

PSY2011 - Forskningsmetode II: Eksperimentell design og statistisk analyse, høst 2016.

Fredag 28. oktober, 09:00 (3 timer).

Det er 4 oppgaver - noen med deloppgaver. Alle oppgaver skal besvares.

Alle beregninger gjøres eksplisitte. Ettersom vi ikke har kalkulator tilgjengelig holder det at man beregner tallene "tilnærmet lik". Trenger dere for eksempel å beregne en ratio mellom tallene 16.35 og 4.12 holder det i massevis med: $16.35/4.12$ er ca. 4. Det vil imidlertid alltid være et pluss at man klargjør hvordan beregningene er utført.

En liste med relevante formler (på engelsk) er vedlagt på slutten av oppgaven, s. 13-14.

OBS! Tekst på begge sider av arkene.

Oppgave 1.

En forsker ønsker å undersøke en mulig effekt av et nytt spesialpedagogisk undervisningsprogram som skal øke barnas motivasjon for læring. Han velger ut 3 barn som får dette intensive undervisningstilbudet ("eksperimentgruppe") - og 3 barn som fortsetter med det gamle undervisningstilbudet ("kontrollgruppe"). Etter to måneder måles alle barna på et instrument som skal måle barnas motivasjon.

1.1. *Hva kaller man gjerne et forskningsdesign som dette?*

Resultatene ble slik:

Gruppe	Motivasjon
Kontroll	4
Kontroll	2
Kontroll	6
Eksperiment	8
Eksperiment	4
Eksperiment	6

1.2. *Gjennomfør en variansanalyse og fyll ut så mange som mulig av tallene som mangler i tabellen under. Sett den opp på nytt i besvarelsen. Trenger ikke være så pent bare det er leselig.....*

Kilde	SS	df	MSS	F	p
Between				1.5	0.288
Within					
Total					

Forskeren beregnet også η^2 (R^2). Den ble her **.27**.

1.3. *Er effekten "statistisk signifikant"?*

1.4. *Ut fra vanlig anvendte kriterier: er denne effekten sterk eller svak?*

1.5. *Hvordan vil du generelt konkludere med hensyn til effekten av intervensjonen?*

Forskeren konkluderte med at: "resultatene tyder på at det nye undervisningstilbudet ikke gir noen gevinst i forhold til det eksisterende".

1.6. *Hvorfor tror du han konkluderte slik? Og er dette en rimelig konklusjon?*

1.7. *Pek kort på noen viktige svakheter ved dette forskningsdesignet.*

Oppgave 2.

En annen forsker er lite tilfreds med denne konklusjonen, og bestemmer seg for å etterprøve denne ved å designe et bedre eksperiment. Han plukker ut 24 barn. Disse deles på samme måte som i forrige eksperiment i to like store grupper: en "kontrollgruppe" som bare fortsetter med det opprinnelig undervisningsopplegget og en "eksperimentgruppe" hvor det nye undervisningsopplegget implementeres. Motivasjonen måles imidlertid for alle barna både før intervensjonen implementeres (pre) og to måneder senere (post).

2.1. *Hva kaller man gjerne et slikt design?*

Resultatene ble nå slik:

Gjennomsnittlige motivasjonsskårer.

	Pre	Post	
Eksperiment	5	8	6.5
Kontroll	4	5	4.5
	4.5	6.5	5.5

Variansanalyse.

Effekt	SS	df	MSS	F	p
Tid	48	1	48	24.00	0.000
Tid*Gruppe	12	1	12	6.00	0.023
Error ₁	44	22	2		
Gruppe	48	1	48	9.60	0.005
Error ₂	110	22	5		

2.2. *Hvilken av "effektene" i variansanalysetabellen mener du er av størst interesse for forskeren her? Forklar hvorfor.*

2.3. *Hvordan vil du konkludere med hensyn til effekten av intervensjonen?*

2.4. *Pek på noen forhold som vil gjøre dette til et mye bedre forskningsdesign enn det som ble benyttet i oppgave 1.*

Oppgave 3.

