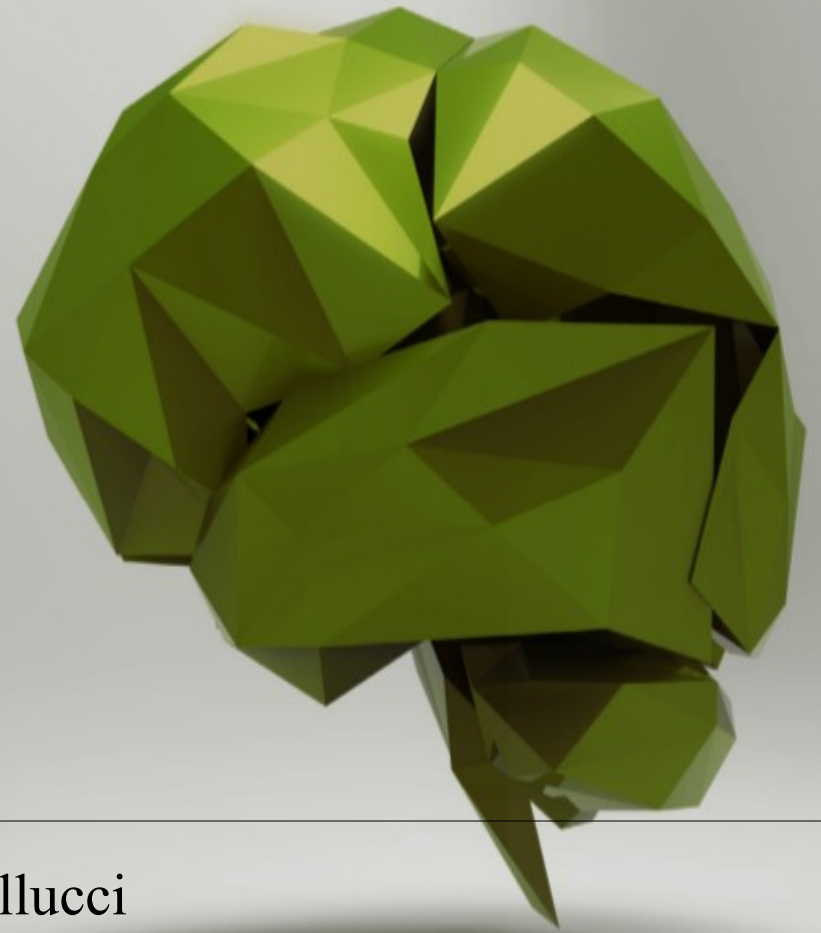


Linear mixed models

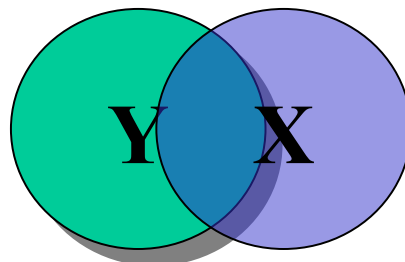
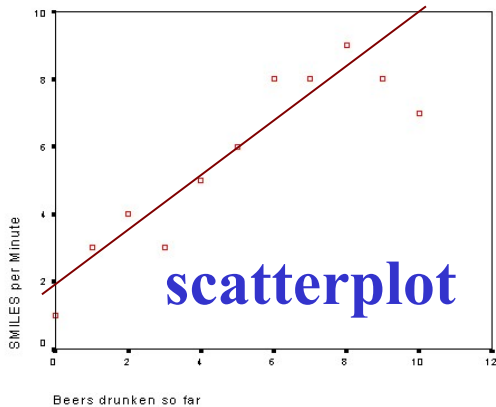
Part I



Marcello Gallucci
University of Milan-Bicocca

Introduction

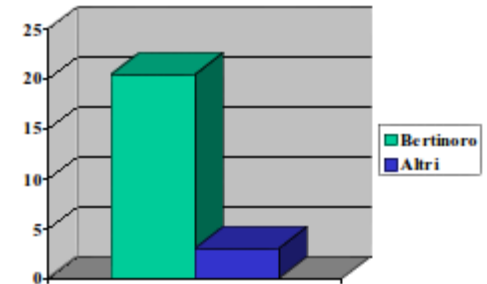
A simple **statistical model** is an efficient and concise representation of the data describing an empirical phenomenon



Path Diagram



Difference in mean

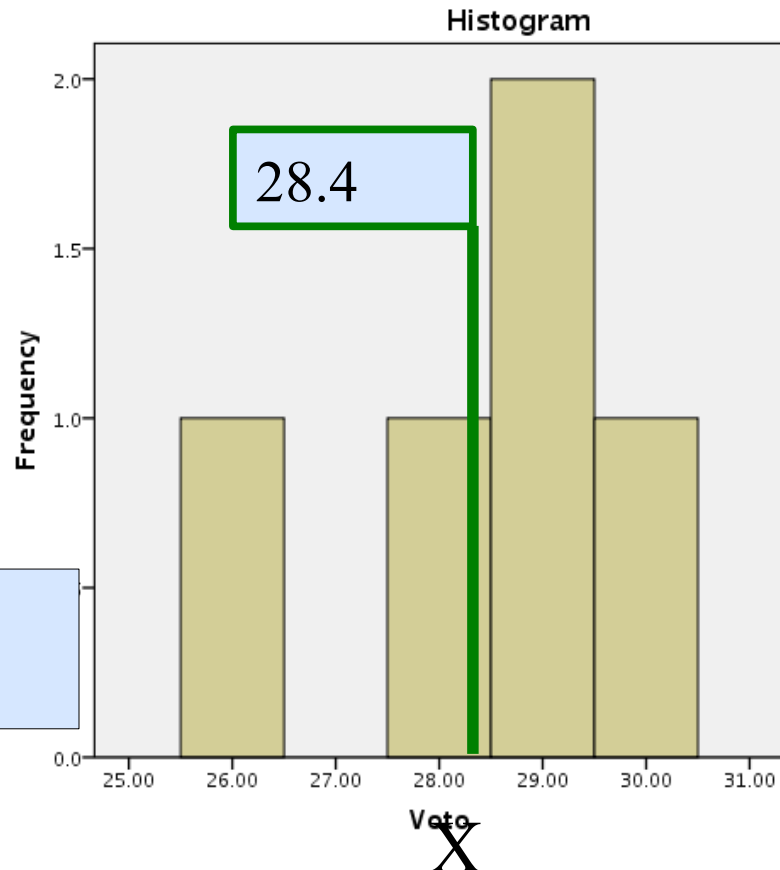


Introduction

The statistical model and its representation is aimed at:

- Efficient and Concise Description
- Prediction of future
- Generalization to a population

That is: understanding of main trends in data



GLM

- The majority of the statistical technique that we use belong to one single **general model**

General Linear Model

$$y_i = a + b_1 \cdot x_{1i} + b_2 \cdot x_{2i} + \dots + b_k \cdot x_{ki} + e_i$$

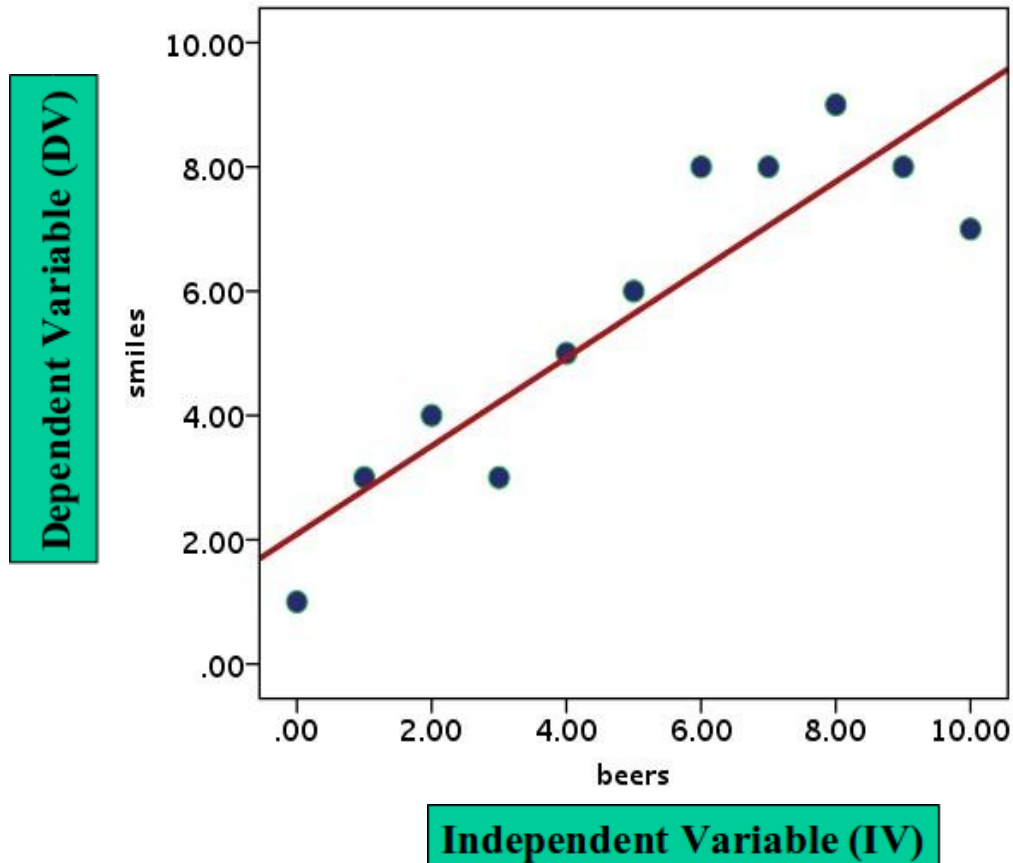
Dependent variable

Independent variables

Errors

GLM Example

The aim of regression analysis is to fit the data using a function



For most applications, we just need a linear function: straight line

$$y_i = a + b \cdot x_i + e_i$$

$$\hat{y}_i = a + b \cdot x_i$$

GLM

pros

- It allows to estimate the effects of one or several IVs on a DV
- It can be applied in many different research problems
- It allows to estimate many different types of effects

cons

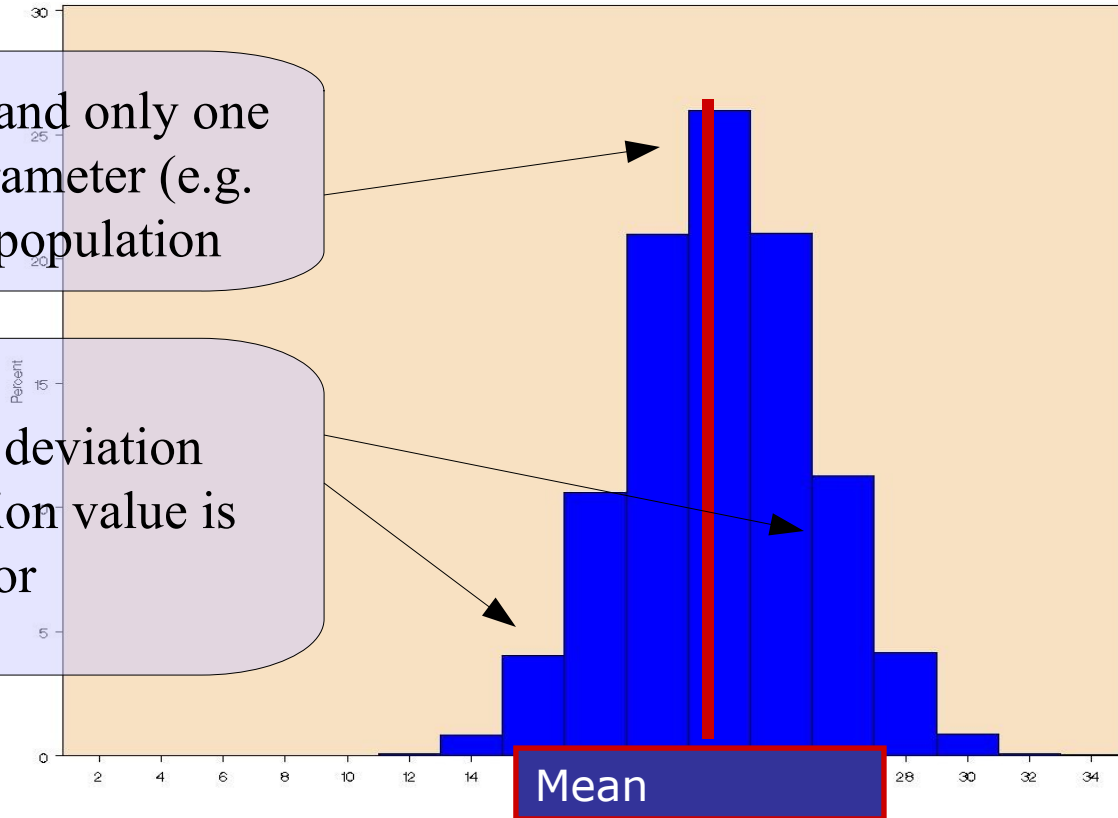
- It assumes data have a very simple structure
- It applies on a limited amount of dependent variables (continuous or semi-continuous DV)

GLM Assumptions

$$y_i = a + e_i$$

1) There exists one and only one value of each parameter (e.g. the mean) in the population

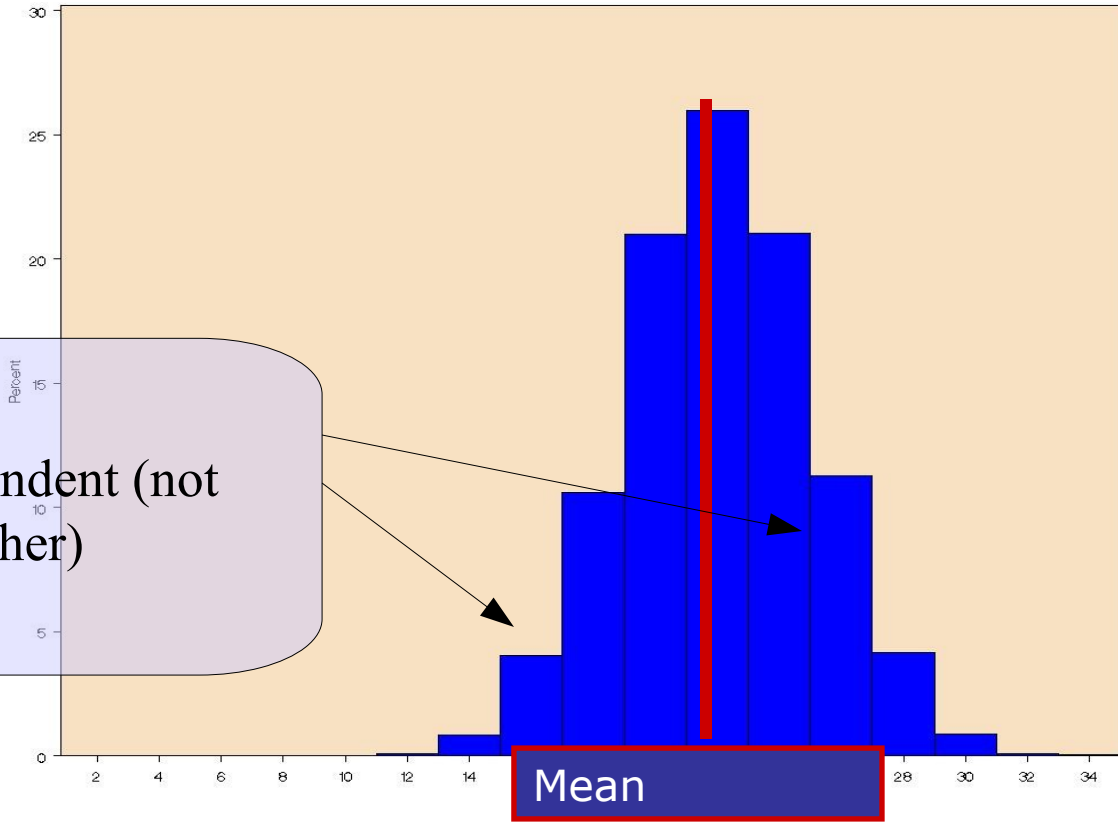
2) Any observed deviation from the population value is deemed to be error



