

PSY2014 - Kvantitativ metode

Skriftlig skoleeksamen onsdag 21. mai, 09:00 (3 timer).

Kalkulator uten grafisk display og tekstlagringsfunksjon er tillatt. En liste av relevante formler og en tabell av t-fordelingen er gitt på slutten av oppgaven. Forsøk å besvare alle oppgavene.

Digitale håndtegninger

- Du kan svare med digital håndtegning på denne eksamenen.
- Du kan kun bruke skisseark du får utdelt.
- Du kan bruke flere ark per oppgave.
- Se instruksjon for utfylling av skisseark (vedlagt).
- Du må bruke **blå** eller **svart penn**.
- Du får IKKE ekstra tid til å fylle ut informasjonsboksene på skisseark (engangskoder, kand.nr. o.l.).

Oppgave 1: Tekstforståelse og bruk av mobile enheter (65%)

I denne oppgaven vil vi se på ulike prediktorer for tekstforståelse hos grunnskolebarn (5-12 år).

TEKSTFORSTÅELSE: Høyere verdier antyder bedre forståelse av sammenhenger i lengre tekstavsnitt, og dette vil være den avhengige variabelen i alle analyser.

Følgende uavhengige variabler vil bli brukt i regresjonsanalysene:

MOBILTIMER: Antall timer i uken barnet bruker mobile enheter (smarttelefon/nettbrett).

ALDER: Alder i år

ALDER2: En variabel gitt av $ALDER * ALDER$ (altså alder kvadrert).

KJØNN: Gutter kodet 0, jenter kodet 1.

IQ: Mål på generell intelligens

AKTIVITET: Gjennomsnittlig antall timer fysisk aktivitet barnet bedriver i uken.

SØVN: Gjennomsnittlig antall timer barnet sover per natt.

a) I den første regresjonsmodellen er MOBIL den eneste uavhengige variabelen.

1. Hvordan ville du oppsummere forholdet mellom MOBILTIMER og TEKSTFORSTÅELSE basert på modell 1?

Av det som kan trekkes frem er blant annet:

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,036 ^a	,001	-,004	4,660766078	,001	,253	1	198	,615

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	14,447	,810		17,828	,000	12,849	16,045
	MOBILTIMER	-,120	,238	-,036	-,503	,615	-,590	,350

- Punktestimat og tolkning av \hat{b}_0 (14.447) og \hat{b}_1 (-0.120). De som skårer 0 på MOBILTIMER har en forventet TEKSTFORSTÅELSE på 14.447. Altså en negativ

sammenheng mellom MOBILTIMER og TEKSTFORSTÅELSE som ikke er signifikant ($p=0.615$), noe som også kommer til uttrykk i at konfidensintervallet rommer 0, i en ikke-signifikant F-verdi, og en ikke-signifikant økning i R^2 fra en modell uten uavhengige variabler. Modellen forklarer 0.001% av variasjonen i tekstforståelse, Pearson korrelasjonen er 0.036, (justert R^2 er <0.0).

- I oppsummering, lite evidens for en sammenheng mellom variablene.
2. Hva er forventet skåre på TEKSTFORSTÅELSE for et barn som bruker mobile enheter i gjennomsnitt 5 timer i uken?

$$Y=14.447+(-0.120)*5=13.847$$

3. Kolonnen til høyre for koeffisientene i "Coefficients"-tabellen har betegnelsen "Std.Error". Hva refererer tallene i denne kolonnen til, og hvilke forhold kan påvirke Std.Error til MOBILTIMER?

Std.Error (standardfeilen) er et estimat på standaravviket i samplingdistribusjonen av koeffisientene. StdError for MOBILTIMER er således et estimat på standardviket i fordelingen av estimater på effekten av mobiltimer vi ville sett ved gjentatte samplinger fra den samme populasjonen.

Uttrykket for standardfeil av stigningskoeffisienten i en bivariat regresjon er gitt i formelarket. Den er bestemt av tre forhold:

- Antall observasjoner (n)
- Residualvariansen i tekstforståelse (som bestemmes av påvirkning av alle ikke målte forhold, inklusiv målefeil i tekstforståelse).
- Variansen i prediktoren (MOBILTIMER). Større variabilitet i MOBILTIMER ville gi mindre standardfeil for b_1 .

b) I modell 2 blir KJØNN og ALDER lagt til som uavhengige variabler.