"Stroop-effekten" er velkjent fra kognitiv psykologi (går jeg ut fra). I en studie av denne benyttet man stimuli presentert på en skjerm under tre ulike betingelser - og de samme 11 respondentene skulle i alle betingelser respondere med å oppgi den fargen som stod skrevet på skjermen.

Eksperimentbetingelser:

Kongruent - respondentene ble presentert for ord som beskrev en farge, og denne var skrevet i samme farge som den som skulle identifiseres (for eksempel ordet "rød" i rød skrift).

Inkongruent - respondentene ble presentert for ord som beskrev en farge, men dette var skrevet i en annen farge enn den som skulle identifiseres (for eksempel ordet "rød" i blå skrift).

Nøytral - alle ord i svart, og bare den navngitte fargen skulle identifiseres.

Under hver betingelse fikk personene en skåre som uttrykte antall riktig rapporterte farger.

	Eksperimentbetingelse				
	Kongruent	Inkongruent	Nøytral		
Gjennomsnittlig antall riktige:	8.29	5.00	6.90	6.73	
Effekt	SS	df	MSS	F	p
Eksperimentbetingelse	40.00	2	20.00	4.00	0.04
Person	60.00	10	6.00		
Eksperimentbetingelse*Person	100.00	20	5.00		

Kontraster	Differanse	SE	t	p	Hedge's G
Kongruent - Inkongruent	3.29	0.953	3.451	0.003	1.24
Kongruent - Nøytral	1.39	0.953	1.458	0.160	0.52
Inkongruent - Nøytral	-1.90	0.953	-1.993	0.060	-0.72

- 3.1. *Hvordan vi du raskt oppsummere resultatene over?*
- 3.2. *Vis ved å sette tall inn i formlene (se vedlegg) hvordan tallene i kolonnen SE og t er beregnet.*
- 3.3. *Hvordan bør man tolke verdiene i kolonnen "p" i tabellen over kontraster?*

Oppgave 4.

Forklar hva man i statistikken forstår med begrepet "samplingfordeling" (sampling distribution). Bruk noen av resultatene fra oppgave 3 til å illustrere og tydeliggjøre dette.

PSY2011 - Forskningsmetode II: Eksperimentell design og statistisk analyse, haust 2016.

Fredag 28. oktober, 09:00 (3 timer).

Det er 4 oppgåver – nokre med deloppgåver. Du skal svare på alle oppgåvene.

Alle utrekningar gjerast eksplisitte. Etersom vi ikkje har kalkulator tilgjengeleg held det at du reknar ut tala «tilnærma lik». Treng de til dømes å rekne ut ratio mellom tala 16.35 og 4.12 er det tilstrekkeleg med: $16.35/4.2$ er ca. 4. Det vil likevel alltid være eit pluss at du gjer greie for korleis du har gjort utrekninga.

Ei liste med relevante formelar (på engelsk) er lagt ved på slutten av oppgåva, s. 13-14.

OBS! Tekst på begge sidene av arket.

Oppgave 1.

Ein forskar ynskjer å undersøkje ein mogleg effekt av eit nytt spesialpedagogisk undervisningsprogram som skal auka borna sin motivasjon for læring. Han vel ut 3 born som får dette intensive undervisningstilbodet («eksperimentgruppa») – og 3 born som fortsett med det gamle undervisningstilbodet («kontrollgruppa»). Etter to månadar målar ein alle borna på eit instrument som skal måle borna sin motivasjon.

1.1. Kva kallar ein gjerne eit forskingsdesign som dette?

Resultata vart som følgjer:

Gruppe	Motivasjon
Kontroll	4
Kontroll	2
Kontroll	6
Eksperiment	8
Eksperiment	4
Eksperiment	6

1.2. Gjennomfør ei variansanalyse og fyll ut så mange som mogleg av tala som manglar i tabellen under. Sett den opp på nytt i svaret ditt. Treng ikkje være pent berre det er leseleg.

