

Eksamensl sninger - HSTAT1101 - desember 2022

Oppgave 1

a)

En simultan sansynnligheter er n r to begivenheter skjer samtidig, $P(A \cap B)$. Et slikt eksempel er   trekke en bl  kule, hvor kule og bl  er de to begivenhetene.

b)

Den marginale sansynnligheten er summen av alle de simultane sansynnlighetene for begivenhetene av interesse inng r. Man kan alts  trekke b de gr nne kuber og gr nne kuler, som gir f lgende marginal sansynnlighet, $P(\text{gr ntelement}) = P(\text{gr nn} \cap \text{kule}) + P(\text{gr nn} \cap \text{kube}) = \frac{10}{34} + \frac{8}{34} = \frac{18}{34} = 52,9\%$.

c)

$$P(\text{kule}) = P(\text{gr nn} \cap \text{kule}) + P(\text{bl } \cap \text{kule}) = \frac{10}{34} + \frac{5}{34} = \frac{15}{34} = 44,1\%.$$

d)

Basert p  informasjonen i oppgaven og tidligere svar kan vi sette opp f lgende tabell:

Table 1:

	P(kule)	P(kube)	
P(gr�nn)	$\frac{10}{34}$	$\frac{8}{34}$	$\frac{18}{34}$
P(bl�)	$\frac{5}{34}$	$\frac{11}{34}$	
	$\frac{15}{34}$		

Ved bruk av sansynnlighetsreglene (komplementregelen) kan vi f lle ut resten av tabellen og dermed f  ut resten av de marginale sannsynnlighetene.

Table 2:

	P(kule)	P(kube)	
P(gr�nn)	$\frac{10}{34}$	$\frac{8}{34}$	$\frac{18}{34}$
P(bl�)	$\frac{5}{34}$	$\frac{11}{34}$	$\frac{16}{34}$
	$\frac{15}{34}$	$\frac{19}{34}$	1

e)

$$P(\text{kube}) \cdot P(\text{bl } \cap \text{kule}) = \frac{19}{34} \cdot \frac{5}{33} = 8,4\%$$

f)

Ettersom det er med tilbakelegging er sansynnligheten for   trekke et element likt hver gang. Vi vet at det er flere bl  elementer enn det er kuler. Alts  er det st rst sansynnlighet for   trekke bare bl  elementer.

g)

$$P(\text{rett}) = \frac{1}{8} = 12,5\%$$

h)

$$P(\text{rett}) \cdot P(\text{rett}) \cdot P(\text{rett}) = \left(\frac{1}{8}\right)^3 = 0,19\%$$

i)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\text{torett})}{P(\text{forsterrett})} = \frac{(\frac{1}{8})^2}{\frac{1}{8}} = \frac{1}{8}$$

j)

For at to begivenheter skal være statistisk uavhengig så må $P(A|B) = P(A)$. I vårt eksempel er det like stor sannsynlighet for å svare rett på en oppgave ($\frac{1}{8}$), som det er å svare rett på oppgave nr. 2, betinget på at du har svart rett på oppgave 1 ($\frac{1}{8}$). Altså, begivenhetene er uavhengige.

Oppgave 2

a)

Vi angir gruppen med nytt legemiddel som Y, og de med standardbehandling X. Hypotesene blir derfor:

$$H_0 : \mu_x = \mu_y$$

$$H_1 : \mu_x \neq \mu_y$$

Alternativhypotesen kunne også være en ensided test, hvor $H_1 : \mu_x > \mu_y$

b)

- N_x - størrelsen på X-populasjonen
- N_y - størrelsen på Y-populasjonen
- σ_x^2 - variansen i X
- σ_y^2 - variansen i Y

c)

- $H_0 : \mu_x = \mu_y$ vs. $H_1 : \mu_x \neq \mu_y$
- $60 = N_x = N_y$
- $7.5 = \sigma_x^2$
- $8.5 = \sigma_y^2$
- $\alpha = 0.05$
- $Z_{\text{kritisk}, \alpha/2} = 1.96$

$$Z_{\text{observert}} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(14 - 12.5)}{\sqrt{\frac{7.5}{60} + \frac{8.5}{60}}} = 2.90$$

Ettersom den observerte Z-verdien er større enn den kritiske verdien forkaster vi nullhypotesen.

d)

Vi bruker et signifikansnivå på 1% som gir en kritisk verdi på 2.58.

- $(\bar{X} - \bar{Y}) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_X^2}{N_X} + \frac{\sigma_Y^2}{N_Y}}$
- $(14 - 12, 5) \pm 2.58 \sqrt{\frac{7.5}{60} + \frac{8.5}{60}}$
- (0.1677, 2.8323)

e)

Om et konfidensintervall inkluderer null konkluderer vi med at det ikke er en statistisk signifikant forskjell, gitt det signifikansnivået som ble brukt. Konfidensintervallet over inneholder ikke 0, så vi konkluderer at det er en statistisk signifikant forskjell på 1% signifikansnivå.

f)

I dette tilfellet samsvarer konklusjonen fra hypotesetesten og konfidensintervallet. Det behøver det ikke å gjøre om man bruker ulike signifikansnivå, men om det blir brukt samme signifikansnivå vil konklusjonene være likt.

g)

- Symmetrisk rundt gjennomsnitt
- Gjennomsnitt, median og typetallet er det samme
- Karakterisert ved hjelp av gjennomsnitt og varians alene
- Pga sentralgrenseteoremet så går går fordelingen til tilfeldige variabler mot en normalfordeling
- Kan bli brukt som tilnærming for andre fordelinger når utvalgene blir store nok (f.eks binomisk fordeling, t-fordeling)

For hver av disse punktene man nevnte fikk man ett poeng (maks 4 poeng).

Oppgave 3

Oppgaven spør etter standardavviket i et utvalg, vi bruker derfor følgende formel for å regne ut standardavviket $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})}$.

$$\bar{X} = 3.667 \text{ (gjennomsnitt)}$$

$$s^2 = 7.467 \text{ (varians)}$$

$$s = 2.733 \text{ (standardavvik)}$$

Oppgave 4

$$P(Z \geq 1.98) = 1 - P(Z \leq 1.98) = .0239 \rightarrow .0239 * 2 (\text{da det er en tosidig test}) = 0.0478$$

Oppgave 5

Ved bruk av den kumulative binomiske fordelingen

$$P(6 \leq X) = 0.961$$

$$P(9 \leq X) = 0.999$$

$$0.999 - 0.961 = 0.038$$

PS: Grunnet ulik avrunding kunne man få 0.039 på denne oppgaven.

Oppgave 6

a)

Ved bruk av den kumulative Poisson-fordelingen

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.8795 = 0.1205$$

b)

Ved bruk av den kumulative Poisson-fordelingen

$$P(X \leq 3) = 0.8747$$

c)

Den er best tilpasset diskrete fordelinger