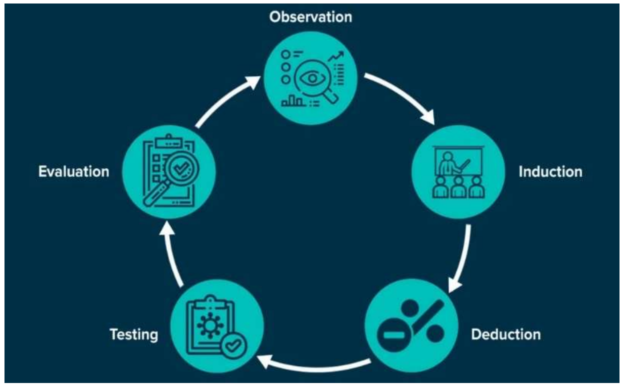
Empirical research in management and economics

Inferential data analysis II

Thorsten Pachur

Technical University of Munich
School of Management
Chair of Behavioral Research Methods



Recap from last week

- What is the difference between probability and non-probability sampling? Give two subtypes for each approach.
- What is the sampling distribution? What is the standard error?
- What is a null hypothesis? What is a p-value?
- What role does the alternative hypothesis play for the assumed effect size?
- What are the key differences between the Fisherian and the Neyman-Pearson schools to hypothesis testing?

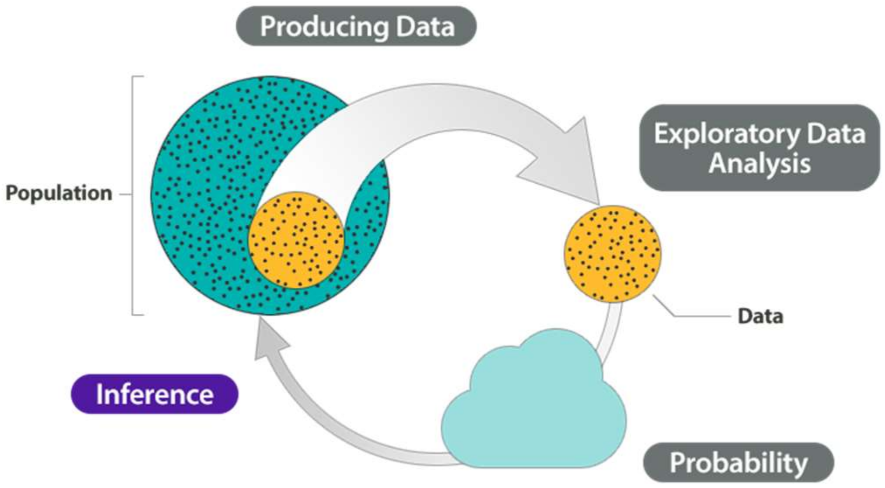
  → H_1 ⇒ Effect size, Power
  "officially" significant p-value
- What is the difference between one-tailed and two-tailed testing?
- What does a confidence interval show?
- What is statistical power? What factors influence statistical power and which is the most relevant factor for a researcher designing a study?

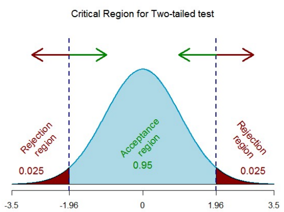
Agenda for the semester

Session	Date	Topic
1	13 October	Introduction
2	20 October	Descriptive data analysis
3	27 October	Hypothesis development and measurement
4	3 November	Inferential data analysis I
5	**10 November**	**Inferential data analysis II**
6	17 November	Simple regression
7	24 November	Multiple regression
8	1 December	Logistic regression
9	8 December	Factor analysis
10	15 December	Cluster analysis
11	12 January	Conjoint analysis
12	19 January	The replication crisis and open science
13	26 January	Summary and questions
	11 February	Exam

