

# Utsatt eksamen PSY1010 - PSYC1100

## Forskningsmetode I - våren 2018

### Bokmål

*Skriftlig skoleeksamen, 8. august (3 timer)*

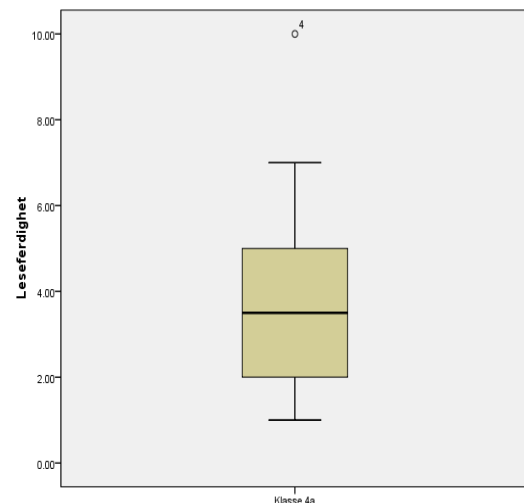
*Besvar tre – 3 – av de følgende fire – 4 – oppgavene*

#### Oppgave 1

Du er ansatt som forskningsansvarlig på en behandlingsinstitusjon for ungdom og blir bedt om å finne ut om det er noe som skiller pasientene fra vanlig ungdom bortsett fra problemet de er innlagt for. Du lurer på om du skal benytte et kvalitativt eller kvantitativt design. Diskuter fordeler og ulemper ved de to metodetilnærmingene. Bruk gjerne den beskrevne situasjonen som eksempel.

#### Oppgave 2

Du har gjennomført en undersøkelse av leseferdigheten i en klasse. For å få en oversikt over resultatet foretar du en deskriptiv dataanalyse. Resultatet finner du i boksplottet til høyre.



- Hvilke mål på spredning og sentraltendens kan du lese av figuren? (Oppgi omtrentlig verdier)
- Elevene i klassen har fått følgende resultat: (3, 5, 1, 10, 7, 3, 1, 4, 2, 4). Regn ut gjennomsnittet og bruk formelen under til å beregne variansen.

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

- Noen ganger deler man på «n – 1» istedenfor «N», som i formelen over. Hvorfor gjør man dette?
- Standardavviket for klassen er 2,79. Hva betyr dette og hva er sammenhengen mellom varians og standardavvik?
- I boksplottet finner du også en uteligger (outlier). Hva er en uteligger og hvordan påvirker uteliggere målene på sentraltendens?

#### Oppgave 3

Du finner at antall voldsepisoder og en persons IQ skåre ser ut til å ha en sammenheng med hverandre.

- a) Korrelasjonen mellom de to variablene er  $r = -0.50$ . Hva betyr det?
- b) Hvorfor skal man være forsiktig med å trekke slutninger om årsak og virkning fra korrelasjonsdata?
- c) I mange sammenhenger er vi mer tjent med å oppgi korrelasjonen som delt varians (coefficient of determination). Hva forstår du med delt varians (Coefficient of determination), og hvor stor er den?
- d) I dette eksempelet er «Pearsons  $r$ » benyttet. Kjenner du andre korrelasjonskoeffisienter?

#### **Oppgave 4**

Overføringseffekter (Carryover effect) blir sett på som en trussel mot validiteten i repeterte (within-subjects) design.

- a) Hva menes med repeterte design?
- b) Hva menes med ytre og indre validitet i eksperimenter?
- c) Hva er overføringseffekter?
- d) Hva kan man gjøre for å redusere faren for overføringseffekter?

Lykke til!