GLM Assumptions

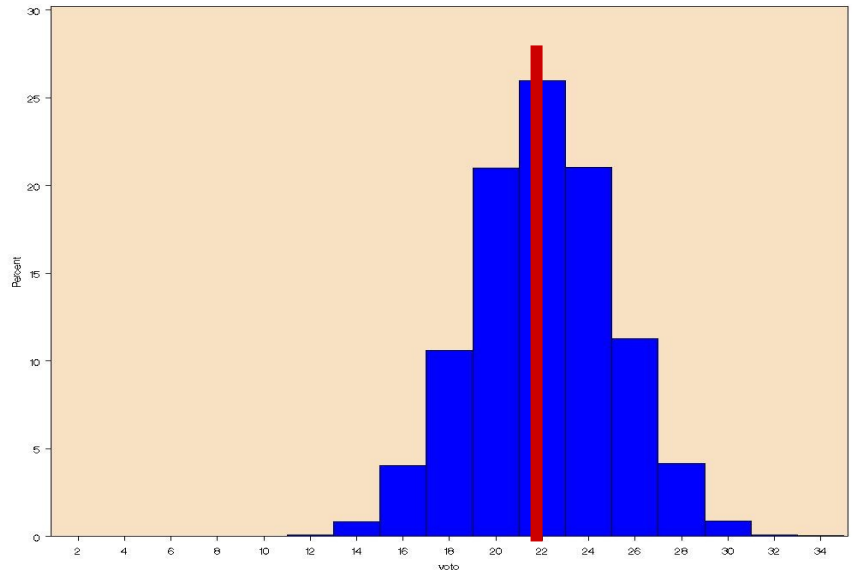
$$y_i = a + e_i$$

Errors are independent (not related to each other)



GLM Assumptions

The estimated value is a **fixed** parameter

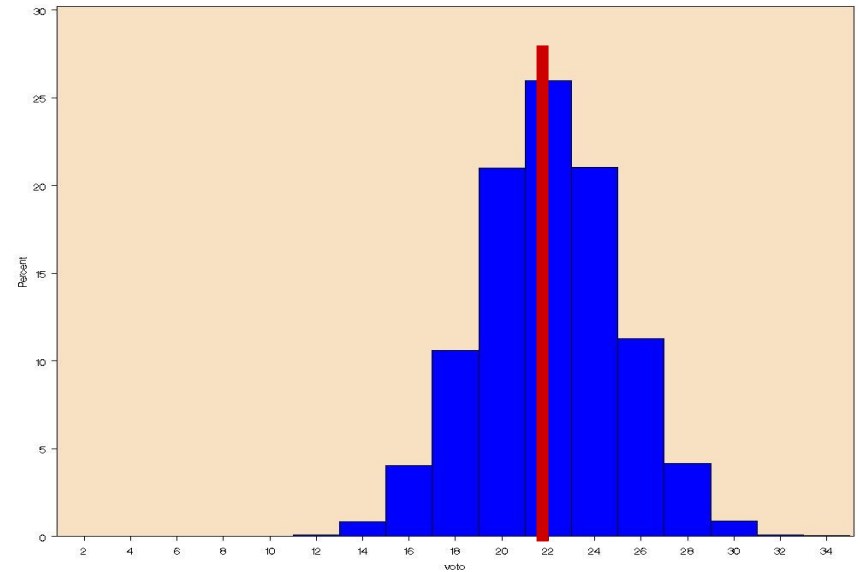


$$y_i = a + e_i$$
$$\text{corr}(e_i, e_j) = 0$$

Random variations are uncorrelated

GLM Assumptions

The estimated value is a **fixed** parameter



$$y_i = a + e_i$$
$$\text{corr}(e_i, e_j) = 0$$

Random variations are uncorrelated

GLM

When the assumptions are NOT met because the data, and thus the errors have more complex structures, we generalize the GLM to the Linear Mixed Model

Linear Mixed Model

GLM

Regression

T-test

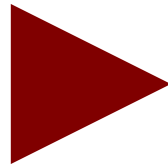
ANOVA

ANCOVA

Moderation

Mediation

Path Analysis



LMM

Random coefficients models

Random intercept regression models

One-way ANOVA with random effects

One-way ANCOVA with random effects

Intercepts-and-slopes-as-outcomes models

Multi-level models

Software

SPSS



R



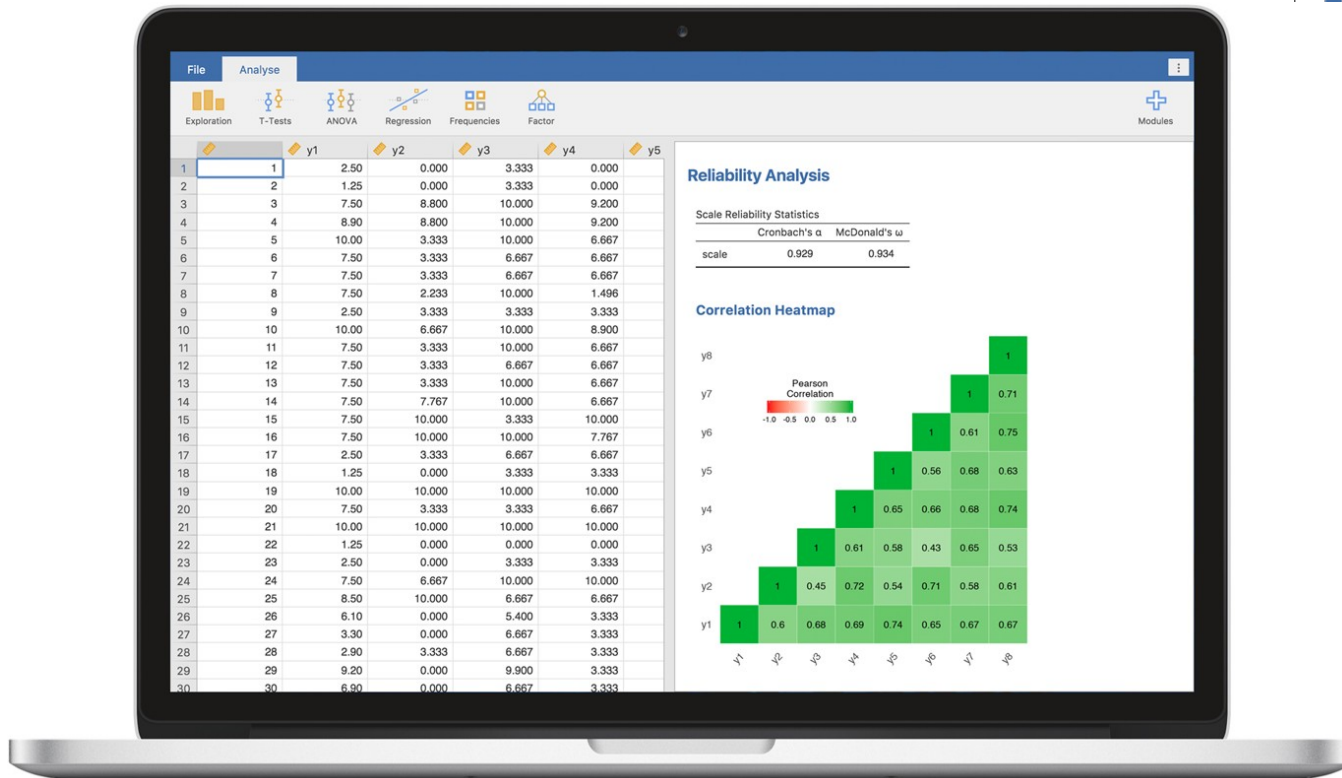
jamovi Stats.
Open.
Now.



Jamovi

www.jamovi.org

iamovi Stats.
Open.
Now.



Example “beers”

Let's consider the case where the beer-smile research was conducted by gathering data in several different bars

bar

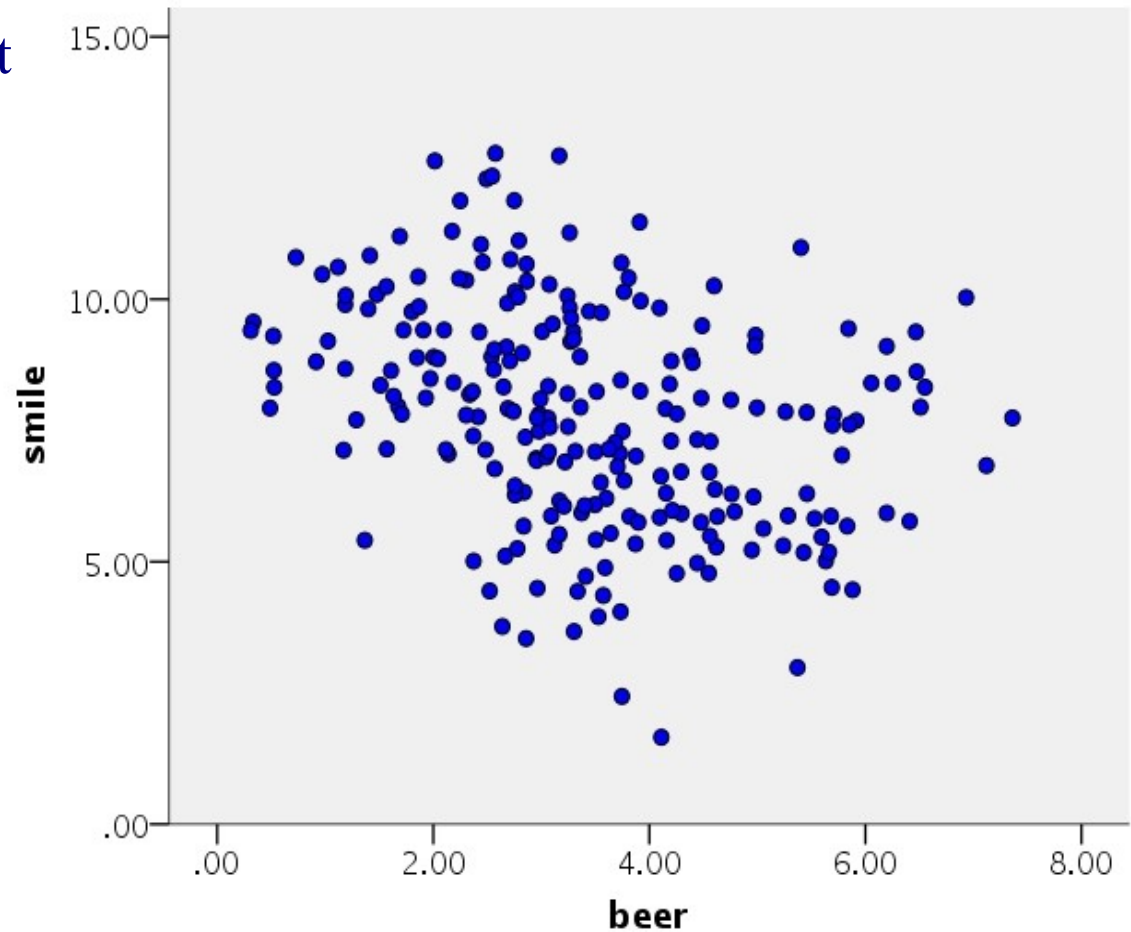
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid a	3	1.3	1.3	1.3
b	14	6.0	6.0	7.3
c	22	9.4	9.4	16.7
d	21	9.0	9.0	25.6
e	14	6.0	6.0	31.6
f	20	8.5	8.5	40.2
g	24	10.3	10.3	50.4
h	12	5.1	5.1	55.6
i	16	6.8	6.8	62.4
l	22	9.4	9.4	71.8
m	21	9.0	9.0	80.8
n	15	6.4	6.4	87.2
o	16	6.8	6.8	94.0
p	11	4.7	4.7	98.7
q	3	1.3	1.3	100.0
Total	234	100.0	100.0	

For each participant
we measured # of
beers and # of smiles

For a total of 234 participants

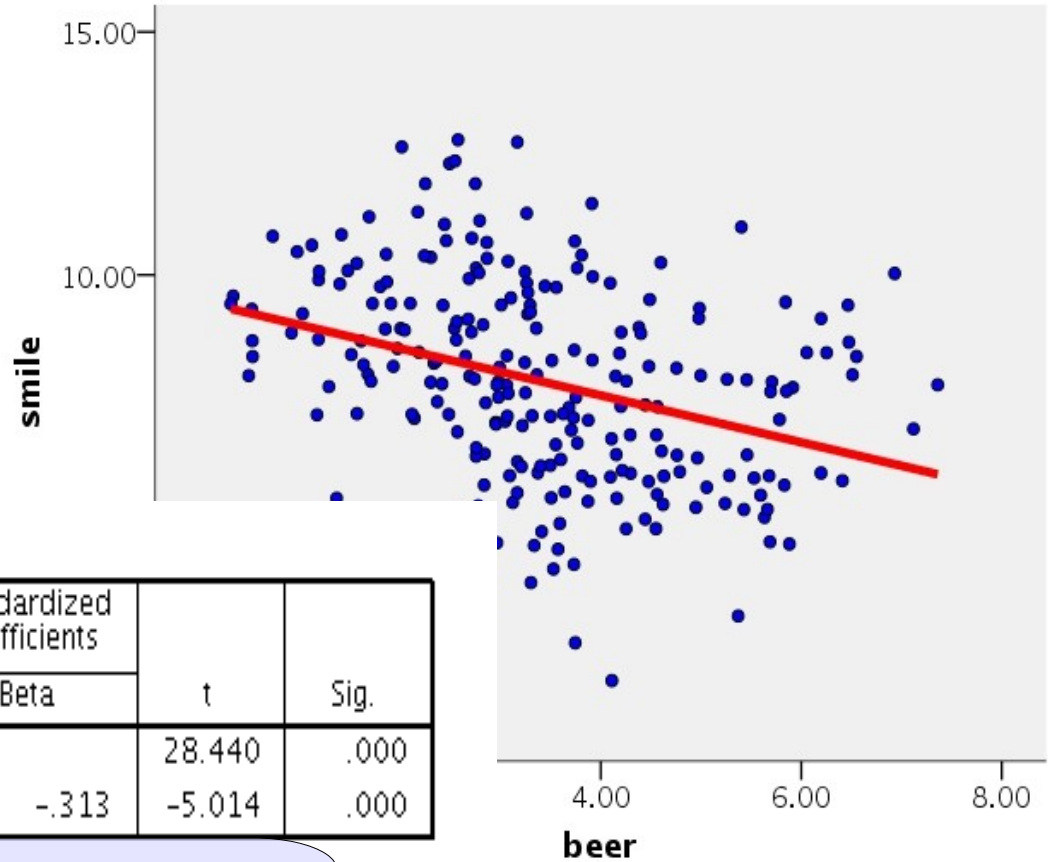
Example “beers” 2

As compared with the example with a few participants, now we have a very different scatterplot



Example “beers” 2

A simple regression confirms that results are indeed different



Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	9.303	.327		28.440	.000
	beer	-.432	.086	-.313	-5.014	.000

a. Dependent Variable: smile

Negative effect

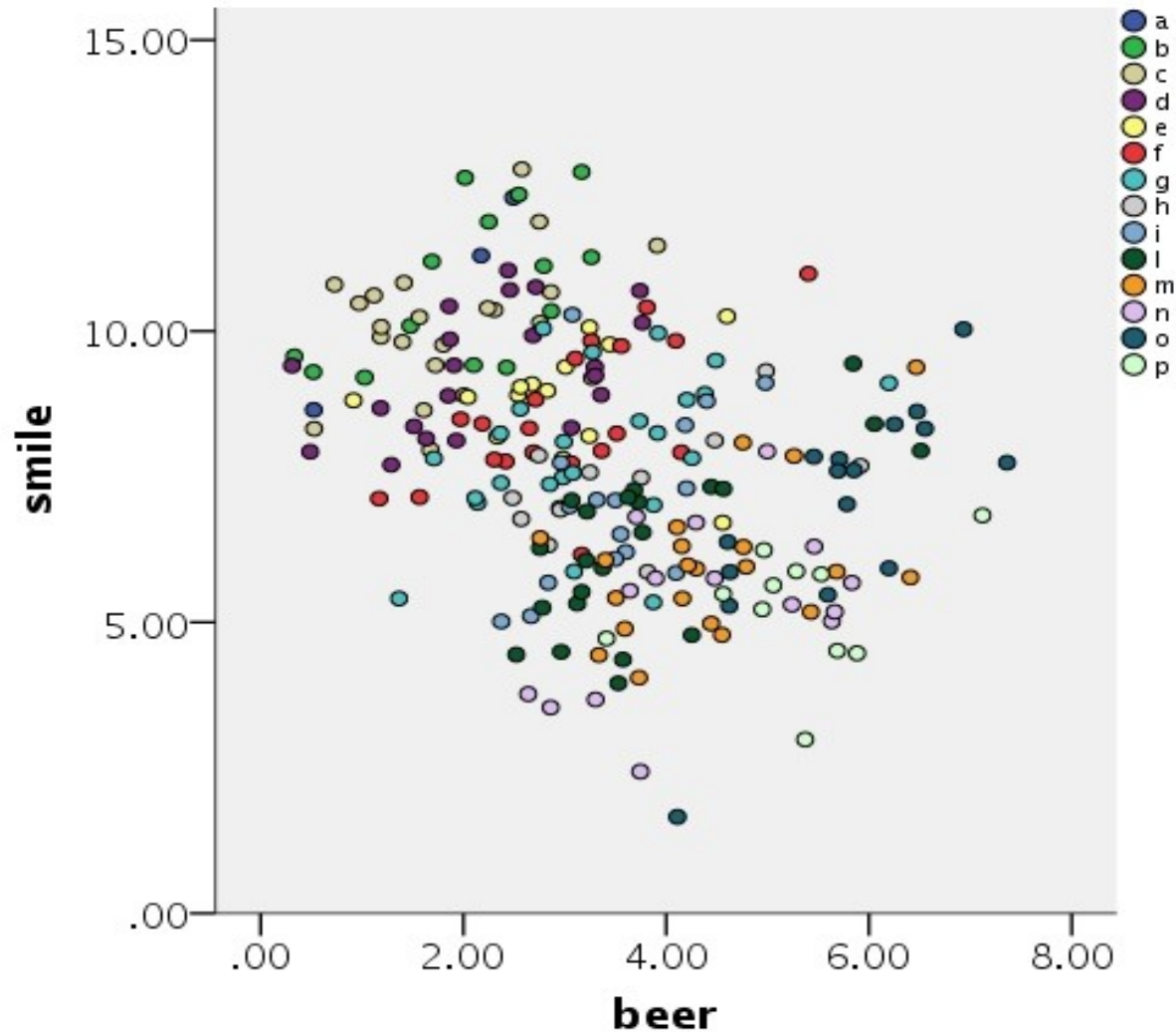
Why

Results may be biased by a mis-specification of the model, where the structure of the data is not taken into account

