

Eksamen er todelt. Første del inneholder oppgaver som skal besvares basert på informasjon i dette oppgavesettet. Andre del inneholder spørsmål omkring analyser av datasettet som ble levert ut to uker før eksamen. Denne prøveeksamenen inneholder omkring halvparten så mange oppgaver som den kommende eksamen. Oppgavene i denne prøveeksamenen er representative med tanke på tematikk og vanskegrad.

Del 1

Personlighetstester blir ofte brukt i rekruttering i arbeidslivet. Formålet er å finne ut hvilke personlighetstrekk en jobbkandidat har. En stor bedrift lyste ut ti ledige stillinger som kundekonsulent. Totalt var det $n = 295$ søkere til stillingene, og alle tok en personlighetstest. Testen resulterte blant annet i scorer på følgende personlighetstrekk:

- *Medmenneskelighet*: Et mål på hvor stor tillit man har til medmennesker, og hvor stor omtanke man har for medmennesker. Måles på en kontinuerlig skala fra 0-100.
- *Nevrotisisme*: Et mål på følelsesmessig ustabilitet med tilbøyelighet til angst, irritasjon og depresjon. Måles på en kontinuerlig skala fra 0-100.

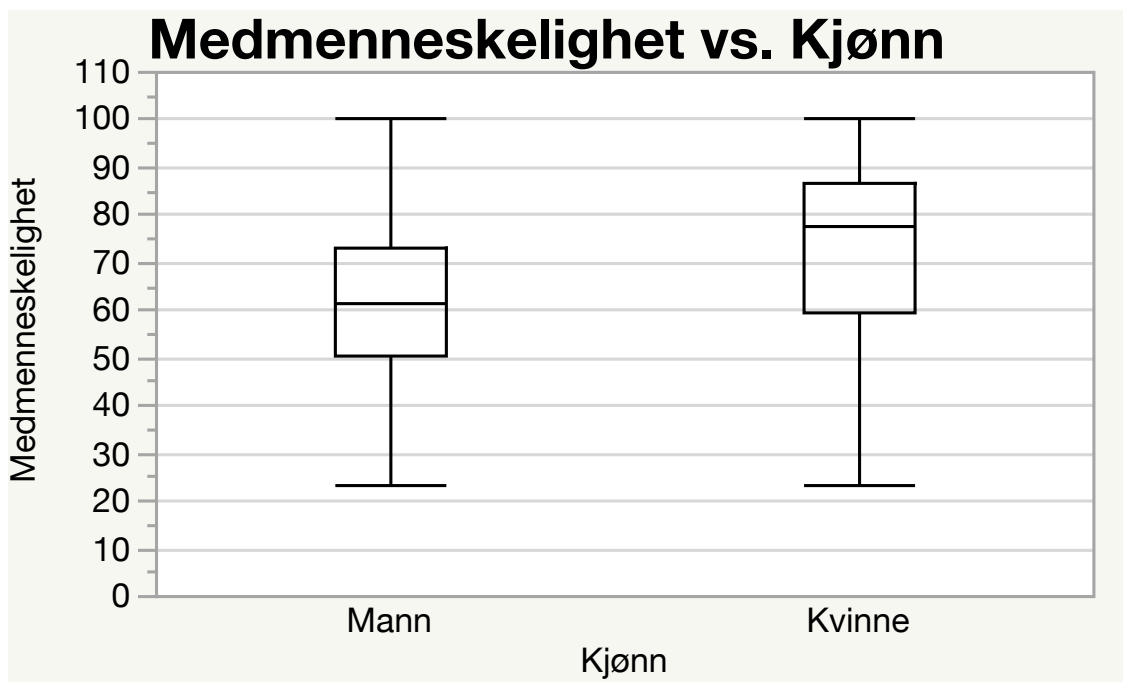
I tillegg skal vi se på følgende variabler: *Kjønn*, *Alder* og *Utdanningsnivå*. Sistnevnte variabel har følgende fem kategorier: Ungdomsskole, Videregående skole, Årsstudium innen høyere utdanning, Bachelorgrad og Mastergrad.

Oppgave 1 (4 poeng)

- (a) På hvilket målenivå er variabelen *Utdanningsnivå*?
- (b) Krysstabellen mellom *Kjønn* og *Utdanningsnivå* er som følger:

	Ungdomsskole	Vid.g. skole	Årsstudium	Bachelor	Master
Mann	13	11	43	?	18
Kvinne	11	22	102	35	29

- Hvor mange menn hadde Bachelor som sitt utdanningsnivå?
- (c) Hvor stor prosentandel av søkerne hadde Mastergrad?
- (d) Lag en graf som viser fordelingen til *Utdanningsnivå* for de kvinnelige søkerne.



Figur 1: Oppgave 2

Oppgave 2 (4 poeng)

Figur 1 viser fordelingen til *Medmenneskelighet* hos menn og kvinner.

- Hva kalles den type graf som vises i figuren?
- Hvilket kjønn scorer høyest på medmenneskelighet?
- Anslå medianverdien til medmenneskelighet hos kvinner.
- Anslå interkvartilbredden (IK) i medmenneskelighet hos menn.

Oppgave 3 (2 poeng)

La oss anta at *Nevrotisisme* hos en tilfeldig valgt søker er normalfordelt med $\mu = 50$ og standardavvik $\sigma = 15$.

- Lars har en nevro-tisisme-score på 80. Hva er z -verdien til denne scoren?
- Hva er sannsynligheten for at en tilfeldig valgt søker scorer mellom 20 og 80 i nevro-tisisme? Skisser sannsynligheten som arealet under en graf.

