

Skriftlig eksamen i: MET 34311 Statistikk
 Eksamensdato: 01.06.11, kl. 09.00-14.00
 Tillatte hjelpemidler: Alle + BI-definert eksamenskalkulator :
 TEXAS INSTRUMENTS BA II Plus TM
 Innføringsark: Ruter
 Totalt antall sider: 5

Oppgave 1

I løpet av en dag har et massasjelinstituttt 131 klienter. Sekretæren noterer følgende informasjon for hver klient: *Dato*, *Navn*, *Kjønn*, *Personnummer*, *Behandlingsform* og *Beløpet* som pasienten skal betale. Det er fire behandlingsformer: klassisk massasje, idrettsmassasje, triggerpunktsmassasje og bindevevs-massasje.

- På hvilket målenivå er variabelen *Behandlingsform*?
- På slutten av dagen legger sekretæren inn all innsamlet informasjon i en ny JMP datafil. Hvor mange rader og søyler(kolonner) har datafila?
- Histogrammet i figur 1 på side 4 viser fordelingen til *Beløp*. Omtrent hvor mange klienter betalte mindre enn 300 kroner?
- Gjennomsnittsbeløpet er 314.2 kroner mens medianbeløpet er 293 kroner. Kan du kort forklare ut fra histogrammet hvorfor gjennomsnittet er såpass mye større enn medianen?
- Sekretæren setter opp en kryss-tabell for sammenhengen mellom behandlingsform og kjønn:

	Behandlingsform			
	A	B	C	D
Mann	22	3	15	36
Kvinne	12	24	9	?

Hvor mange kvinner fikk behandlingsform D? Hvor mange prosent av klientene var menn?

- Sekretæren sender et spørreskjema til tre tilfeldig valgte klienter fra datafila. Hva er sannsynligheten for at alle de tre utvalgte klientene er menn?

Oppgave 2

For å få et innblikk i kundenes alders sammensetning spør ekspeditøren i en butikk hver tiende kunde hvilken alder hun/han har. Resultatet ble:

11, 15, 23, 26, 19, 22, 15, 25, 24

- (a) Hvilket type utvalg er det ekspeditøren har framskaffet?
- (b) Regn ut gjennomsnitt og standardavvik til alderen basert på stikkprøven. Vil du si at en kunde på 12 år har en 'uvanlig' alder? Begrunn svaret kort.
- (c) Beregn nedre og øvre kvartil (Q_1 og Q_3), og finn interkvartilbredden $Q_3 - Q_1$ til alderen. Forklar også kort i hvilke situasjoner interkvartilbredden kan være et bedre mål for variasjon enn standardavviket.

Oppgave 3

- (a) Du er i BI Nydalen og skal ta heisen. Anta at *antall sekunder* som går fra du trykker på heisknappen til heisen kommer er normalfordelt med $\mu = 22$ og $\sigma = 8$ sekunder. Hva er sannsynligheten for at du må vente mellom 25 og 30 sekunder på heisen? Skisser sannsynligheten som arealet under en graf.
- (b) Anta at *antall minutter* en BI student bruker på å finne en ledig parkeringsplass er normalfordelt med gjennomsnitt μ og standardavvik σ . Foreslå verdier for μ og σ som du tror kan passe for ditt BI studiested. Hva er sannsynligheten for at en tilfeldig valgt student må bruke mer tid enn to standardavvik over gjennomsnittet for å finne parkeringsplass? Skisser denne sannsynligheten som arealet under en graf.

Oppgave 4

Direktørene for de forskjellige BI studiestedene ønsker å vite hvor mye deres studenter jobber per uke utenom studiene.

- (a) I en stikkprøve på 52 studenter ved BI Nydalen finner man at gjennomsnittlig antall timer på jobb er $\bar{x} = 9.74$ timer med et standardavvik på $s = 6.86$ timer. Konstruer et 90 % konfidensintervall for populasjonsgjennomsnittet μ .
- (b) Direktøren ved BI Stavanger spør 11 tilfeldig valgte studenter på sitt studiested hvor mye de jobber. Basert på disse svarene konstruerer han et konfidensintervall for μ . Diskuter kort om det er noen svakheter ved direktørens metode. Hvilken betingelse må være oppfylt for at direktørens konfidensintervall skal være gyldig?
- (c) Direktøren ved BI Drammen henter inn en tilfeldig stikkprøve på 63 studenter. Han beregner konfidensintervallet for ukentlig antall timer på jobb:

$$5.263 < \mu < 11.658$$

Finn stikkprøvegjennomsnittet \bar{x} og feilmarginen E som ble brukt for å konstruere dette konfidensintervallet.

Oppgave 5

- (a) I hypotesetesting, hva vil det si å begå en type I feil? Anta at du tester på signifikansnivå $\alpha = 0.05$, og at H_0 er sann. Hva er sannsynligheten for å begå en type I feil?
- (b) I en stikkprøve på 157 studenter ved BI Bergen våren 2011 hadde 13 studenter mobiltelefon av merket Samsung. Test påstanden om at mindre enn ti prosent av studentene ved BI Bergen har Samsung mobiltelefon. Skriv opp nullhypotesen og alternativhypotesen. Foreta testen på et $\alpha = 0.05$ signifikansnivå og formuler konklusjonen i et lettfattelig språk.
- (c) Du har tilgang til stikkprøver av ukentlige timer på jobb for 60 mannlige studenter og for 99 kvinnelige studenter. Er de to stikkprøvene med menn og kvinner uavhengige eller relaterte? Test hypotesen om at mannlige studenter gjennomsnittlig jobber mer enn kvinnelige studenter basert på de to stikkprøvene vist i Figur 2. Skriv opp nullhypotesen og alternativhypotesen og bruk signifikansnivå $\alpha = 0.1$. Skisser p-verdien som arealet under en graf.