Kilde	SS	df	MSS	F	p
Between				1.5	0.288
Within					
Total					

Forskaren reknar også ut η^2 (R^2). Den vart her **27**.

1.3. Er effekten "statistisk signifikant"?

1.4. Ut frå vanleg brukte kriterier : er denne effekten sterk eller svar?

1.5. Korleis vil du generelt konkludere med omsyn til effekten av intervensjonen?

Forskaren konkluderte med at: «resultata tyder på at det nye undervisningstilbodet ikkje gir noko gevinst i forhold til det eksisterande».

1.6. Kvifor trur du han konkluderte slik? Og er dette ein rimeleg konklusjon?

1.7. Peik kort på nokre viktige svakheiter ved dette forskingsdesignet.

Oppgave 2.

Ein annan forskar er lite tilfreds med denne konklusjonen, og bestemmer seg for å etterprøve denne ved å designe eit betre eksperiment. Han plukkar ut 24 born. Desse blir delt på same måte som i førre eksperiment i to like store grupper: Ei «kontrollgruppe» som berre fortsett med det opphavlege undervisningsopplegget og ei «eksperimentgruppe» der det nye undervisningsopplegget blir implementert. Motivasjon blir no målt for alle borna både før intervensjonen blir implementert (pre) og to månadar seinare (post).

2.1. Kva kallar ein gjerne eit slikt desgin?

Resultata var no som følgjer:

Gjennomsnittlige motivasjonsskårer.

	Pre	Post	
Eksperiment	5	8	6.5
Kontroll	4	5	4.5
	4.5	6.5	5.5

Variansanalyse.

Effekt	SS	df	MSS	F	p
Tid	48	1	48	24.00	0.000
Tid*Gruppe	12	1	12	6.00	0.023
Error ₁	44	22	2		
Gruppe	48	1	48	9.60	0.005
Error ₂	110	22	5		

2.2. Kva av «effektane» i variansanalysetabellen meiner du er av størst interesse for forskaren her? Forklar kvifor.

2.3. Korleis vil du konkludere med omsyn til effekten av intervensjonen?

2.4. Peik på nokre forhold som vil gjere dette til eit mykje betre forskingsdesign enn det som vart nytta i oppgave 1.

Oppgave 3.

«Stroop-effekta» er velkjend frå kognitiv psykologi (går eg ut frå). I ei studie av denne nytta ein stimuli presentert på ein skjerm under tre ulike føresetnadar – og dei same 11 respondentane skulle i alle føresetnadane respondere med å oppgje den fargen som stod skriven på skjermen.

Eksperimentføresetnadar:

Kongruent – respondentane vart presentert for ord som skildra ein farge, og denne var skriven i same farge som den som skulle identifiserast (til dømes «raus» i raud skrift).

Inkongruent – Respondenten vart presentert for eit ord som skildra ein farge, men dette var skriven i ein annan farge enn den som skulle identifiserast (til dømes «raud» i blå skrift).

Nøytral – Alle ord i svart, og berre den namngjevne fargen skulle identifiserast.

Under kvar føresetnad fekk personane ein skåre som uttrykte talet på rette rapporterte fargar.

	Eksperimentbetingelse				
	Kongruent	Inkongruent	Nøytral		
Gjennomsnittlig antall riktige:	8.29	5.00	6.90	6.73	
Effekt	SS	df	MSS	F	p
Eksperimentbetingelse	40.00	2	20.00	4.00	0.04
Person	60.00	10	6.00		
Eksperimentbetingelse*Person	100.00	20	5.00		

Kontraster	Differanse	SE	t	p	Hedge's G
Kongruent - Inkongruent	3.29	0.953	3.451	0.003	1.24
Kongruent - Nøytral	1.39	0.953	1.458	0.160	0.52
Inkongruent - Nøytral	-1.90	0.953	-1.993	0.060	-0.72

3.1. Korleis vil du raskt summere resultatata over?

3.2. Vis ved å setje tal inn i formlane (sjå vedlegg) korleis tala i kolonnen SE og t er utrekna.

3.3. Korleis bør ein tolke verdiane i kolonne «p» i tabellen over kontrastar?

Oppgave 4.