- In fact:
 - Subjects are sampled in clusters specified by **bars**
 - Each bar may have specific characteristics (quality, entertainment, etc) that may affect the measured variables
 - Subjects within the same bar may be more similar than across bars

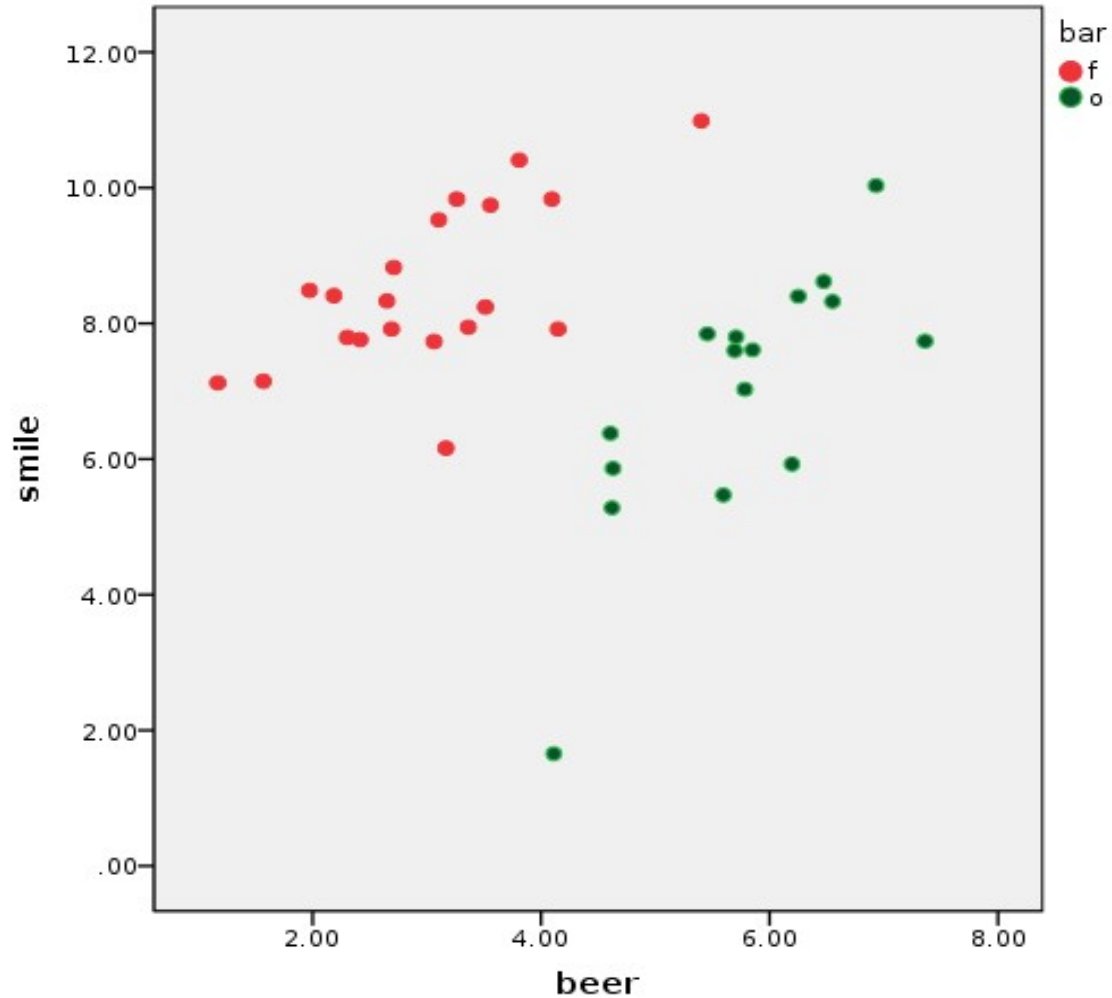
Scatterplot by Bar

Let's see the data broken down by bar



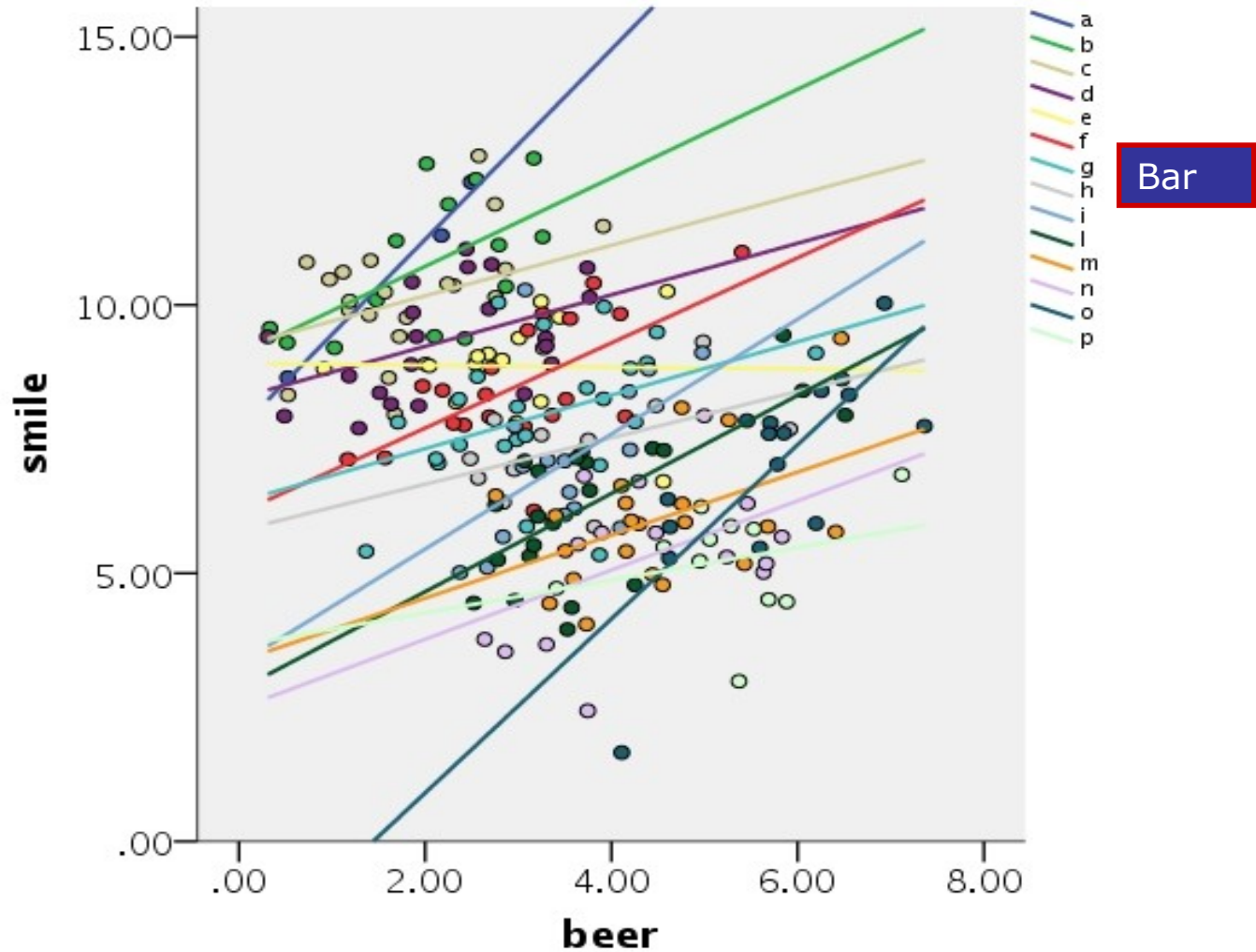
Scatterplot by Bar

Let's see the data only for bar “f” and “o”

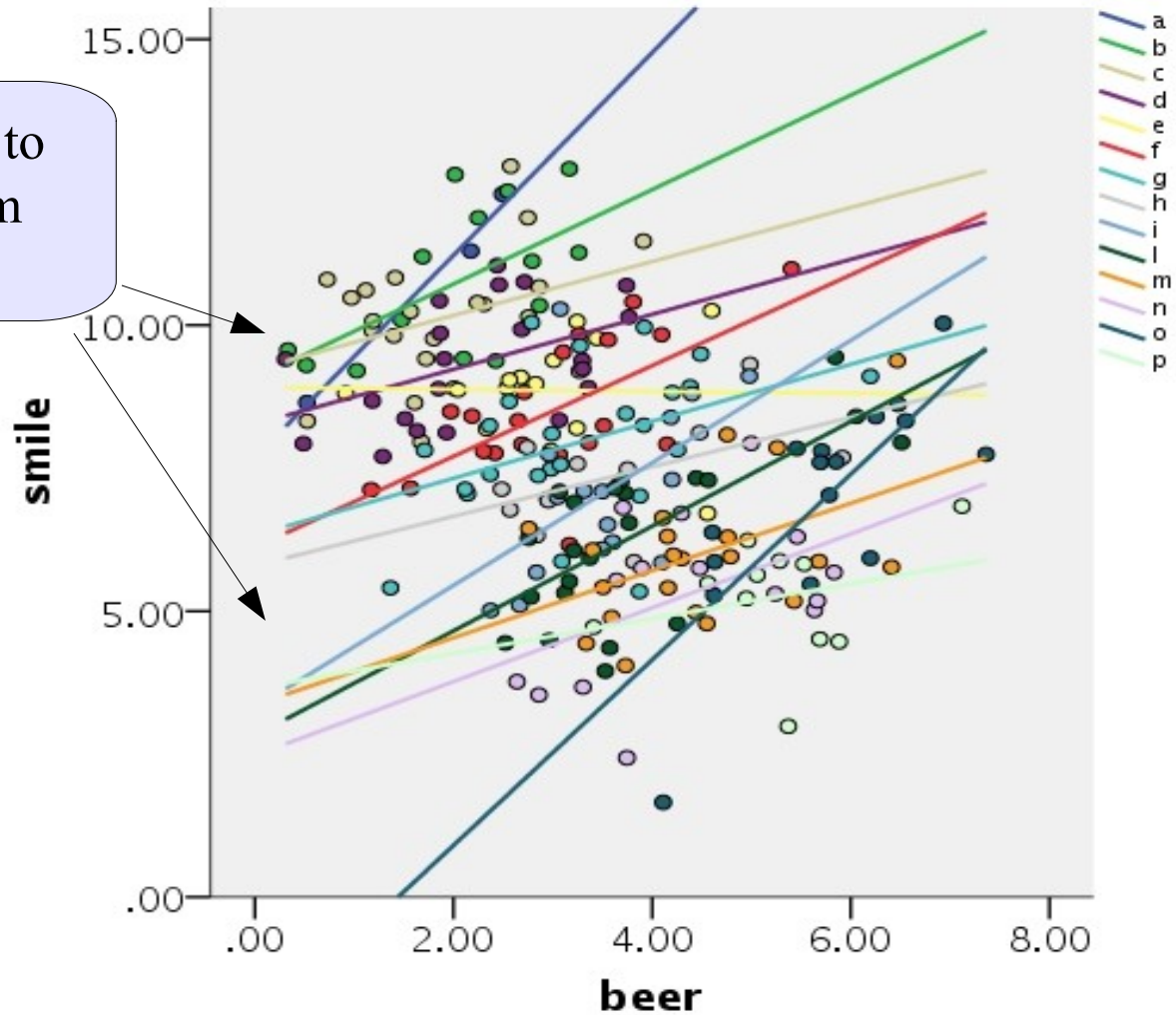


Scatterplot by Bar

It seems that the relations between IV and DV is positive, but within each bar



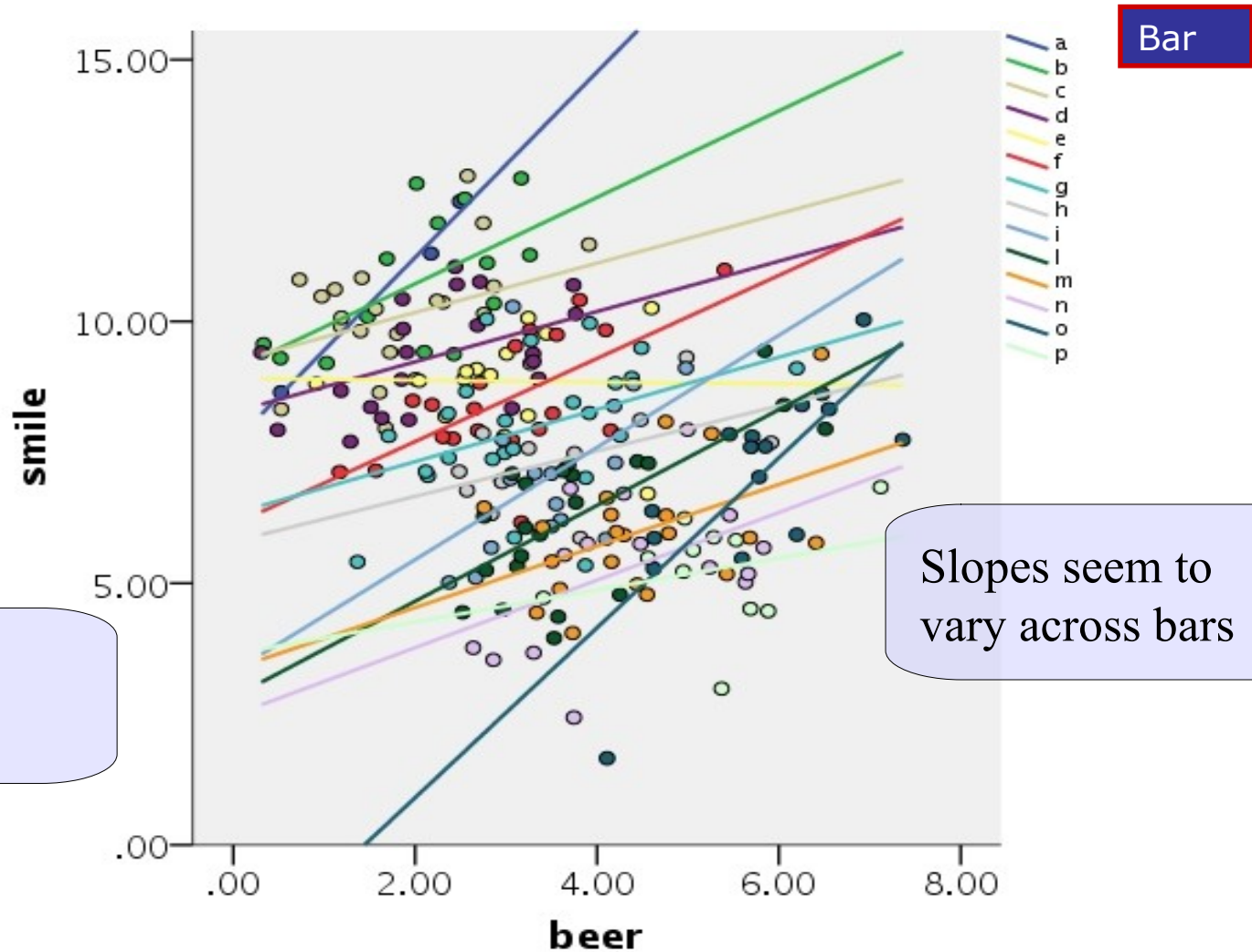
Scatterplot by Bar



Bar

Intercepts seem to be different from bar to bar

Scatterplot per Bar



Slopes are all positive

Slopes seem to vary across bars

The Model

- It seems that considering the participants as all equivalent and independent one each other (GLM assumption) does not fit our data
- It seems that a better model should allow each bar (each cluster) to have a different regression line (a different intercept and **b** coefficient)