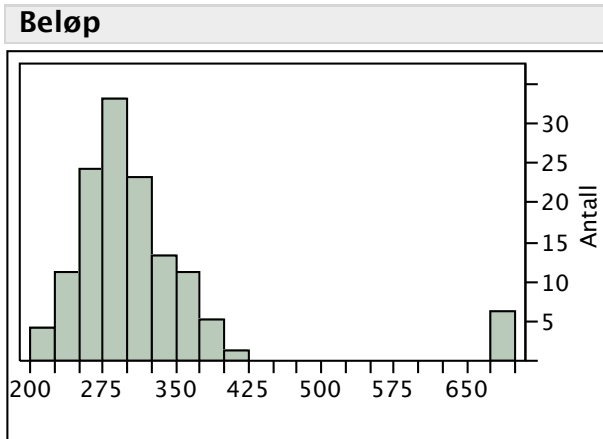
Oppgave 6

- (a) Figur 3 viser et scatterplott av timer på jobb vs studietimer per uke for en tilfeldig valgt stikkprøve av åtte BI studenter våren 2011. Hvor mange studenter i stikkprøven oppga at de studerte 30 timer i uka? Hvor mye jobbet hver av disse studentene i uka?
- (b) Korrelasjonskoeffisienten for dataene i Figur 3 er $r = 0.73$. Test på $\alpha = 0.05$ nivået om det er en korrelasjon mellom timer på jobb og timer brukt på studier. Skriv opp H_0 og H_1 .
- (c) Gi et eksempel på to kontinuerlige variable som du tror er signifikant negativt korrelerte.

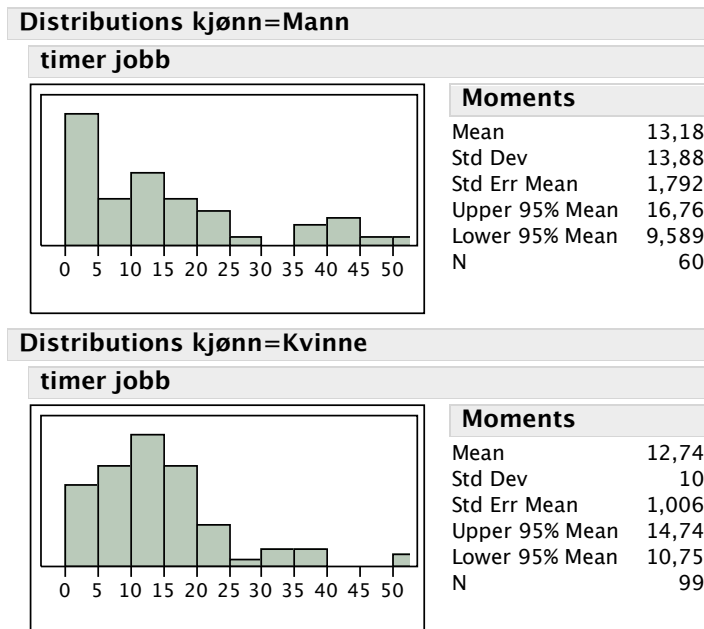
Oppgave 7

Figur 4 gir krysstabellen for sammenhengen mellom kjønn og studium i en tilfeldig stikkprøve av 291 studenter fra BI Trondheim. I hver celle oppgis det faktiske og det forventede antallet.

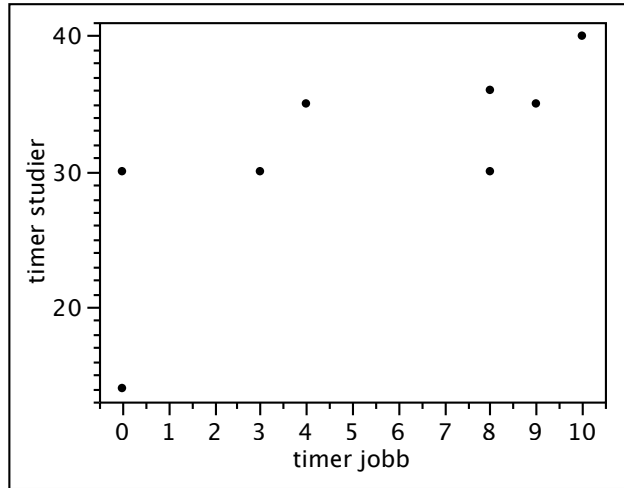
- (a) Hvilken hypotesetest kan du bruke for å avgjøre om det er en sammenheng mellom kjønn og studium i populasjonen av studenter ved BI Trondheim?
- (b) Foreta en test på 1% nivået og konkluder i et lettfattelig språk.



Figur 1: Oppgave 1



Figur 2: Oppgave 5



Figur 3: Oppgave 6

		kjønn		
Count		Mann	Kvinne	
Expected				
Studium	Økonomi	92	69	161
		79,6701	81,3299	
Markedsføring		52	78	130
		64,3299	65,6701	
		144	147	291

Figur 4: Oppgave 7