Forklar kva ein i statistikken forstår med omgrepet «samplingfordeling» (sampling distribution). Bruk nokre av resultatata frå oppgave 3 til å illustrere og gjere dette tydeleg.

PSY2011 - Research Methods II: Experimental design and statistical analysis, fall 2016.

Friday October 28, 09:00 (3 timer).

There are 4 tasks - some with sub-task. All should be solved.

All computations should be made explicit. As electronic calculators are not allowed, approximate calculations will be accepted. To compute a ratio between numbers 16.35 and 4.12 an answer like: "16.35/4.12 is approximately 4" is acceptable. That computations are made explicit and clarified is more important than the exact results of calculations.

Relevant formulas are attached at the end of the exam, p. 13-14.

Notice: Text on both sides of sheets.

Task 1.

A researcher wants to investigate a possible effect of a new special-education program to increase children's motivation for learning. Three children are receiving this intensive teaching program (the "experimental group") - and three children continue with the old teaching program (the "control group"). After two months all children are measured on an instrument that is supposed to measure children's motivation.

1.1. *What is usually a research design like this called?*

Results came out like this:

Group	Motivation
Control	4
Control	2
Control	6
Experiment	8
Experiment	4
Experiment	6

1.2. *Perform an analysis of variance, and fill in as many as possible of the figures missing in the table below. The table needs not look "pretty" - as long as it's readable.....*

Source	SS	df	MSS	F	p
Between				1.5	0.288
Within					
Total					

The researcher also computed η^2 (R^2). It came out as .27.

1.3. *Is the estimated effect of treatment "statistically significant"?*

1.4. *Judging from usually applied criteria: is this effect strong or weak?*

1.5. *How would you in general conclude regarding the effect of the treatment?*

The researcher concluded that: "the results imply that the new teaching program does not have any advantages when compared to the existing (old) program".

1.6. *Why do you think he came to this conclusion? And is this a sensible conclusion?*

1.7. *Comment on some important short-comings of this research design.*

Task 2.

Another researcher is not satisfied with this conclusion and decides to examine this by designing a better experiment. Twenty-four children are selected. In the same way as in the previous experiment, they are split into two equal sized groups: a "control group" that continues the original teaching program, and an "experimental group" where the new teaching program is implemented. For all children, motivation is measured both before the intervention is implemented (pre) and two months later (post).

2.1. *What is usually a research-design like this called?*

The results came out like this:

Mean motivation scores.

	Pre	Post	
Experiment	5	8	6.5
Control	4	5	4.5
	4.5	6.5	5.5

Analysis of variance.

Effect	SS	df	MSS	F	p
Time	48	1	48	24.00	0.000
Time*Group	12	1	12	6.00	0.023
Error ₁	44	22	2		
Group	48	1	48	9.60	0.005
Error ₂	110	22	5		

- 2.2. *Which of the "effects" displayed in the Anova-table would in your opinion be of primary interest for the researcher? Explain why.*
- 2.3. *How would you conclude regarding the effect of the intervention (treatment)?*
- 2.4. *Comment on some aspects that makes this a way better research design than the one employed in task 1.*

Task 3.

The "Stroop-effect" is well known from cognitive psychology (I assume). In a study of this, stimuli were presented on a screen under three different conditions - and the same 11 respondents were in all conditions asked to specify the color that was written on the screen.