1. Beskriv kort ulike grunner man kan ha for å inkludere mer enn en uavhengig variabel i en regresjonsmodell.

Her er det mange ting som kan nevnes:

- Bedre prediksjon av den avhengige variabelen.
- Redusere standardfeilen til en parameter ved å legge til andre uavhengige variabler som reduserer den uforklarte variansen i den avhengige variabelen.
- Bedre forståelse av effekten av en eller flere uavhengige variabler gjennom kontroll for konfunderende variabler.
- Modellering av mer komplekse sammenhenger mellom variablene (mediering, interaksjon).
- Modellere effekten av en kategorisk variabel ved å kode den til et sett dummyvariabler.

2. Hvorfor har den estimerte effekten av MOBILTIMER endret seg når ALDER og KJØNN er lagt til?

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	14,447	,810		17,828	,000	12,849	16,045
	MOBILTIMER	-,120	,238	-,036	-,503	,615	-,590	,350
2	(Constant)	15,103	,783		19,282	,000	13,558	16,648
	MOBILTIMER	-,522	,217	-,155	-2,400	,017	-,951	-,093
	ALDER	,848	,114	,481	7,418	,000	,623	1,074
	KJØNN	1,328	,585	,142	2,271	,024	,175	2,481

MOBILTIMER blir signifikant når ALDER og KJØNN legges til.

- ALDER forklarer en del av variabiliteten av MOBILTIMER ($r=0.257$), og ALDER er også korrelert med TEKSTFORSTÅELSE ($r=0.435$). Således er ALDER en konfunderende variabel. Teknisk sett er ALDER også en supressor, da effekten av MOBILTIMER blir sterkere og mer signifikant når ALDER kontrolleres for.
 - KJØNN er korrelert med TEKSTFORSTÅELSE, men ikke med MOBILTIMER. Det vil ha lite påvirkning på TEKSTFORSTÅELSE å kontrollere for KJØNN, men siden det vil redusere den uforklarte variabiliteten i utfallsvariabelen å kontrollere for KJØNN, kan det gjøre standardfeilen til MOBILTIMER mindre (og p-verdien lavere).
3. Hva er tolkningen av den partielle regresjonskoeffisienten for kjønn? Ville de partielle regresjonskoeffisientene eller R^2 endre seg dersom jenter var kodet 0 og gutter 1? (begrunn svaret).

- Tolkningen: Siden KJØNN er en dummyvariabel, og gutter er kodet 0/jenter 1, er tolkningen «*hvor mye høyere/mindre jenter skårer på tekstforståelse*» (1.328 enheter høyere).
- Dersom kodingen var reversert ville \hat{b}_0 endre seg (øke til nivået for jenter), mens \hat{b}_1 ville endre fortegn (til hvor mye lavere gutter skårer på tekstforståelse). R^2 ville ikke endre seg, da modellen ville passe like godt (heller ikke justert R^2).

c) I modell 3 blir den uavhengige variabelen ALDER2 lagt til. Hvordan ville du beskrive sammenhengen mellom alder og TEKSTFORSTÅELSE basert på estimatene fra modell 3?

- Dette er en regresjonsmodell med et annengradsledd for alder. En slik modell (polynomisk regresjon) kan fange opp ikke-lineære (her kvadratiske) sammenhenger mellom alder og tekstforståelse. Siden estimatet for ALDER2 er negativ betyr dette at stigningen i tekstforståelse som funksjon av alder flater ut ettersom barnet blir eldre. På et tidspunkt vinner annengradsleddet over det lineære leddet, og forventet tekstforståelse begynner å synke.

d) I modell 4 blir de uavhengige variablene AKTIVITET, SØVN, IQ lagt til. Noen av tallene i SPSS-utskriften fra denne modellen er sladdet.

1. Hvilke av de uavhengige variablene i modell 4 er statistisk signifikant assosiert med TEKSTFORSTÅELSE på et 0,05-nivå? (Begrunn svaret)

