

**Noen momenter ved vurdering av eksamen
PSY1010 – PSYC1100 våren 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 anses eksamen som ikke bestått. (Det samme gjelder dersom svaret på en av deloppgavene er så mangelfullt at man kan anse den for ubesvart). 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.

OPPGAVE 1

Tenk deg at du skal gjennomføre en studie på effekten | av fysisk trening på sosial angst. I forbindelse med dette ønsker du også å utvikle et spørreskjema som måler symptomer på sosial angst.

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

Fysisk trening er uavhengig variabel, symptomer på sosial angst er avhengig variabel. Ufyllende svar kan gjerne honoreres.

b) Hvorfor er det viktig å undersøke reliabilitet når man utvikler tester, og hvordan kan man gjøre dette?

Oppgave kan besvares på flere måter. Kandidaten bør kunne si noe om viktigheten av stabilitet i målingen for at resultatet skal kunne betraktes som et godt mål på et fenomen. Det er et pluss om kandidaten tar med forholdet mellom reliabilitet og validitet. (Kan også diskuteres under c) Har kandidaten forstått at en test kan være stabil uten at den måler det man mener at den skal måle? Pluss hvis kandidaten har forstått at man ikke opererer med absolutter. eks: en vekt kan være et valid måleinstrument for høyde hvis man bare er ute etter grove sammenhenger. Ulike former for reliabilitet kan også diskuteres og når dette gjøres relevant for spørsmålet bør det honoreres.

c) Hvorfor er validitet viktig når man utvikler tester og hvilke former for validitet bør man vurdere?

Kandidaten bør diskutere viktigheten av man faktisk måler det man skal måle. Følgende former for validitet er nevnt i pensum. Face validity, Kriteriebasert validitet, samtidig validitet, prediktiv validitet, innholdsvaliditet og begrepsvaliditet. Det er også i orden om kandidaten nevner indre og ytre validitet, men det er knyttet til slutninger, og ikke tester, og derfor ikke så viktig her.

d) Hvilke utfordringer er knyttet til det at deltakerne kjenner studiens formål? Diskuter.

Her forventes det at kandidaten diskuterer forhold knyttet til indre validitet. Spesielt vil det være viktig om reaktivitet, og demand characteristics nevnes og diskuteres.

OPPGAVE 2

Du gjennomfører en studie av effekten av fysisk trening på sosial angst. Det er totalt 40 deltakere i studien. Halvparten ($n_1=20$) følger ikke noe spesielt program, mens den andre halvparten ($n_2=20$) gjennomfører et opplegg med hard fysisk trening tre ganger i uka. Etter en måned svarer de på et spørreskjema som måler symptomer på sosial angst. Hver deltaker får en skåre fra 1 (få symptomer) til 7 (mange symptomer).

I gruppe 1, som IKKE gjennomførte treningen, var gjennomsnittet (\bar{x}_1) 3,0 og variansen (s_1^2) 2,25.

I gruppe 2, som gjennomførte treningen, var gjennomsnittet (\bar{x}_2) 2,5 og variansen (s_2^2) 2,25.

Under er noen formler du kan få bruk for i denne oppgaven:

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

$$SD = \sqrt{s^2}$$

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

$$SE(D) = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t = \frac{D}{SE(D)} = \frac{\bar{x}_1 - \bar{x}_2}{SE(D)}$$

a) Hvilken statistisk test ville du benyttet for å teste forskjellen mellom gruppene? Gi en kort begrunnelse.

Her bør man si at man skal bruke en uparet t-test og at det er fordi man undersøker to grupper, og at det er fordi det er ulike personer i de to gruppene. I tillegg er det fint om kandidaten nevner at vi bør bruke en tohalet test.

b) Forklar kort hva varians (s^2), standardavvik (SD) og standardfeil (SE) betyr og regn ut standardavviket og standardfeilen for de to gruppene.

Her trengs kun korte definisjoner, 1-3 setninger per begrep.

Varians: Kandidaten bør vise forståelse for at man at dette er mål på spredningen av skårene. Oppgaven etterspør en kort forklaring, så mye mer er derfor ikke nødvendig, men ulike ting det går an å ta med (uten at man trekker om dette ikke er med): Varians kan aldri være negativ; varians består av ulike komponenter – sann variasjon og varians som skyldes målefeil, som igjen kan bestå av systematisk og tilfeldig målefeil; man kan studere ulike kilder til varians, som i variansanalyse, forklart varians osv.