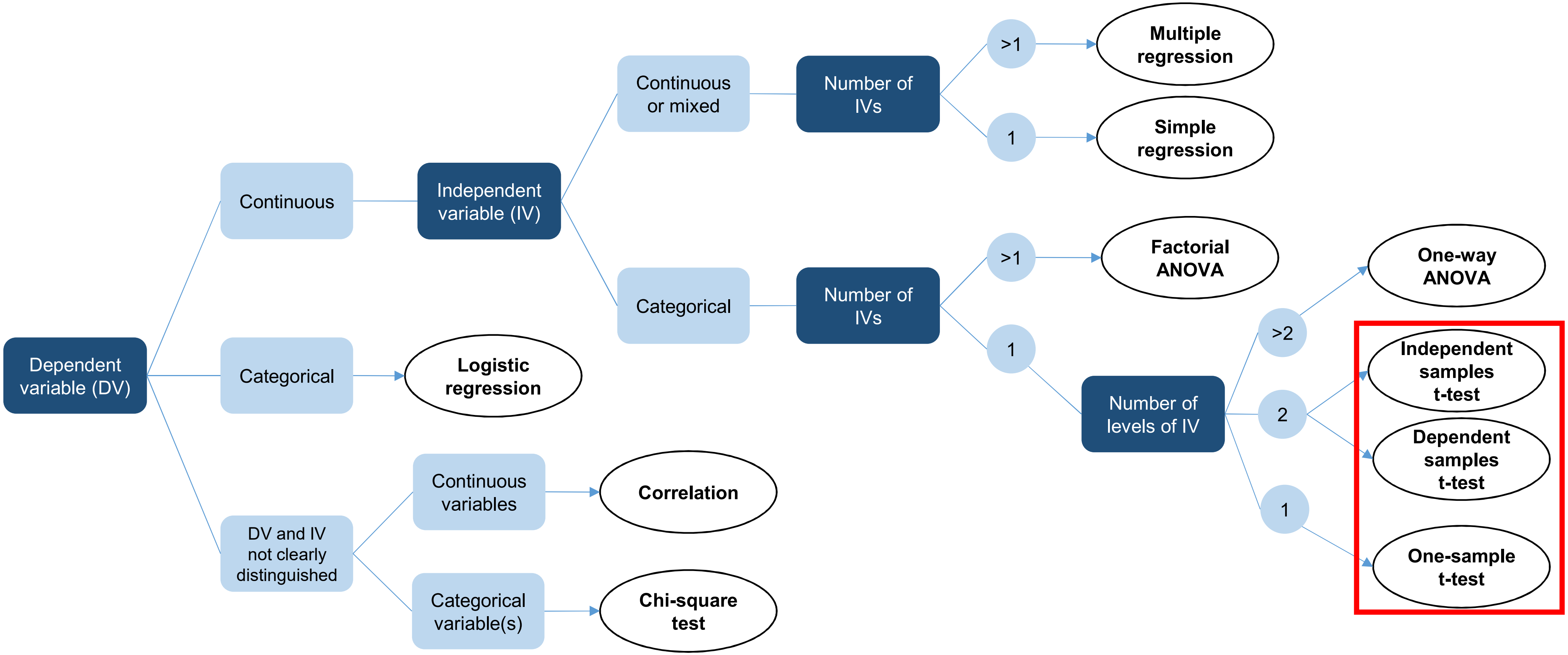
Goals for this week

- You are famililar with common statistical inference tests
  - Comparing two means: *t*-test
  - Comparing more than two means: Analysis of variance (ANOVA)
  - Testing an association between two nominal variables: Chi-square test
- You know how to the test statistic for each test is computed
- You can interpret the results of each test to make an inference about the null hypothesis (*p*-value)
- You know effect size measures for each test



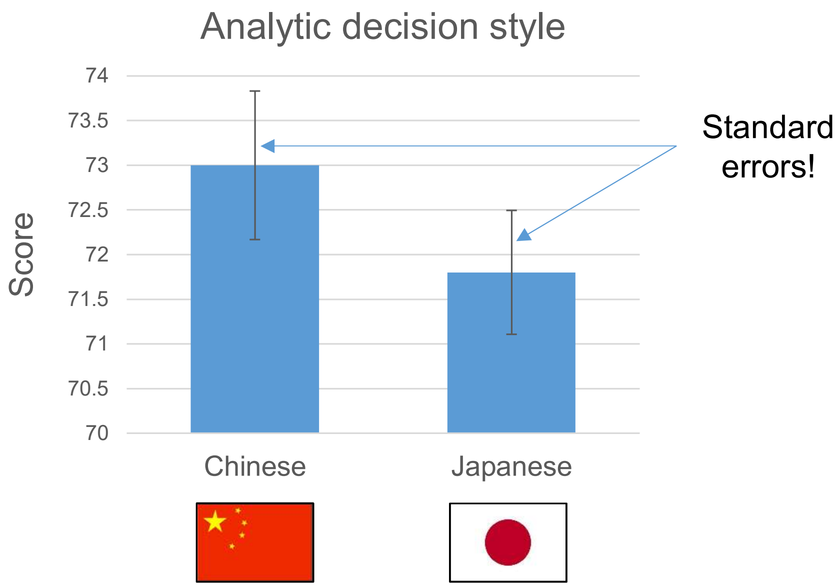


Statistical inference tests


Dependent variable (DV)
Continuous
Independent variable (IV)
Continuous or mixed
Number of IVs
>1
Multiple regression
1
Simple regression
Categorical
Number of IVs
>1
Factorial ANOVA
1
Number of levels of IV
>2
One-way ANOVA
2
Independent samples t-test
Dependent samples t-test
1
One-sample t-test
Categorical
Logistic regression
DV and IV not clearly distinguished
Continuous variables
Correlation
Categorical variable(s)
Chi-square test

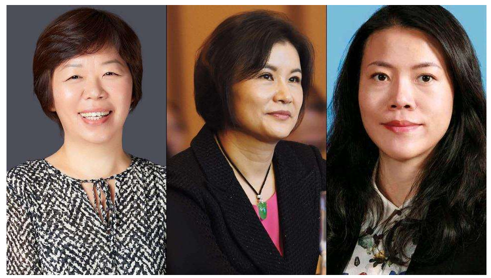
Comparing two means

Do Chinese business leaders have a stronger preference for an **analytic decision style** than Japanese business leaders?



Chinese (n = 88)
M = 73.0, SD = 7.8

Japanese (n = 82)
M = 71.8, SD = 6.3



Methods for comparing means

1 or 2 levels → t-test

1 factor → How many **levels**?

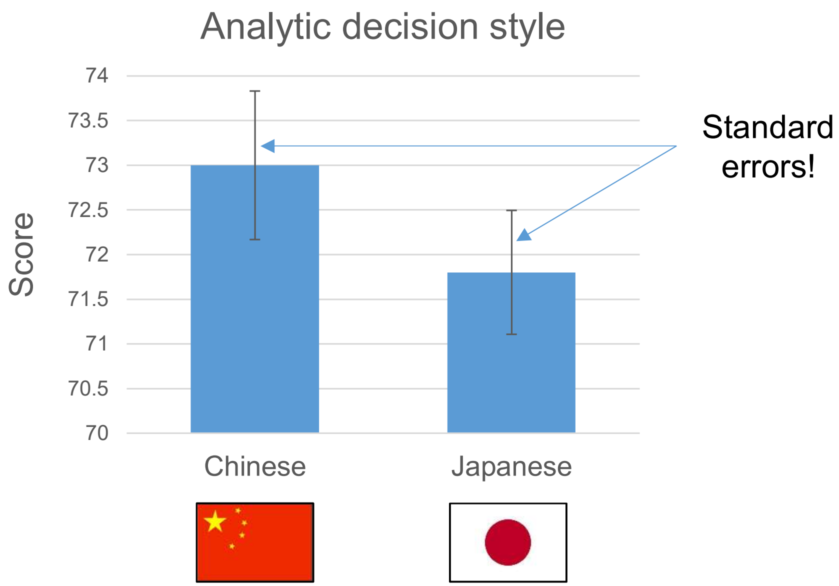
levels = different groups

How many **factors**?

