer i utvalget som svarer, "det som står på kortet stemmer", kan antas å være binomisk fordelt.

B. La q være den relative andelen som ville ha svart, "det som står på kortet stemmer", dersom forskeren hadde intervjuet hele den voksne befolkningen. La p være andelen i befolkningen som har kjøpt smuglersprit. Det kan vises (jfr. punkt **D**) at sammenhengen mellom p og q er gitt ved

(2) $$q = \frac{2-p}{3}$$

La X 1være antall som svarer "det som står på kortet stemmer" i et tilfeldig utvalg på n voksne. Forklar hvorfor

$$\hat{p} = 2 - 3\frac{X}{n}$$

er en forventningsrett estimator for p. Vis at standardavviket (SD) for $\hat{p}$ er (uttrykt ved q):

$$\text{SD}(\hat{p}) = 3\sqrt{\frac{q(1-q)}{n}}$$

<< **Svar:** Opplagt. >>

C. (i) Konstruer et 95% konfidensintervall for p basert på $\hat{q} = \frac{X}{n}$.

[Hint: Ta utgangspunkt i et 95% konfidensintervall for q, som vi kan skrive (A, B), der A og B er stokastiske variable som oppfyller $P(A \le q \le B) = 0{,}95$ (tilnærmet). Dette intervallet kan du så overføre til et intervall for p ved å vise at begivenheten $(A \le q \le B)$ er ekvivalent med begivenheten $(2 - 3B \le p \le 2 - 3A)$.]

<<**Svar:** KI for q: $\hat{q} \pm (1{,}96)\sqrt{\frac{\hat{q}(1-\hat{q})}{n}}$

KI for p: $2 - 3\hat{q} \pm 3 \cdot (1{,}96)\sqrt{\frac{\hat{q}(1-\hat{q})}{n}}$ >>

(ii) Hvor mange observasjoner trengs for at lengden, $L = B - A$, på konfidensintervallet for q ikke skal overskride 0,06? Beregn også hvor mange observasjoner som trengs for at lengden på intervallet for p ikke skal overskride 0,06.

<<**Svar:** For q: Formel i boka der vi ikke vet noe om q: $n \geq \left(1{,}96/0{,}06\right)^2 = 1067$

For p: Her er lengden

$L = 3(B - A) = 0{,}06 \Rightarrow B - A = 0{,}02 \Rightarrow n \geq \left(1{,}96/0{,}02\right)^2 = 9604$ >>

D. Vis relasjonen (2) i punkt **B**.

[**Hint:** Betrakt intervju-situasjonen med en tilfeldig trukket respondent og innfør begivenhetene:
$K =$ "Respondenten har kjøpt smuglersprit i år"
$S =$ "Respondenten svarer at det som står på kortet stemmer"
$J =$ " Det står JA på kortet som respondenten trekker"
$N =$ " Det står NEI på kortet som respondenten trekker"
Gjør rede for og utnytt at $S = (K \cap J) \cup (\overline{K} \cap N)$, som er en disjunkt union.
Merk også at det som står på kortet, åpenbart er uavhengig av om respondenten har kjøpt smuglersprit eller ikke.]

<< **Svar:**

$$q = P(S) = P(K \cap J) + P(\overline{K} \cap N) = P(K)P(J) + P(\overline{K})P(N) = p\frac{1}{3} + (1-p)\frac{2}{3} = \frac{2-p}{3}$$

>>