psyc3010 lecture 6

blocking & analysis of covariance

last week: power
next week: regression

1

last week → this week

- Questions on Assignment 1?
- last week we talked about the importance of maximising power in our studies – i.e., maximising the likelihood that we will correctly identify an effect that exists in the population and reject the null hypothesis
- this week we consider some ways of doing so:
  - Blocking
  - ANCOVA

2

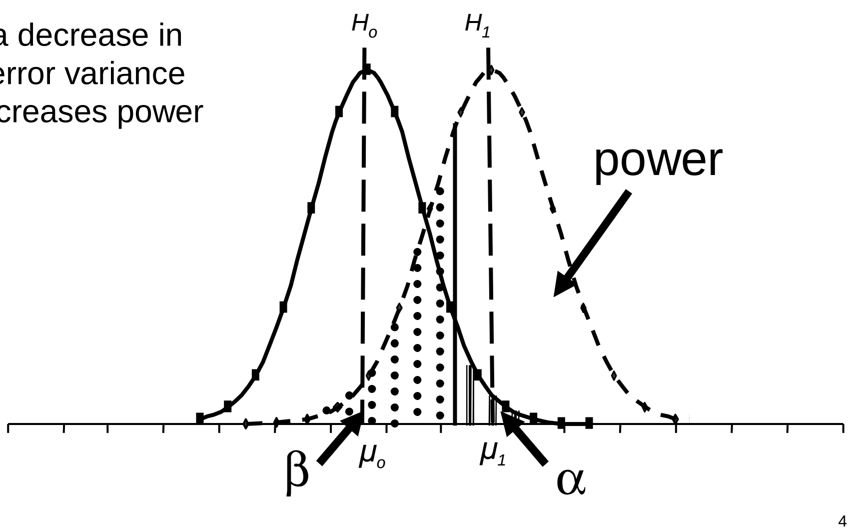
reducing error variance

- other than increasing sample size (N), reducing error (in anova, MS_{error}) variance is a typical method to increase power
- what is “error variance” ?
  – anything left over after accounting for systematic (measured / manipulated IV) variance
  – i.e., it is *residual* variance
  – could reflect **actual error** or simply **unmeasured influences** on the DV

3

Recall…

a decrease in error variance increases power



4

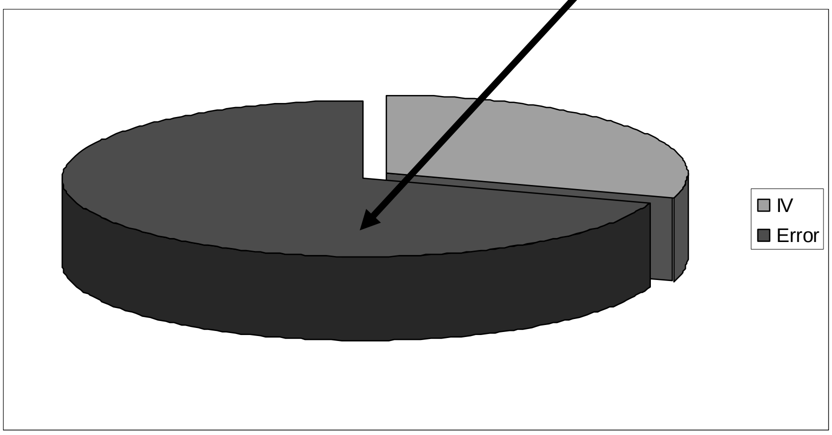
from last week

- recall the power analyses from our "distraction" study…
  - when we treated the study as a 1-way anova (i.e., collapsed across the levels of distraction)
    - MS_{error} = 125.21
    - Power = .91
  - when we treated the study as a 2-way anova (i.e., included the distraction factor in our design)
    - MS_{error} = 83.02
    - Power = .99

5

1-way anova . . . .

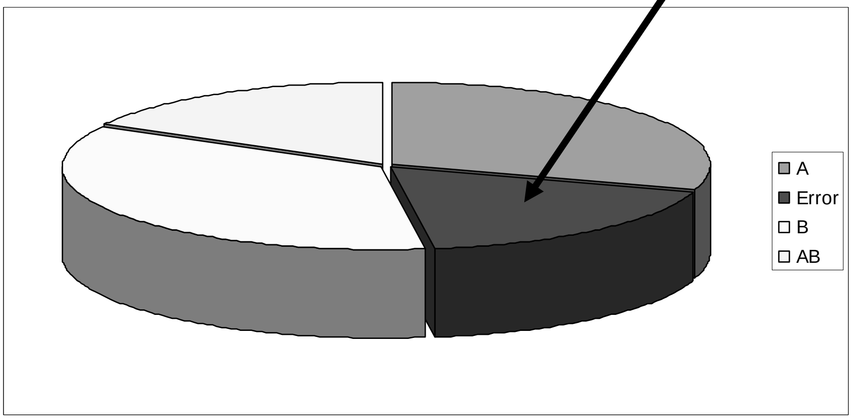
Large "error" or residual variance means low power for test



6

2-way anova . . . .

Error / residual variance decreased by partitioning variability due to blocking factor



7

blocking

- introduce a control or ***concomitant variable***
  - a variable which is correlated with the DV, reflecting some additional source of variation or pre-existing difference on the DV score

Examples:

- problem-solving -- IQ, GPA
- perception -- sensory acuity
- educational achievement -- SES
- social behaviour -- attitudes
- clinical symptoms -- support

- **'blocking'** is a technique for increasing the sensitivity of ANOVA by including a known concomitant variable as a 2nd factor
  - Take advantage of the power of factorial designs

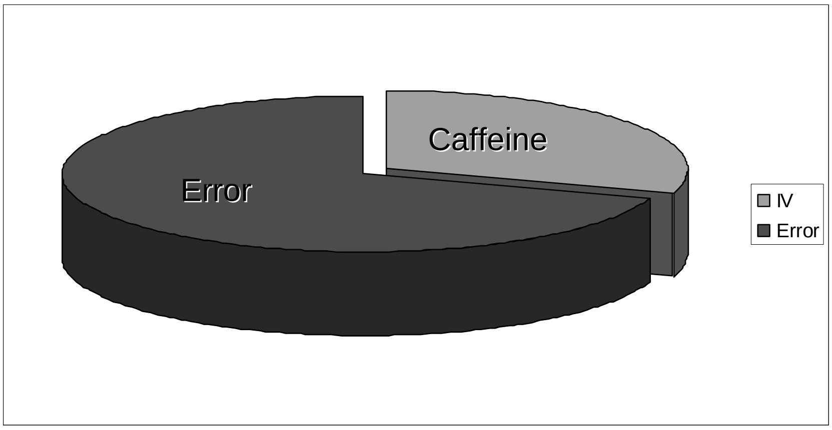
8

blocking

- an example:
  - 1-way anova
    - DV = problem solving time
    - IV e.g. 3 different amounts of caffeine
  - we could introduce another IV that was related to the DV, giving us a 2-way anova, i.e.,
    - DV = problem solving time
    - *factor 1* (IV) caffeine dose
    - *factor 2* (blocking variable) is something related to the DV, e.g., IQ, GPA, general reaction time measure

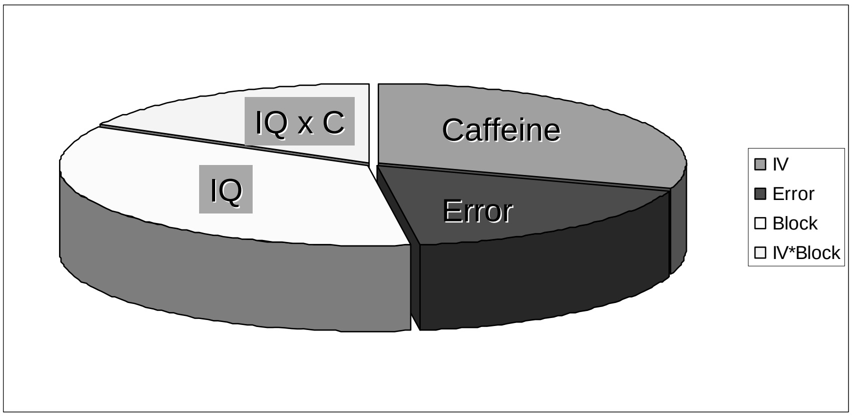
9

1-way anova . . . .



10

2-way anova . . . .



11

12

blocking – how to do it

- homogenous blocks are created with levels of blocking factor (participants are “matched” within levels of the blocking factor)
  - e.g., might first divide people into IQ groups (hi, med, low)
- participants within each block are then randomly assigned to IV conditions (“stratified random assignment”)
  - e.g., within each of three IQ groups assign equally to caffeine vs control condition

	Low IQ	Mid IQ	High IQ
Control			
Caffeine			

13

blocking – how to do it

- different to normal factorial designs, which are fully randomised
- Called “randomized block design”, also stratification or matched samples design

14

an example . . .