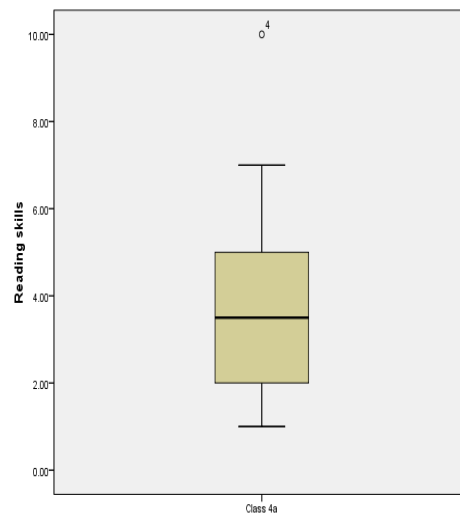
Please answer three – 3 – of the following four – 4 – assignments

### Assignment 1

You are employed as a research manager at an institution for adolescents and are asked to find out if there is something that differentiates patients from non-patient adolescents, apart from the problem for which they are hospitalized. You wonder if you should use a qualitative or a quantitative design. Discuss the advantages and disadvantages of the two methodological approaches. Feel free to use the described situation as an example.

### Assignment 2

You have conducted a study on reading skills in a school class. To get an overview of your data, you have done a descriptive data analysis. The result is shown in the box-plot to the right.



- Describe measures of spread and central tendency that you can read from the figure? (Provide approximate values)
- The pupils have obtained the following results: (3, 5, 1, 10, 7, 3, 1, 4, 2, 4). Use the following formula to calculate the variance.

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

- Sometimes "n – 1" is used instead of "N", as in the formula above. Why do you do that?
- The standard deviation for the class is 2.79. What does this mean and what is the relation between the standard deviation and the variance?
- The box-plot indicates that there is an outlier in the sample. What do you understand by outliers and how do outliers affect the different measures of central tendency?

### Assignment 3

You find that the frequency of violent episodes seems to relate with a person's IQ-score.

- The correlation between the two variables was  $r = -0.50$ . What does this mean?
- Why should you be careful when drawing conclusions about cause and effect from correlational data?
- What do you understand by shared variance (Coefficient of determination) in this context and how large is it?
- In this example we use «Pearson's r». Do you know any other correlation coefficients?

#### **Assignment 4**

Carryover effects are regarded as a threat against the validity in within-subjects design (repeated design).

- a) What do you understand by a within-subjects (repeated) design?
- b) What can you say about internal and external validity in experimental designs?
- c) What do you understand by carryover effects?
- d) What can be done to reduce the risk for carryover effects?

Good luck!

## Noen momenter ved vurdering av utsatt eksamen PSY1010 – PSYC1100 våren 2018.

### Generelt:

Denne veiledningen peker på noen enkeltelementer som kan diskuteres i oppgavene. Den er ikke dekkende eller ment som en fullstendig liste, og skal ikke betraktes som absolutte krav. Oppgavene kan løses på mange måter og i tilfeller der kandidaten har benyttet en uortodoks løsning blir sensor nødt til å vurdere hvorvidt løsningen kan regnes som akseptabel.

Som en regel skal kandidaten belønnes for å vise forståelse av faget og de prinsipielle sidene av metodestoffet. I tilfeller der kandidaten trekker inn kunnskap fra andre områder av faget eller tilgrensende fagområder enn det som direkte spørres etter i det enkelte spørsmål må man vurdere relevansen. Er stoffet gjort relevant for spørsmålet skal dette honoreres. I tilfeller der stoffet ikke er særlig relevant og mer bærer preg av "Jeg kan også dette" skal man se bort fra dette.

Legg vekt på forståelse og oversikt av stoffet når dere leser oppgavene. Kandidater som viser god oversikt og forståelse bør honoreres. Det er også flott hvis kandidaten klarer å illustrere poengene sine med gode eksempler.

Ved regneoppgaver bør man i tillegg til svaret også vurdere måten utregningen er gjort på. I tilfeller der galt svar åpenbart bare skyldes enkle regnefeil bør man ta hensyn til dette i vurderingen av svaret. Merk at for de som ikke har med kalkulator, er det viktig at kandidaten viser utregningene og forståelse for fremgangsmåten.

Alle fire oppgaver teller like mye og kandidaten får ikke trekk eller bonus for å velge noen framfor noen andre. Det skal heller ikke krediteres om kandidaten har svart på alle fire, i stedet for tre.