Standardavvik: Det bør framgå at kandidaten forstår at standardavviket er et mål på spredning som kan overføres direkte til den skalaen målingen er gjort på, og bra om det nevnes at standardavviket er roten av variansen. Mange vil nok forklare dette som (absolutt) gjennomsnittlig avvik fra gjennomsnittet – en slik tolkning er ok å ha med, men bonus til de som forstår at SD ikke faktisk helt sammenfallende med den tolkningen.

Standardfeil: Kandidaten bør vise at han/hun har forstått at standardfeilen er knyttet tilfeldig variasjon/hvor langt unna det 'sanne' gjennomsnittet vi kan forvente å komme/at det er et mål på feilmargin. Videre skal det honoreres om kandidaten redegjør for som at standardfeilen er beregnet på bakgrunn av standardavviket i utvalget og utvalgsstørrelsen og kan forstås som standardavviket i utvalgsfordelingen.

Standardavviket er $\sqrt{s^2} = \sqrt{2,25} = 1,5$
Standardfeilen er $SD/\sqrt{n} = 1,5/\sqrt{20} = 0,335$

c) Regn ut standardfeilen for differansen mellom gruppene ($SE(D)$) og t . Forklar kort hva disse begrepene viser til. Obs: Siden variansen er lik i gruppene, er $s_p = SD$, som du regnet ut i oppgave 2b.

Standardfeil for differanse: $SE(D) = s_p \cdot \sqrt{1/n_1 + 1/n_2} = 1,5 \cdot \sqrt{1/20 + 1/20} = 1,5 \cdot \sqrt{1/10} = 0,473$

Regn ut t : $(x_1 - x_2) / SE(D) = (3,0 - 2,5) / 0,473 = 1,054$ [1,05 – 1,06 avhengig av avrunding underveis]

Standardfeil for differansen: Hvor stor forskjell vi kan forvente mellom gruppene selv om det ikke finnes noen forskjell i virkeligheten (under nullhypotesen).

t : Plassering langs t -fordelingen; avvik fra forventet verdi (0) målt i standardfeil. Hvor ekstrem skåren er. Jo høyere skåre desto mindre sannsynlig å oppstå når det ikke finnes noen sammenheng i virkeligheten.

d) Du slår opp i en tabell og finner at t -verdien i dette tilfellet tilsvarer en p -verdi på 0,295. Hva betyr dette? Hvordan ville du konkludert om studien på bakgrunn av dette?

Dette vil si at dersom nullhypotesen er sann, vil man observere resultater like eller mer ekstreme i 0,295 (altså 29,5%) av tilfellene. Dette betyr at resultatene ikke er statistisk signifikante. (Alfa er ikke oppgitt, men studentene bør vite at den aldri settes så høyt at dette ville vært signifikant.) Det vil si at vi ikke forkaster nullhypotesen, og at vi ikke kan konkludere med at det er noen forskjell i symptomnivå mellom de som gjennomførte eller ikke gjennomførte trening. Hvis studentene trekker inn styrke/power og at man kan øke utvalgsstørrelsen vil det også være en relevant betraktning.

OPPGAVE 3

En nyhetssak i VG viser at røyking og skolefravall i ungdomsskolen ser ut til å ha en statistisk sammenheng med hverandre.

- Korrelasjonen mellom de to variablene er $r = 0.40$. Hva betyr det?
Kandidaten bør forklare prinsippet om samvariasjon. I tillegg til å svare at det er positiv korrelasjon mellom variablene er det et pluss hvis kandidaten sier at ved positiv korrelasjon vil variablene gå i samme retning, at både x og y øker eller minker sammen. Kandidaten kan selv velge hva som er x og y , da sammenhengen kan gå begge veier ved korrelasjon.
- Forklar hvorfor vi skal være forsiktige med å gi sammenhengen mellom røyk og skolefravall en kausal fortolkning her.

Selv om man har dokumentert den statistiske sammenhengen, altså korrelasjonen, skal man være forsiktig med å gi sammenhengen en kausal fortolkning fordi:

1) Kausaliteten kan gå motsatt vei: For eksempel kan det være at de som sliter på skolen begynner å røyke, fordi det roer ned og virker avstressende. Eller at man blir sosialt usikker av å ikke mestre skolen og begynner å røyke for å øke selvfølelsen.