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
IV	199.850	1	$\boxed{199.850}$	2.464	.124
Error	3487.500	43	$\boxed{81.105}$		
Total	8966.667	44			

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
IV	199.85	1	$\boxed{199.85}$	5.323	.026
Block	1902.290	2	951.145	25.332	.000
IV*Block	120.870	2	60.435	1.610	.212
Error	1464.340	39	$\boxed{37.547}$		
Total	8966.667	44			

blocking - some issues

- doesn't the introduction of a second factor reduce the size of *df*$_{error}$, leading to a *loss* of power? (see *F* table in Howell)
  - –that loss is easily compensated by the reduction in the size of the error term

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
IV	199.850	1	199.850	2.464	.124
Error	3487.500	$\boxed{43}$	$\boxed{81.105}$		
Total	8966.667	44			

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
IV	199.85	1	199.85	5.323	.026
Block	1902.290	2	951.145	25.332	.000
IV*Block	120.870	2	60.435	1.610	.212
Error	1464.340	$\boxed{39}$	$\boxed{37.547}$		
Total	8966.667	44			

16

blocking - some issues

- what if the blocking variable *doesn't* reduce the error term?
  - then you have chosen the blocking variable poorly! blocking variables are selected because of their ***known*** association with the dependent variable
- how does the blocking variable fit in with predictions and reporting of findings?
  - It doesn't! unlike a main effect in normal factorial ANOVA, the effect of the blocking factor is not usually of interest → it is only factored in to reduce error and increase the power of the test for the focal IV

17

a second application of blocking

- **detecting potential confounds**
  - not *directly* relevant to power, but improves the general sensitivity of the experiment
- example:
  - each condition in your study has several different experimenters
  - effects may therefore be due to the experimenter and not due to the treatment conditions
  - Or maybe the effect of the treatment depends on who the experimenter was

18

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Task	232.000	4	58.000	6.520	.000
Error	400.000	45	8.890		
Total	632.000	49			

a significant task*experimenter interaction would mean that task and experimenter are confounded – treatment differences due to task would vary with who was conducting the experiment

Source	of Squares	df	Mean Square	F	Sig.
Task	232.000	4	58.000	8.790	.000
Experimenter	49.000	4	12.250	1.860	.149
Task x Experimenter	186.000	16	11.630	1.760	.168
Error	165.000	25	6.600		
Total	632.000	49			

19

Controls, Confounds, and IVs

- Label depends entirely on theory / researcher
  - **Confound** is significant when you don't want it to be
  - **Control** is significant but you only care because that reduces error or residual variability
  - **IV** is significant and of interest to you
- Control and confound in blocking
  - Main effect of blocking = sign of good control variable.
    - Shows systematic variability due to blocking factor which has been removed from "error" variance
    - Increases power of test for focal IV
  - Blocking factor x IV interaction = sign of confound
    - Increases power to detect focal IV main effect (b/c systematic variability due to interaction removed from "error")
    - But that positive outcome is outweighed by negative outcome: interaction means that effect of focal IV changes depending on blocking factor.
    - significant Block x IV interaction shows failure of treatment IV effect to generalise across levels of Block

20

21

Effects of Task (T) on motor skills of elderly: **CONTROL**

One-way ANOVA

Source	SS	*df*	MS	*F*
Task	213.78	2	106.89	0.71
Error	2263.33	15	150.89	
Total	2477.11	17		

Block on age (e.g., 60-69, 70-79, 80+)

Source	SS	*df*	MS	*F*
Task	213.78	2	106.89	4.18*
Age	1933.78	2	966.89	37.83*
TA	99.55	4	24.89	0.97
Error	230.00	9	25.56	
Total	2477.11	17		

- observe same treatment effects for 3 age groups (ns interaction -> generalisability) - GOOD
- reduction in *df*error (9 vs. 15), BUT this loss of power is compensated for by smaller error term (25.56 vs. 150.89)

22

Effects of length of exercise (E) on students' flexibility:

One-way ANOVA

Source	SS	*df*	MS	*F*
Exercise	1065.50	2	532.75	18.42*
Error	1301.75	45	28.93	
Total	2367.25	47		

Block on gender (*G*):

Source	SS	*df*	MS	*F*
Exercise	1065.50	2	532.75	36.02*
Gender	330.75	1	330.75	22.36*
EG	350.00	2	175.00	11.83*
Error	621.00	42	14.79	
Total	2367.25	47		

- finding that treatment effects differ according to gender (significant interaction) – potential confound – NOT GOOD – or new IV (moderator) – interesting!

Evaluating blocking

advantages:

- may equate treatment groups better than completely randomized design (equal n for levels of Blocking factor)
- greater power (smaller error term)
- check interactions of treatments and blocks (do effects of treatments generalise?)

disadvantages:

- cost of introducing blocking factor
- need blocking variables that are highly correlated with DV
- loss of power if blocking variable is poorly correlated with DV (r < .20), because fewer *df*error
- treats blocking factor as having discrete levels; some variables must be artificially grouped for analysis (lose information)

25

ancova – analysis of covariance

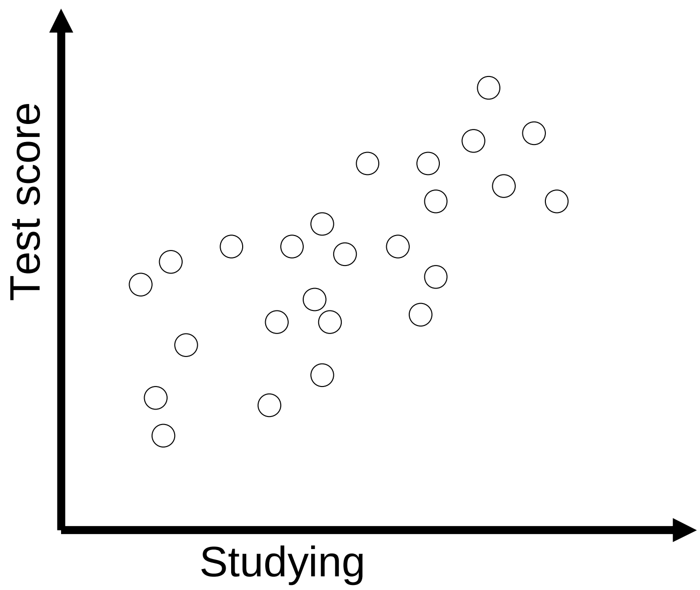
- has the same goal as blocking but works differently:
  - blocking works at the ***level of design*** – the reduction in the size of the error term is a consequence of including a factor that explains a good proportion of variance in the DV
  - With ancova the error term is adjusted ***statistically***

26

"remind me…what is covariance?"

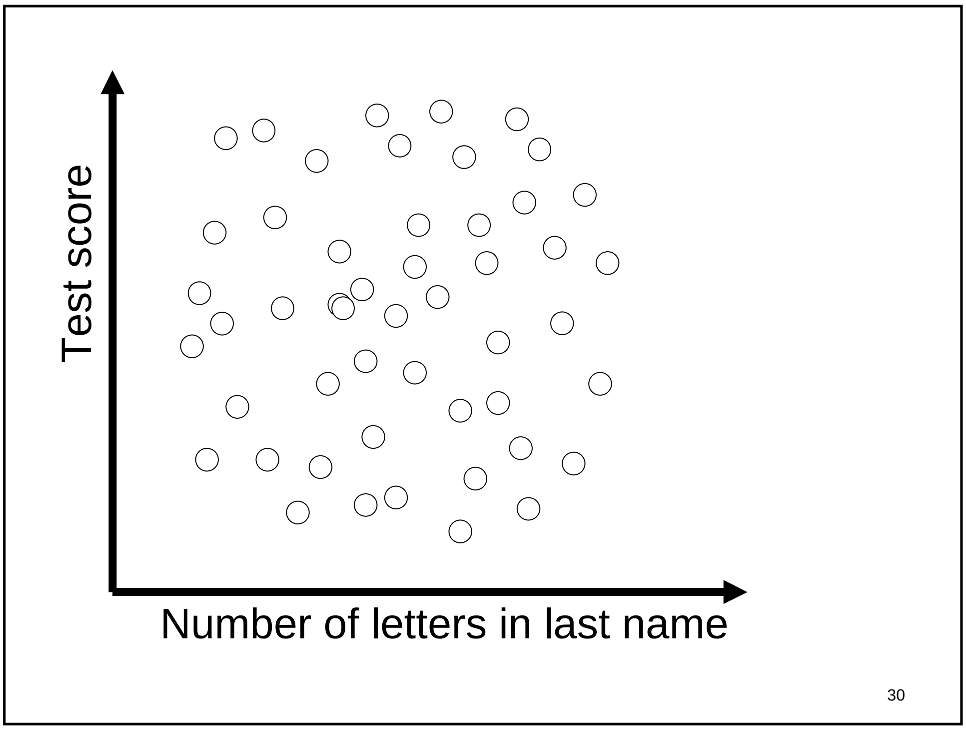
- **variance** is the tendency for scores to vary around some mean value $\sum\left(X_{ij}-\overline{X}\right)^2 / N-1$
- ***co*-variance** is the tendency for two scores to vary *together* $$\frac{\sum\left(X-\overline{X}\right)\left(Z-\overline{Z}\right)}{N-1}$$
  - *If a participant's score on one variable deviates from the mean, the score on the other (covarying) variable also deviates*
  - *positive covariance = both deviate in the same direction* [Review pp. 252-3 of Howell (5th) or pp.236-8 (6th) if nec]
- a ***covariate*** is like the control variable used for blocking, with a couple of differences:
  - the covariate is a *continuous* variable and treated as such (i.e., participants are not matched at discrete levels)
  - in ancova, the covariate is used to remove error from both the error term and treatment effect