Vedrørende deloppgavene, er det med hensikt at noen oppgaver er vanskeligere enn andre. Dette gjør det lettere å sette karakter, og det er ikke meningen at sensor skal være like streng på alle.

### Oppgave 1

Et universitet hyrer deg inn for å kartlegge alkoholbruk blant studentene og hvordan det påvirker resultatene de oppnår. Data blir samlet inn ved at du deler ut spørreskjema til studenter som er tilstede på utvalgte forelesninger.

De 14 første studentene oppga at de drakk følgende antall alkoholenheter per uke:

0, 0, 0, 1, 2, 3, 3, 4, 4, 5, 7, 7, 8, 12

$$SD = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

a) Hva er uavhengig og avhengig variabel i denne undersøkelsen?

Kan besvares kort med alkoholbruk som uavhengig variabel og oppnådde resultater som avhengig variabel. Ufyllende svar kan gjerne honoreres.

b) Hva slags utvalg har vi med å gjøre? Hvordan påvirker utvalgsteknikken tolkningen av resultatene?

Sammenheng mellom populasjon og utvalg – utvalget skal være en «miniatyrkopi» av populasjonen. Kandidaten kan godt definere populasjon, sampling frame og utvalg og nevne ulike utvalgsteknikker. Ifølge oppgaveteksten er vi interessert i alkoholbruk blant studentene ved et universitet. Dette er et ikke-tilfeldig utvalg (nonrandom sample). (Dette er dekket i pensumboka, fra side 165.) Kan også kalles studentutvalg, men ikke-tilfeldigheten bør diskuteres. Bonus til studenter som kaller det bekvemmelighetsutvalg (convenience sample), eller argumenterer for at det er en slags cluster-sampling (dette er sagt på forelesning). Det at utvalget er ikke-tilfeldig kan føre til bias ved at ikke alle har samme sannsynlighet for å delta. Det bør komme tydelig fram at dette ikke er et vanlig randomisert utvalg. Sampling frame, selektivt frafall og responsrate kan påvirke resultatene. Bias ved at kun de fremmøtte kan delta i studien. Kandidaten kan også fortelle om alternative utvalgsteknikker og i så fall skal det belønnes om kandidaten forklarer hvordan de ville gitt en annen/bedre tilnærming til problemstillingen.

c) Regn ut tre mål på sentraltendens (gitt tallene over) og forklar kort hva de betyr.

Gjennomsnitt 4, median 3,5, modus/modalverdi 0. Likt fordelt på alle, drikker de 4 enheter, midterste verdi er 3,5 (snitt av 3 og 4 ettersom partall antall verdier), vanligste verdi er 0. Kandidaten bør kommentere hvorfor vi får ulike verdier, og kommentere på hva dette betyr. Avvik mellom median, gjennomsnitt og modus finner vi ved skjeivfordelte variabler. Aritmetisk gjennomsnitt påvirkes sterkest av uteliggere.

d) Regn ut standardavviket (SD). Hva forteller det oss?

Standardavviket er 3,53009. Ved 3,5 skal dette anses som riktig, ved andre verdier må det vurderes hvor i utregningen det har gått feil. Noen kandidater hadde ikke med kalkulator, hvilket gjør det vanskelig å regne ut kvadratrot. Variansen er 12,46, så uten kalkulator bør man kunne gjette seg til at det er et sted mellom 3 og 4. Men i alle fall kandidaten vise utregningen. Viktigere er det likevel at kandidaten forstår hva standardavviket er. Kandidaten må forstå at standardavviket er et mål på spredning, og det bør framkomme at SD kan overføres direkte til den skalaen målingen er gjort på (positivt om det nevnes at det til forskjell fra varians, som er kvadrert). Standardavviket tolkes som verdiens gjennomsnittlige avstand fra gjennomsnittet. Teller positivt om kandidaten kommenterer vs. gjennomsnittlig avvik fra gjennomsnittet og gjennomsnittlig absolutt avvik til gjennomsnittet. Det skal honoreres om kandidaten kommenterer forholdet mellom SD og varians og/eller forholdet mellom SD og standardfeil. Honorer hvis kandidaten viser til normalfordelingskurven og den prosentvise fordelingen ligger innenfor standardavviket (for eksempel at 68,2% av fordelingen vil ligge +/- 1 SD fra gjennomsnittet for normalfordelte variabler). Også fint om kandidaten forklarer hvorfor n-1 er nevner.