4	(Constant)	12,413	2,852		4,353	,000	6,788	18,038
	MOBILTIMER	-,511		-,152	-2,399			
	ALDER	,891		,505	7,830			
	KJØNN	1,267	,572	,136				
	ALDER2	-,109		-,149				-,018
	IQ	,054		,180		,004	,018	,090
	AKTIVITET	-,022	,154	-,009			-,325	,281
	SØVN	-,193		-,055		,366		

a. Dependent Variable: TEKSTFORSTÅELSE

4	(Constant)	12,413	2,852		4,353	,000	6,788	18,038
	MOBILTIMER	-,511	,213	-,152	-2,399	,017	-,930	-,091
	ALDER	,891	,114	,505	7,830	,000	,667	1,116
	KJØNN	1,267	,572	,136	2,217	,028	,140	2,395
	ALDER2	-,109	,046	-,149	-2,373	,019	-,200	-,018
	IQ	,054	,018	,180	2,948	,004	,018	,090
	AKTIVITET	-,022	,154	-,009	-,141	,888	-,325	,281
	SØVN	-,193	,213	-,055	-,906	,366	-,614	,228

a. Dependent Variable: TEKSTFORSTÅELSE

- **MOBILTIMER:** T-verdien er mindre enn -2. Utvalget er så stort at vi kan bruke de kritiske verdiene for en Z-fordelt variabel (+2, eller -1.96 for å være helt nøyaktig). **SIGNIFIKANT**
- **ALDER:** T-verdien >> 2, sterkt **SIGNIFIKANT**.
- **KJØNN:** Vi er ikke gitt t-verdien, men kan regne den ut ved $t = b_i / SE(b_i)$. $1.267 / 0.572 = 2.215$. Dette er >2, så kjønn er **SIGNIFIKANT**
- **ALDER2:** Øvre grense for konfidensintervallet er <0. Intervallet inneholder ikke 0, så effekten er **SIGNIFIKANT**
- **IQ:** P-verdien er gitt direkte (0.004). Dette er mindre enn 0,05, så IQ er **SIGNIFIKANT**.
- **AKTIVITET:** Konfidensintervallet rommer 0 så **IKKE SIGNIFIKANT**
- **SØVN:** P-verdien gitt direkte, (0.366), så **IKKE SIGNIFIKANT**

2. Dersom du lagde en figur av de standardiserte residualene fra modell 4, hva ville du lete etter som tegn på at en underliggende antagelse ved lineær regresjon var brutt?

Vi har sett på scatter plots og histogrammer av residualer, (og i noen grad vist til QQ-plots). I plottene leter vi etter:

- Ikke-linearitet (brudd på antagelse om at forholdet mellom TEKSTFORSTÅELSE og de ulike uavhengige variablene er lineært).
- At residualene er normalfordelte (best vurdert gjennom et histogram).
- Konstant varians i residualene

- Tegn på avhengighet mellom residualene (liket i verdien på residualene med liknende verdi på den uavhengige variabelen).
- Uteliggere

e) Vi skal nå se på to sett med regresjonsanalyser (sett 1 og sett 2), der SPSS-utskriften er vedlagt som pdf. I det ene settet er de uavhengige variablene svært lavt korrelert, mens i det andre er de svært høyt korrelert. Oppgaven her er å vise hvor i utskriften du kan se om de uavhengige variablene er høyt eller lavt korrelert.

For hvert sett er det først gjennomført en bivariat analyse med hver av de to uavhengige variablene enkeltvis, og så en multippel regresjonsanalyse med begge de uavhengige variablene inkludert.

De uavhengige variablene i sett 1 er ukorrelerte. Dette kan vi se ved:

- At R^2 i den multiple regresjonsmodellen er tilnærmet lik summen av R^2 til hver de bivariate analysene (ukorrelerte uavhengige variabler forklarer separate proporsjoner av variansen).
- Standardfeilen til både $X1a$ og $X1b$ i den multiple regresjonsmodellen er mindre enn den er i de bivariate regresjonsmodellene.
- VIF verdiene er lave

De uavhengige variablene i sett 2 er høyt korrelerte. Dette kan vi se ved:

- At R^2 i den multiple regresjonsanalysen er veldig mye mindre enn summen av R^2 i de to bivariate analysene.
- Standardfeilen til både $X2a$ og $X2b$ i den multiple regresjonsmodellen er mye større enn den er i de bivariate regresjonsmodellene.
- De diagnostiske statistikkene for multikollinearitet (VIF verdiene) er meget høye (godt over 5).
- Verken $X2a$ eller $X2b$ er individuelt signifikante, men F-verdien i ANOVA tabellen antyder at hypotesen om at begge stigningskoeffisientene er 0 i populasjonen kan forkastes.

Oppgave 2: Effekten av fargen på smertestillende tabletter (35%)

Å senke hånden ned i et kar med isvann oppleves av de fleste som svært smertefullt, og fordi det ikke er forbundet med vevsskade er dette en vanlig måte å indusere smerte i eksperimentelle studier.