= independent variables

Comparing two means

Do Chinese business leaders have a stronger preference for an **analytic decision style** than Japanese business leaders?



Chinese (n = 88)
M = 73.0, SD = 7.8

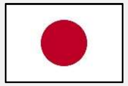
Japanese (n = 82)
M = 71.8, SD = 6.3

Comparing two means

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}_{\bar{X}_1 - \bar{X}_2}}$$

$$\hat{\sigma}_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\hat{\sigma}^2_{pooled}}{n_1} + \frac{\hat{\sigma}^2_{pooled}}{n_2}}$$

$$\hat{\sigma}^2_{pooled} = \frac{\hat{\sigma}^2_1 \cdot (n_1 - 1) + \hat{\sigma}^2_2 \cdot (n_2 - 1)}{(n_1 - 1) + (n_2 - 1)}$$

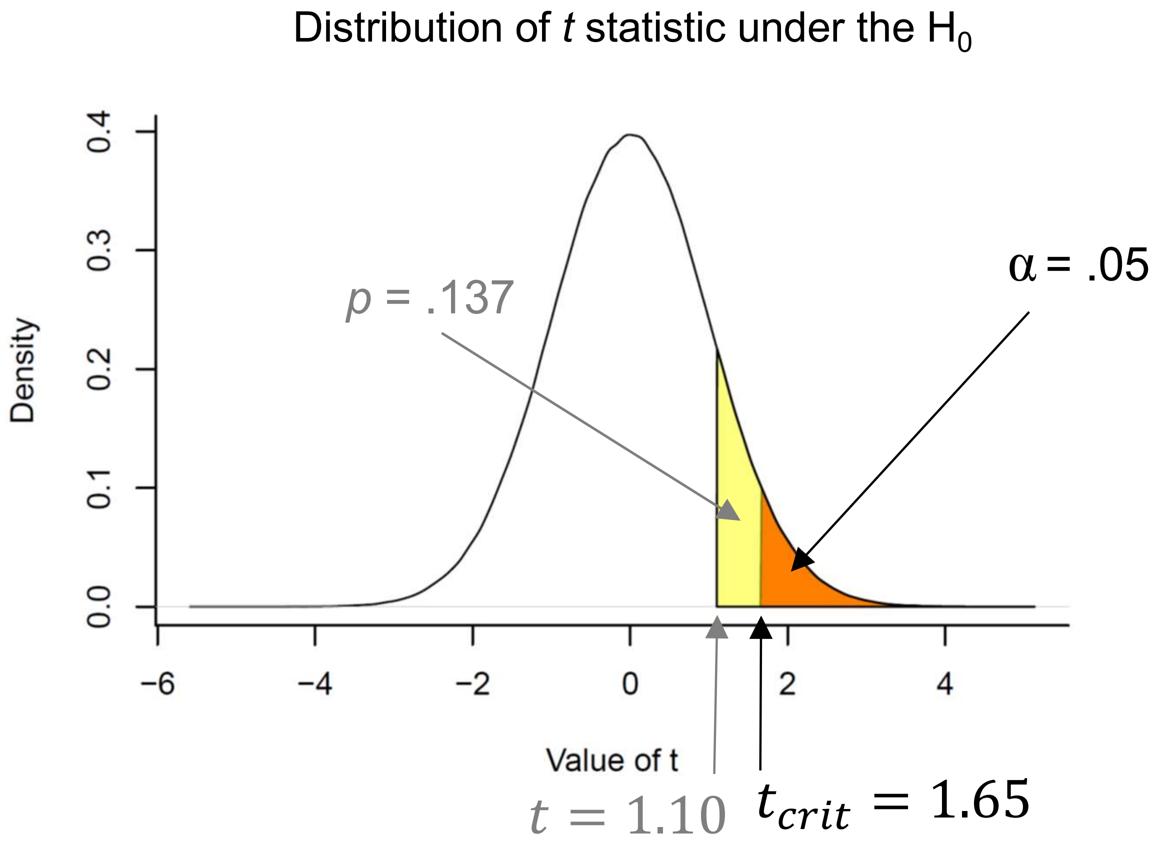
$$t = \frac{73.0 - 71.8}{1.09} = 1.10$$

$$\hat{\sigma}_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{50.64}{88} + \frac{50.64}{82}} = 1.09$$

$$\hat{\sigma}^2_{pooled} = \frac{7.8^2 \cdot (88 - 1) + 6.3^2 \cdot (82 - 1)}{(88 - 1) + (82 - 1)} = 50.64$$

Comparing two means

Degrees of freedom (df)
t(df = n_1+n_2-2
= 88+82-2
= 168)

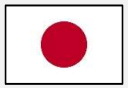


Effect size for t-test

Cohen‘s *d* (Cohen, 1988)

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\hat{\sigma}^2_{pooled}}}$$

→ How large is the difference between the group means in terms of the pooled standard deviation?

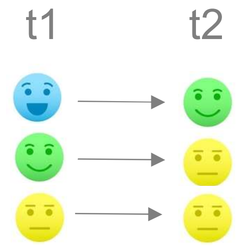
 

$$d = \frac{73.0 - 71.8}{\sqrt{50.64}} = 0.17$$

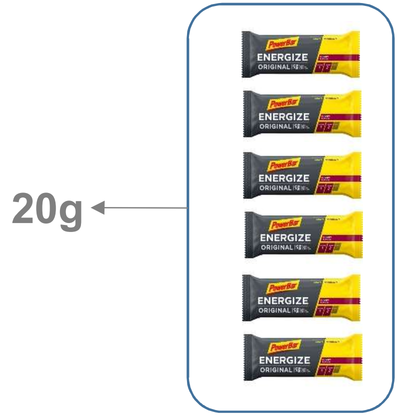
Value of *d*	Interpretation
.2	Small effect
.5	Medium effect
.8	Large effect