## **Oppgave 2**

I et stort utvalg tester du om det er slik at studieretning er assosiert med alkoholforbruk. Totalt sammenligner du 5 forskjellige studieretninger. Du gjennomfører en signifikanstest på forskjellen mellom gruppene og finner en p-verdi på 0,026 ( $p = ,026$ ). Alfa er 0,05.

a) Hva menes med nullhypotese, alternativ hypotese, type 1- og type 2-feil?

Forklare prinsippet bak  $H_0$  og  $H_1$ . Forklare at det er 4 ulike beslutninger man kan ta i forhold til å beholde eller å avvise  $H_0$ , hvorav to fører til riktige beslutninger mens de to andre er feilslutninger. Vise at kandidaten har forstått at dette er forhold som gjelder forskjell i gjennomsnitt mellom gruppene i populasjonen. Risiko for type 1 feil kan justeres med alfa, sannsynlighet for type 1 og type 2 feil påvirker hverandre gjensidig. Kandidaten kan diskutere antagelser om årsaksforklaring.

b) Hvilken type test ville du benyttet for å regne ut signifikansen i denne undersøkelsen? (Begrunn svaret ditt).

Her er flere svar mulig, men her bør kandidaten vise til at det er fem gjennomsnitt som skal vurderes og at man derfor bør bruke ANOVA. Kandidaten bør kommentere på hvorfor vi ikke gjør en rekke parvise t-tester (samlet feilrate). (ANOVA er godt dekket i pensum, så svarer kandidaten at vi burde brukt en rekke (uparede) t-tester bør det gi en noe uttelling.) Det er et pluss hvis kandidaten viser til at det er uavhengige grupper og at man derfor skal benytte enveis mellomgruppe variansanalyse. Noen kandidater kan her gå inn på en dypere forklaring av variansanalyse. Vurder svaret og om det er gjort relevant og viser forståelse av prinsippet om partisjonert varians.

c) Forklar hva som menes med  $p = ,026$  i denne sammenhengen? Er resultatet statistisk signifikant, og hva betyr i så fall det?

Resultatet er signifikant, og man tolker dette som at det er en forskjell i gjennomsnittlig alkoholbruk mellom de 5 ulike studieretningene. Kandidaten bør få fram at signifikans dreier seg om sannsynlighet for at et gitt resultat i utvalget skal oppstå tilfeldig, gitt at det ikke er noen forskjell i populasjonen. Positivt hvis kandidaten eksplisitt nevner at sannsynlighet under nullhypotesen ikke er det samme som sannsynlighet for  $H_0$  eller  $H_1$  som sådan. Også positivt om kandidaten nevner forskjell mellom assosiasjon og kausalitet, og at dette ikke er en randomisert studie, som ikke sier noe om årsak. Kriteriet for å avvise  $H_0$  kalles for alfa nivået. Som en konvensjon brukes et sannsynlighetsnivå på 0,05 eller 0,01 (her 0,05). Utregningen av signifikansnivået er avhengig av det resultatet som er oppnådd i utvalget sammen med andre forhold som spredning og utvalgets størrelse. Noen kandidater kan her vise til at sannsynligheten beregnes ut fra samplingsfordelingen. Hvis dette er forstått bør dette honoreres ekstra.

d) Du kan også legge inn variabelen kjønn i analysen. Hvorfor kan det være lurt og hva kaller vi en slik analyse?