27



28

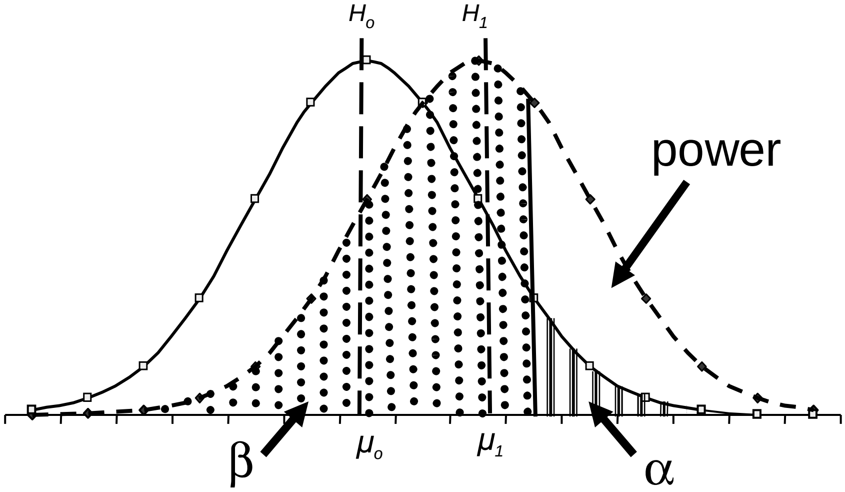

Test score
Drunkeness
29


Test score
Number of letters in last name
30

Analysis of Covariance—ANCOVA

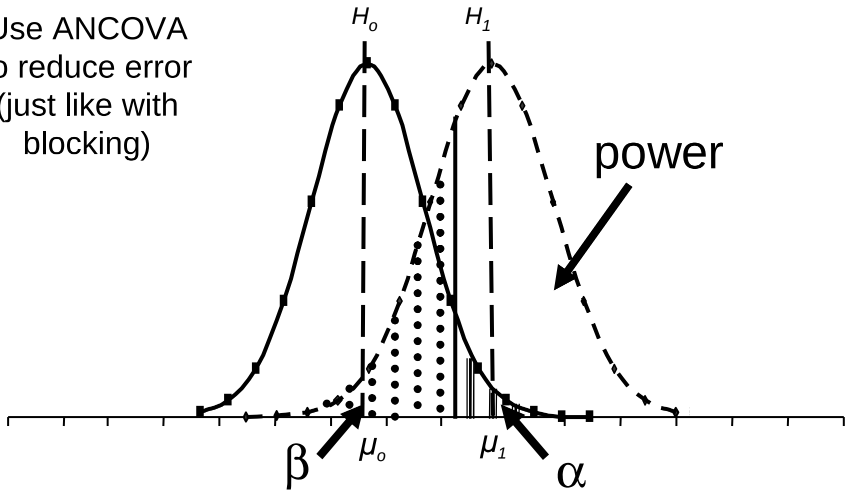
- Originally a technique for analysing experiments and removing nuisance variation
- Attempt to <u>reduce error term</u> by measuring another variable and estimating its parameters
  - if the variable affects the DV and it is not part of the statistical model for the ANCOVA, the variable is in the unmeasured 'error'…

31



32

Use ANCOVA to reduce error (just like with blocking)



33

ANCOVA

- All forms of ANOVA can be performed with a covariate (or several)
- A covariate is another IV/predictor in the model
  - but continuous (ordered scores, not discrete groups)
- Can reduce *error* term—*if* it is related to the DV
- if unrelated you lose DF (lose power) without compensatory reduction in error (i.e., bad trade-off)

34

Uses of ANCOVA

1. To control unwanted variation that would otherwise inflate the error with which we test our models (classical usage)
2. To control for group differences, esp. in the analysis of clinical trials or other pre/post designs (controversial, see Howell 16.5)

35

Oneway ANCOVA structural model

X (the DV) for participant I in Jth group

Grand mean

2nd IV – score on variable Z multiplied by a fixed weight (beta)

$$X_{ij} = \mu + \alpha_j + \beta Z_{ij} + \varepsilon_{ij}$$

Error

1st IV – factor A – group j

Covariate is just another source of variance

- Use the term βZ_{ij} because of continuous nature (will address more next week)
- Implicitly, we have specified *no interaction* between covariate and the IV
  - the presence of such an interaction is a violation of ANCOVA assumptions
    - stats software normally provides output to check as a default
    - Howell includes interaction in the model

36

the structural models

$$X_{ij} = \mu + \tau_j + e_{ij}$$

One-way ANOVA model

β_s are not the same!!

$$X_{ijk} = \mu + \alpha_j + \beta_k + \alpha\beta_{jk} + e_{ijk}$$

Factorial ANOVA model

$$X_{ij} = \mu + \alpha_j + \beta_j Z_{ij} + e_{ijk}$$

One-way ANCOVA model

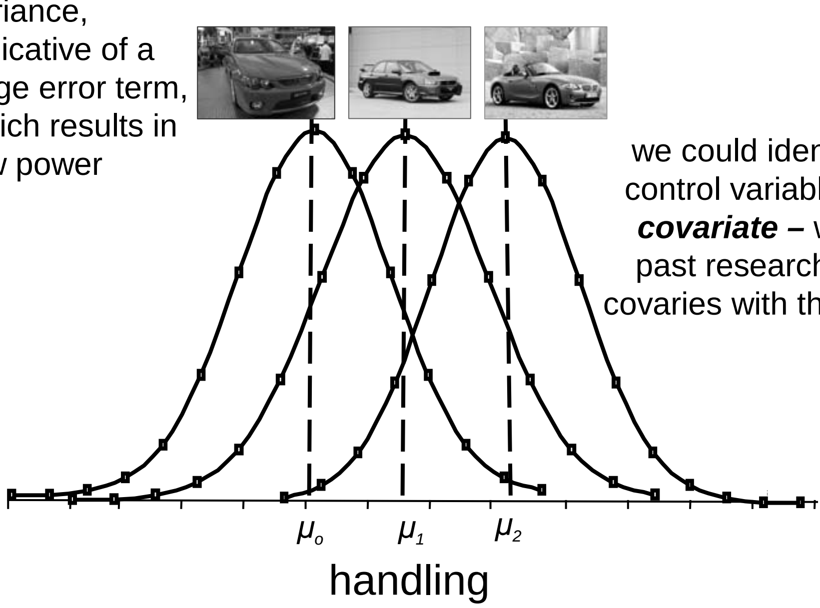
37

an example from Howell

- **comparison of driving performance with three different car sizes – are smaller cars easier to handle?**
- easily addressed using 1-way anova
  - 3 performance cars are driven around a set course 10 times to get an average handling rating (DV)
    - BMW Z3 (small)
    - Subaru WRX (medium)
    - Ford GTP (large)

38

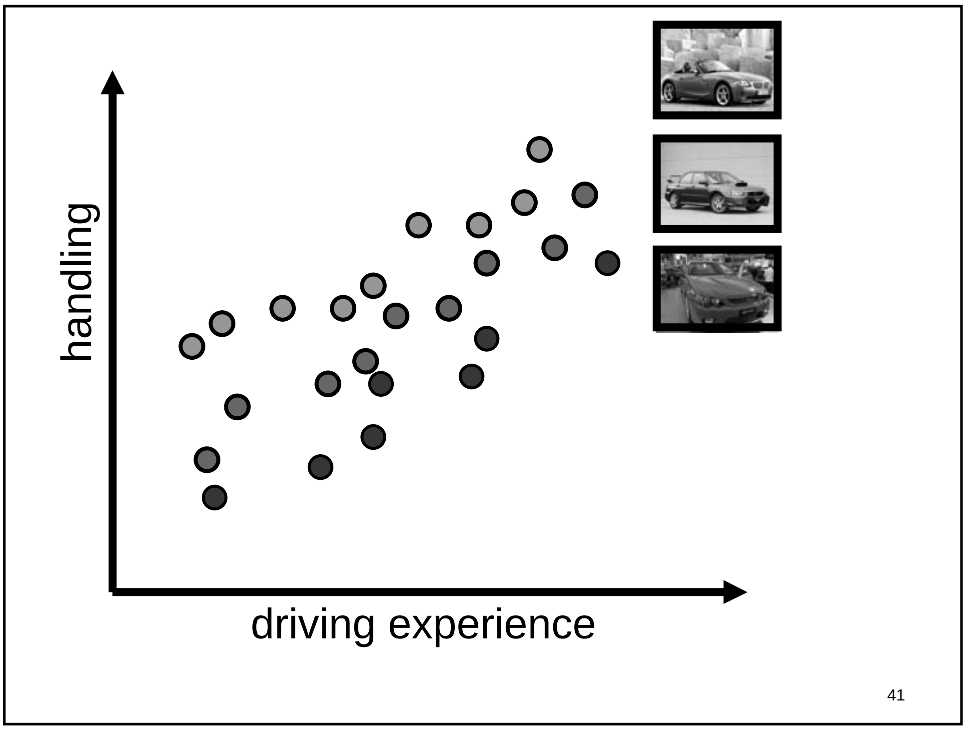
results show lots of overlapping variance, indicative of a large error term, which results in low power

