Eksamensløsninger - HSTAT1101 - desember 2023

Oppgave 1

Oppgaven spør etter standardaviket i en populasjonen, vi bruker derfor følgende formel for å regne ut standardaviket $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)}$.

$$\mu = 5.33 \text{ (gjennomsnitt)}$$
$$\sigma^2 = 8.888 \text{ (varians)}$$
$$\sigma = 2.981 \text{ (standardavvik)}$$

Oppgave 2

a)

- Begge bruker gjennomsnitt, standardavvik og størrelsen på utvalget/populasjonen
- I begge utregningene må du sette et signifikansnivå (α)
- Konfidensintervallet sier noe om usikkerheten rundt utvalgsestimatet og hvorvidt vi tror det sanne populasjonsgjennomsnittet ligger i dette intervallet.
- I en hypotesetest setter man to hypoteser opp imot hverandre og regner på sannsynligheten for at en av dem er sann gitt visse forutsetninger. Basert på utfallet av hypotesetesten, avgjør man om man forkaster eller beholder en av hypotesene som sann.

b)

Gjennomsnitt, standardavvik, størrelsen på utvalget/populasjonen, og signifikansnivå (α).

c)

Når man har ukjente populasjonsvarianser, så må vi bruke t-fordelingen istedenfor z-fordelingen. Om man har kjente populasjonsvarianser, så kan vi bruke z-fordelingen.

d)

I store utvalg/populasjoner, så kan vi bruke normalfordelingen som en tilnærming, da vi vet at mange fordelinger går mot en normalfordeling når størrelsen er stor nok

e)

- Hvert Bernoulli-forsøk har to utfall (suksess og ikke-suksess). De er altså gjensidig utelukkende og fulstendig dekkende.
- P(suksess) er den samme i alle Bernoulli-forsøkene
- De enkelte Bernoulli-forsøkene er (stokastisk) uavhengige

1 Oppgave 3

a)

P(rødt ess) = $\frac{2}{52}$

b)

P(tallkort) = $\frac{36}{52}$

c)

P(ett par) = $\frac{52}{52} \cdot \frac{3}{51} = \frac{3}{52}$

Spørsmålet var sansynnligheten for å trekke et par, ikke å trekke er kongepar. Kongepar-eksempelet i oppgaven var bare et eksempel.

d)

Sannsynligheten for å trekke en konge blant 52 kort: $\frac{4}{52}$
Sannsynligheten for å trekke et rødt kort blant 52 kort: $\frac{26}{52}$ Men vi har uten tilbakelegging, som betyr at du kan trekke en rød konge ved første trekk, som betyr at sannsynligheten for å trekke et rødt kort på trekk nr. 2 er: $\frac{25}{51}$
Disse to trekkene (da vi har uten tilbakelegging) er avhengige av hverandre. Det betor at $P(A|B) = P(A)$ ikke stemmer i dette tilfellet.

Oppgave 4

$P(Z \geq 1.85) = 1 - P(Z \leq 1.85) = 1 - 0.9678 = 0.0322 \rightarrow .0322 * 2(\text{da det er en tosidig test}) = 0.0644$

Grunnen avrunding var riktig svar i Inspera 0.663.

Oppgave 5

Ved bruk av den kumulative binomiske fordelingen

$$P(3 > X) = 0.073$$
$$P(6 \geq X) = 0.613$$
$$0.613 - 0.073 = 0.54$$

Oppgave 6

a)

Ved bruk av den kumulative Poisson-fordelingen

$$P(2 \geq X) = 0.4337$$
$$P(5 \geq X) = 0.9868$$
$$0.9868 - 0.4337 = 0.5531$$

b)

Ved bruk av den kumulative Poisson-fordelingen

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - 0.6990 = 0.301$$

Oppgave 7

Formelen for konfidensintervallet for et gjennomsnitt er.

$$X \pm Z_{\alpha/2} \cdot \sqrt{\frac{\sigma^2}{N}}$$
$$4.4 \pm 2.576 \cdot \frac{0.7}{\sqrt{68}}$$
$$(4.182, 4.618)$$

Oppgave 8

a)

Basert på informasjonen i oppgaven opp følgende tabell:

Table 1:

	P(jobb)	P(ikke jobb)	
P(lunsj)	=0.65*0.25=0.1625		0.45
P(ikke lunsj)			
	0.25		1

Ved bruk av sansynnlighetsreglene (komplementregelen) kan vi fylle ut resten av tabellen og dermed få ut de marginale og simultane sansynnlighetene.

Table 2:

	P(jobb)	P(ikke jobb)	
P(lunsj)	=0.65*0.25=0.1625	0.45-0.1625=0.2875	0.45
P(ikke lunsj)	0.25-0.1625 = 0.0875	0.75-0.2875=0.4625	1-0.45 = 0.55
	0.25	1-0.25 = 0.75	1

b)

Variansen er σ^2 Ettersom vi vet at $\sigma = 230$, så blir variansen $\sigma^2 = 230^2 = 52,900$

c)

$$X \pm Z_{\alpha/2} \cdot \sqrt{\frac{\sigma^2}{N}}$$
$$640 \pm 1.96 \cdot \frac{230}{\sqrt{195}}$$
$$(607.718, 672.282)$$

d)

Vanskelig å få tak i hele populasjonen av praktiske (kostnader, logistikk etc) og etiske årsaker (kan fks ikke ta kontakt med hele populasjonen). Vi kan derfor bruke utvalgsdata og statistiske metoder for å kunne si noe om populasjonen.

e)

I gjentatte utvalg vil fordelingen til utvalgsgjennomsnittet gå mot en normalfordeling når antallet gjentatte utvalg blir stort nok.

f)

- Y= UiO
- X= UiB
- $H_0 : \mu_x = \mu_y$ vs. $H_1 : \mu_x < \mu_x$
- $195 = N_x = N_y$
- $\mu_Y = 640$
- $\mu_X = 1155$
- $230 = \sigma_y$
- $510 = \sigma_x$
- $\alpha = 0.05$
- $Z_{kritisk,\alpha/2} = 1.645$

$Z_{observert} = \frac{(\overline{x}-\overline{y})-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}}} = \frac{(1155-640)}{\sqrt{\frac{510^2}{195}+\frac{230^2}{195}}} = 12.85$

Ettersom den observerte Z-verdien er større enn den kritiske verdien forkaster vi nullhypotesen. I oppgaven ble ikke antallet på UiB oppgitt, alle antagelser for UiB-antallet ble derfor godkjent.