

Noen momenter ved vurdering av eksamen PSY1010 – PSYC1100 høsten 2018.

Generelt:

Denne veiledningen peker på noen elementer 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. I slike tilfeller ser man bort fra den siste oppgaven i besvarelsen.

Dersom kandidaten kun har besvart 2 oppgaver eller dersom en oppgave er klar stryk, skal hele eksamen anses som ikke bestått. Ellers kan total karakteren settes som et gjennomsnitt av karakterene på deloppgavene.

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.

Svar på 3 av 4 oppgaver

Oppgave 1

Se for deg at du ønsker å teste effekten av en ny behandlingsmetode for sosial angst. Femti pasienter som står på venteliste for behandling trekkes tilfeldig ut og får prøve den nye metoden. De som ikke trekkes ut fortsetter å stå på venteliste inntil videre. Etter tolv uker måler du symptomer på sosial angst blant de femti som har mottatt behandling og blant femti kontroller som fortsatt står på venteliste og sammenligner angstsymptomer i de to gruppene.

a) Kan man trekke kausale slutninger på bakgrunn av studien? Hvorfor eller hvorfor ikke?

Nja. Det finnes argumenter for og mot, og det viktige her er at kandidaten drøfter disse. I utgangspunktet er dette lagt opp som et eksperiment med tilfeldig tildeling til gruppe, altså eksperimentell manipulering. Altså, all systematisk variasjon mellom gruppene skyldes den tilfeldige treningen. Det er likevel ikke et helt optimalt design. Det viktigste grunnen er at eksperimentet ikke er blindet. Både pasienten og behandleren vet at får behandling, og ikke står på venteliste. I tillegg kan det være ulikheter i hvorvidt personer i de ulike gruppene faller fra studien.

b) Hva er nullhypotesen og den alternative hypotesen? Hvordan bruker vi disse?

Nullhypotesen er at gjennomsnittet i angstsymptomer i de to gruppene er likt. Den alternative hypotesen er at de er ulike. Vi vurderer sannsynligheten av å data under scenariet at nullhypotesen er sann. Dersom sannsynligheten er tilstrekkelig lav, forkaster vi nullhypotesen. Flere trinn/begrep i hypotesetesting kan trekkes inn: Nullfordistribusjon, alfa, p-verdi, kritisk grense, konfidensintervaller osv.

c) Vil du gjøre en enhalet eller tohalet test? Begrunn svaret.

Vi er ute etter en tohalet test. Ved en enhalet test kan man ikke oppdage effekter som går i motsatt retning av det man forventet.

d) Hva slags statistisk test ville du brukt for å avgjøre om forskjellen mellom gruppene er statistisk signifikant?

Uparet t-test. Det er to grupper, og det er ikke de samme personene i de to gruppene.

e) Vi innfører en tredje gruppe, som mottok medikamentell behandling mens de ventet på psykoterapi. Hvordan tester vi nå om det er forskjeller mellom de tre gruppene?

Variansanalyse/ANOVA. Nærmere bestemt en-faktor mellomgruppe-variansanalyse. Grunnen til at vi bruker variansanalyse er at vi da kan gjøre én test i stedet for tre tester. Det reduserer den samlede feilraten.

Oppgave 2

Les først teksten til oppgave 1.

Symptomer på sosial angst måles på en skala som går fra 0 (ingen angst) til 40 (maksimal skåre). Ved studiens slutt er gjennomsnittet for angstsymptomer blant de som mottok behandling $\bar{x}_1 = 18$ og standardavviket $SD_1 = 6$. Blant de som ikke mottok behandlingen var gjennomsnittet for angstsymptomer $\bar{x}_2 = 22$ og standardavviket $SD_2 = 6$. Siden $SD_1 = SD_2$ er $SD_p = 6$. Som nevnt er $n_1 = n_2 = 50$. Under er noen formler du kanskje får bruk for:

$$SE = \frac{SD}{\sqrt{n}}$$

$$SE_D = SD_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t = \frac{\bar{x}_2 - \bar{x}_1}{SE_D}$$

a) Forklar hva standardavvik og standardfeil er. Hva er standardfeilen (SE) i hver av gruppene?
Standardavvik: 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 kan tolkes som verdien gjennomsnittlige avstand fra gjennomsnittet.

Standardfeil: Standardfeilen er et estimat av standardavviket i utvalgsfordelingen. Det viser dermed presisjonen i estimatet av gjennomsnittet (hvor langt unna det sanne gjennomsnittet vi kan forvente å komme). Honorer hvis kandidaten viser til normalfordelingskurven og den prosentvise fordelingen som ligger innenfor bestemte avstander til forventningen (for eksempel at 68,2% av fordelingen vil ligge +/- 1 SD fra gjennomsnittet for normalfordelte variabler).

Standardfeilen er $6/\sqrt{50} = 0,8485$. Dette er likt i begge gruppene (som har lik SD og lik n).

b) Hva er standardfeilen for forskjellen mellom gruppene (SE_D)? Hva betyr dette?
 Formelen over viser at standardfeilen for forskjellen er $SE_D = 6 \cdot \sqrt{1/50 + 1/50} = 1,2$.