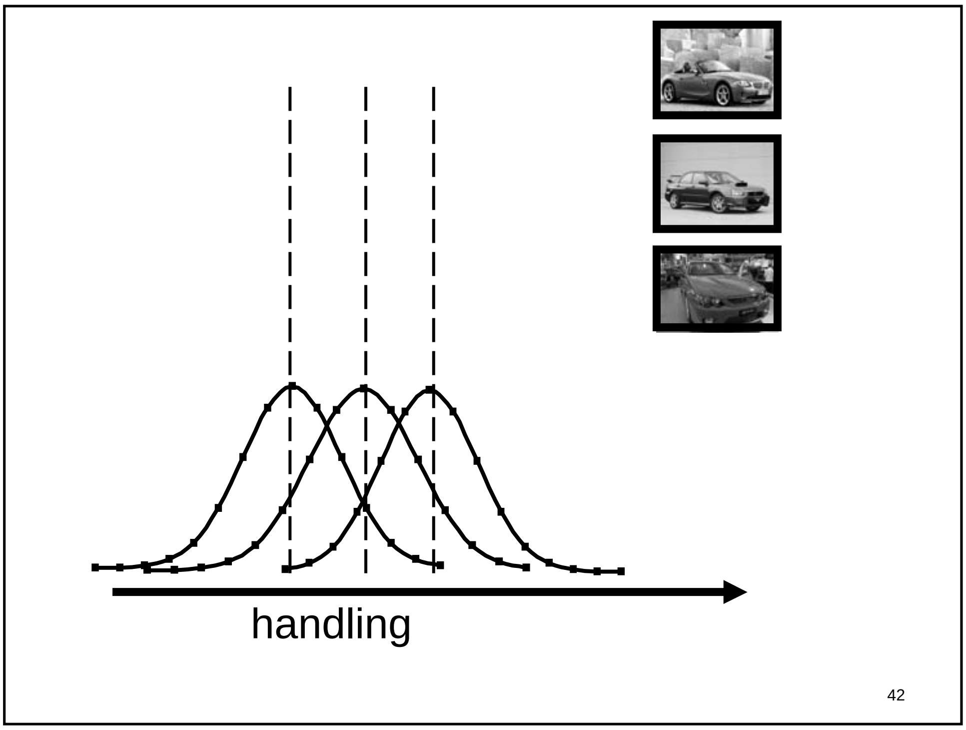


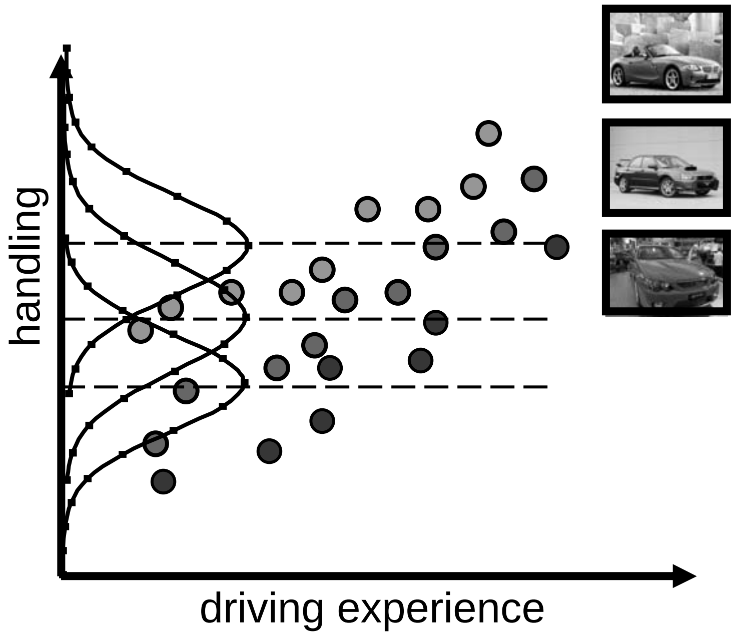
we could identify a control variable – a ***covariate*** **–** which past research tells covaries with the DV

39

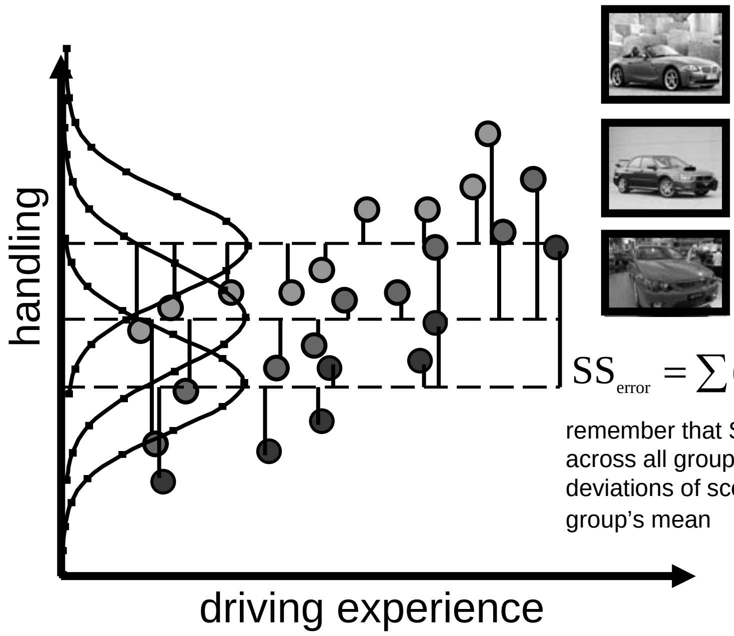
40


handling
driving experience
41


handling
42



43

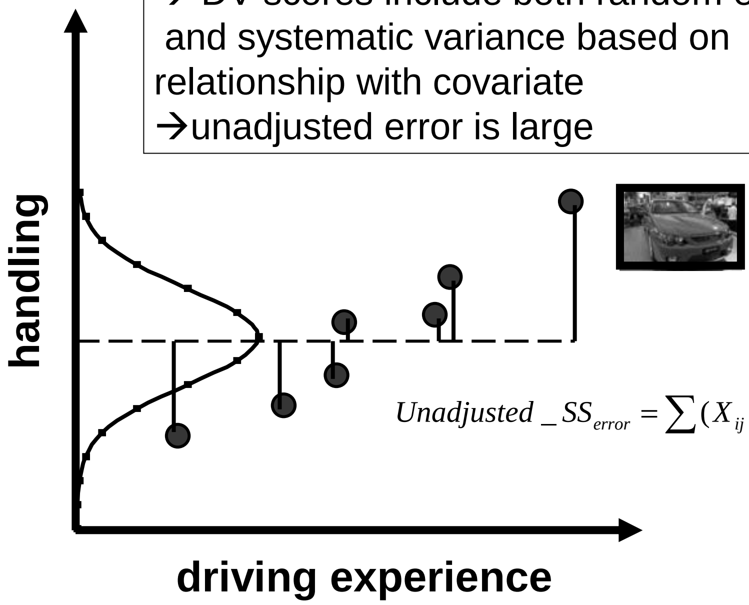


$$SS_{error} = \sum (X_{ij} - \overline{X}_j)^2$$

remember that SS_{error} is the sum across all groups of the squared deviations of scores around their group's mean

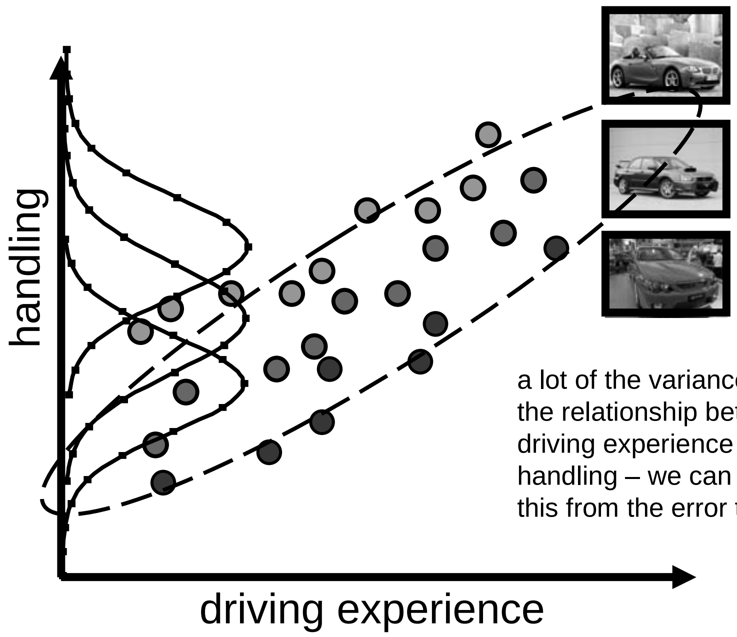
44

→ DV scores include both random error and systematic variance based on relationship with covariate
→unadjusted error is large