The Model

- Let's define a model with a regression line for each cluster

 y_{ij}

Smiles of subject i in the cluster j

$$\hat{y}_{ia} = a_a + b_a \cdot x_{ia}$$

$$\hat{y}_{ib} = a_b + b_b \cdot x_{ib}$$

$$\hat{y}_{ic} = a_c + b_c \cdot x_{ic}$$

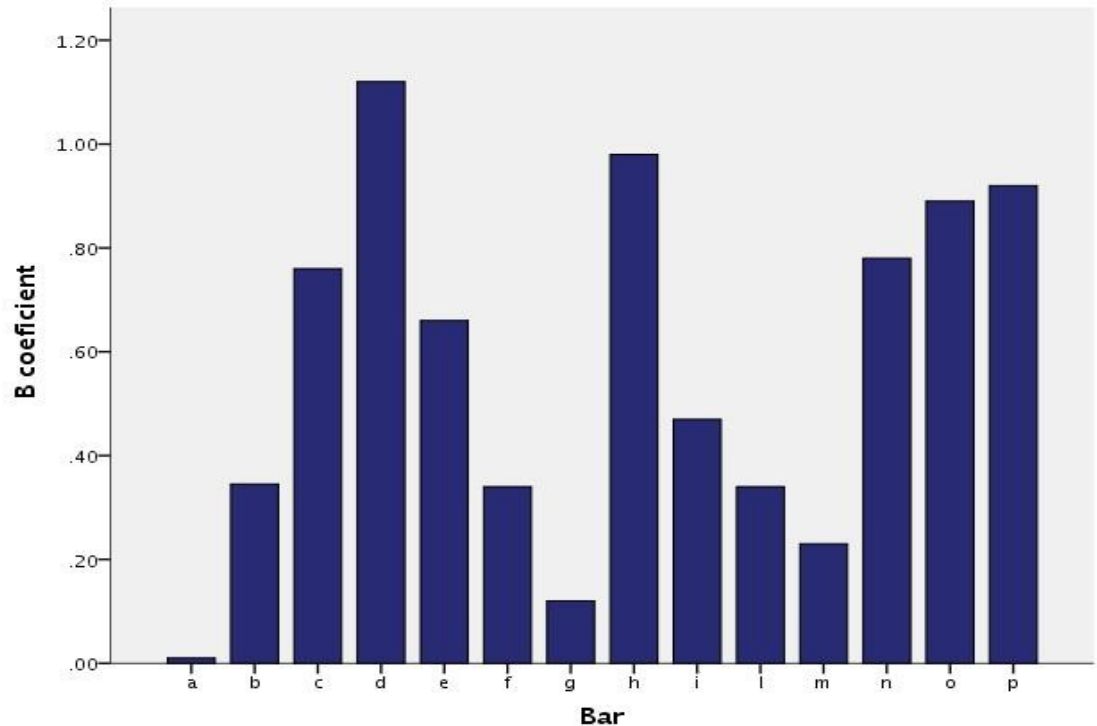
$$\hat{y}_{ij} = a_j + b_j \cdot x_{ij}$$

In these regressions the coefficients may vary from cluster to cluster: **they are not Fixed**

Varying coefficients

- If coefficients may vary, they will have a distribution

A possible distribution of coefficients **b** estimated for different clusters

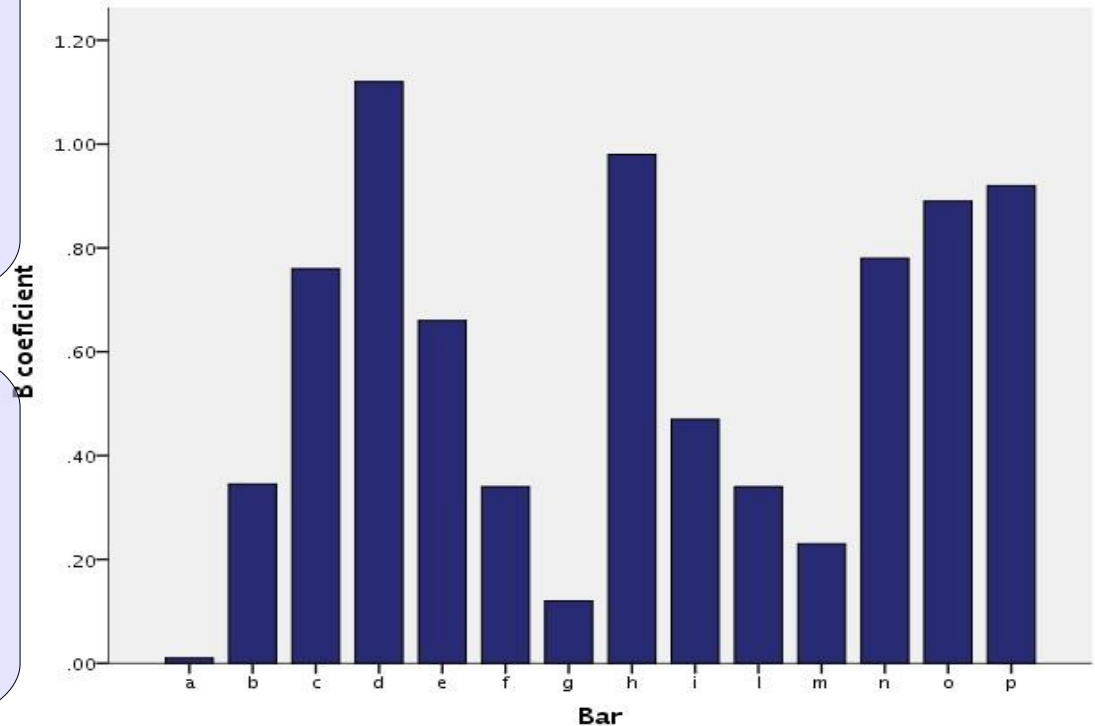


Random coefficients

- Varying coefficients are called random coefficients

Coefficients will exhibit variability

That is: in the **population** there exist different coefficients, a sample of which we estimated using the clustered data

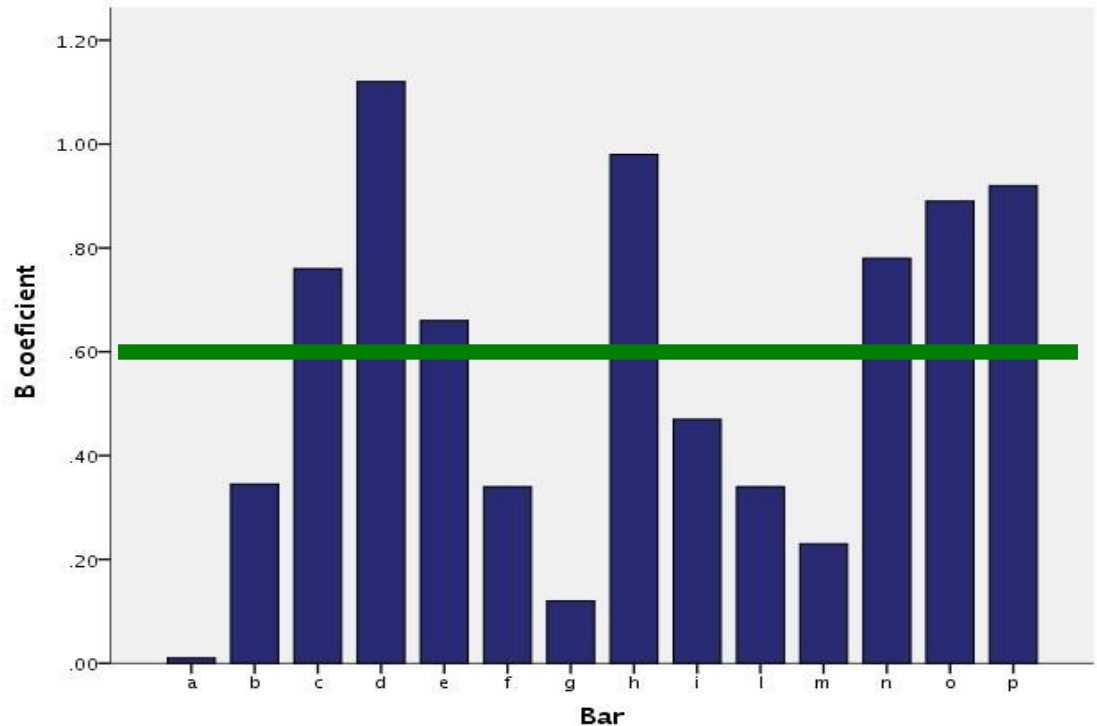


Average of the coefficients

- If coefficients vary as a variable in the population, they will have a mean and a variance, that we can estimate in our data

Mean

$$\bar{b} = \frac{\sum_j b_j}{k}$$



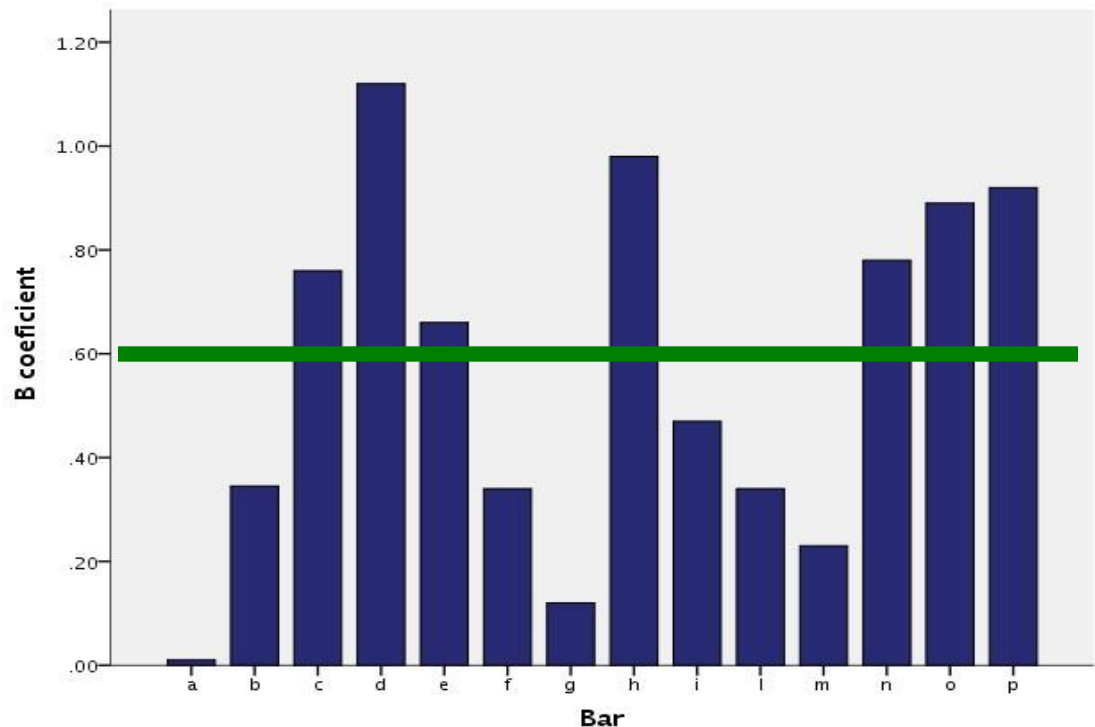
Fixed coefficients

- If coefficients vary as a variable in the population, they will have a mean and a variance, that we can estimate in our data

Mean

$$\bar{b} = \frac{\sum_j b_j}{k}$$

Recall the mean is a fixed parameter for a distribution, and so is the mean of the coefficients: it is a **fixed effect**



The Model

- We can now define a model with a regression for each cluster and the mean values of coefficients

One regression per cluster

$$\hat{y}_{ij} = a_j + b_j \cdot x_{ij}$$

Each coefficient is defined as the deviation from the mean coefficient

$$b'_j = b_j - \bar{b}$$

Overall model

$$\hat{y}_{ij} = a_j + b'_j \cdot x_{ij} + \bar{b} \cdot x_{ij}$$

The Model

- We can now define a model with a regression for each cluster and the mean value of coefficients

Overall model

$$\hat{y}_{ij} = a_j + b'_j \cdot x_{ij} + \bar{b} \cdot x_{ij}$$

Random coefficients

Fixed coefficient

The mixed model

- The same goes for the intercepts

One regression per cluster

$$\hat{y}_{ij} = a_j + b_j \cdot x_{ij}$$

Intercepts as deviations from the average intercept

$$a'_j = a_j - \bar{a}$$

Overall model

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$

The mixed model

- We can now define a model with a regression for each cluster and the mean values of coefficients

Overall model

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$

Random coefficients

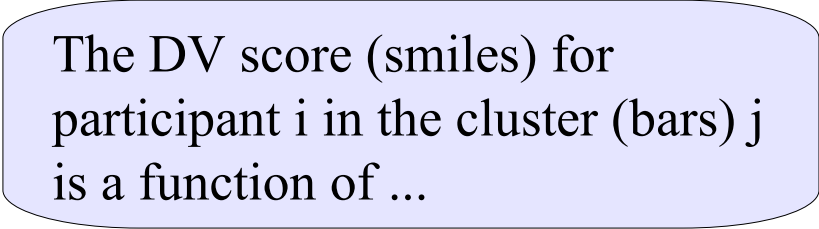
Fixed coefficients

A GLM which contains both fixed and random effects is called a Linear Mixed Model

The mixed model

- Interpretation

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$



The DV score (smiles) for participant i in the cluster (bars) j is a function of ...

The mixed model

Interpretation: y_{ij} is a function of...

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$

The average of the expected value for $x=0$, across clusters

That is: For $x=0$, how big is y on average

The mixed model

Interpretation: y_{ij} is a function of...

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$


Each expected value of y for $x=0$, in each cluster (bar) as a deviation from the overall intercept

For $x=0$, how much to add or subtract to the average intercept for that particular cluster

The mixed model

Interpretation: y_{ij} is a function of...

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$




The average effect of x on y,
averaged across clusters

On average, the expected change
in the DV for a unit change in
the IV

The mixed model

Interpretation: y_{ij} is a function of...

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$



The deviation from the average effect of x for cluster j

How much the effect of x increases (or decreases) in that particular cluster

GLM as a special case

It is clear that everything we know for the GLM applies here: the GLM is in fact a special case of the LMM, where there are not random effects

LMM

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$

GLM

$$\hat{y}_{ij} = \bar{a} + \bar{b} \cdot x_{ij}$$

Notation

For clarity, it's better to use the following notation

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + b_j \cdot x_{ij} + e_{ij}$$

y_{ij}, x_{ij}

Scores for case i in cluster j

\bar{a}, \bar{b}

Fixed effects

a_j, b_j

Effetti random as deviation from their
mean

e_{ij}

Error associated to the case i

Variance

For clarity, it's better to use the following notation

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + b_j \cdot x_{ij} + e_{ij}$$

 σ_a

Coefficients a variance

 σ_b

Coefficients b variance

 σ

Error variance (residual)

 σ_{ab}

Covariance between a and b coefficients

The mixed model

- In practice, mixed models allow to estimate the kind of effects we can estimate with the GLM, but they allow the effects to vary across clusters.
- Effects that vary across clusters are called **random effects**
- Effects that do not vary (the ones that are the same across clusters) are said to be **fixed effects**

The mixed model

- To specify a correct model, we only need to understand if there are **clusters of cases** (measures or subjects) and decide which coefficients (intercepts or b coefficients) may vary across those clusters
- The fixed effects of the model are interpreted like in the GLM (regression/ANOVA)
- **Random effects** are generally not interpreted, but we can look at their variance to decide to keep them as random (variance >0) or fix them.
- In this way we take into the account the dependence among data

Building a model

To build a model in a simple way, we need to answer very few questions:

- What is (are) the cluster variable(s)?
- What are the fixed effects?
- What are the random effects?