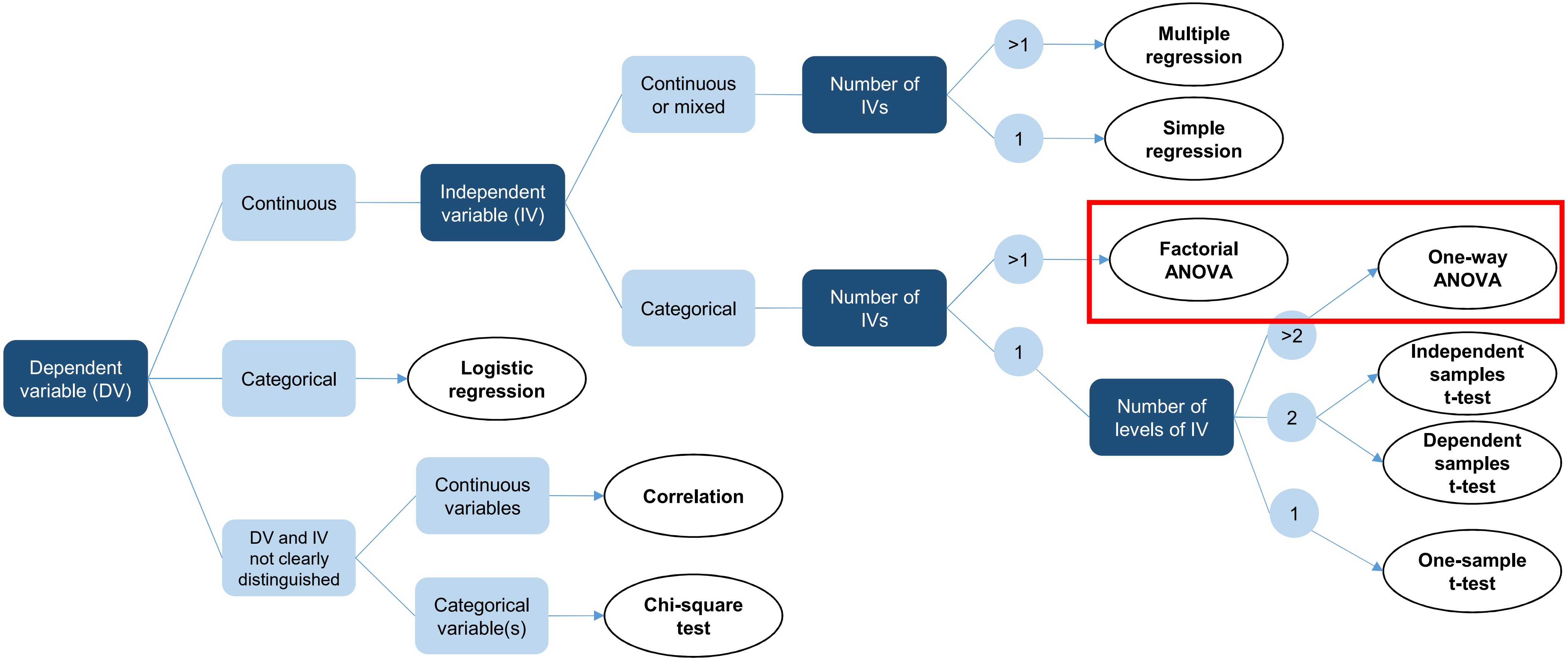
Different types of t-tests

- Independent samples t-test
  → Comparing the means of two independent groups
- Dependent samples t-test
  → Comparing the means of two related sets of observations
  Example: Measuring happiness of a group of people before and after an event (i.e., each person is measured twice)



- One-sample t-test
  → Comparing the mean of one group against a single value
  Example: Comparing the protein content of a sample of packages of an energy bar to a reference value of 20g (the amount indicated on the package)