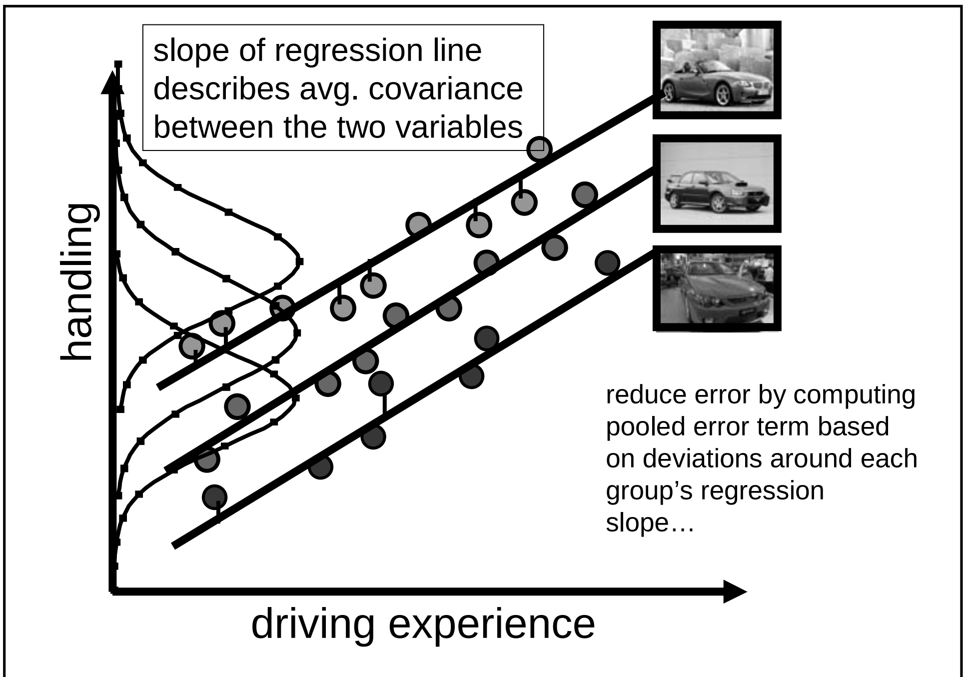
$$Unadjusted\_SS_{error} = \sum (X_{ij} - \overline{X}_j)^2$$

45

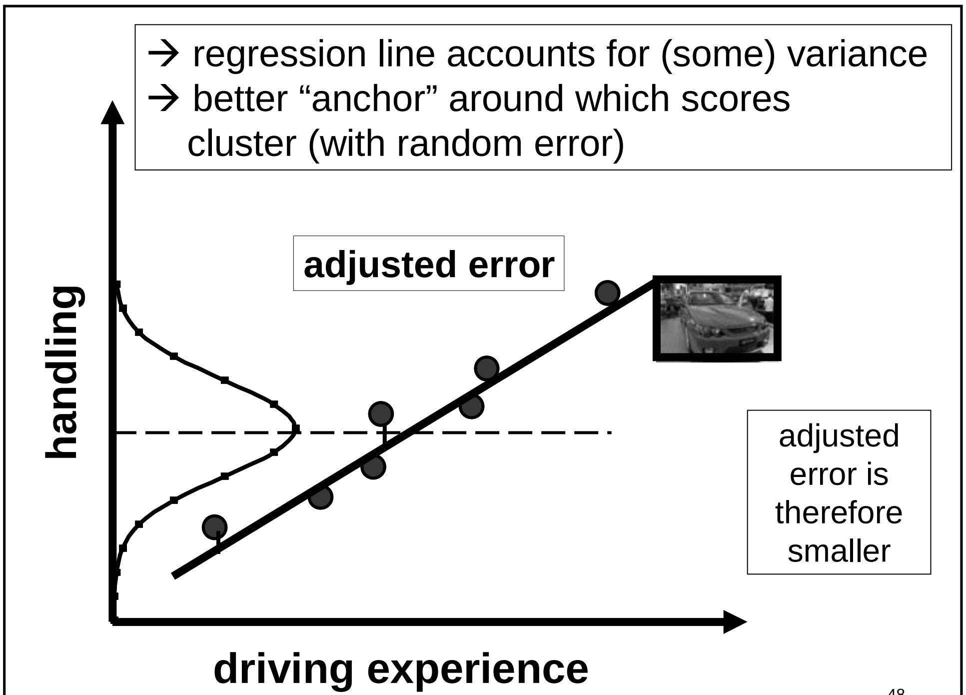


a lot of the variance is due to the relationship between driving experience and handling – we can remove this from the error term *first*

46


slope of regression line describes avg. covariance between the two variables
handling
reduce error by computing pooled error term based on deviations around each group's regression slope…
driving experience

47


→ regression line accounts for (some) variance
→ better "anchor" around which scores cluster (with random error)
adjusted error
handling
adjusted error is therefore smaller
driving experience

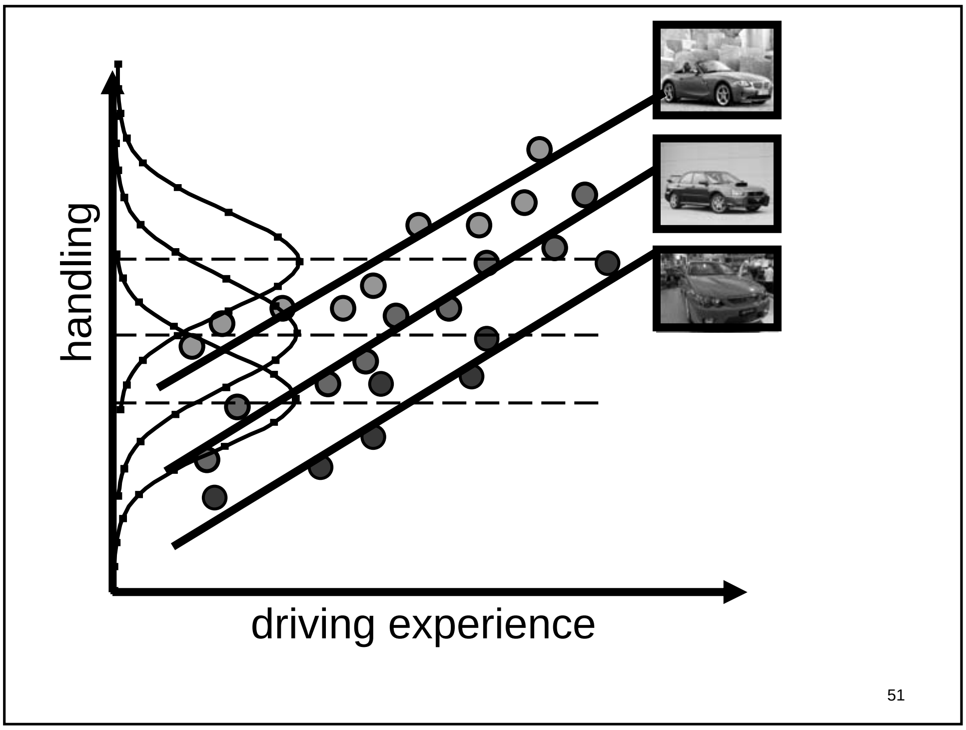
48

49

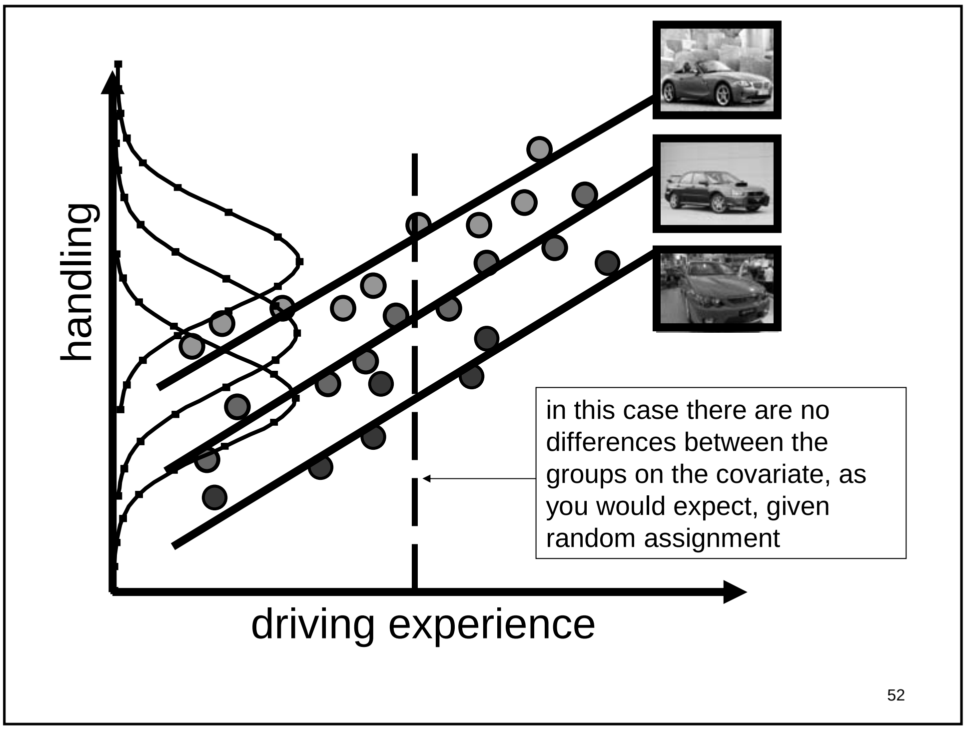
how does that do anything different to blocking?

