

Sensorveiledning:		MET 34311		Statistikk	
Eksamensdato:	30.05.2012	kl.	09.00-14.00	Totalt antall sider:	3
Se oppgavesettet for tillatte hjelpemidler					
Se oppgavesettet for vektning av oppgavene					
Ansvarlig institutt: Institutt for samfunnsøkonomi					

Oppgave 1

- (a) $n = 1100$. Stikkprøven er stratifisert: for hvert bilmerke (stratum) har man valgt ut hundre kunder.
- (b) Alder måles på forholdstallnivå.
- (c) Ja, noenlunde normalfordelt (evt. litt flat på toppen). Vi ser at 34 tilsvarer 10% persentilen, så det er $0.1 \cdot 1100 = 110$ som er yngre enn 34 år.
- (d) Ordinalnivå. Gjennomsnitt og standardavvik er to observatorer som er meningsfulle for variabler på intervallnivå, men ikke på lavere nivåer (ordinal og nominal).
- (e) Binomisk forsøksrekke. Fem suksesser på rad har sannsynlighet $0.5^5 = 0.03125$. Fem suksesser på ti forsøk har sannsynlighet $P(x = 5) = \binom{10}{5} \cdot 0.5^5 \cdot (1 - 0.5)^5 \approx 0.246$.

Oppgave 2

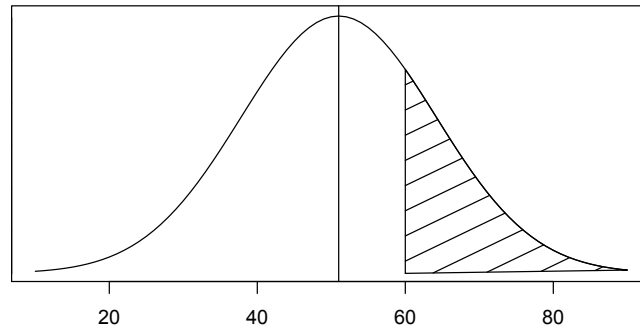
- (a) Standardavviket er

$$s = \sqrt{\frac{(3-6)^2 + (7-6)^2 + (7-6)^2 + (8-6)^2 + (5-6)^2}{5-1}} = \sqrt{\frac{16}{4}} = 2.$$

- (b) Medianen er 6.5. Q_1 befinner seg i posisjon $L = 0.25 \cdot 10 = 2.5$ som vi runder opp til $L = 3$, så $Q_1 = 5$.
- (c) Gjennomsnittet av de 15 verdiene er $\bar{L} = 6.2$.

$$I_L = \frac{6.2 - 1}{9} \cdot 100 \approx 57.8$$

- (d) Opel har både størst variasjonsbredde og interkvartilbredde, og har derfor størst variasjon i lojalitet. Bilmerker med median lik 8 er Audi, Ford, Opel, Peugeot, Nissan, Volkswagen.



Oppgave 3

- (a) Vi standardiserer 60: $z = \frac{60-51}{13.2} = 0.68$. Så $P(\text{Alder} > 60) = P(z > 0.68) = 1 - 0.7517 = 0.2483$.
- (b) Begge over 60 år har sannsynlighet $0.2483 \cdot 0.2483 = 0.062$. Gjennomsnittet \bar{x} basert på $n = 2$ observasjoner er selv normalfordelt med $\mu_{\bar{x}} = 60$ og standardavvik $\sigma_{\bar{x}} = 13.2/\sqrt{2}$. 60 år har da z -verdien $z = \frac{60-51}{13.2/\sqrt{2}} = 0.96$. $P(\bar{x} > 60) = P(z > 0.96) = 1 - 0.8315 = 0.1685$.
- (c) For en normalfordelt variabel så tilsvarende 5% persentilen en z -score på $z = -1.645$, altså 1.645 standardavvik under gjennomsnittet. Ole Johan er da $1.645 \cdot 13.2 = 21.7$ år yngre enn gjennomsnittet, altså $51 - 21.7 = 29.3$ år.

