ECON 2130 EKSAMEN 2009 VÅR

TALLSVAR

Oppgave 1

Fra 8 mannlige kandidater velges 6 ut ved loddtrekning (slik at alle har samme sjanse for å bli trukket ut) til å delta i et TV-program. Arne og Bjørn er to av kandidatene. La A være begivenheten at Arne trekkes ut, og B at Bjørn trekkes ut.

A. **(i)** Hvor mange ikke-ordnete utvalg på 6 fra 8 er det i alt?

(ii) Utvalget trekkes en og en (altså i alt 6 trekninger). Hva er sannsynligheten for at Arne blir trukket ut som nr. 1?

(iii) Etter at 5 er trukket ut viser det seg at Arne fortsatt ikke er kommet med i utvalget. Hva er da sannsynligheten for at han blir trukket ut som sistemann i den 6. trekningen?

<<**Svar: (i)** $\binom{8}{6} = \binom{8}{2} = 28$ **(ii)** $\frac{1}{8}$ **(iii)** $\frac{1}{3}$ >>

B. **(i)** Vis at $P(A) = \frac{3}{4}$ og $P(A \cap B) = \frac{15}{28}$. [**Hint.** Bruk for eksempel hypergeometriske sannsynligheter].

(ii) Finn $P(A \cup B)$. Er A, B disjunkte og/eller uavhengige begivenheter? Begrunn svaret.

<<**Svar: (i)** $P(A) = \frac{\binom{1}{1} \cdot \binom{7}{5}}{\binom{8}{6}} = \frac{21}{28} = \frac{3}{4}$, $P(A \cap B) = \frac{\binom{2}{2} \cdot \binom{6}{4}}{\binom{8}{6}} = \frac{15}{28}$

(ii) $P(A \cup B) = \frac{3}{4} + \frac{3}{4} - \frac{15}{28} = \frac{27}{28}$. Hverken disjunkte eller uavhengige. >>

C. Etter trekningen viste det seg at både Arne og Bjørn kom med. De 6 utvalgte blir nå stilt opp ved siden av hverandre i en rekke på en slik måte at alle mulige ordninger i rekkefølge er like sannsynlige. Finn sannsynligheten for at Arne og Bjørn havner ytterst på hver sin side i rekken (altså Arne ytterst til venstre og Bjørn ytterst til høyre eller omvendt).

<< **Svar:** $\dfrac{2\cdot 4!}{6!}=\dfrac{1}{15}$ >>

Oppgave 2

Innledning. TV-Norge kjører for tiden en program-serie som heter "Jakten på den 6. sans" der en gruppe antatt sensitive personer blir utsatt for forskjellige typer av oppgaver som utfordrer den 6. sansen. I hver episode blir en av deltakerne eliminert på grunnlag av hvor lite vedkommende får til, slik at det til slutt i siste program blir stående igjen en vinner.

En av utfordringene i programmet, som vi skal se på, er som følger: Seks personer av samme kjønn blir stilt opp ved siden av hverandre i en rekke. Forsøkspersonen, som vi kaller FP, får utdelt seks navneskilt som skal henges på personene i riktig rekkefølge slik at hver person i rekken får sitt eget navn hengt på seg. FP får så poeng etter hvor mange riktige tilordninger av navn til person FP klarer basert på sin intuisjon.

La X være antall riktige tilordninger av navn til person som FP klarer. Vi betrakter X som en stokastisk variabel med sannsynlighetsfordeling som generelt vil avhenge av hvor sensitiv FP er overfor utfordringer av denne typen. I det spesielle tilfellet at FP ikke er sensitiv i det hele tatt, slik at hver tilordning av navn til person er en ren gjetning, kan det vises (som du slipper) at sannsynlighetsfordelingen for X blir som angitt i tabell 1.

Tabell 1 Fordeling for X i tilfelle FP ikke er sensitiv.

x	0	1	2	3	4	5	6
$P(X=x)$	$\frac{53}{144}=0,368$	$\frac{11}{30}=0,367$	$\frac{3}{16}=0,188$	$\frac{1}{18}=0,056$	$\frac{1}{48}=0,021$	0	$\frac{1}{720}=0,001$

(Brøkene er eksakte sannsynligheter mens desimaltallene er avrundet til tre desimaler)

A. **(i)** Gjør rede for de to siste sannsynlighetene i tabell 1, det vil si

$$P(X=5)=0 \text{ og } P(X=6)=\frac{1}{720}$$

(ii) Vis at forventningen er $E(X)=1$ (eller nær 1 om du regner med de avrundete sannsynlighetene).

<<**Svar:** **(i)** $X = 5$ er umulig. Det er $6! = 720$ ordninger hvorav bare en gir $X = 6$.

(ii) $E(X)=1\cdot\dfrac{11}{30}+2\cdot\dfrac{3}{16}+\cdots=\dfrac{11}{30}+\dfrac{3}{8}+\dfrac{1}{6}+\dfrac{1}{12}+\dfrac{1}{120}=\dfrac{44+45+20+10+1}{120}=1$ >>

B. **(i)** Anta vi har en gruppe på $n=30$ antatt sensitive personer (forsøkspersoner) som alle blir utsatt for navneprøven beskrevet i innledningen. La X_i være antall riktige tilordninger av navn til person som forsøksperson *i* klarer. Vi antar $X_1,X_2,\ldots,X_n$ er stokastisk uavhengige med $E(X_i)=\mu_i$, der $\mu_i\geq 1$ for $i=1,2,\ldots,n$. X_i-ene er altså ikke nødvendigvis identisk fordelte. Innfør gjennomsnitts-forventningen $\mu = \frac{1}{n}(\mu_1 + \mu_2 + \cdots + \mu_n)$ som ny (ukjent) parameter. Forklar hvorfor $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ er en forventningsrett estimator for μ.