- at this stage it does not…
  - the effects of the covariate are subtracted from the error term, making it smaller
- the next thing ancova does is quite different to blocking
    - ***treatment means*** are adjusted to account for differences on the covariate
    - random assignment to IV conditions normally prevent differences in covariate means (confounds should be designed out)
    - But in case covariate does differ across groups, ANCOVA effectively ***partials out*** the effects of the covariate from the ***focal IV as well as the error term***

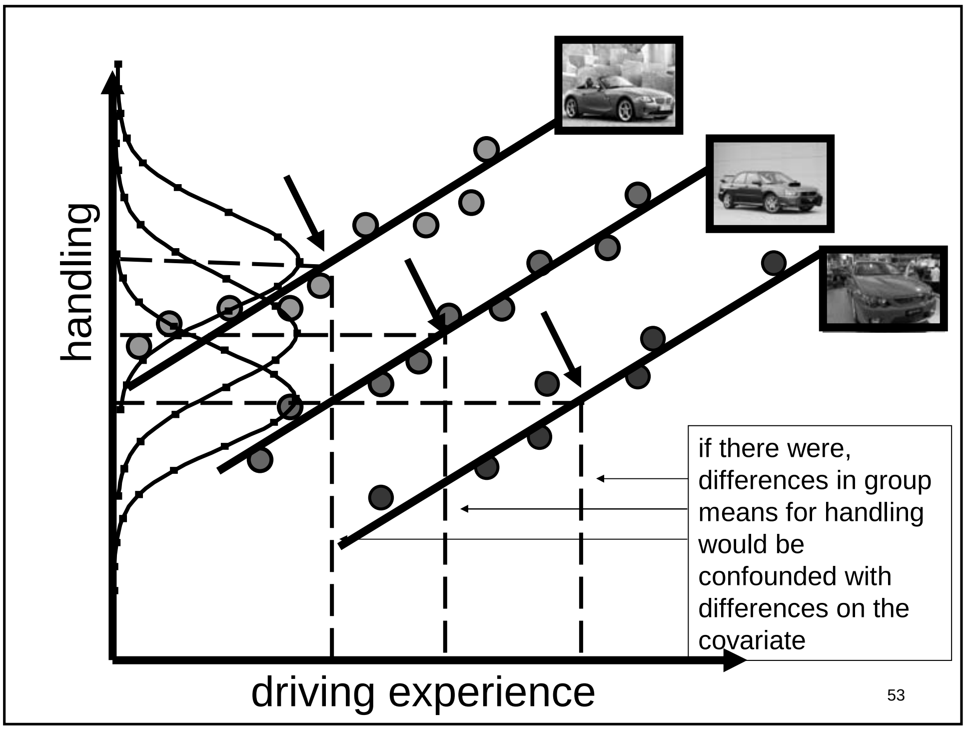
50

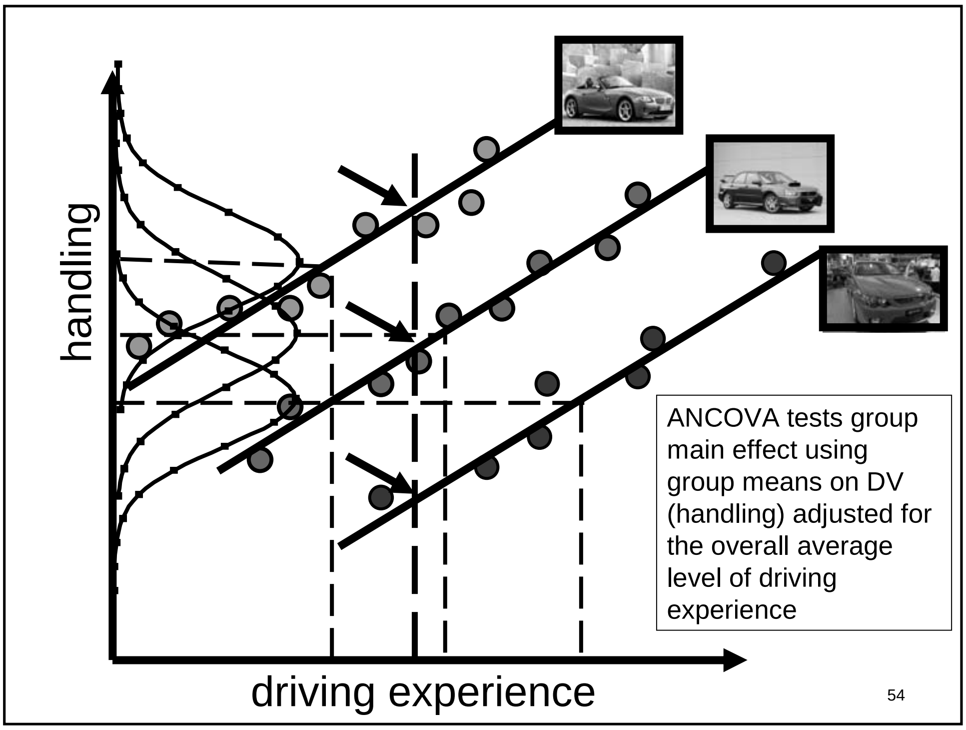

handling
driving experience

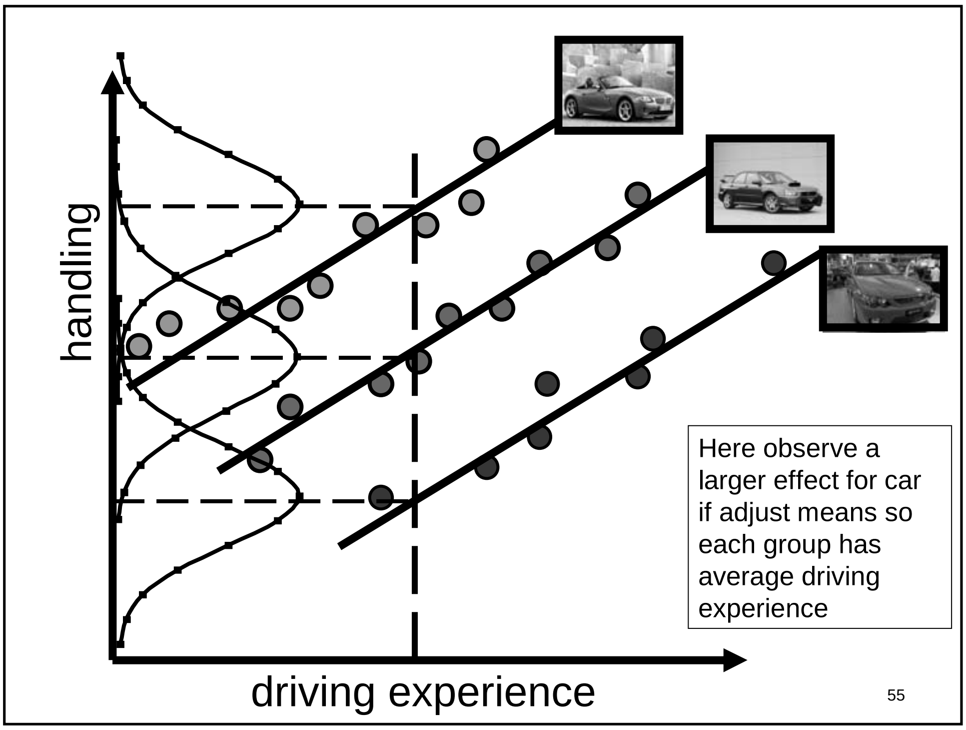
51

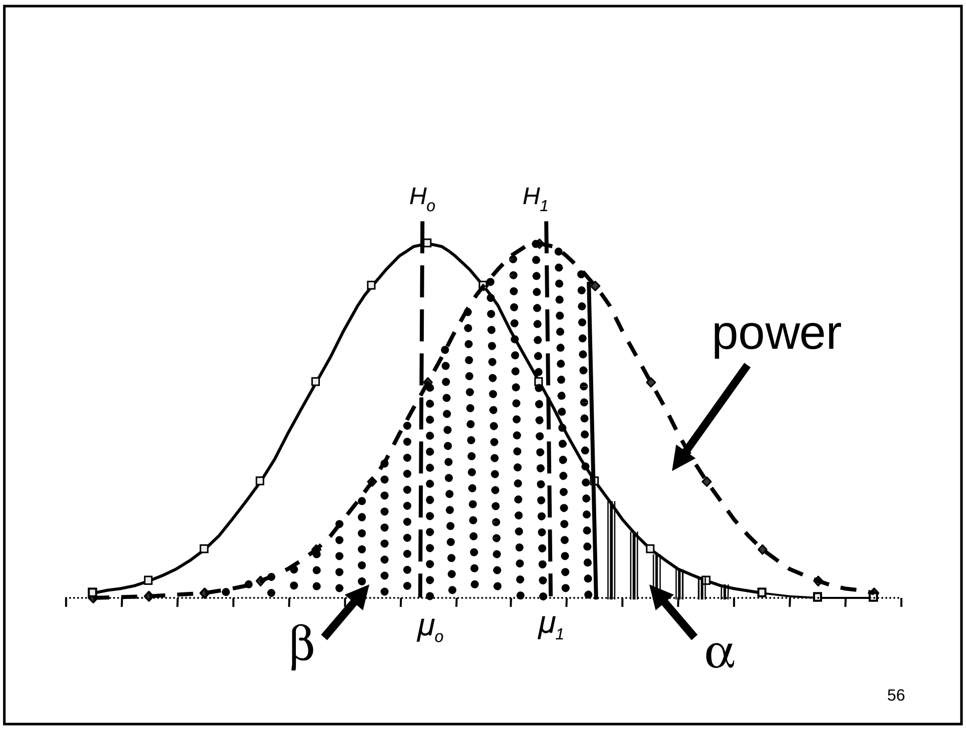

handling
driving experience
in this case there are no differences between the groups on the covariate, as you would expect, given random assignment