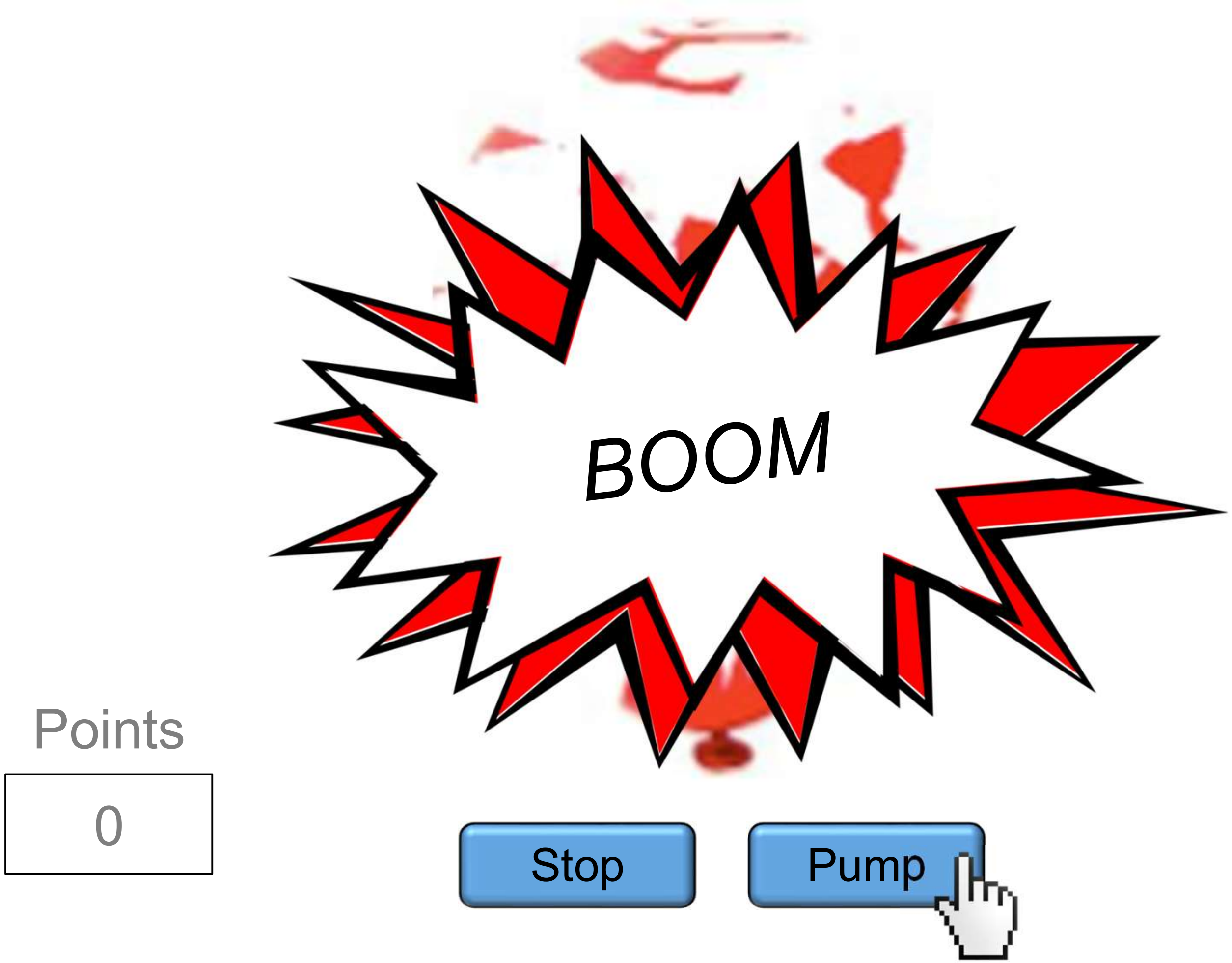

Dependent variable (DV)
Continuous
Categorical
DV and IV not clearly distinguished
Independent variable (IV)
Logistic regression
Continuous variables
Categorical variable(s)
Correlation
Chi-square test
Continuous or mixed
Categorical
Number of IVs
Number of IVs
>1
1
>1
1
Multiple regression
Simple regression
Factorial ANOVA
One-way ANOVA
Number of levels of IV
>2
2
1
Independent samples t-test
Dependent samples t-test
One-sample t-test

Methods for comparing means

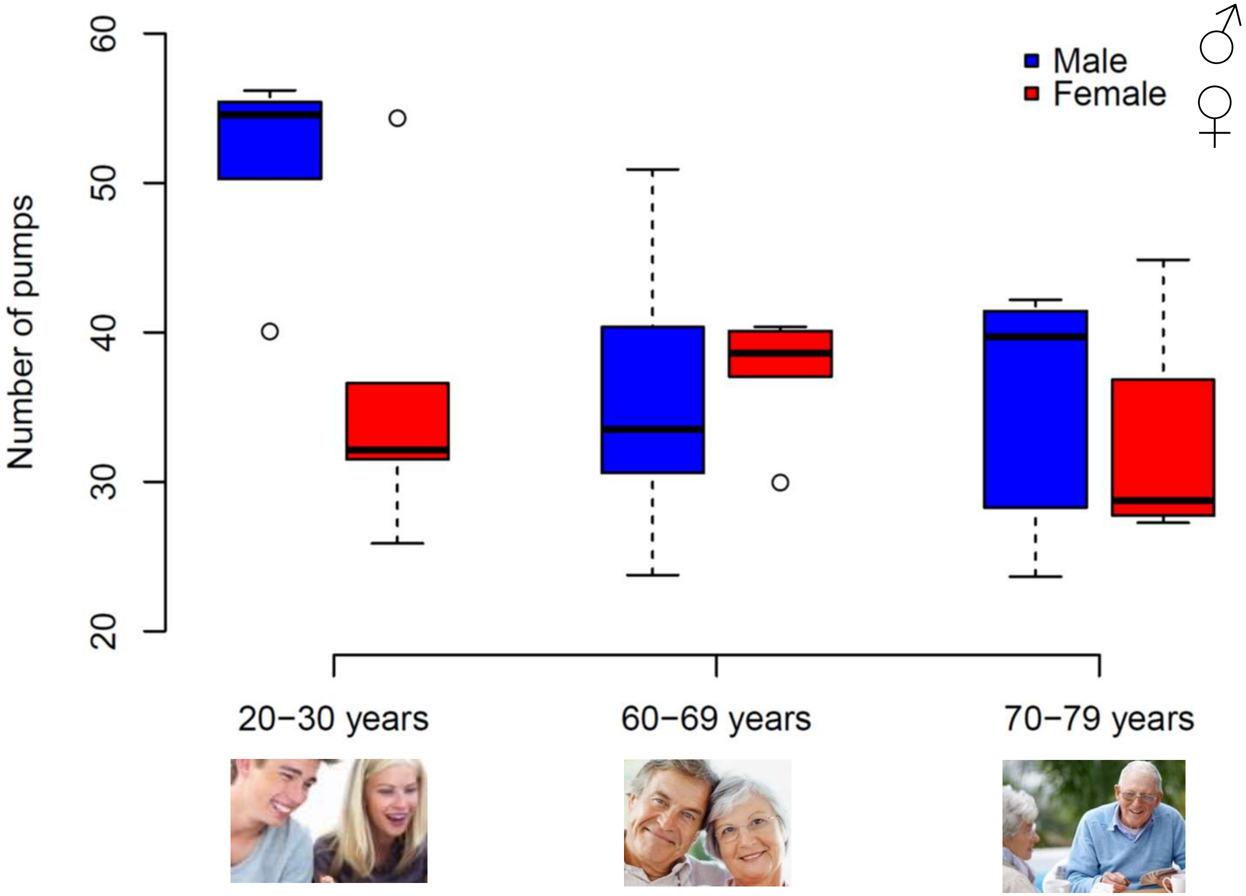
How many **factors**?