Experimental conditions:

Congruent - the respondents were presented the name of a color written in the same color as the color to be identified (for example the word "red" written in the color of red).

Incongruent - the respondents were presented the name of a color written in another color than the color to be identified (for example the word "red" written in the color of blue).

Neutral - all words were written in black and only the color named should be identified.

In each condition the respondents were given a score indicating the number of colors correctly identified.

	Experimental condition				
	Congruent	Incongruent	Neutral		
Mean number of correct responses:	8.29	5.00	6.90	6.73	
Effect	SS	df	MSS	F	p
Experimental condition	40.00	2	20.00	4.00	0.04
Subject	60.00	10	6.00		
Experimental condition*Subject	100.00	20	5.00		

Contrasts	Difference	SE	t	p	Hedge's G
Congruent - Incongruent	3.29	0.953	3.451	0.003	1.24
Congruent - Neutral	1.39	0.953	1.458	0.160	0.52
Incongruent - Neutral	-1.90	0.953	-1.993	0.060	-0.72

- 3.1. *How would you in short describe the results displayed above?*
- 3.2. *By putting figures (numbers) in formulas (see attachment), show how the numbers in columns "SE" and "t" are computed.*
- 3.3. *How are the numbers in column "p" in the table of contrasts interpreted?*

Task 4.

Explain the statistical concept "sampling distribution". Use some of the results from task 3 to illustrate and explicate this.

Some relevant formulas.

Standardscores (Z):

$$ZX_i = \frac{X_i - \bar{X}}{SD_x}$$

$$SS_{\text{total}}: \quad SS_{\text{total}} = \sum (y_i - \bar{y})^2$$

For one-way Anova (between-subjects *design*):

$$SS_{\text{between}}: \quad SS_b = \sum_{j=1}^g \sum_{i=1}^{n_j} (\bar{y}_j - \bar{y})^2 = \sum_{j=1}^g n_j (\bar{y}_j - \bar{y})^2$$

$$SS_{\text{within(error)}}: \quad SS_w = \sum_{j=1}^g \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2$$

Degrees of freedom:

Between (effects): Number of groups - 1

Within (error): n - number of groups

For one-way Anova (within-subjects design, *repeated measures*):

$$SS_{\text{between}}: \quad SS_b = \sum_{j=1}^g \sum_{i=1}^{n_j} (\bar{y}_j - \bar{y})^2 = \sum_{j=1}^g n_j (\bar{y}_j - \bar{y})^2$$

SS_{error} : SS for subject*effect (treatment) interaction

Degrees of freedom:

Between (Effects): number of conditions - 1

Error: (number of conditions-1)*(number of subjects-1)

For the "standard error" (SE) of a difference between two means, we apply:

$$SE_{\text{diff}} = \sqrt{\frac{2 * MSS_w}{n}} \quad \text{where } n = \text{number of subjects in each group (or in each condition)}$$

and

$$t = \frac{X_1 - X_2}{\sqrt{\frac{2 * MSS_w}{n}}} \quad \text{with degrees of freedom (df) from } MSS_w \text{ (} MSS_{\text{error}} \text{)}$$

Cohen's standard of "effect-sizes" (also apply to Hedge's G):

Cohen's Standard	d	r	r ²
	2.00	0.71	0.50
	1.90	0.69	0.47
	1.80	0.67	0.45
	1.70	0.65	0.42
	1.60	0.63	0.39
	1.50	0.60	0.36
	1.40	0.57	0.33
	1.30	0.55	0.30
	1.20	0.51	0.27
	1.10	0.48	0.23
	1.00	0.45	0.20
	0.90	0.41	0.17
LARGE	0.80	0.37	0.14
	0.70	0.33	0.11
	0.60	0.29	0.08
MEDIUM	0.50	0.24	0.06
	0.40	0.20	0.04
	0.30	0.15	0.02
SMALL	0.20	0.10	0.01
	0.10	0.05	0.00
	0.00	0.00	0.00