I denne studien tar deltagerne en smertestillende tablett 10 min før de senker hånden ned i isvannet. Denne kommer i en av tre farger; Rød, gul eller grønn. Den avhengige variabelen i denne oppgaven er grad av smerte rapportert av deltagerne 30 sekunder etter at de senket hånden ned i vannet.

a) Først ble det gjennomført en enveis variansanalyse av farge av tablettene og smerte. Du finner SPSS-utskriften fra modell 1 øverst i det vedlagte PDF-dokumentet.

1. Gjør rede for rasjonalet bak F-testen, slik den benyttes i variansanalyse.

- Variabiliteten i de observerte skårene (TSS) kan deles inn i to komponenter, SSM og SSE. Hver har et assosiert antall frihetsgrader (df), og «mean squares» er gitt av SS/df . Når H_0 er sann, vil MSM og MSE begge være forventningsrettede estimater på populasjonsvariansen, og F-verdien vil ha en forventet verdi på 1. F-verdien er således gitt av ratioen av variabiliteten mellom gruppene over variabiliteten innad i gruppene. Dersom det er forskjeller på gjennomsnittsverdiene mellom gruppene i populasjonen vil variabiliteten mellom gruppene være større enn variabiliteten innad i gruppene, og F-verdien bli større enn 1. Den kritiske verdien for å forkaste H_0 med en gitt alfa kan finnes gjennom en tabell.

2. Fyll ut de sladdede verdiene i tabellen.

ANOVA

Smerte

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	238,576	2	119,288	2,864	,061
Within Groups	4873,643	117	41,655		
Total	5112,219	119			

- Vi bruker at $TSS = SS_b + SS_w$,
- Tilsvarende bruker vi at $df_b = \text{antall grupper} - 1$, altså 2, og $df_t = df_b + df_w$.
- $F = MS_b / MS_w$

3. Variansanalyse og lineær regresjon er begge medlemmer av klassen *generelle lineære modeller*. Forklar hvordan du kunne gå frem for å studere sammenhengen mellom fargen på tablettene og smerte ved hjelp av lineær regresjon.

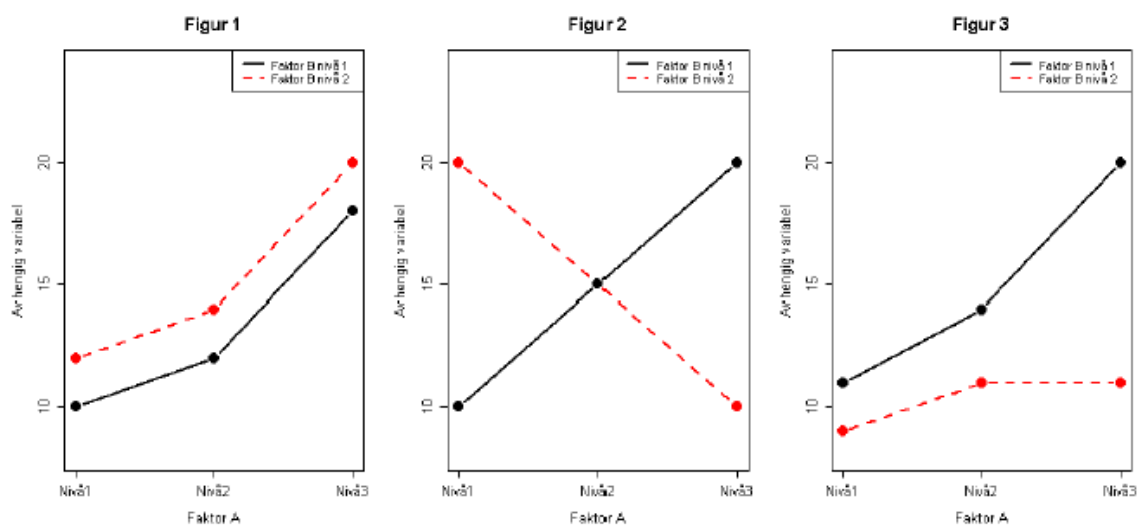
Farge er en kategorisk variabel. Denne kan ikke inkluderes som uavhengig variabel direkte i en regresjonsanalyse, men kan rekodes til et sett med dummyvariabler. F-verdien i ANOVA tabellen fra en modell med disse dummyvariablene som uavhengige variabler tester samme H_0 som variansanalyse, nemlig at dataene er trukket fra en populasjonen der gjennomsnittet er det samme i alle gruppene.