- 1 factor → How many **levels**?
  - 1 or 2 levels → *t*-test
  - > 2 levels → One-way analysis of variance (ANOVA)
- > 1 factors → Factorial analysis of variance (ANOVA)

Measurement of individual risk propensity: Balloon Analogue Risk Task (BART)



Lejuez et al. (2002)



2 factors

- Age group (3 levels)
- Sex (2 levels)

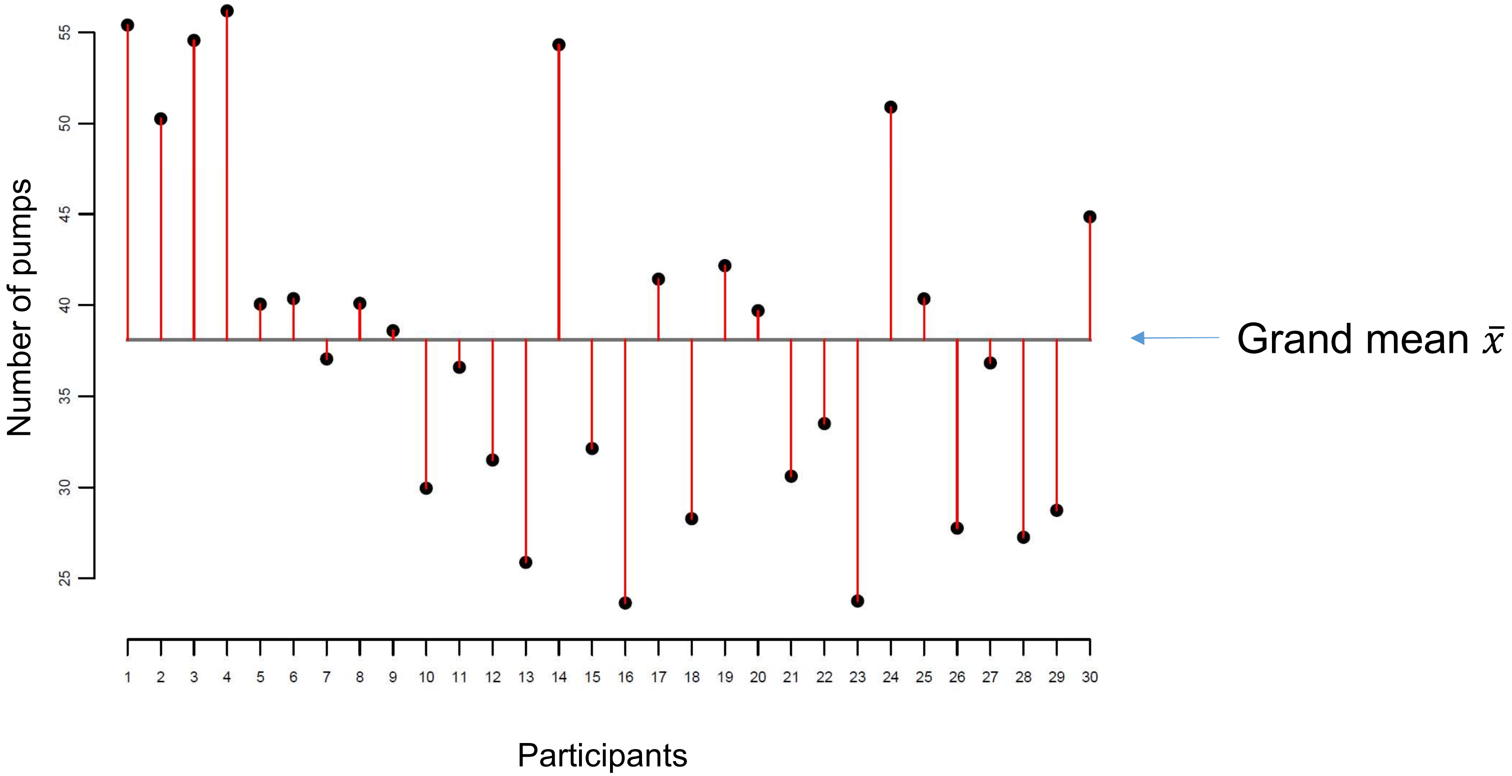
→ Factorial ANOVA

Lighthall, Mather, & Gorlick (2009)