Oppgave 4 (3 poeng)

Tabellen under inneholder nøkkeltall for score på medmenneskelighet for menn og kvinner.

Means and Std Deviations						
Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
Mann	96	59.78	20.6	2.103	55.61	63.96
Kvinne	199	73.43	18.79	1.332	70.8	76.06

Figur 2: Oppgave 5.

- (a) Er tallene i tabellen observatorer eller parametre?
- (b) Det påstås at kvinnelige søkere i gjennomsnitt er mer medmenneskelige enn mannlige søkere. Skriv opp hypoteser H_0 og H_A for å teste denne påstanden. Bruk symbolene μ_K og μ_M .
- (c) Når testen utføres så blir p-verdien $< .001$. Hva blir konklusjonen på testen? Formuler deg i et lettfattelig språk.

Oppgave 5 (5 poeng)

Vi ser på regresjonen mellom den uavhengige variabelen $x = \text{Alder}$ og den avhengige variabelen $y = \text{Medmenneskelighet}$. Regresjonstabellen for linja

$$\widehat{\text{Medmenneskelighet}} = b_0 + b_1 \cdot \text{alder}$$

er

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	62.06	3.38	18.36	0.00
alder	0.24	0.11	2.19	0.03

- (a) Lise er en ny søker, og hun er 50 år gammel. Hva er prognosen for hennes medmenneskelighets-score?
- (b) Gi en tolkning av koeffisienten $b_1 = 0.24$.
- (c) Det påstås at søkere har høyere score på medmenneskelighet jo eldre de er. Skriv opp H_0 og H_A for å teste denne påstanden. Bruk symbolet β_1 .

- (d) Har vi tilstrekkelig støtte til å hevde at scoren på medmenneskelighet øker med alder? Bruk signifikansnivå $\alpha = 0.05$.
- (e) Nevn tre forutsetninger for at vi kan stole på resultatet fra en lineær regresjonsmodell.

Oppgave 6 (4 poeng)

Fram til nå har vi antatt at nevrotisisme er målt på en skala fra 0-100. I forskning brukes skalaer for å måle nevrotisisme, blant annet som del av big 5. Forskerne Weinstock og Whisman lanserte i 2006 en skala for nevrotisisme som består av 8 item, som måles på 4-punkts Likert skala. Itemene kalles (tense, nervous, temperamental, irritable, envious, unstable, insecure, emotional).

- (a) Forklar hvorfor forskerne bruker flere enn ett item.
- (b) Reliabiliteten til skalaen var $\alpha = .86$ i et representativt utvalg. Definer reliabilitet og gi en tolkning av α verdien.
- (c) Forklar hva validitet er.
- (d) Forskerne oppgir at deres skala er korrelert med $r = 0.7$ med nevrotisisme skalaen som ofte brukes i big five instrumentet (NEO Personality Inventory, Costa og McCrae, 1992). Forklar hvordan dette styrker validiteten. Hva kaller vi den type validitet som her styrkes?

Oppgave 7 (3 poeng)

- (a) Forklar hva en latent variabel er for noe. Bruke gjerne nevrotisisme som eksempel.
- (b) Hva er formålet med EFA (eksplorerende faktor analyse). Bruke gjerne nevrotisisme som eksempel.
- (c) Gi et eksempel på at en skala kan bestå av to subskalaer. Bruke gjerne nevrotisisme som eksempel.

1 Del 2

Oppgave 8 (10 poeng)

- (a) Hvor mange respondenter hadde missing på en eller flere variabler?
- (b) Alle spørsmål fra nå av antar at alle respondenter med missing er fjernet fra datasettet. Hvor mange prosent av respondentene var kvinner?

- (c) Hva var interkvartilbredden for alder?
- (d) Hva vil du si må til for å kalle en respondent i dette datasettet for uvanlig ung?
- (e) Du ønsker å teste om engasjementet er forskjellig i de to kjønnene. Skriv opp H_0 og H_A ved hjelp av symbolene μ_K og μ_M .
- (f) Utfør testen med signifikansnivå $\alpha = 0.1$ og konkluder i et lettfattelig språk.
- (g) En arbeidstaker skårer i snitt 3 på Autonomi. Bruk regresjon til å predikere hvor engasjert hun er.
- (h) Hvilken av de tre forutsetningene for regresjon ser ut til å være minst til stede for regresjonsmodellen i forrige oppgave?
- (i) Hvor mye av variasjon i engasjement kan forklares av varierende autonomi?
- (j) Er det viktig å også kontrollere for alder og kjønn i denne regresjonen?

Oppgave 7 (10 poeng)

- (a) I korrelasjonsplottet for indikatorene for de 14/15 konstruktene, hvilket konstrukt ser ut til å være minst korrelert med de andre konstruktene?
- (b) Hva er reliabiliteten til engasjement?
- (c) Hvilket item for engasjement ville du droppet dersom du skulle redusere til 5 items? Begrunn og diskuter om dette kan endre validitet.
- (d) Sammenlikne korrelasjon mellom autonomi og engasjement når det er beregnet ut fra gjennomsnittskåre og ut fra latent variabel modellering. Kommenter hvorfor de ikke er like.
- (e) Vil du anbefale å splitte utbrenthet i to sub-faktorer (såkalte fasetter)?
- (f) Hvor mange faktorer (konstrukter) ender man opp med dersom man blindt stoler på parallel-analyse?
- (g) Hvilken andre metode(r) for å bestemme antall faktorer kunne vært aktuelle å bruke?