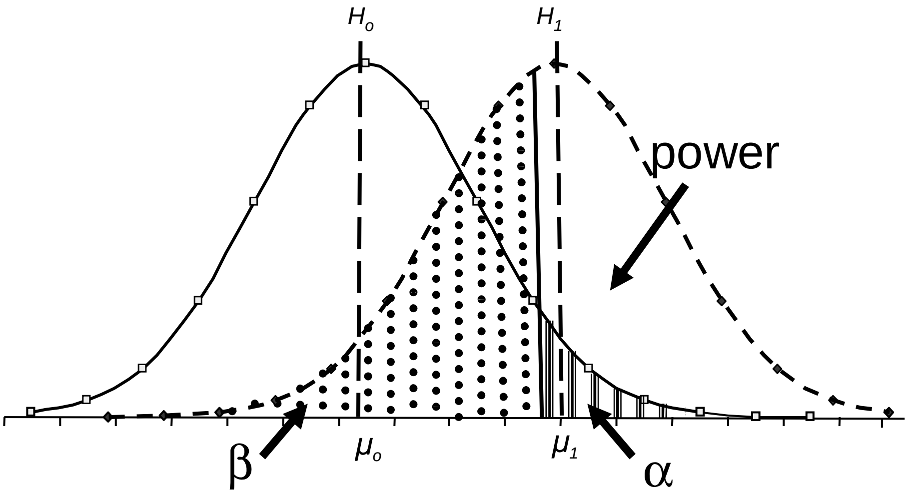
52


handling
if there were, differences in group means for handling would be confounded with differences on the covariate
driving experience
53


handling
ANCOVA tests group main effect using group means on DV (handling) adjusted for the overall average level of driving experience
driving experience
54


handling
driving experience
Here observe a larger effect for car if adjust means so each group has average driving experience
55


H_o
H_1
power
β
μ_o
μ_1
α
56



57

logic of ancova

- *“would groups have differed on the DV if they had been equivalent on a covariate?”*
- ***refines error term*** by subtracting variation that is predictable from covariate
  - larger adjustment when covariate-DV relationship is strong (e.g. pre-test / post-test)
- ***refines treatment effect*** to adjust for any between group differences on covariate that existed before experimental treatment
- **useful** in two situations:
  - covariate related to IV and DV (confound)
  - covariate related to DV only

58

59

comparison of
1-way anova, blocking and 1-way ancova

Tests of Between-Subjects Effects

[redacted]

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Car	231.780	2	106.890	.710	.506
Error	2263.330	15	150.890		
Total	2477.110	17			

1-way anova

effect is not significant

60

comparison of 1-way anova, blocking and 1-way ancova

Tests of Between-Subjects Effects

D[redacted]ence – e.g., no training, some training, professional)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Car	213.780	2	106.890	4.180	.052
Experience	1933.780	2	966.890	37.830	.000
Car x Experience	99.550	4	24.890	.970	.469
Error	230.000	9	25.560		
Total	2477.110	17			

blocking, using factorial anova

effect is *marginally* significant due to reduction of error term from 150.89 to 25.56

61

comparison of 1-way anova, blocking and 1-way ancova

Tests of Between-Subjects Effects

DV = handling (experience as a *continuous* scale, included as a covariate)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Car	252.040	2	126.020	8.697	.003
Regression	1833.780	1	1833.780	126.555	.000
Error	202.880	14	14.490		
Total	2477.110	17			

ancova

reduction of error term from 150.89 to 14.49

increase in treatment effect from 106.89 to 126.02

effect is now significant!

62

ancova vs blocking

- **blocking**
  - – conceptually simpler
  - – requires fewer assumptions
- **ancova**
  - – easier to administer
  - – can use continuous covariate
  - – preservation of df
  - – remove effect from error term <u>and</u> IV
  - – requires specific assumptions…

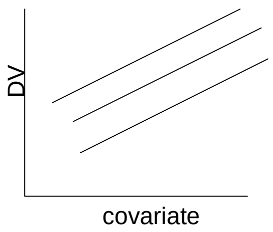
63

ancova - assumptions

- **all your regular ANOVA assumptions:**
  - – Homogeneous variance
  - – Normal distribution
  - – Independence of errors
- Plus:
  - – relationship between covariate and DV is linear
  - – relationship between with covariate and DV is linear *within each group*
  - – relationship between DV and covariate is equal across treatment groups (***homogeneity of regression slopes***)

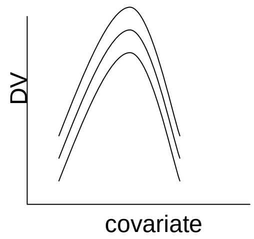
64

re: assumption 1



Linear relationships

☺



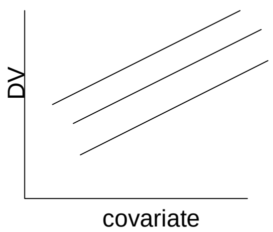
Non-linear relationships

☹

Non-linear relationships generally cannot be detected with ANCOVA – degrades power.

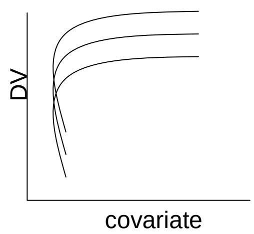
65

re: assumption 1



Linear relationships

☺



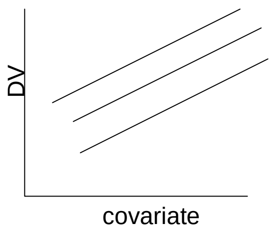
Non-linear relationships

☹

Non-linear relationships generally cannot be detected with ANCOVA – degrades power.

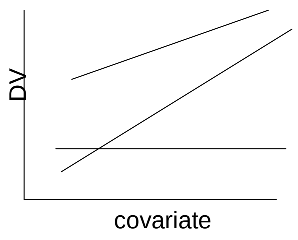
66

re: assumption 3



homogeneity of regression slopes

☺



heterogeneity of regression slopes

☹

homogeneity of regression slopes is important because adjustments to treatment means are based upon an average within-cell regression coefficient

67

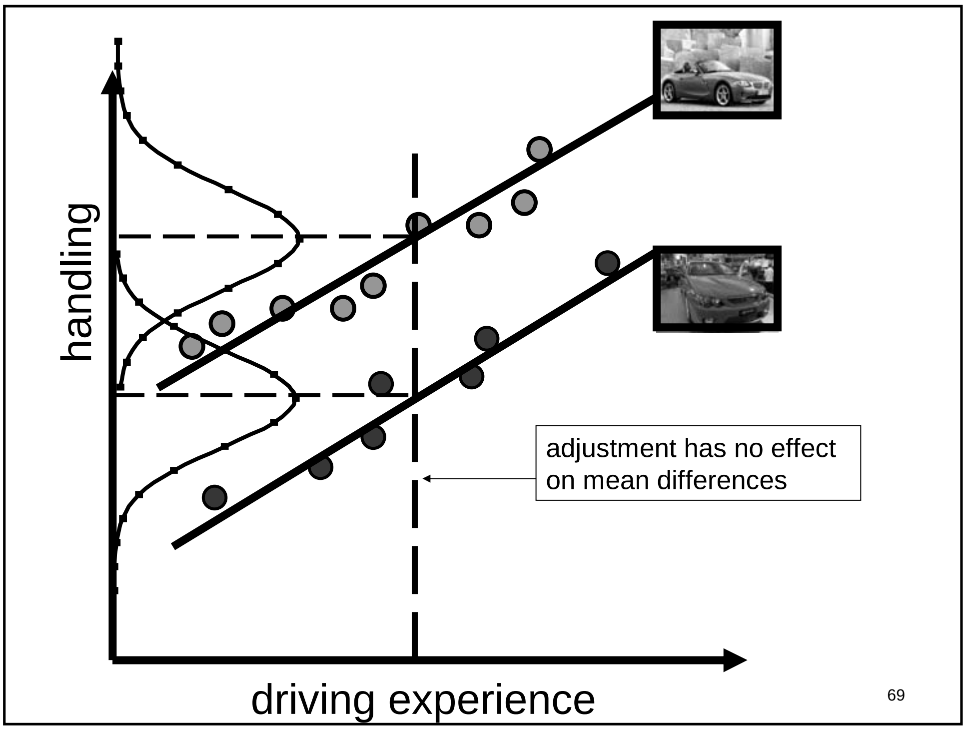
adjusting treatment effects
the fine print