Total variability

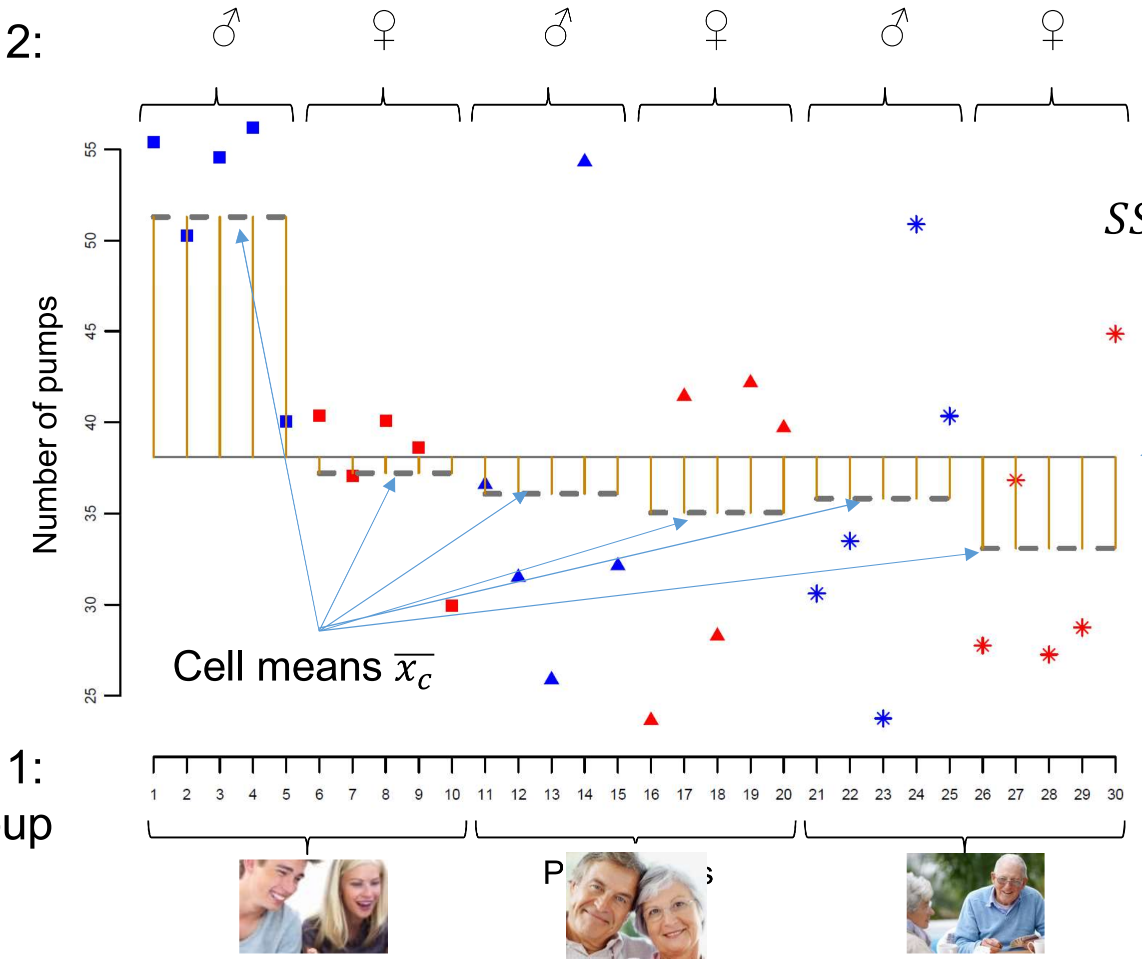
$$SS_{Total} = \sum_{i=1}^{N} (x_i - \bar{x})^2$$

"Sum of squares"



Model variability

Factor 2:
Sex



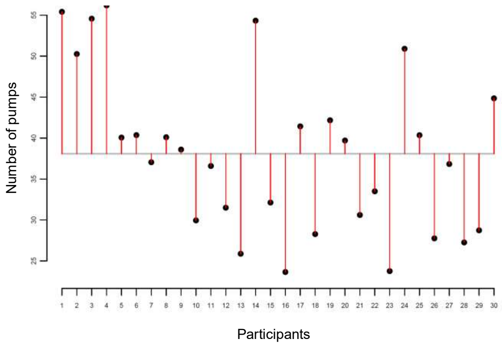
$$SS_{Cells} = \sum_{c=1}^{C} n_c(\overline{x_c} - \bar{x})^2$$

Grand mean $\bar{x}$

Factor 1:
Age group

Sum of squares (SS)

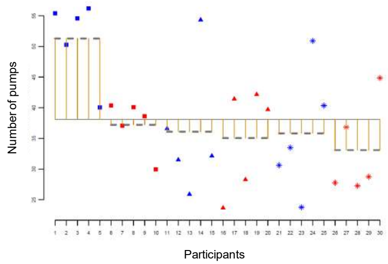
Total variability



SS_{Total}

$= 2767.4$

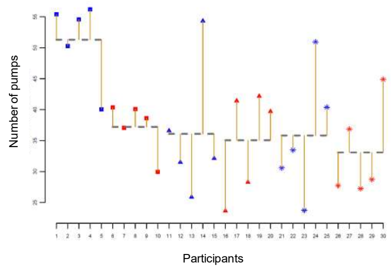
Model variability



SS_{Cells}

$= 1093.3$

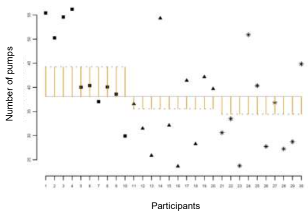
Residual variability



$SS_{Residual}$

$= 1674.1$

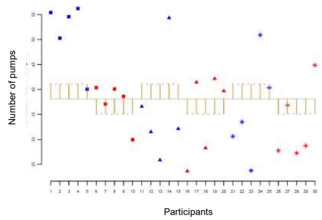
Variability Factor 1 (Age group)



$SS_{Factor\ 1}$

$= 575.8$

Variability Factor 2 (Sex)



$SS_{Factor\ 2}$

$= 265.9$

Variability Factor 1 × Factor 2

= Model variability - Variability Factor 1 – Variability Factor 2

$SS_{Factor\ 1\ \times\ Factor\ 2}$

$= 251.6$

Analysis of variance (ANOVA)

Mean squares (MS)