Denne oppgaven er vanskelig, og er ment å skille mellom gode kandidater. Her kan kandidaten svare flere ting. Kandidatene må reflektere rundt hvordan kjønn kan påvirke assosiasjonen mellom studieretning og alkoholbruk, for eksempel at kjønn kan være korrelert med avhengig og uavhengig variabel. Når det gjelder hva analysen heter er det riktigste å bruke en ANCOVA, og må telles som bonus om kandidaten klarer dette. Det kan også argumenteres for å bruke en toveis variansanalyse. Ved ANCOVA kan kandidaten si noe om at kjønn korrelerer med studieretning og at kovariansanalyse kontrollerer for denne effekten. Argumenteres det for toveis ANOVA bør kandidaten si noe om hvordan man tenker seg kjønn benyttet som uavhengig variabel.

### Oppgave 3

Du finner at antall voldsepisoder og en persons IQ skåre ser ut til å ha en sammenheng med hverandre.

a) Korrelasjonen mellom de to variablene er  $r = -0.50$ . Hva betyr det?

Kandidaten bør forklare prinsippet om samvariasjon. Bra om kandidaten viser til at korrelasjon er basert på naturlig variasjon til forskjell fra eksperimentet som benytter påført variasjon under kontrollerte betingelser. I tillegg til å svare at det er negativ korrelasjon mellom variablene er det et pluss hvis kandidaten sier hvorfor og diskuterer styrkeforholdet til korrelasjonen. Naturlig å tenke at hvis voldsepisoder er  $x$  og IQ  $y$ , vil en negativ korrelasjon være at når  $x$  (voldsepisoder) går ned, går  $y$  (IQ) opp. Variablene går i motsatt retning. En  $r$  på  $0.50$  er på forelesning vurdert som en moderat korrelasjon. Er negativ, positive og nullkorrelasjon også forklart er det et pluss selv om det ikke spørres eksplisitt etter dette i oppgaven.

b) Hva forstår du med delt varians (Coefficient of determination) i denne sammenhengen, og hvor stor er den?

Kandidaten bør ha forstått at delt varians er å ta  $r^2$  ( $0.5 \times 0.5 = 0.25$ ). Dette betyr at de to variablene deler 25 % varians. I noen tilfeller er selve utregningen feil, men det skal ikke tillegges for stor vekt så lenge prinsippet er forstått. Det er et pluss hvis kandidaten sier noe om overgangen fra ordinalnivå til intervallnivå når vi går fra  $r$  til  $r^2$ , men det er ikke forventet.

c) Hvorfor skal man være forsiktig med å trekke slutninger om årsak og virkning fra korrelasjonsdata? Her står kandidaten fritt til å forklare med egne ord. Vi har snakket mye om å skille mellom korrelasjon og kausalitet, samt korrelasjonsmetoden begrensninger: det at årsaksretning kan gå begge veier, tredjevariabler og spuriøse sammenhenger. Ved antagelser om årsak må rivaliserende hypoteser utelukkes da disse kan forklare korrelasjonen, og det er vanskelig å gjennomføre i praksis når man studerer sammenhenger som opptrer i en naturlig setting.

d) I dette eksempelet er «Pearsons  $r$ » benyttet. Kjenner du andre korrelasjonskoeffisienter?

Flere korrelasjonskoeffisienter er nevnt i pensum. Mest diskutert er Spearman's rho eller Spearman Rank Order korrelasjon (rho). Andre koeffisienter nevnt i pensum er Point-Biserial korrelasjon og Phi-koeffisienten. Det er pluss hvis studenten forklarer litt hva korrelasjonskoeffisienten brukes til. Rho benyttes når man studerer variabler som minst er på ordinalnivå, og når man ønsker å fastlegge om forholdet mellom variablene er monotont. Har man et datasett rangordner man først skårene (fra minst til størst verdi) Man lager en ny variabel der den laveste skåren får verdien 1, nest laveste verdien 2, osv. Deretter kan man de nye variablene i utregningen av rho. Utregning kan gjøres på samme måte som for Pearsons  $r$  og tolkes på tilsvarende måte. Point Biserial benyttes når man studerer samvariasjonen mellom en dikotom variabel og en kontinuerlig variabel. Ved å omkode den dikotome variabelen til 0 for den ene kategorien og 1 for den andre, kan formelen for utregning av Pearsons  $r$  benyttes. Inneholder de to kategoriene av den dikotome variabelen ulikt antall personer, må man være klar over at maksimumsverdien for korrelasjonskoeffisienten blir lavere enn 1. Koeffisienten reduseres i tillegg av at det bare er to verdier på den dikotome variabelen. Det er derfor lite lurt å omkode en kontinuerlig variabel til en dikotom variabel før man regner ut samvariasjonen. Phi benyttes når man studerer