Tolkes som andre standardfeil. Viser presisjon i estimatet av forskjellen mellom gruppene. Er standardavviket i utvalgsfordeling av forskjell. (Kan f.eks. eksemplifiseres med simulering eller repetisjon av tilsvarende studie forventer å observere en forskjell mellom gruppene som ligger ca. 1,2 fra den sanne forskjellen.)

c) Regn ut t-verdien og fortell hva den betyr.

$$t = (22 - 18) / 1,2 = 3,33$$

Tilsvarende en z-fordeling (normalfordeling), men har tyngre haler. Ettersom n er godt over 30 er avvik fra normalfordelingen ubetydelig. Vil si at den observerte forskjellen er 3,33 ganger større enn standardfeilen for forskjell. Kan konverteres til p-verdi ved å angi antall frihetsgrader (her 98).

d) t-verdien tilsvarende en p-verdi på 0,0012. Hva forteller dette oss?

Sannsynligheten for å observere like eller mer ekstreme resultater dersom nullhypotesen er sann. Hvis nullhypotesen var sann ville man i 0,12 % av tilfellene funnet like stor eller større forskjell mellom gruppene. (Noen kommer kanskje til å svare at dette er sannsynligheten for at nullhypotesen er feil. Et slikt svar skal man trekke for.) Alfa er ikke oppgitt. Dette gir kandidaten mulighet til å reflektere rundt p- og alfa. Her kan det trekkes inn at konvensjonelt er alfa 0,05 eller 0,01, og at resultatene er statistisk signifikante ved disse konvensjonene. Det vil si at vi forkaster nullhypotesen og antar at det ER en forskjell i angstsymptomer mellom de som mottok og de som ikke mottok behandling.

OPPGAVE 3

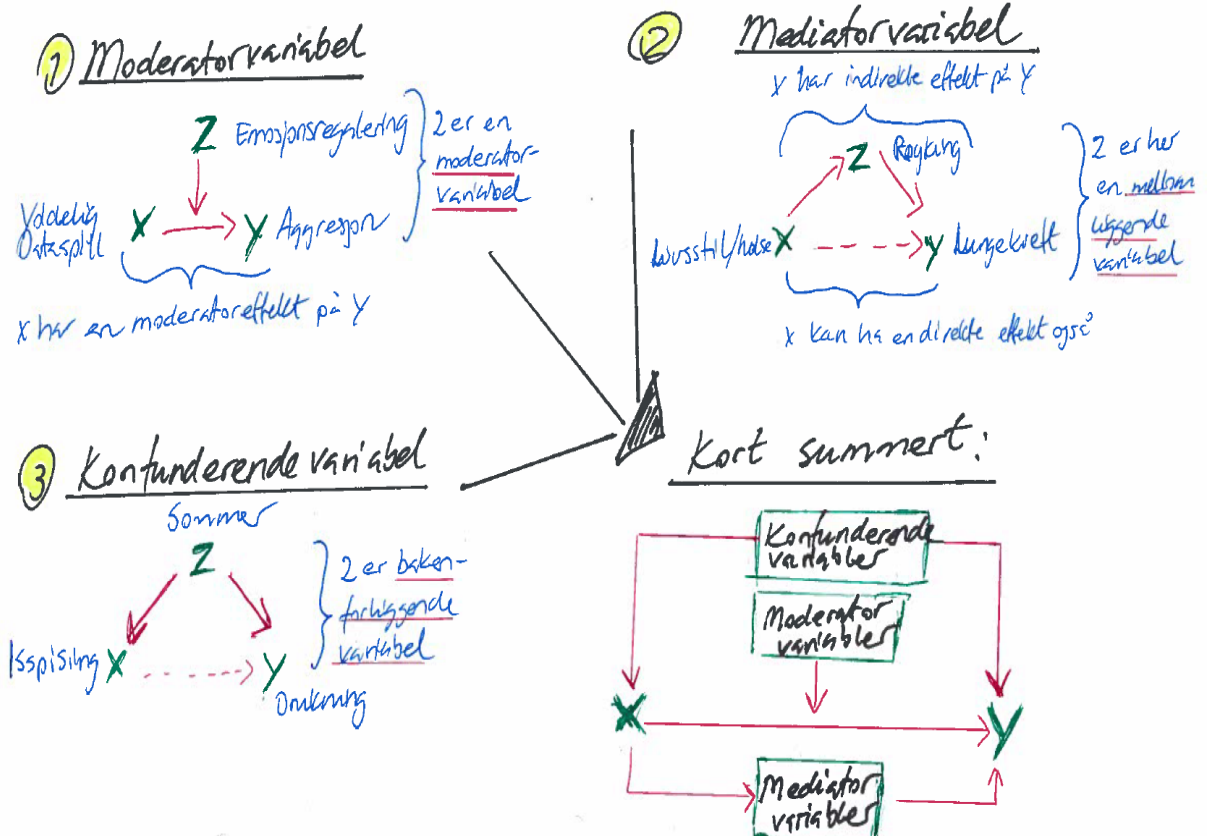
Et dansk studie viser at det er en sammenheng mellom storkebestanden i et område og antall barn som fødes i Danmark.