A clustering variable

- **What is (are) the cluster variable(s)?**
- What are the fixed effects?
- What are the random effects?
 - Any variable that groups observations (cases or measurements) such that scores may be more similar within each group than across groups
 - Any variable whose levels (groups) are a sample of a larger population of levels (groups)
 - Example: bars created groups of scores (participants) that may be more similar within the bar than across bars

Fixed effects

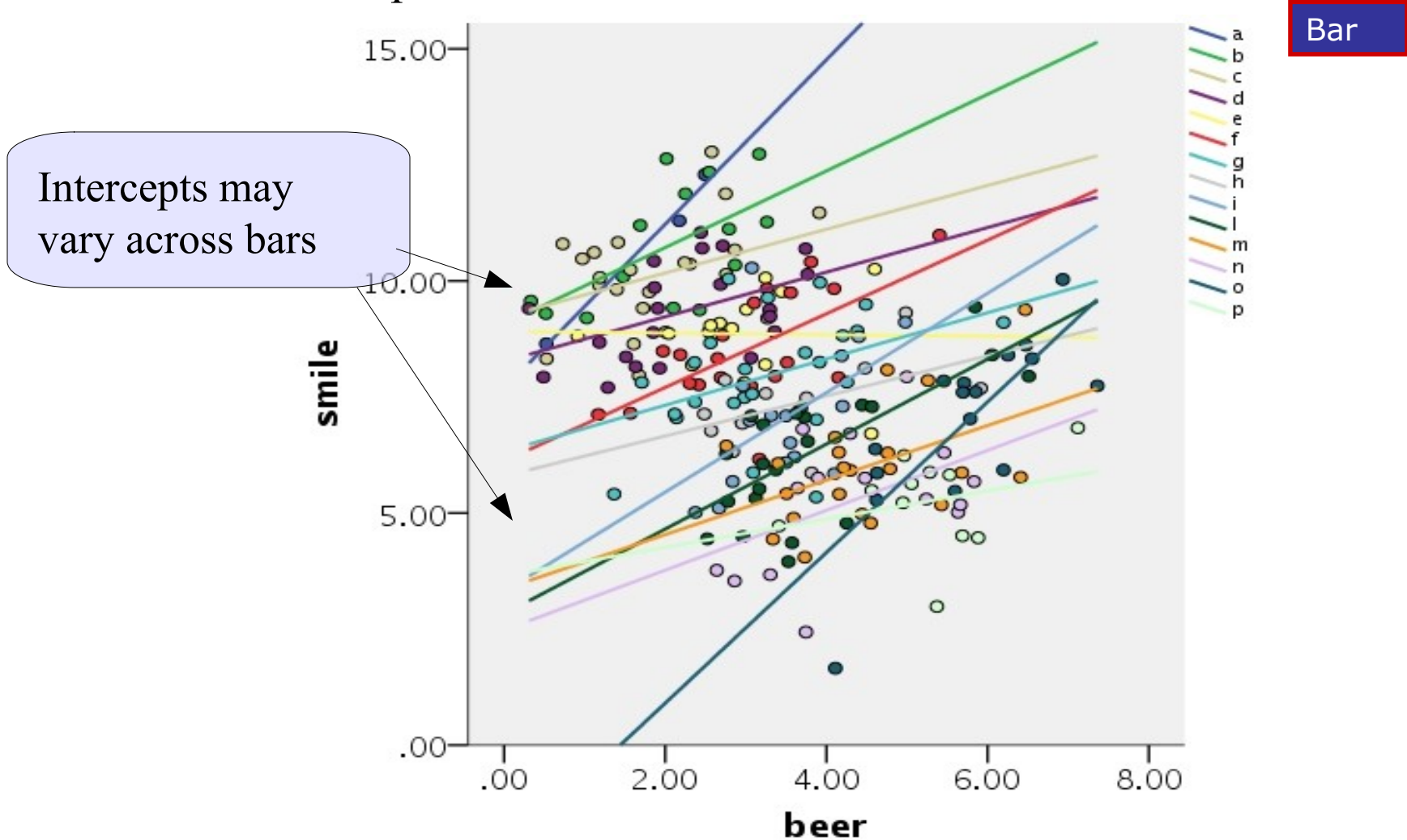
- What is (are) the cluster variable(s)?
- **What are the fixed effects?**
- What are the random effects?
 - Any effect that we are interested in on average (as in a standard ANOVA/Regression)
 - Example: the effect of beer on smiles in general

Fixed effects

- What is (are) the cluster variable(s)?
- What are the fixed effects?
- **What are the random effects?**
 - Any effect that may vary from cluster to cluster
 - (Thus:) **Any effect that can be computed within each cluster**
 - Example: the intercepts and the effect of beer on smiles each bar

Beers at the bar

We start with a simple model



Beers at the bar

We define a model where each cluster is allow to have a different intercept, the rest of the model is like a standard regression

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + e_{ij}$$

- Fixed effects? Intercept and beer effect
- Random effects? Intercepts
- Clusters? bar

Authors and books may call this model:
Random-intercepts regression
or
Intercepts-as-outcomes model

SPSS Input

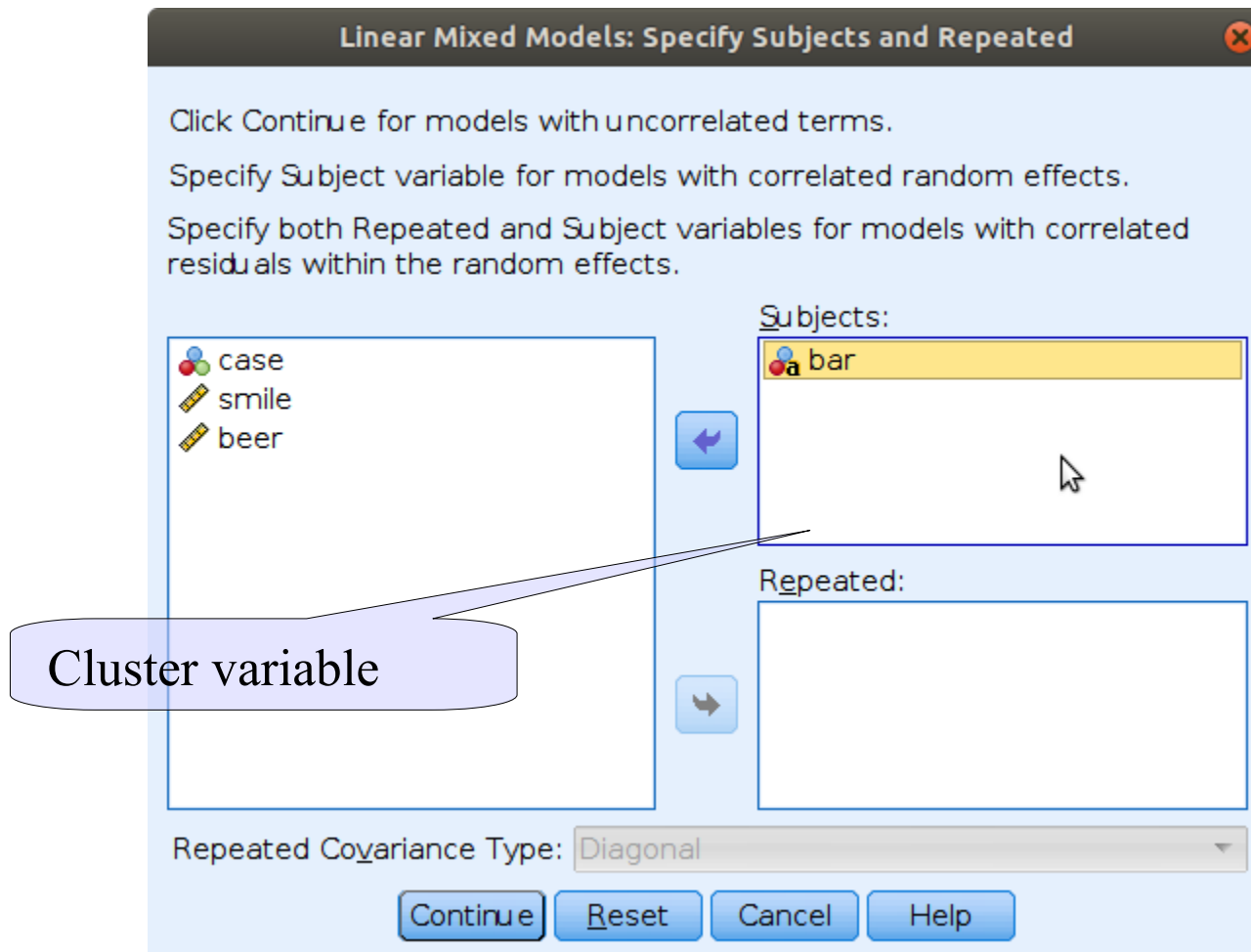
Analyze → Mixed Models → Linear

The screenshot shows the IBM SPSS Statistics Data Editor interface. The menu path 'Analyze → Mixed Models → Linear' is selected. The data table contains 16 rows and 4 columns. The 'case' column has values 1.00 and 1.00. The 'beer' column has values .52, 2.17, 2.01, 1.69, 1.03, 3.26, 2.43, 1.48, 2.10, .33, 2.55, 2.87, 2.25, .52, and 2.70. The 'var' columns are empty.

case	beer	var	var
1	1.00 a	.52	
2	1.00 a	2.17	
3	1.00 a		
4	1.00 b	2.01	
5	1.00 b	1.69	
6	1.00 b	1.03	
7	1.00 b	3.26	
8	1.00 b	2.43	
9	1.00 b	1.48	
10	1.00 b	2.10	
11	1.00 b	.33	
12	1.00 b	2.55	
13	1.00 b	2.87	
14	1.00 b	2.25	
15	1.00 b	.52	
16	1.00 b	2.70	

SPSS Input

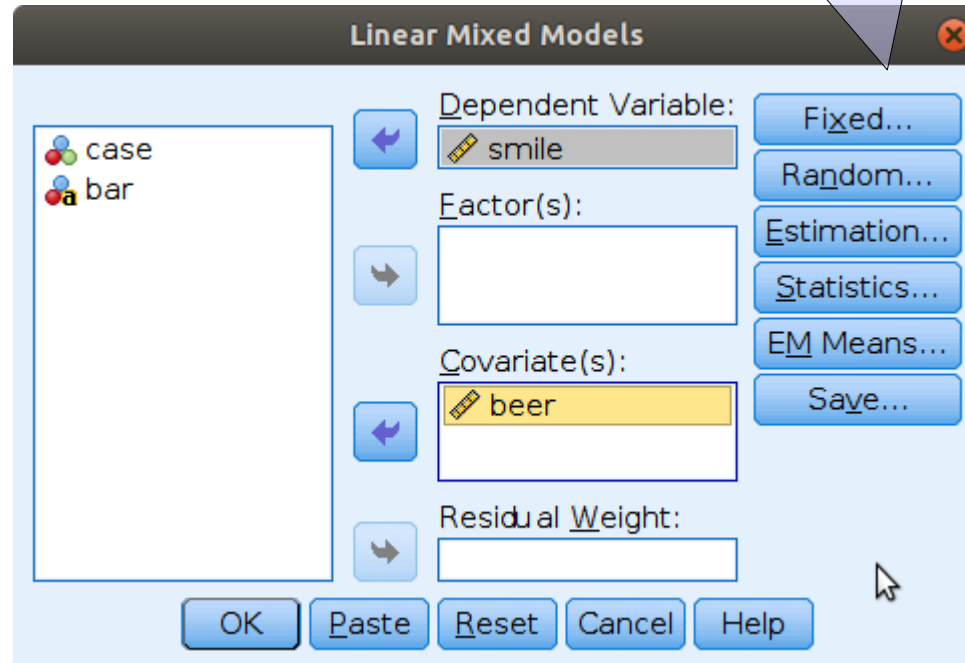
Analyze → Mixed Models → Linear



SPSS Input

Analyze → Mixed Models → Linear

Then select Fixed



SPSS Input

Analyze → Mixed Models → Linear

The screenshot shows the 'Linear Mixed Models: Fixed Effects' dialog box in SPSS. The window title is 'Linear Mixed Models: Fixed Effects'. Under the 'Fixed Effects' section, the 'Build terms' radio button is selected. The 'Factors and Covariates' list contains 'beer'. The 'Model' list also contains 'beer'. A 'Factorial' dropdown menu is positioned between the two lists. Below the lists are buttons for 'By*', '(Within)', 'Clear Term', 'Add', and 'Remove'. At the bottom, the 'Include intercept' checkbox is checked, and the 'Sum of squares' dropdown is set to 'Type III'. The 'Build Term:' text box is empty. At the very bottom are 'Continue', 'Cancel', and 'Help' buttons.

Linear Mixed Models: Fixed Effects

Fixed Effects

Build terms Build nested terms

Factors and Covariates: Model:

beer beer

Factorial

↓ By* (Within) Clear Term Add Remove

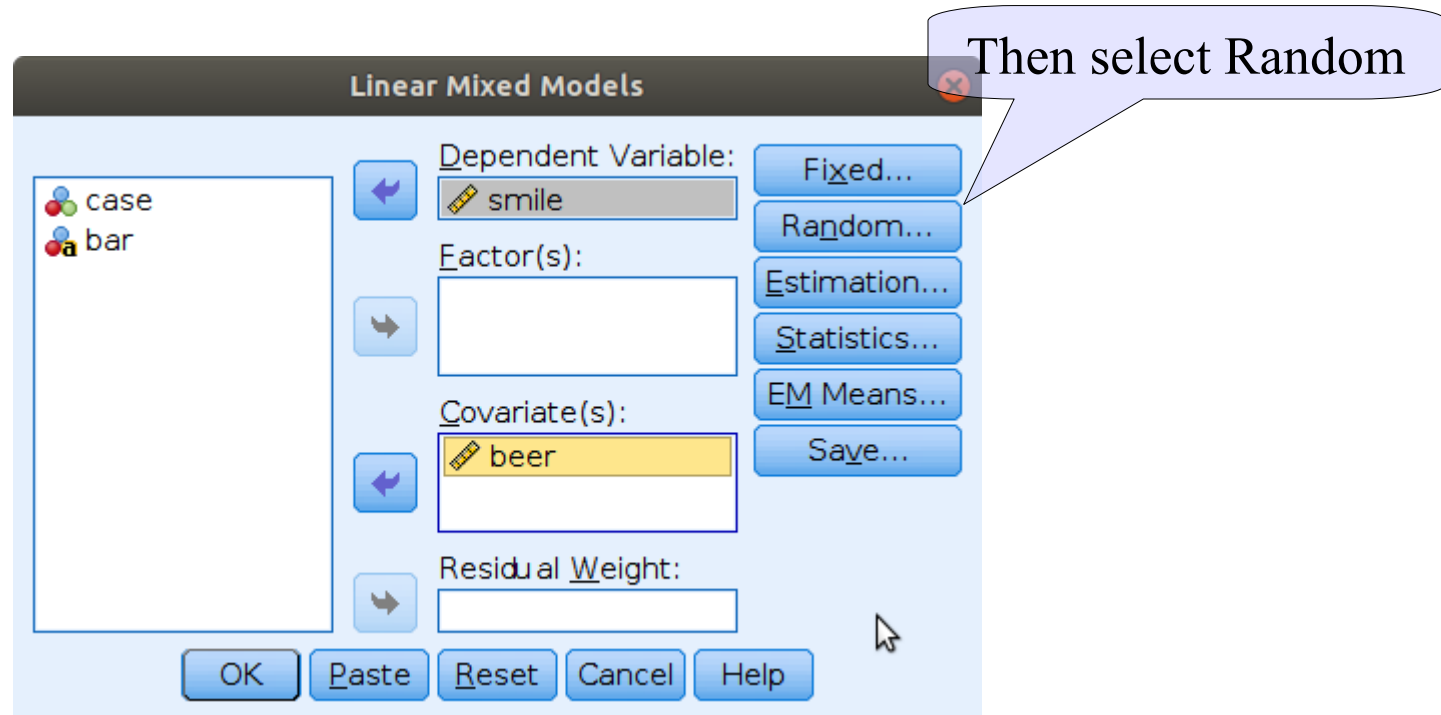
Build Term:

Include intercept Sum of squares: Type III

Continue Cancel Help

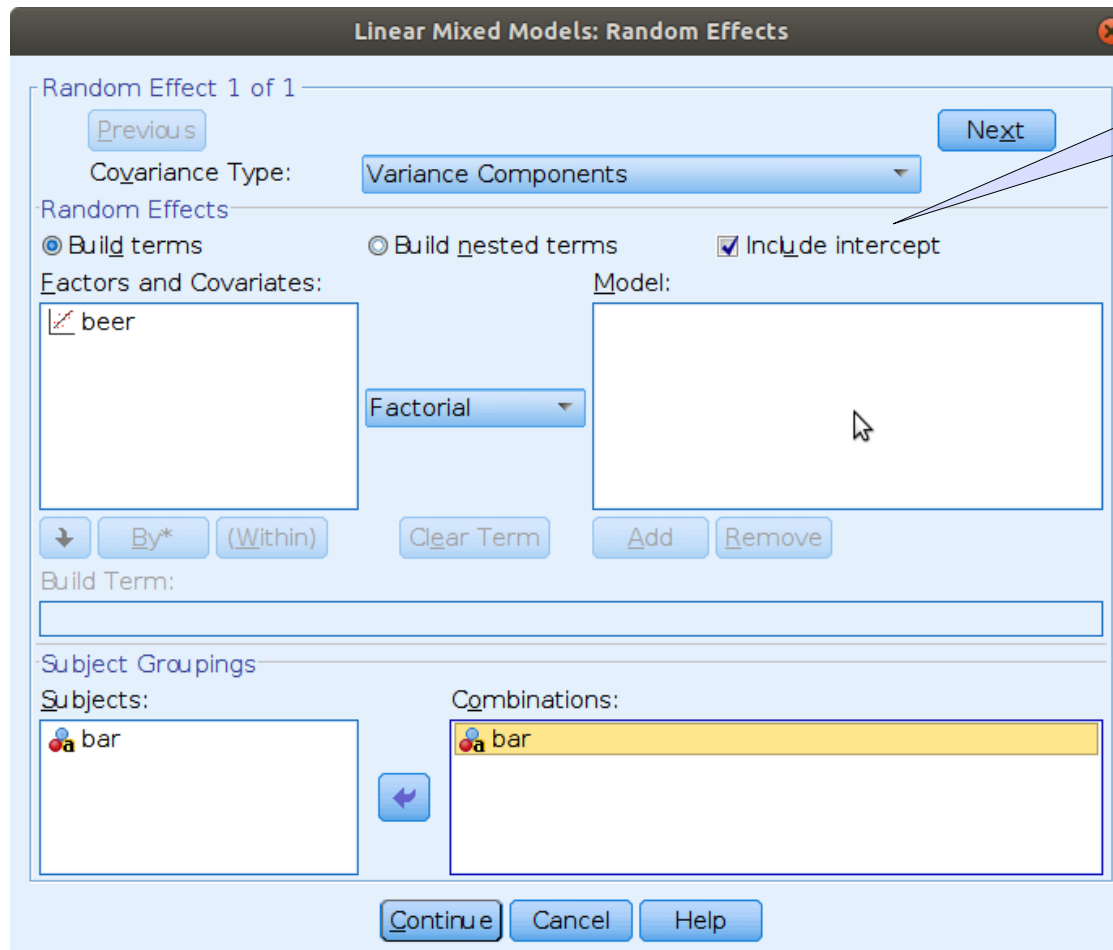
SPSS Input

Analyze → Mixed Models → Linear



SPSS Input

Analyze → Mixed Models → Linear

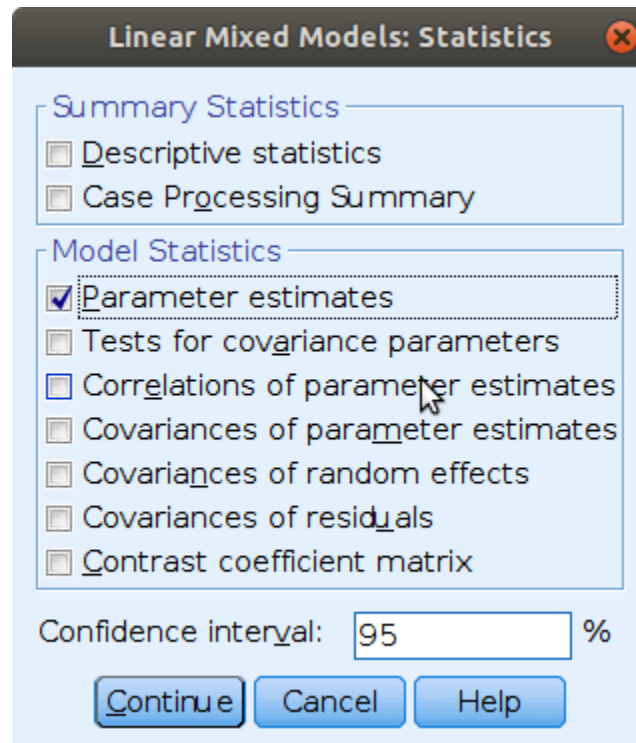


The image shows the 'Linear Mixed Models: Random Effects' dialog box in SPSS. The window title is 'Linear Mixed Models: Random Effects'. The 'Random Effect 1 of 1' section has 'Previous' and 'Next' buttons. The 'Covariance Type' is set to 'Variance Components'. Under 'Random Effects', 'Build terms' is selected, and the 'Include intercept' checkbox is checked. The 'Factors and Covariates' list contains 'beer'. The 'Model' list is empty. The 'Build Term' field is empty. The 'Subject Groupings' section has 'Subjects' containing 'bar' and 'Combinations' containing 'bar'. At the bottom are 'Continue', 'Cancel', and 'Help' buttons.

Random Intercept

SPSS Input

Analyze → Mixed Models → Linear



Print coefficients

SPSS syntax

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + e_{ij}$$

```
MIXED smile WITH beer by bar
```

```
/CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0,  
ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
```

```
/FIXED=beer| SSTYPE(3)
```

```
/METHOD=REML
```

```
/print solution TESTCOV
```

```
/random intercept | SUBJECT(bar) COVTYPE(un).
```

Fixed effects (intercept is included by default)

Random effects

Cluster variable

SPSS Output

Let's see if the model is how intended

Model Dimension^a

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	bar
	beer	1		1	
Random Effects	Intercept	1		1	
Residual				1	
Total		3		4	

a. Dependent Variable: smile.



OK!

SPSS Output

We then check the variability of the random effects. If there is variability across bars, it means we were right to model the coefficients as random

Covariance Parameters

Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1.284725	.123256	10.423	.000	1.064501	1.550509
Intercept [subject = bar] Variance	6.531614	2.584158	2.528	.011	3.007824	14.183668

a. Dependent Variable: smile.

Variance greater than 0

SPSS Output

If everything is fine, we interpret the fixed effects as in any other GLM (regression)

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	5.551071	.724190	18.295	7.665	.000	4.031359	7.070783
beer	.638704	.077690	227.901	8.221	.000	.485621	.791787

a. Dependent Variable: smile.

Intercept: On average, for zero beers we expect 5.5 smiles

SPSS Output

If everything is fine, we interpret the fixed effects as in any other GLM (regression)

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	5.551071	.724190	18.295	7.665	.000	4.031359	7.070783
beer	.638704	.077690	227.901	8.221	.000	.485621	.791787

a. Dependent Variable: smile.

B coefficient: On average, for each beer more, we expect .638 more smiles

R syntax

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + e_{ij}$$

```
random_beers.R* x
Source on Save
1 setwd("/home/marcello/Skinner/Teaching/Phd/MIB")
2 library(lme4)
3 library(lmerTest)
4 library(foreign)
5 dat<-read.spss('data/regression_beers_bars.sav',to.data.frame = T)
6 head(dat)
7 mm1<-lmer(smile~1+beer+(1|bar),data=dat)
8 summary(mm1)
9 |
```

Load the required libraries

Random effects

Cluster variable

Fixed effects (intercept can be omitted as ti is included by default)

R Output

Let's see if the model is how it was intended

```
> summary(mm1)
Linear mixed model fit by REML ['merModLmerTest']
Formula: smile ~ 1 + (1 | bar) + beer
Data: dat

REML criterion at convergence: 786.7
```



OK!

R Output

We then check the variability of the random effects. If there is variability across bars, it means we were right to model the coefficients as random

```
Random effects:
  Groups   Name      Variance Std.Dev.
  bar      (Intercept) 6.532    2.556
  Residual                1.285    1.133
Number of obs: 234, groups: bar, 15
```

Variance greater than 0

R Output

If you need a test for the random variances (Likelihood Ratio Test) run this:

```
dat<-read.spss('data/regression_beers_bars.sav',to.data.frame = T)
head(dat)
mm1<-lmer(smile~1+beer+(1|bar),data=dat)
rand(mm1)
```

```
> rand(mm1)
Analysis of Random effects Table:
  Chi.sq Chi.DF p.value
bar    201     1 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

OK!

R Output

If everything is fine, we interpret the fixed effects as in any other GLM (regression)

```
Fixed effects:
      Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  5.55107    0.72419  18.29000   7.665 4.03e-07 ***
beer         0.63870    0.07769 227.88000   8.221 1.53e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Intercept: On average, for zero beers we expect 5.5 smiles

Note that without library(lmerTest) you do not get the p.values!

R Output

If you prefer the F-test, use `anova()`



Default DF

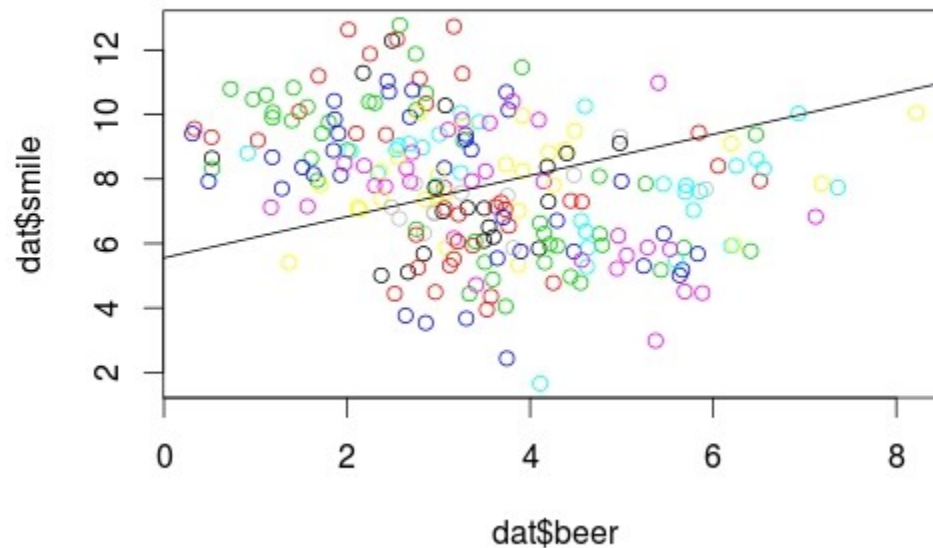
```
> anova(mm1)
Analysis of Variance Table of type 3 with Satterthwaite
approximation for degrees of freedom
      Sum Sq Mean Sq NumDF  DenDF F.value    Pr(>F)
beer  86.831  86.831     1 227.88  67.588 1.532e-14 ***
```

R Output

Plot fixed effects:

```
2 summary(mm1)
3 plot(dat$smile~dat$beer,col=dat$bar)
4 abline(fixef(mm1))
5
```

fixef extracts fixed coefficients from the model

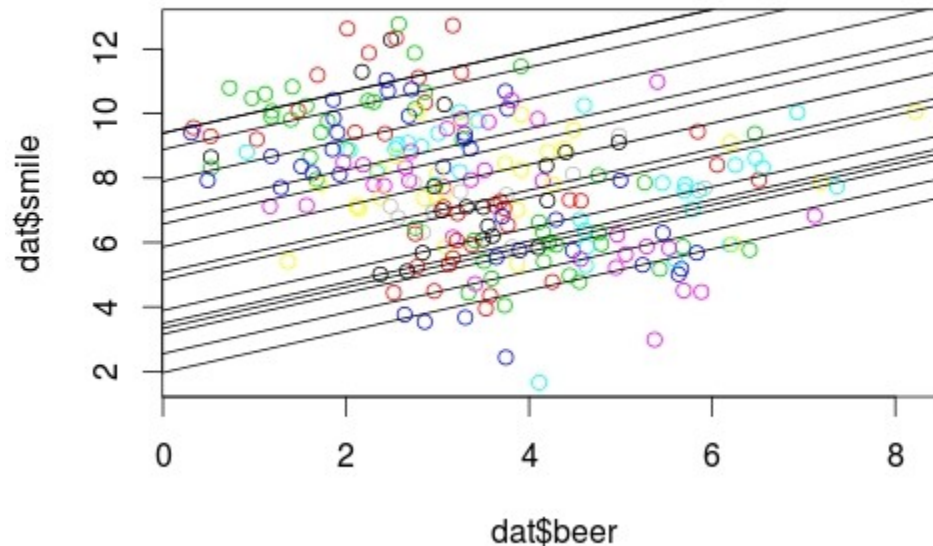


R Output

Plot random effects effects:

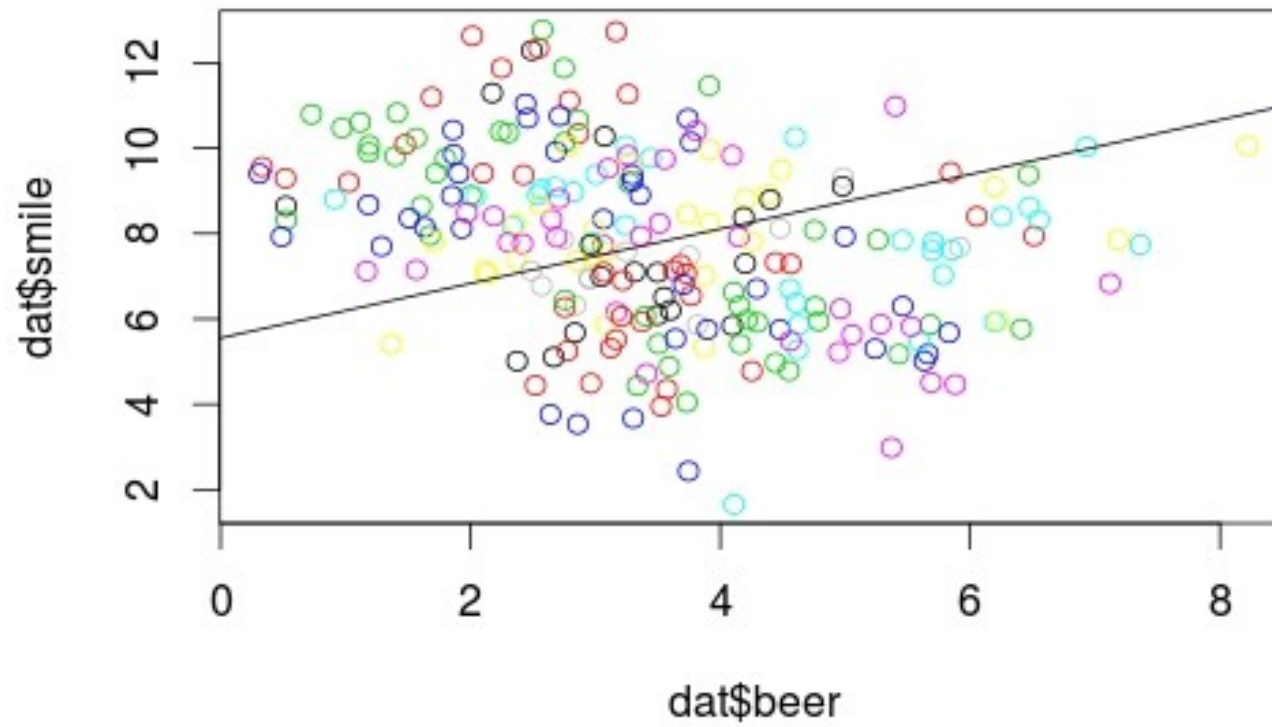
```
16  
17 plot(dat$smile~dat$beer,col=dat$bar)  
18 apply(coef(mm1)[[1]],1,abline)  
19 |
```

coef extracts random coefficients from the model



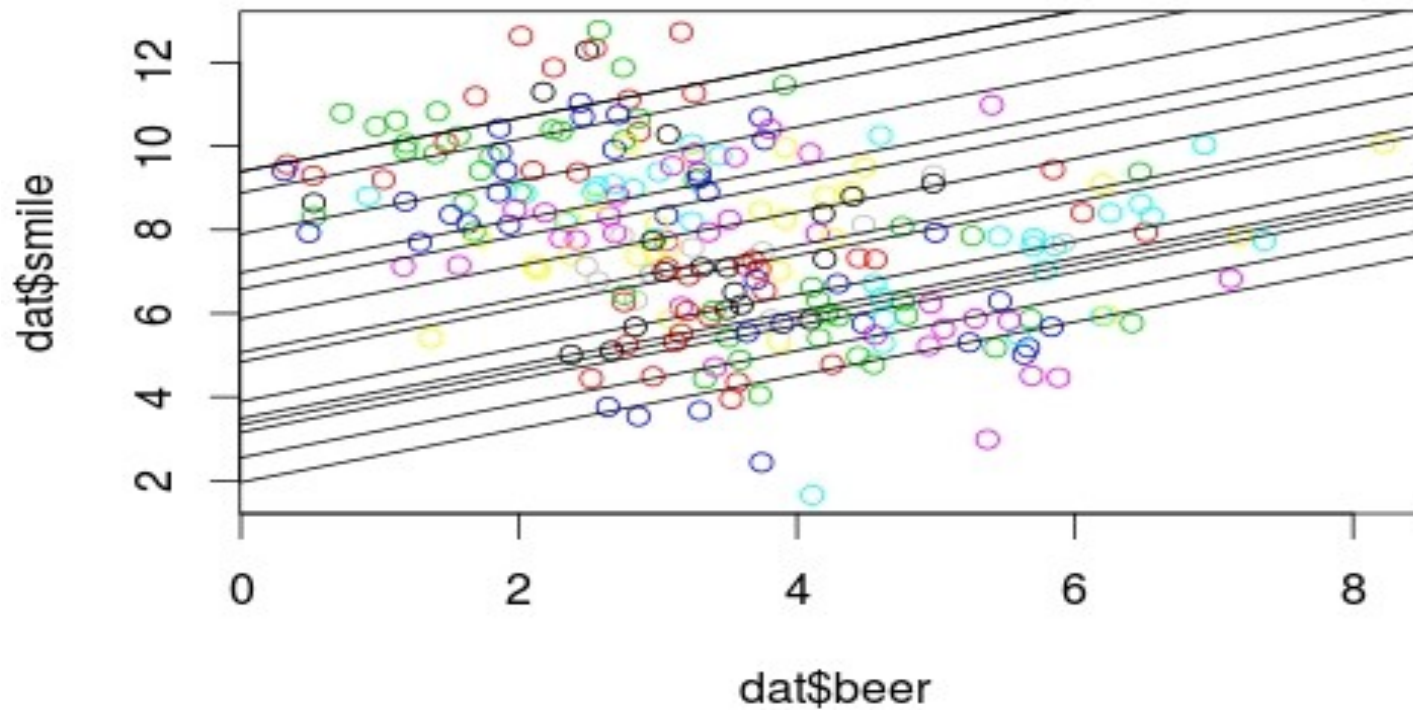
Plots

- Plotting fixed effects is simply plotting the straight line as in any linear model