sammenhengen mellom to dikotome variabler. På samme måte som for Point Biserial korrelasjon kan formelen til Pearsons r benyttes hvis man omkoder kategoriene til 0 og 1.

#### **Oppgave 4**

Disse deloppgavene er nært knyttet til hverandre, og besvarelsene kan lett gli over i hverandre. Vis forståelse hvis kandidaten hopper noe frem og tilbake i fremstillingen, eller forteller at dette er noe de allerede har besvart i forrige deloppgave.

Overføringseffekter (Carryover effect) blir sett på som en trussel mot validiteten i repeterte (within-subjects) design.

##### a) Hva menes med repeterte design?

Hovedkjernen er at dette er et design med samme person/gruppe i alle betingelser. Det er et pluss om kandidaten er inne på forskjellen mellom det at samme person/gruppe inngår i samtlige betingelser og på den måten inngår som kontroll for seg selv, kontra at forskjellige personer inngår i ulike betingelsene som i mellomgruppedesign med uavhengige grupper. Innengruppedesign brukes når det er kostbart å bruke ulike forsøkspersoner, eller ved mangel på forsøkspersoner. Fordelene er at det blir mindre variasjon som skyldes variasjon mellom personer, dvs. lettere å oppdage eksperimentell effekt.

##### b) Hva menes med ytre og indre validitet i eksperimenter?

Her skal de kun nevne validitet for eksperimenter. De må forklare at indre validitet handler om slutninger om årsaksbetingelser i et eksperiment. Ytre validitet er om resultatene har en generaliseringsverdi, at det også er gyldig for «verden utenfor». Pluss hvis de nevner at dette kan være motstridende i et eksperiment, jo strengere kontroll over confunderende variabler jo mindre systematisk feilvarians og bedre indre validitet, men det kan bidra til at generaliserbarheten blir dårligere fordi testsituasjonen kan bli kunstig ved ønske om høy kontroll (som i noen laboratoriestudier). Her er det et stort pluss om kandidaten ser at mellomgruppedesign med uavhengige grupper lettere oppnår god ytre validitet mens innengruppedesign har en fordel mht indre validitet.

##### c) Hva er overføringseffekter?

Ved repeterte design er det en risiko for at tidligere måling har smitteeffekt ved at erfaring fra en betingelse overføres til neste betingelse. Det påvirker resultatet i studien. Pluss hvis dette sees i sammenheng med confunderende variabler. Ulike typer av overføringseffekter bør være nevnt. Pensum nevner 6 stk spesifikt: treningseffekter, uttrøttingseffekter, habituering, sensitivisering, kontraster og adaptasjon. Overføringseffekter gir opphav til systematisk feilvarians som er et problem for indre/intern validitet.

##### d) Hva kan man gjøre for å redusere faren for overføringseffekter?

Her bør studenten kunne si noe om full motbalansering og delvis motbalansering, og problemet som oppstår hvis overføringen fra variabel A til variabel B ikke er den samme som fra B til A.

Boka diskuterer latin square design og muligheten for å omgjøre rekkefølgen av betingelser som en uavhengig variabel, samt hvordan man kan minimere faren for effekter gjennom gode design, for eksempel ved å gi treningsrunder før man samler inn data. I forhold til problemer med longitudinelle design er det et stort pluss hvis kandidaten har tatt opp

såkalte kohort- sekvensielle design (kombinasjon av longitudinelle og tverrsnittdesign), men det er ikke forventet.

Lykke til med sensuren av eksamensoppgavene!

Beste hilsen,

Maria Teresa Grønning Dale og Fartein Ask Torvik