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

Kandidaten bør forklare prinsippet om samvariasjon. I dette eksempelet er det naturlig å diskutere kausalitet vs. korrelasjon. 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 positiv korrelasjon mellom variablene er det et pluss hvis kandidaten sier hvordan (at begge går i samme retning: at begge øker eller minker) og diskuterer styrkeforholdet til korrelasjonen. Kandidaten kan selv velge hva som er x og y , da sammenhengene kan gå begge veier ved korrelasjon. En r på 0,50 er på forelesning vurdert som moderat korrelasjon, men dette er kun konvensjoner.

- b) En observert sammenheng mellom variabel x og variabel y kan skyldes innvirkning fra en tredje variabel: z . Forklar hvordan z kan virke inn i eksempelet med antall storker og antall barn som fødes i et område.

Ved antagelser om årsak må rivaliserende hypoteser utelukkes da de kan forklare korrelasjonen. Kandidaten bør her derfor vise forståelse for forholdet til tredjevariabler. Her er det snakk om at z (det å bo på landet) virker som en konfunderende variabel (se figur 3). Det kan være at familier på landet får flere barn enn familier i byene, og det korrelerer med at det er på landet storkene også oppholder seg. Det er også mulig at sammenhengene mellom x og y er helt vilkårlig og uavhengig av hverandre, da snakker vi om en spuriøs sammenheng.



- c) Hvorfor kan regresjonsanalyse ses på som en videreføring av korrelasjon?
Regresjonsanalyse sier noe om forholdet mellom variabler, hvor det er det mulig å predikere verdien på en variabel ved hjelp av en annen (MEN ikke nødvendigvis kausalitet). Her er prediktor den uavhengige variabelen x og kriteriet den avhengige variabelen y , hvor man spør: Hva skjer med y når x øker med én?
- d) Hva menes med prediktorvariabel og kriterievariabel i en regresjonsanalyse? Vis med eksempler.
Prediktorvariabelen er uavhengig variabel x , og kriteriet er den avhengig variabel y . Her kan alle tenkelige variabler fungere som eksempel.

OPPGAVE 4

Singel-Subject Design ($N=1$ design) benytter én (1) eller svært få deltagere.

- a) Kjenner du til noen eksempler på Singel-Subject Design innen psykologisk forskning.
Her er det fort gjort å hente eksempler fra kvalitative kasusstudier (observasjon i naturlige omgivelser) og ikke fra psykologiske eksperimenter. De kan for eksempel nevne eksperimentene til Thorndike, Ebbinghaus, Skinner og Pavlov.
- b) Hvilke fordeler og ulemper har dette designet i forhold til indre- og ytre validitet?
 $N=1$ design ligner within-subject design (alle subjekter i eksperimentet gjennomgår alle nivåer av en variabel), men her sammenlignes ikke gj.snitt med andre subjekter. For single-subjects designs benyttes gjennomsnittet av mange målinger på samme person/dyr hvor subjektet er sin egen kontrollgruppe. Slik reduseres feilvarians. Stabilitetskriteriet hos forsøkspersonen er viktig, og god kontroll på ytre omgivelser for å trekke riktige konklusjoner om årsak-virkning (god indre validitet). Ulempe er carry-over effekter og problemer med gjentatte målinger over tid. Dette designet har problemer med generaliserbarheten (ytre validitet).
- c) Variasjon i et eksperiment er essensielt: Hva menes med planlagt systematisk varians?
I eksperimenter manipulerer man en betingelse og undersøker virkningen av dette med hensikt å påvise hvordan en variabel (x) påvirker en annen variabel (y). Dersom man finner variasjon, dvs endringer i målingene, ønsker man at variasjonen kun skal skyldes vår planlagte systematisk manipulasjon, som er den uavhengige variabelen (x). Her kan kandidaten også diskutere hvordan dette er forskjellig fra annen varians som ikke skyldes vår manipulerings: usystematisk varians (oppstår tilfeldig) og systematisk feilvarians (andre ukontrollerte ytre variabler).
- d) Hva er et ABAB-design?
 $N=1$ design som brukes mye innen læringspsykologi. Her kan kandidaten utdype og komme med egne eksempler. Det er viktig å definere hva fasene betyr og hva som er hensikten med dette designet.
A: Baseline fase, den naturlige forekomsten av det vi studerer.
B: Intervensjonsfasen, her ser vi hvordan manipuleringen endrer adferden.

A (2.gang): Tilbakeholder intervensjon

B (2.gang): Introduserer intervensjon/manipulering på nytt

Lykke til med sensuren av eksamensoppgavene!

Beste hilsen,

Maria Teresa Grønning Dale og Fartein Ask Torvik