Det kan eventuelt vises til at antagelsene i de to modellene også er de samme (lik varians innad i hver gruppe, normalfordelte residualer, og uavhengige observasjoner).

b) I tillegg til å motta piller med ulik farge, var deltagerne i en av to grupper. Den ene mottok et medisinsk uvirksomt stoff (placebogruppen), mens den andre mottok en pille med identisk utseende, men som inneholdt det smertestillende stoffet paracetamol (eksperimentgruppen). Det ble så utført en toveis variansanalyse med faktorene FARGE og GRUPPE. Du finner resultatene i det vedlagte PDF-dokumentet.

- I vedlegget finner du tre linjediagrammer som viser eksempler på ulike scenarier som kan komme til uttrykk i 2-faktorielle eksperimenter (faktor A og faktor B). Sammenlikn de tre figurene, og beskriv hvilke tanker du gjør deg om de to faktorene A og B.

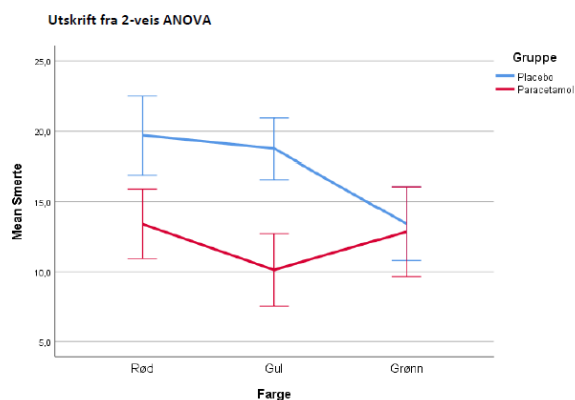
2b) Tre linjediagrammer



Begrepene hovedeffekt og interaksjonseffekt kan med fordel defineres.

- *Venstre: Hovedeffekter av både faktor A og B, ingen interaksjonseffekt.*
 - *Midten: Ingen hovedeffekt av verken A eller B, men en interaksjonseffekt.*
 - *Høyre: Hovedeffekter av både A og B, og antydning til interaksjonseffekt.*
- Under finner du SPSS utskrift fra en toveis-variensanalyse med faktorene FARGE og GRUPPE. Tolk resultatene, og sammenlikn med resultatene fra enveis-variensanalysen.

- *Linjediagrammet kan med fordel kommenteres. Det viser:*
 - *Sterk indikasjon på en hovedeffekt på tvers av GRUPPE (høyere snitt på smerte i placebogruppen enn paracetamolgruppen, og ikke overlappende konfidensintervall for Rød og Gul farge).*
 - *Potensielt en hovedeffekt av farge.*
 - *Potensielt en interaksjonseffekt, der lavere nivåer av smerte ved grønne tabletter stort sett i kun er tilstede i placebogruppen.*



- *Tabellen viser:*
 - *Det er signifikante forskjeller i rapportert smerte på tvers over de ulike fargene. (Signifikant hovedeffekt av farge, $F = 3.651$, $p=0.029$). Vi ser i figuren at smerten er høyest ved rød, og lavest ved grønn.*
 - *Svært signifikant hovedeffekt av GRUPPE ($F=24.606$, $p=0.000$). paracetamolgruppen rapporterer lavere smerte.*
 - *Signifikant interaksjonseffekt. Rapportert smerte på tvers av de ulike fargene er ikke den samme i placebo som paracetamolgruppen ($F=5.28$, $p=0.006$). Utifra figuren kan dette skyldes at grønn farge er forbundet med betydelig reduksjon i smerte relativt til de andre fargene, men kun i placebogruppen.*
- *Av betydning ser vi at FARGE ikke var signifikant i enveis variansanalysen ($p=0.06$), men signifikant i toveis analysen. Dette kan forklares ved at mye av innengruppe variansen i enveis analysen skyldes at noen fikk et reelt smertestillende (paracetamol) mens andre fikk en placebo. I toveis variansanalysen testes variabiliteten i smerte over de ulike fargene ($MS=119.28$) mot en residual MS på 41.655. Fordi GRUPPE forklarer en del av residualvariansen, har vi mer styrke til å påvise en effekt av FARGE i toveis analysen, og her testes $MS=119.28$ mot en residual MS på 32.67, hvilket gir en høyere F-verdi.*