(ii) Anta nå at ingen av de 30 forsøkspersonene er sensitive overfor navneprøven slik at alle X_i-ene har samme sannsynlighetsfordeling gitt i tabell 1. I så fall er $\mu_i = E(X_i) = 1$ for alle *i*, og $\mu = 1$. I tillegg kan man regne ut (du behøver ikke å gjøre det) at $\text{var}(X_i) = 1$ for $i = 1, 2, \ldots, n$. Beregn et tilnærmet 95% spredningsintervall for $\hat{\mu} = \bar{X}$ i dette tilfellet (med $n = 30$). [**Hint.** Finn konstanter c_1, c_2 slik at $P(c_1 \le \bar{X} \le c_2) \approx 0.95$. Sentralgrenseteoremet kan benyttes.]

<<**Svar: (i)** Opplagt. **(ii)** $\bar{X} \sim N\left(1, \frac{1}{\sqrt{30}}\right)$ og $Z = (\bar{X} - 1)\sqrt{30} \sim N(0,1)$ gir

$P(c_1 \le \bar{X} \le c_2) = P\left((c_1 - 1)\sqrt{30} \le Z \le (c_2 - 1)\sqrt{30}\right) = 0{,}95$. Dvs

$(c_j - 1)\sqrt{30} = \pm z_{0{,}025} = \pm 1.96$, eller $c_j = 1 \pm \frac{1.96}{\sqrt{30}} = 0{,}64$ og $1{,}36$ >>

C. Opplegget i **B** kan brukes til å teste, basert på $\bar{X}$, om noen av forsøkspersonene har sensitive evner med hensyn på navneprøven. Vi lar da nullhypotesen, H_0, være at ingen av dem er sensitive slik at modellen for X_i-ene er som beskrevet i **B(ii)** med $\mu = 1$. Den alternative hypotesen, H_1, sier at $\mu > 1$, som innebærer at noen av μ_i-ene må være > 1. En naturlig test vil dermed være: Forkast H_0 hvis $\bar{X} > k$, der *k* er en passende kritisk verdi.

Bestem *k* slik at testen får signifikansnivå (tilnærmet) 5%.

<< **Svar:** $P_{H_0}(\bar{X} > k) = P(Z > (k-1)\sqrt{30}) = 0{,}05$ gir $k = 1 + \frac{1{,}645}{\sqrt{30}} = 1{,}30$ >>

Oppgave 3

VG hadde nylig en artikkel om økende mobbing i norske skoler:

Sitat fra VG 30. april 2009:

Alt tyder på at flere elever enn på lenge blir mobbet i norske skoler. VGs egen rundspørring blant 670 rektorer i Skole-Norge viser at hele 474, eller 71 prosent, har registrert mobbesaker ved sine skoler i løpet av det siste skoleåret. Det er en økning på 13 prosent fra en likelydende spørreundersøkelse i 2005.

La X være antall rektorer som svarer ja på spørsmålet om de har registrert mobbesaker på sine skoler det siste skoleåret i et utvalg på $n = 670$ skoler i Norge. La p betegne andelen av skoler i Norge som helhet der det har vært registrert mobbesaker det siste skoleåret. Som modell antar vi at X er binomisk fordelt, $X \sim \text{bin}(n, p)$, der p er en ukjent parameter.

A. VG sier ingenting om hvordan utvalget er tatt og hvorvidt det kan anses å være representativt. Drøft kort hva som bør være oppfylt for at utvalget skal være representativt og modellen realistisk.

<<**Svar:** Representativt hvis utvalget er et rent tilfeldig utvalg trukket fra populasjonen. (Lite om det i baka, men diskutert flere ganger på forelesningene.) Binomisk modell realistisk hvis populasjonen er stor. >>

B. Beregn et (tilnærmet) 95% konfidensintervall for p basert på de oppgitte data. Hva betyr konfidensgraden 0,95?

<< **Svar:** $\hat{p} = 0{,}71, \quad SE = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{670}} = 0{,}0175,$

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{670}} = 0{,}71 \pm 0{,}03 = [0{,}68,\ \ 0{,}74]$$

Konfidensgraden betyr ….. >>

C. VG fant i 2005 en tilsvarende prosent på 58% av skoler som hadde registrert mobbing.

(i) Har VG, under forutsetning at utvalget er representativt, dekning for sin påstand at (den registrerte) mobbingen i norske skoler har økt i forhold til tallet 58% fra 2005? Bruk signifikansnivå 1%. Formuler passende null-hypotese og alternativ hypotese. Sett opp et 1% forkastningskriterium, gjennomfør testen ut fra de oppgitte data og formuler en konklusjon.

(ii) Er p-verdien for testen din større eller mindre enn 0,01? Begrunn svaret.

(iii) Drøft kort om det at den registrerte mobbingen i norske skoler har økt fra 2005 til 2009 nødvendigvis er det samme som at mobbingen i norske skoler har økt fra 2005 til 2009.

<<**Svar: (i)** $H_0 : p \le 0{,}58$ vs $H_1 : p > 0{,}58$. Forkast H_0 hvis

$$Z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{670}}} > z_{0{,}01} = 2{,}326$$

Observert $z_o = 6{,}818$ som gir forkastning på 1%. Dette medfører **(ii)** at p-verdien er mindre enn 0,01. Merk pga avrunding: $\dfrac{X - np_0}{\sqrt{np_0(1-p_0)}} = 6{,}685$.

(iii) Øket media-fokusering på mobbing etter 2005 kan ha ført at flere mobbetilfeller kommer fram i lyset uten nødvendigvis at mobbingen har økt. Antakelig en mellomting…. >>