2) En tredje variabel kan påvirke både ungdommenes røyking og skolefravall. For eksempel: personlighet, livstil, sosioøkonomisk status som foreldres utdanning/lønn. De som har høy skåre på disse faktorene investerer mer i utdanning, karrierevalg og helse, og opplever samtidig større mestring på skolen.

- c) Du ønsker å predikere ungdomsskoleelevers skoleprestasjon basert på deres IQ ved lineær regresjon. Hva blir x- og y- variabel her?

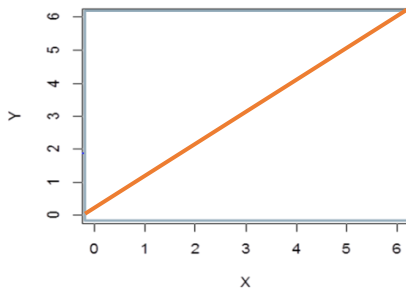
Med regresjonsanalyse er det mulig å fastsette verdien av y når man kjenner verdiene på x-variabelen. Prediktorvariabel er den uavhengige variabelen x, mens kriterievariabel er den avhengige variabelen y. Hvis jeg skal måle om intelligens kan predikere skoleprestasjon, blir skoleprestasjon = y og intelligens = x.

- d) Ligningen for bivariat lineær regresjon kan skrives som: $\hat{Y} = a + bX$.

1. Tegn regresjonslinjen for sammenhengen mellom skoleprestasjon og IQ når:

$$a = 0$$

$$b = 1$$



2. Prediker y ved avlesning når x= 5.

SVAR: y = 5

OPPGAVE 4

I eksperimentell forskning snakker vi om to hoveddesign: mellomgruppe (Between-subjects) og innengruppe- eller repeterte design (Within-subjects).

a) Når benyttes de to designene og hvilke fordeler og ulemper har de?

Her er det mulig å skrive mye, men hovedkjernen er at mellomgruppe er når ulike grupper av deltagere inngår i hver betingelse/manipulering, og innengruppedesign med samme person i alle betingelser. Det er et pluss om kandidaten er inn på at forskjellen mellom det at samme person inngår i samtlige betingelser og på den måten inngår som kontroll for seg selv kontra at forskjellige personer inngår i ulike betingelsene. Mellomgruppedesign med uavhengige grupper brukes når vi har lett tilgang på deltakere og når deltakerne ikke er veldig forskjellige på relevante variabler. 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. Som eksempler på ulemper/fordeler kan også studenten også trekke inn ulike svakheter/trusler mot den indre validiteten som gjelder for hvert av de to ulike designene.

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. Pluss hvis de klarer å trekke inn at vi ønsker å minimere systematisk feilvarians som er et problem for indre/intern validitet. Ytre validitet handler om at resultatene har en generaliseringsverdi. Pluss hvis de nevner at dette kan være motstridende i et eksperiment, jo strengere kontroll over konfunderende variabler jo bedre indre validitet, men det gjør ofte generaliserbarheten dårligere fordi testsituasjonen kan bli kunstig som i laboratoriestudier.

c) Gjør kort rede for hovedforskjeller mellom et korrelasjonsdesign og et eksperiment. Her må kandidaten gjøre rede for hovedforskjellen vedr årsaksforhold og korrelasjonelle forhold, og ta opp naturlig variasjon versus påført variasjon under kontrollerte betingelser. Viktig å trekke frem korrelasjonsmetodens begrensninger: 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 når man studerer sammenhenger som opptrer i en naturlig setting.

d) Forklar hvorfor overføringseffekter (Carryover effects) blir sett på som en trussel mot validiteten i repeterte design (Within-subjects).

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 konfunderende variabler. Ulike typer av overføringseffekter bør være nevnt. Pensum nevner seks spesifikt: treningseffekter, uttrøttingseffekter, habituering, sensitivisering, kontraster og adaptasjon. Overføringseffekter gir opphav til systematisk feilvarians som er et hovedproblem for indre validitet.

Lykke til med sensuren av eksamensoppgavene!

Beste hilsen,

Maria Teresa Grønning Dale og Fartein Ask Torvik