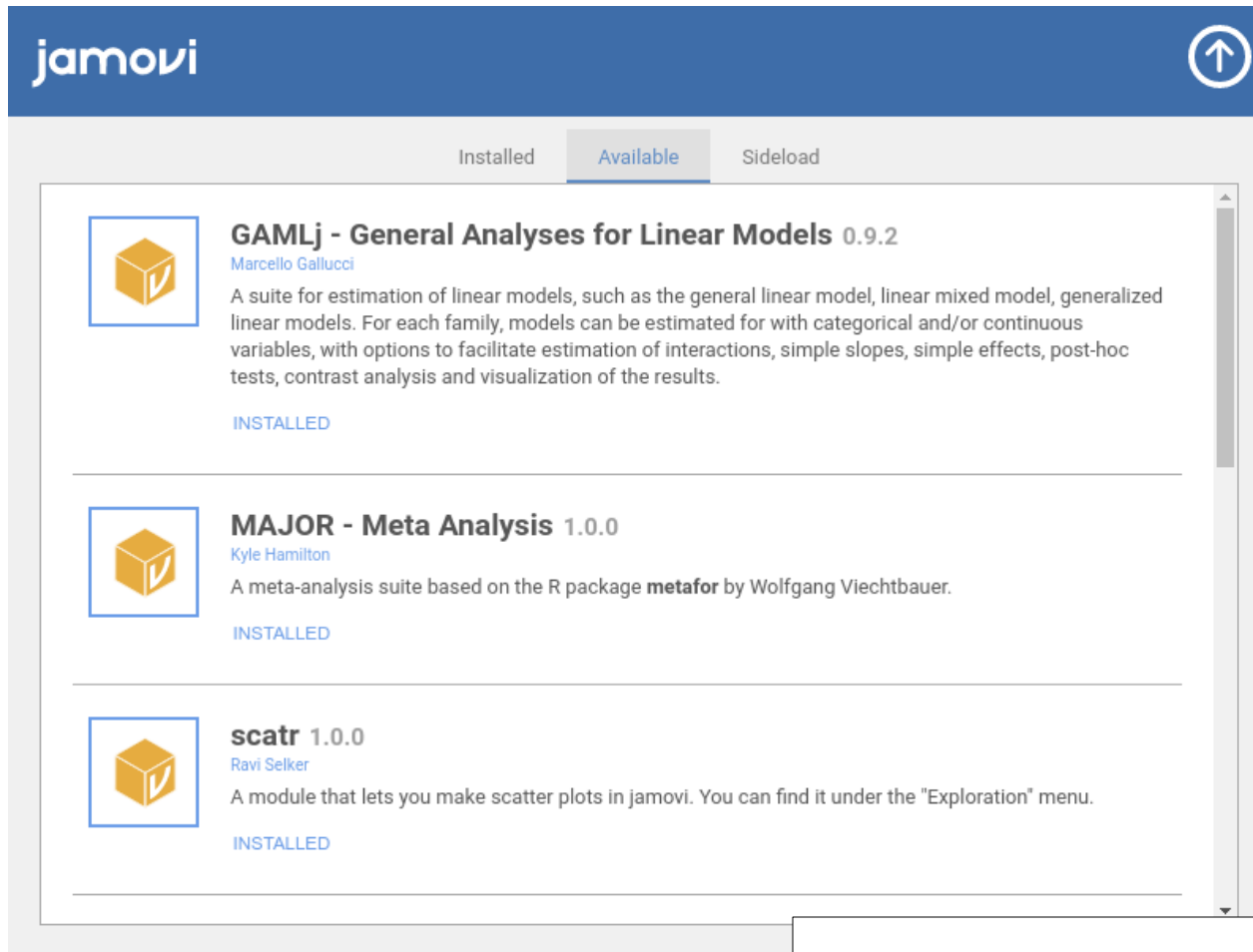
Plots

- Plotting random effects effects is plotting each specific sub-model of each cluster



Jamovi

- In jamovi mixed models can be estimated with the GAMLj module



The screenshot shows the Jamovi software interface. At the top left is the 'jamovi' logo. To the right is a circular icon with an upward-pointing arrow. Below the logo is a navigation bar with three tabs: 'Installed', 'Available', and 'Sideload'. The 'Available' tab is selected. The main content area displays three modules, each with a yellow cube icon containing a white checkmark, indicating they are installed. The first module is 'GAMLj - General Analyses for Linear Models 0.9.2' by Marcello Gallucci. The second is 'MAJOR - Meta Analysis 1.0.0' by Kyle Hamilton. The third is 'scatr 1.0.0' by Ravi Selker. Each module entry includes a brief description and the status 'INSTALLED'.

GAMLj - General Analyses for Linear Models 0.9.2
Marcello Gallucci
A suite for estimation of linear models, such as the general linear model, linear mixed model, generalized linear models. For each family, models can be estimated for with categorical and/or continuous variables, with options to facilitate estimation of interactions, simple slopes, simple effects, post-hoc tests, contrast analysis and visualization of the results.
INSTALLED

MAJOR - Meta Analysis 1.0.0
Kyle Hamilton
A meta-analysis suite based on the R package **metafor** by Wolfgang Viechtbauer.
INSTALLED

scatr 1.0.0
Ravi Selker
A module that lets you make scatter plots in jamovi. You can find it under the "Exploration" menu.
INSTALLED

Docs and examples: https://mcfanda.github.io/gamlj_docs/

Jamovi

- In jamovi mixed models can be estimated with the GAMLj module

The screenshot displays the Jamovi software interface for configuring a Mixed Model. The top menu bar shows 'Data' and 'Analyses' tabs. The toolbar includes icons for Exploration, T-Tests, ANOVA, Regression, Frequencies, Factor, Base R, TOSTER, MAJOR, medmod, Linear Models, and Modules. The 'Mixed Model' panel is active, showing a list of variables (A, case, smile, beer, bar) on the left. The configuration fields include 'Dependent Variable', 'Factors', 'Covariates', and 'Cluster variables'. Below these are sections for 'Estimation' (REML checked) and 'Confidence Intervals' (95% interval). A list of expandable panels is at the bottom, including 'Fixed Effects', 'Random Effects', 'Factors Coding', 'Covariates Scaling', 'Post Hoc Tests', 'Fixed Effects Plots', 'Simple Effects', and 'Estimated Marginal Means'. A callout bubble points to this list with the text 'All options are in expandable panels'.

All options are in expandable panels

Jamovi

- In jamovi mixed models can be estimated with the GAMLj module

Mixed Model

Define the variables role

→

Dependent Variable
smile

Factors

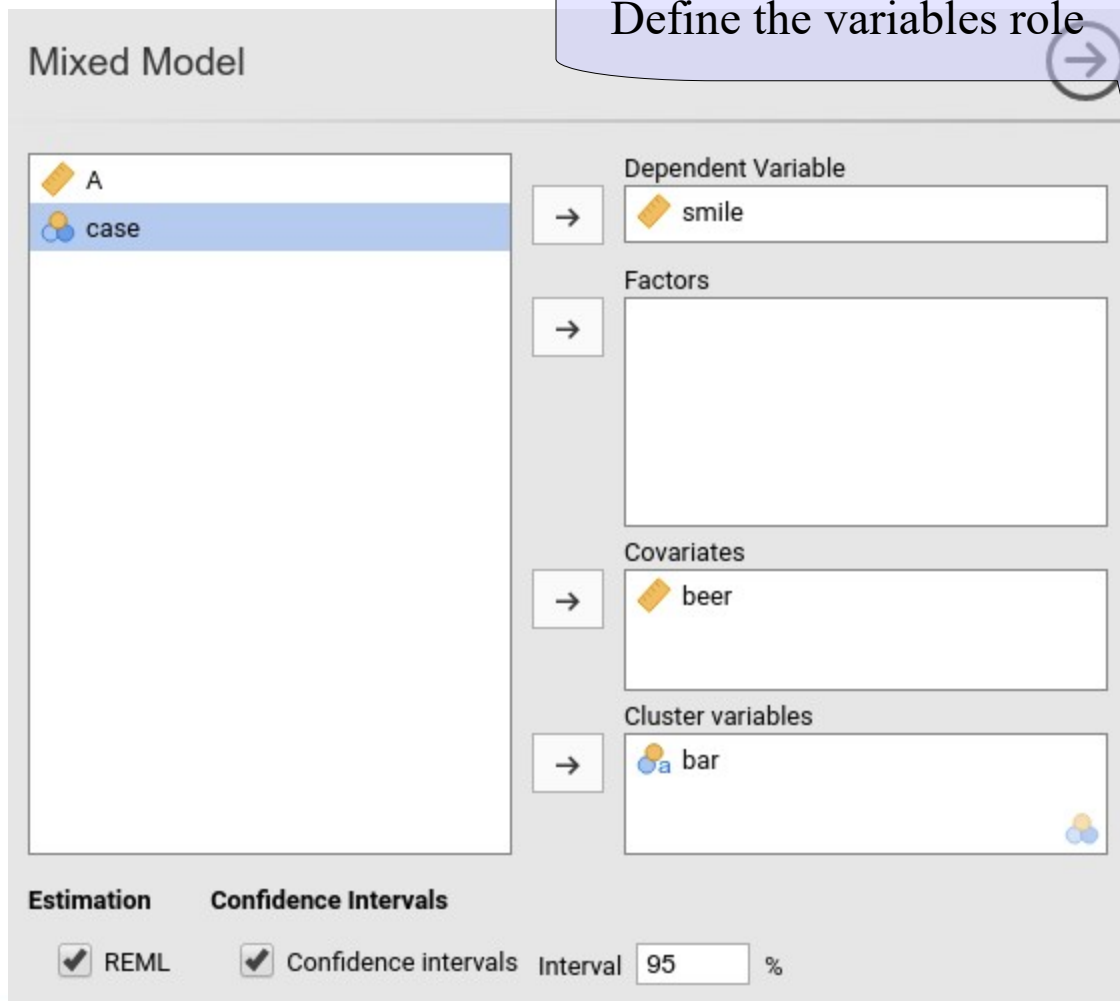
→

Covariates
beer

Cluster variables
bar

Estimation Confidence Intervals

REML Confidence intervals Interval %



Jamovi

- In jamovi mixed models can be estimated with the GAMLj module

Define the fixed effects

The screenshot shows the 'Fixed Effects' dialog box in the Jamovi GAMLj module. The dialog has a title bar with a dropdown arrow and the text 'Fixed Effects'. Below the title bar, there are two main sections: 'Components' on the left and 'Model Terms' on the right. The 'Components' section contains a list box with the word 'beer'. The 'Model Terms' section also contains a list box with the word 'beer'. Between these two sections are two buttons: a right-pointing arrow and a right-pointing arrow with a small downward-pointing arrow. At the bottom left of the dialog, there is a checked checkbox labeled 'Fixed Intercept'.

Jamovi

- In jamovi mixed models can be estimated with the GAMLj module

Define the random component

The screenshot shows the 'Random Effects' dialog box in the GAMLj module of Jamovi. The dialog is titled 'Random Effects' and has a dropdown arrow on the left. It is divided into two main sections: 'Components' on the left and 'Random Coefficients' on the right. In the 'Components' section, the text 'beer | bar' is highlighted in blue. A right-pointing arrow button is located between the two sections. In the 'Random Coefficients' section, the text 'Intercept | bar' is displayed. At the bottom left of the dialog, there is a checked checkbox labeled 'Correlated Effects'.

- As soon as you define the random component, you get the results

Mixed Model

Model Info

Info

Estimate	Linear mixed model fit by REML
Call	smile ~ 1 + (1 bar) + beer
AIC	811.1613
R-squared Marginal	0.0894
R-squared Conditional	0.8172

R-squared Marginal: How much variance can the **fixed effects alone** explain of the overall variance

R-squared Conditional: How much variance can the **fixed and random** effects together explain of the overall variance

Jamovi

- As soon as you define the random component, you get the results

F-test for the main effect of beer

Fixed Effect ANOVA

	F	Num df	Den df	p
beer	46.0	1	229	< .001

Note. Satterthwaite method for degrees of freedom

Jamovi

- As soon as you define the random component, you get the results

coefficients for the main effect of beer

Fixed Effects Parameter Estimates

Effect	Contrast	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	Intercept	7.778	0.6276	6.548	9.008	13.2	12.39	< .001
beer	beer	0.548	0.0808	0.390	0.706	229.4	6.79	< .001

Jamovi

- As soon as you define the random component, you get the results

Random components

Random Components

Groups	Name	SD	Variance
bar	(Intercept)	2.40	5.77
	Residual	1.20	1.45

Note. Numer of Obs: 234 , groups: bar , 15

Jamovi

- Jamovi can plot up to a 3-way interaction

Plot (here just one line)

Fixed Effects Plots

Horizontal axis
→ beer

Separate lines
→

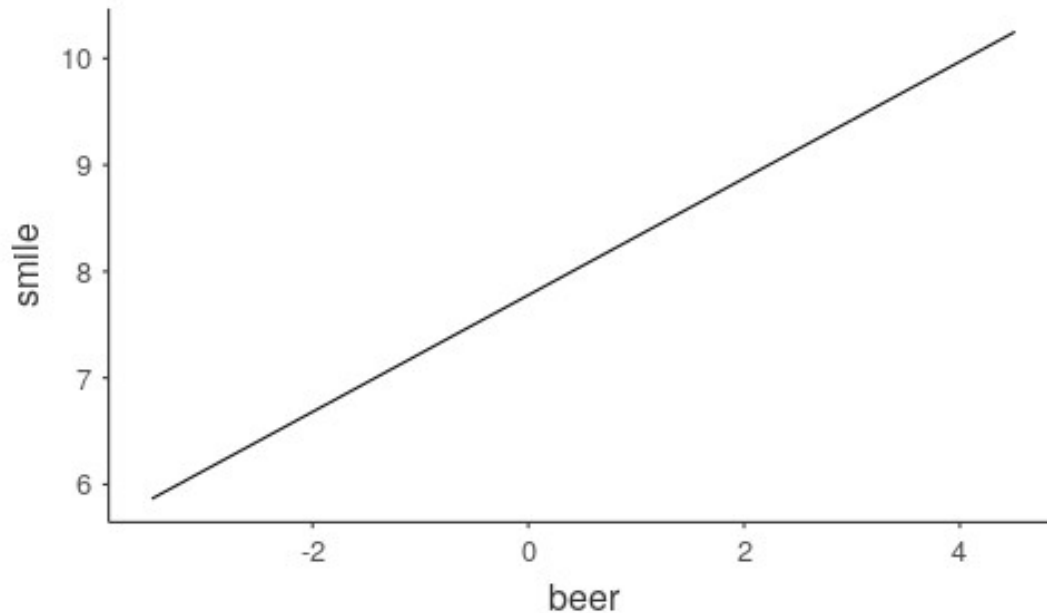
Separate plots
→

Jamovi

- Jamovi can plot up to a 3-way interaction fixed effects

Plot (here just one line)

Fixed Effects Plots

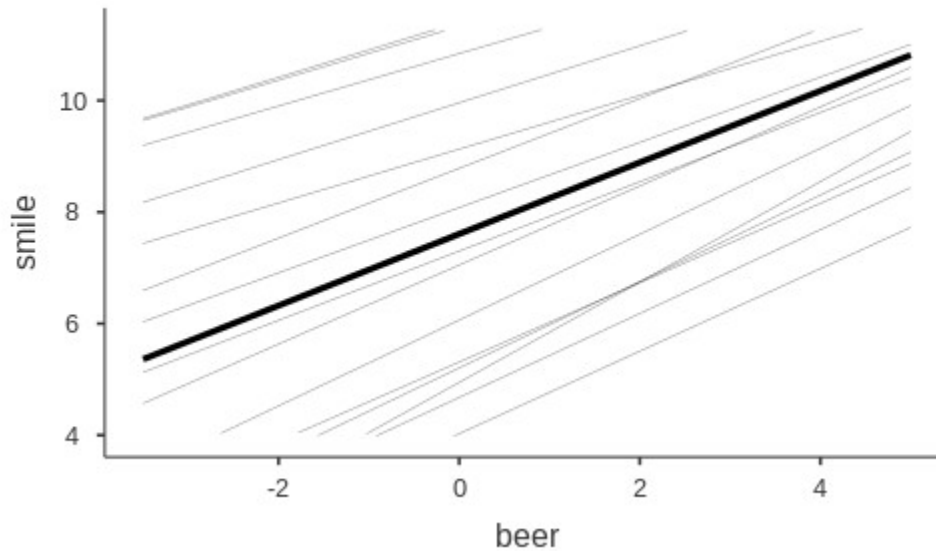


Jamovi

- Jamovi can plot up to a 3-way interaction fixed effects

Random effects

Fixed Effects Plots



Beers at the bar 2

We can now try a model where also the **b** coefficients are allow to vary across clusters

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + b \cdot x_{ij} + e_{ij}$$

- Fixed effects? Intercept and beer effect
- Random effects? Intercepts and b coefficients
- Clusters? bar

Some authors may call this model:
Random-coefficients regression
or
Intercepts- and Slopes-as-outcomes model

SPSS syntax

Now we have all the fixed effects and also the random effects

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + b \cdot x_{ij} + e_{ij}$$

```
MIXED smile WITH beer by bar  
/CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0,  
ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)  
/FIXED=beer| SSTYPE(3)  
/METHOD=REML  
/print solution TESTCOV  
/random intercept beer | SUBJECT(bar) COVTYPE(un).
```

Fixed effects (intercept is included by default)

Random effects

Cluster variable

SPSS Output

Let's see if the model is how intended

Model Dimension^b

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Unstructured	1	bar
	beer	1		1	
Random Effects	Intercept + beer ^a	2		3	
Residual				1	
Total		4		6	

a. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

b. Dependent Variable: smile.