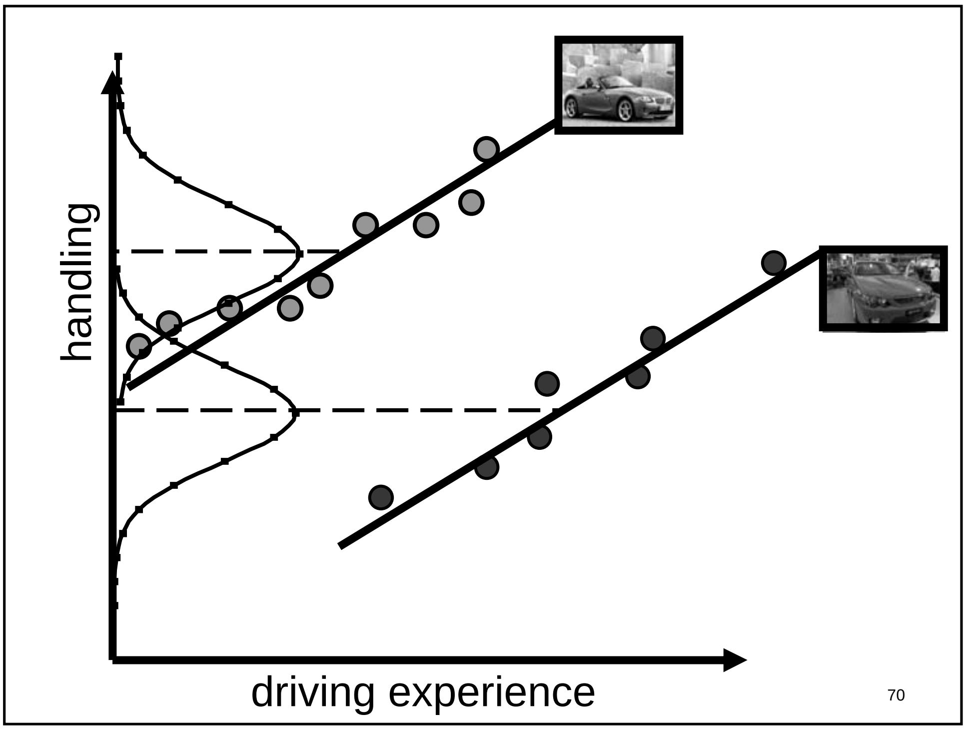
The process is still considered controversial

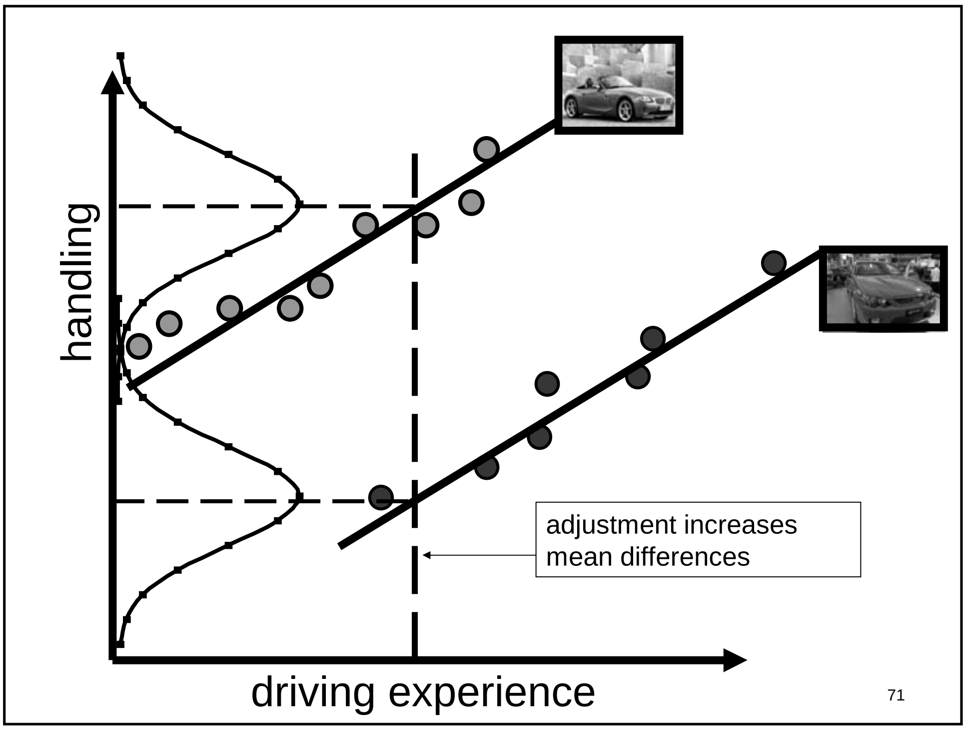
- Some people object to idea of comparing adjusted treatment means at all
  - "real" observed means are not compared
  - comparison means are estimated using regression slope, which may not be reliable
  - If treatment group does affect the covariate as well as the DV, what does adjusted IV mean really tell you?

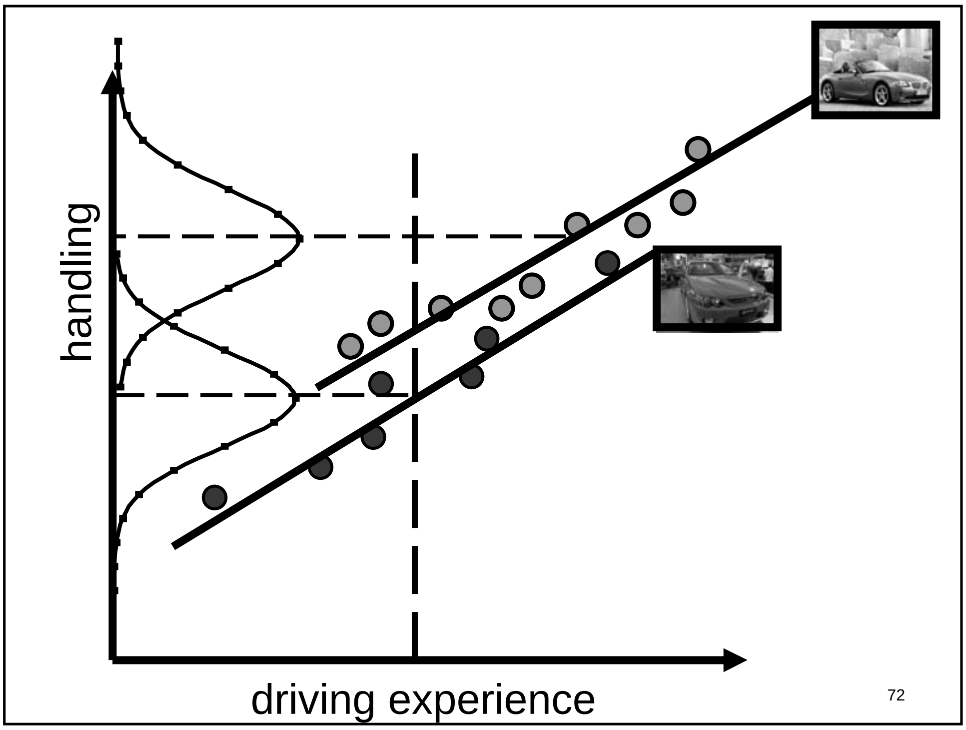
- Some people don't mind adjusted means when the adjustment makes the treatment effect larger
  - But it doesn't always make the treatment effect larger, so it doesn't always work in your favour!

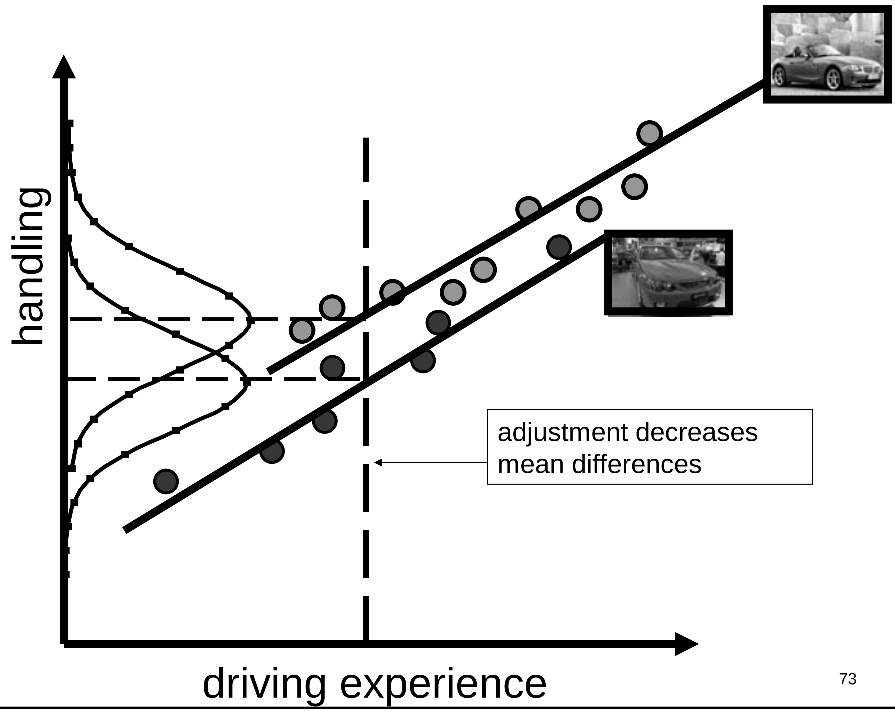
68


handling
adjustment has no effect
on mean differences
driving experience
69


handling
driving experience
70


handling
adjustment increases
mean differences
driving experience
71


handling
driving experience
72



73

final note (for now) on power

- part of the benefit of ancova is the ability to handle continuous data
  - – most psychological variables are continuously distributed, so splitting people into groups is inefficient (lose info) and error prone (mis-categorisation at group boundaries magnifies error)
  - – if your data is continuous, it is best to analyse it using a method which can deal with such data (ANCOVA is more powerful than Blocking)
    - Further techniques next week: correlation and regression

74

Next week in class:

- Correlation and regression

In the tutes:

- This week: Assignment 1 consult
- Next week: Correlational designs, SPSS

readings :

- ancova
  - Howell section 16.5 and 16.6
  - Field, Chapter 9

- review *correlation and regression*:
  - Howell chapter 9, 10 and section 15.1
  - Field, Chapter 4 and Chapter 5: sections 5.1 to 5.4

75