Oppgave 4

- (a) Dataene må komme fra en normalfordelt populasjon.
- (b) Feilmarginen er $E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 1.646 \frac{13.165}{\sqrt{1100}} = 0.68$. Konfidensintervallet blir da 50.99 ± 0.68 , dvs (50.31, 51.67). Vi er 90% sikre på at gjennomsnittsalderen i populasjonen av bilkunder ligger mellom 50.31 år og 51.67 år.
- (c) Feilmarginen er $E = 1.96 \cdot \sqrt{\frac{0.27 \cdot (1-0.27)}{1100}} = 0.026$. Konfidensintervallet blir da 0.27 ± 0.026 , eller (0.244, 0.296).
- (d) Feilmarginen er halvparten av intervallbredden, altså $E = (8.414 - 7.666)/2 = 0.374$. Det betyr at $t_{\alpha/2} \cdot 2.251/\sqrt{100} = 0.374$, altså $t_{\alpha/2} = 0.374 \cdot 10/2.251 \simeq 1.66$. Vi har 99 frihetsgrader og i tabell A3 så finner vi at 1.66 tilsvarende et haleareal på omtrent 0.05 for 100 frihetsgrader. Konfidensnivået er altså $1 - 2 \cdot 0.05 = 0.9$, dvs. 90%.

Oppgave 5

- (a) Dersom H_0 er sann så kan det allikevel hende at vi forkaster den, dette skjer jo med en sannsynlighet α . Så vi har ikke bevist at H_1 er sann.

(b) Siden vi forkaster på $\alpha = 0.05$ nivået så må p-verdien være mindre enn 0.05. Siden vi ikke forkaster på $\alpha = 0.01$ nivået så må p-verdien være større enn 0.01. Altså så må p-verdien ligge mellom 0.01 og 0.05.

(c) $H_0 : p = 0.5$ vs. $H_1 : p > 0.5$. Testobservator er $z = \frac{25/40 - 0.5}{\sqrt{0.5(1-0.5)/40}} = 1.58$.
Kritisk verdi er 1.645. Så vi forkaster ikke H_0 .

(d) $H_0 : \mu_A = \mu_M$ og $H_1 : \mu_A < \mu_M$. Testobservator

$$t = \frac{53.1 - 47.8}{\sqrt{\frac{14.2^2}{100} + \frac{12.4^2}{100}}} = 2.81$$

Kritisk verdi for testen er $t = 2.368$. Vi forkaster H_0 .

Oppgave 6

(a) Avleser på linja at 40 år tilsvarer omtrent $L = 5$ og 70 år tilsvarer omtrent $L = 8$.

(b) Vi ser fra linja at lojaliteten øker fra omtrent 3.9 til omtrent 8.7 når alderen øker fra 35 år til 75 år. Da blir stigningstallet lik $(8.7 - 3.9)/40 = 0.12$. Det betyr at for hvert år en kunde blir eldre så øker lojaliteten med 0.12.

(c) $H_0 : \rho = 0$ vs. $H_1 : \rho > 0$. Testobservator er $t = \frac{0.66}{\sqrt{(1-0.66^2)/13}} = 3.17$. Kritisk verdi er $t_{0.05, df=13} = 1.771$. Vi forkaster H_0 .

(d) Siden det er positiv korrelasjon i populasjonen så forventer vi å forkaste H_0 , iallefall hvis stikkprøven er stor nok. Store stikkprøver øker testens styrke til å forkaste en feilaktig nullhypotese. Derfor forventer vi at p_B er mindre enn p_A .

Oppgave 7

(a) H_0 : Det er ingen sammenheng mellom kjønn og bilmerke.

(b) O er den observerte frekvensen i en celle i krysstabellen, mens E er den forventede verdien i cellen dersom H_0 om ingen sammenheng er sann. Her er $O = 100 - 64 = 36$ og $E = \frac{521}{1100} \cdot 100 = 47.36$.

(c) Her har vi $(11 - 1) \cdot (2 - 1) = 10$ frihetsgrader, så kritisk verdi er $\chi_{df=10}^2 = 18.307$. Vi forkaster H_0 .