OK!

SPSS Output

We then check the variability of the random effects. If there is variability across bars, it means we were right to model the coefficients as random

Covariance Parameters

Estimates of Covariance Parameters^a

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		1.258809	.125848	10.003	.000	1.034814	1.531292
Intercept + beer [subject = bar]	UN (1,1)	9.334205	4.379192	2.131	.033	3.721616	23.411169
	UN (2,1)	-.446262	.434661	-1.027	.305	-1.298181	.405657
	UN (2,2)	.034518	.053446	.646	.518	.001660	.717792

a. Dependent Variable: smile.

Notice that the variance of beer is not significantly different from zero. We do not care

Variance

Variances of random coefficients inform us on the variability of the effects

- Even when they are not significant, we keep them as random
- When variances are exactly zero (and SPSS gives a general warning), effects should be set only as fixed

Covariance Parameters

Estimates of Covariance Parameters^a

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		1.258809	.125848	10.003	.000	1.034814	1.531292
Intercept + beer [subject = bar]	UN (1, 1)	9.334205	4.379192	2.131	.033	3.721616	23.411169
	UN (2, 1)	-.446262	.434661	-1.027	.305	-1.298181	.405657
	UN (2, 2)	.034518	.053446	.646	.518	.001660	.717792

a. Dependent Variable: smile.

SPSS Output

If everything is fine, we interpret the fixed effects as in any other GLM (regression)

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	5.373312	.850910	11.597	6.315	.000	3.512159	7.234465
beer	.641676	.092399	9.336	6.945	.000	.433796	.849555

a. Dependent Variable: smile.

Intercept: On average, for zero beers we expect 5.37 smiles

B coefficient: On average, for each beer more, on average we expect .641 more smiles

Random effect covariance

We noticed that the covariance parameters are 3

Covariance Parameters

Estimates of Covariance Parameters^a

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		1.258809	.125848	10.003	.000	1.034814	1.531292
Intercept + beer	UN (1,1)	9.334205	4.379192	2.131	.033	3.721616	23.411169
[subject = bar]	UN (2,1)	-.446262	.434661	-1.027	.305	-1.298181	.405657
	UN (2,2)	.034518	.053446	.646	.518	.001660	.717792

a. Dependent Variable: smile.

Random effect covariance

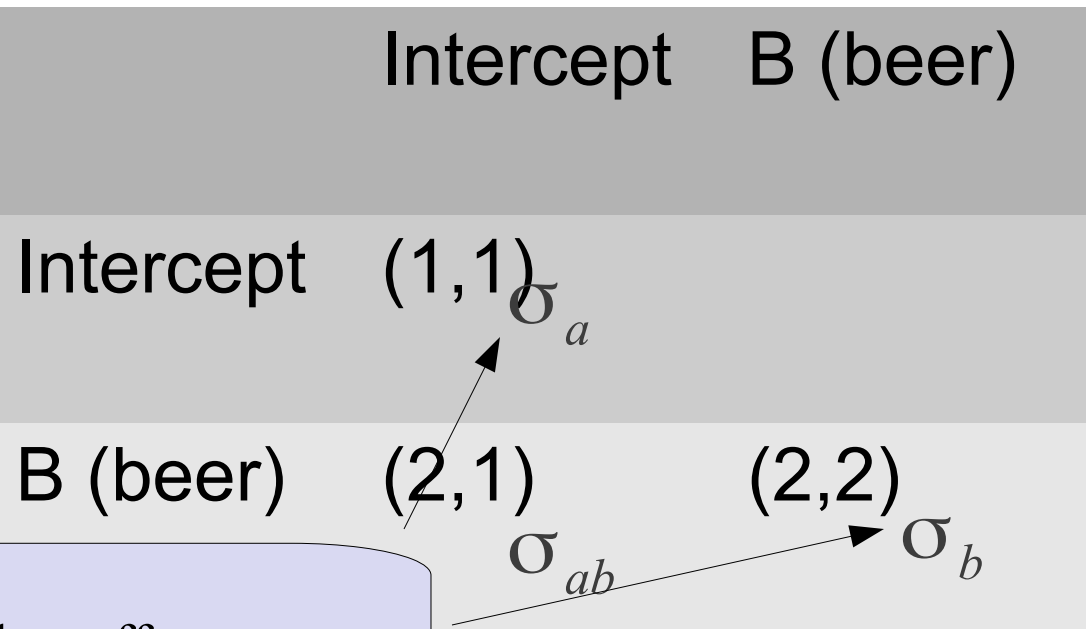
We noticed that the covariance parameters are 3

Covariance Parameters

		Estimates
Parameter		Estimate
Residual		1.258809
Intercept + beer [subject = bar]	UN (1,1)	9.334205
	UN (2,1)	-.446262
	UN (2,2)	.034518

a. Dependent Variable: smile.

Variance of random effects



Random effect covariance

We noticed that the covariance parameters are 3

Covariance Parameters

		Estimates
Parameter		Estimate
Residual		1.258809
Intercept + beer [subject = bar]	UN (1,1)	9.334205
	UN (2,1)	-.446262
	UN (2,2)	.034518

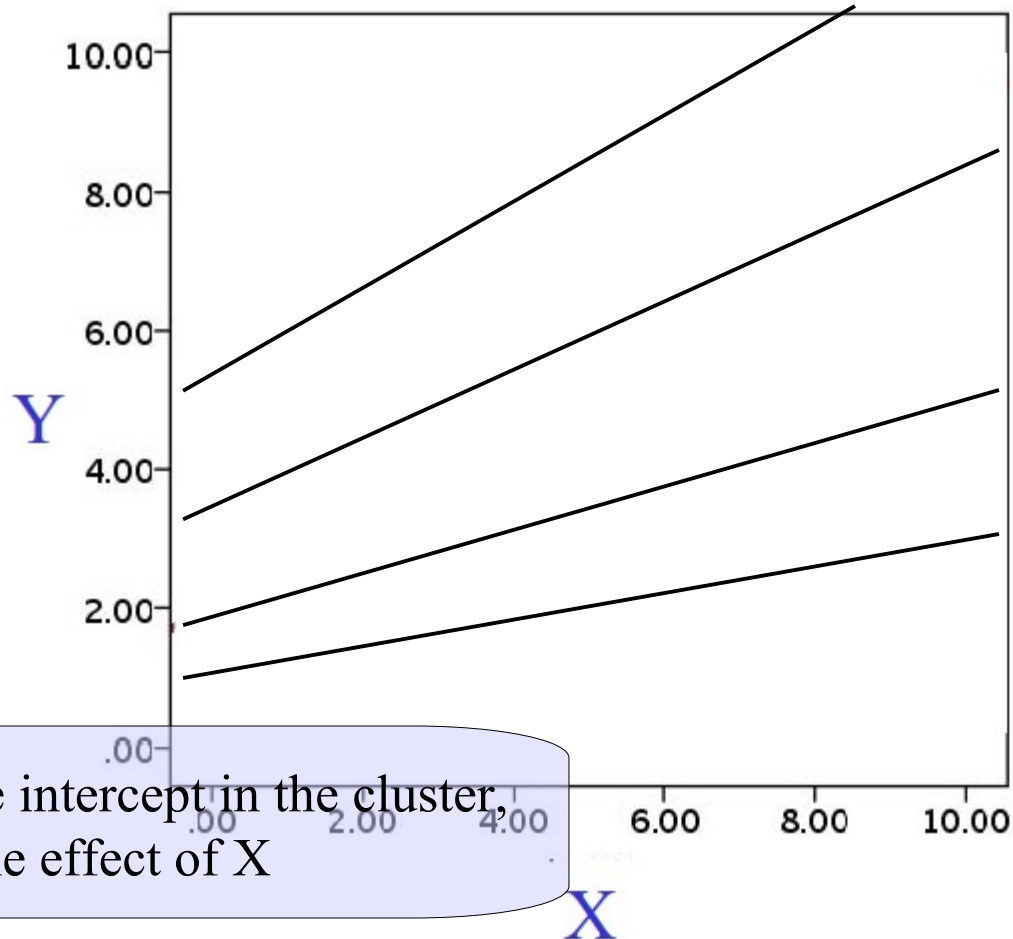
a. Dependent Variable: smile.

	Intercept	B (beer)
Intercept	$(1,1) \sigma_a$	
B (beer)	$(2,1) \sigma_{ab}$	$(2,2) \sigma_b$

Covariance between random coefficients

Covariance

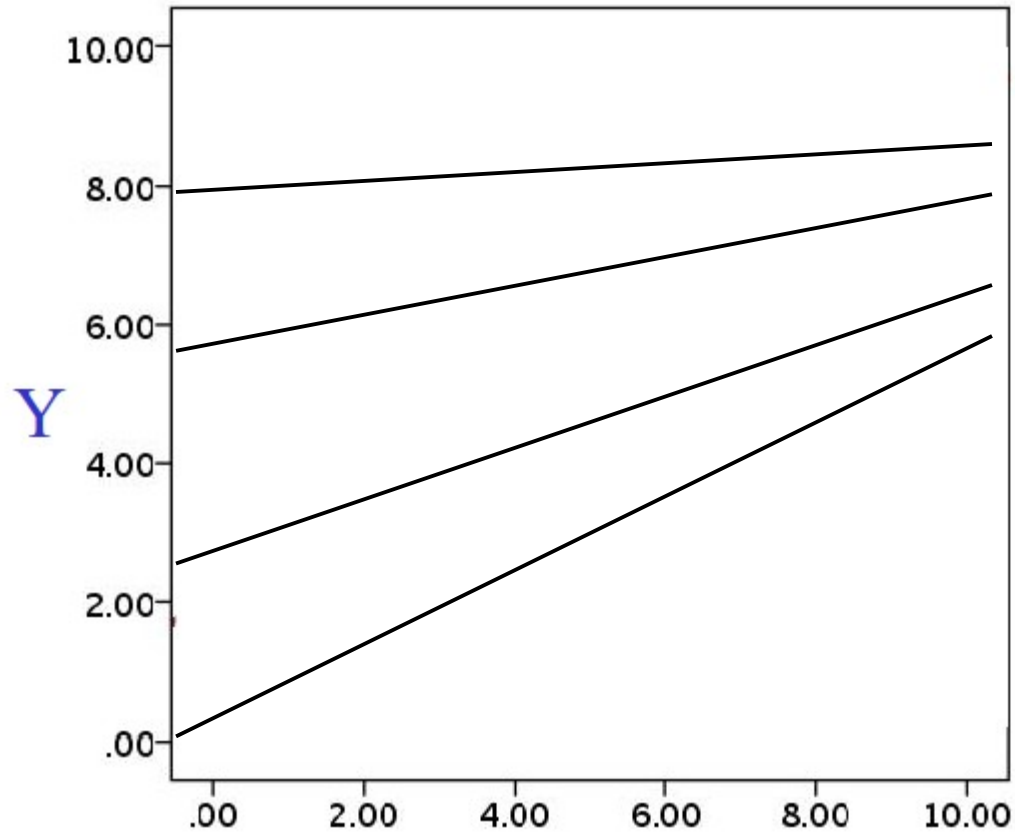
Example of positive covariance between **a** and **b**



The higher the intercept in the cluster, the stronger the effect of X

Covariance

Example of negative covariance between **a** and **b**



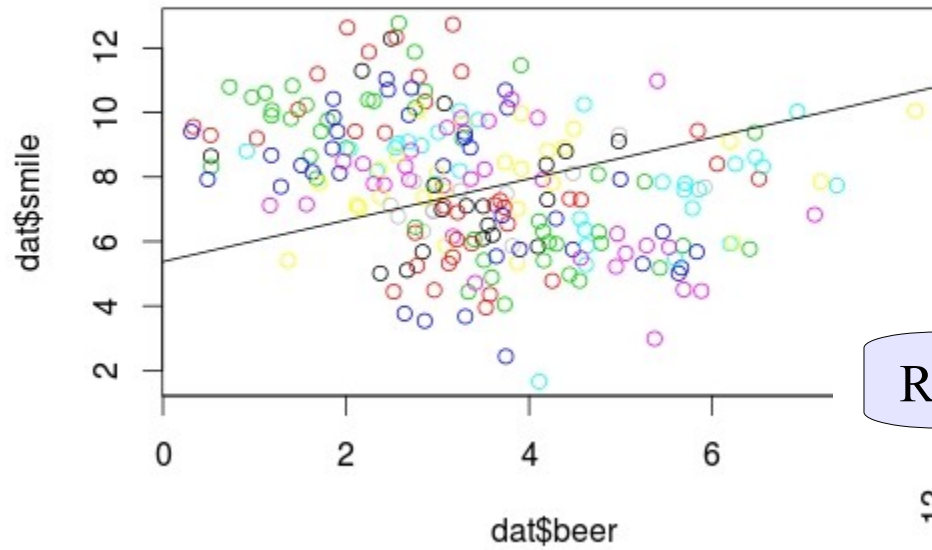
The higher the intercept in the cluster,
the weaker the effect of X

X

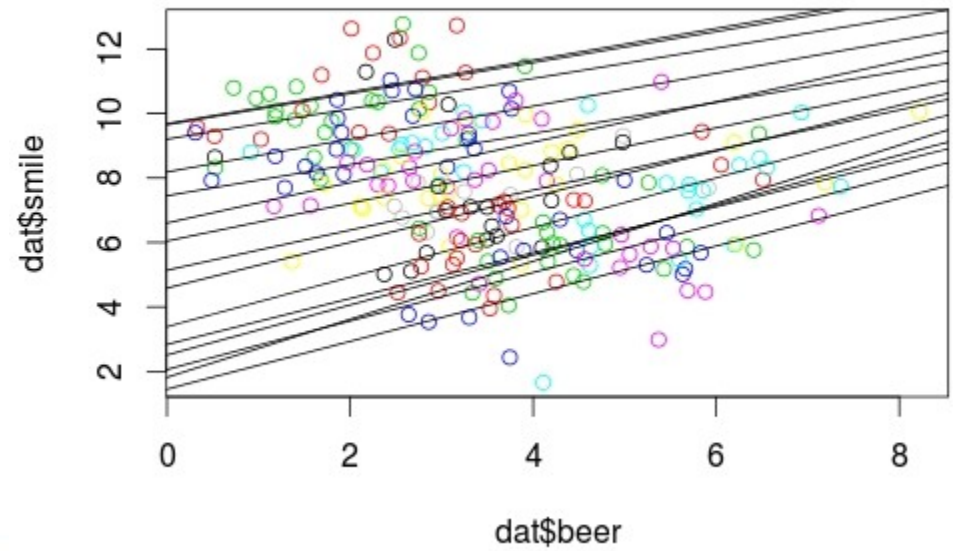
Plots

Plots

Fixed effects



Random effects



Mixed Linear Models

- With the mixed model one can take into the account dependency among measures (within clusters) almost in any situation
- It allows applying the GLM logic to a broader range of designs
- Interactions with any kind of variable
- Efficient handling of missing values
- Continuous repeated measures variables
- Hierarchical organization of the data

Repeated Measures ANOVA as a linear mixed model

Next Meeting