$$MS_{Factor} = \frac{SS_{Factor}}{df_{Factor}}$$

df_{Factor}: Number of $levels_{Factor}$ – 1

$$MS_{Residual} = \frac{SS_{Residual}}{df_{Residual}}$$

$$F = \frac{MS_{Factor}}{MS_{Residual}}$$

$df_{Residual}$: Total number of observations – total number of levels

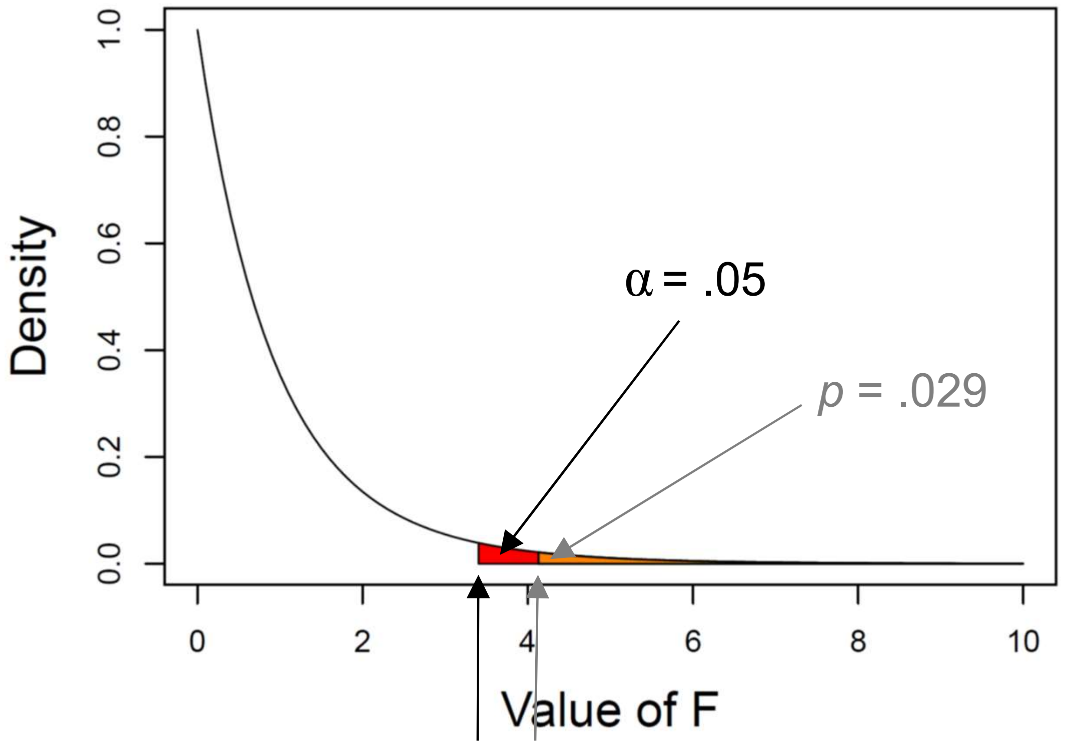
$$MS_{Age} = \frac{575.8}{3-1}$$

$$MS_{Residual} = \frac{1674.1}{30-6}$$

$$F = \frac{287.9}{69.8} = 4.13$$

F distribution

df_{Age} = 2, $df_{Residual}$ = 24



$F_{crit} = 3.40$ $F = 4.13$

Total mean squares

$$MS_{Total} = \frac{SS_{Total}}{df_{Total}}$$

df_{Total}: Number of observations – 1

Model mean squares

Mean squares Factor 1 (Age)

$$MS_{Factor\ 1} = \frac{SS_{Factor\ 1}}{\underbrace{df_{Factor\ 1}}_{\text{Number of levels}_{\text{Factor 1}} - 1}}$$

Mean squares Factor 2 (Sex)

$$MS_{Factor\ 2} = \frac{SS_{Factor\ 2}}{\underbrace{df_{Factor\ 2}}_{\text{Number of levels}_{\text{Factor 2}} - 1}}$$

Mean squares Factor 1 × Factor 2

$$MS_{Interaction} = \frac{SS_{Interaction}}{df_{Factor\ 1} \times df_{Factor\ 2}}$$

Residual mean squares

$$MS_{Residual} = \frac{SS_{Residual}}{df_{Total} - df_{Factor\ 1} - df_{Factor\ 2} - df_{Factor\ 1} \times df_{Factor\ 2}}$$

Total mean squares

$$MS_{Total} = \frac{2767.4}{29} = 95.4$$

Model mean squares

Mean squares Factor 1 (Age)

$$MS_{Facto\quad 1} = \frac{575.8}{2} = 287.9$$

Mean squares Factor 2 (Sex)

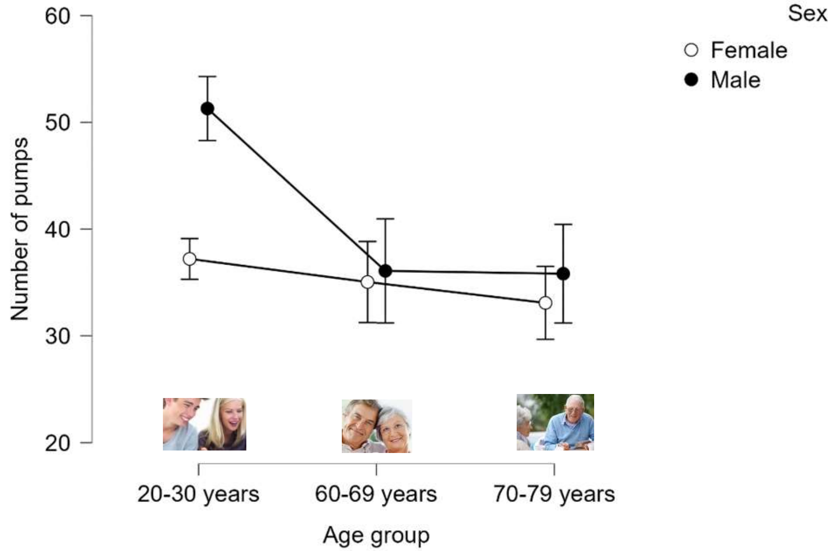
$$MS_{Facto\quad 2} = \frac{265.9}{1} = 265.9$$

Mean squares Factor 1 × Factor 2

$$MS_{Interactio} = \frac{251.6}{2 \times 1} = 125.8$$

Residual mean squares

$$MS_{Residual} = \frac{1674.1}{29 - 2 - 1 - 2 \times 1} = 69.8$$



	df	SS	MS	*F*	*p*
Factor 1 (Age)	2	575.8	287.9	4.13	0.029
Factor 2 (Sex)	1	265.9	265.9	3.81	0.063
Interaction (Age × Sex)	2	251.6	125.8	1.80	0.186
Residual	24	1674.1	69.8		
Total	29	2767.4			

Effect size(s) for ANOVA

$$\eta^2 = \frac{SS_{Factor}}{SS_{Total}}$$

→ Overestimates population value, especially for small *N*

$$\omega^2 = \frac{SS_{Factor} - (df_{Factor})MS_{Residual}}{SS_{Total} + MS_{Residual}}$$

→ Less biased, preferred

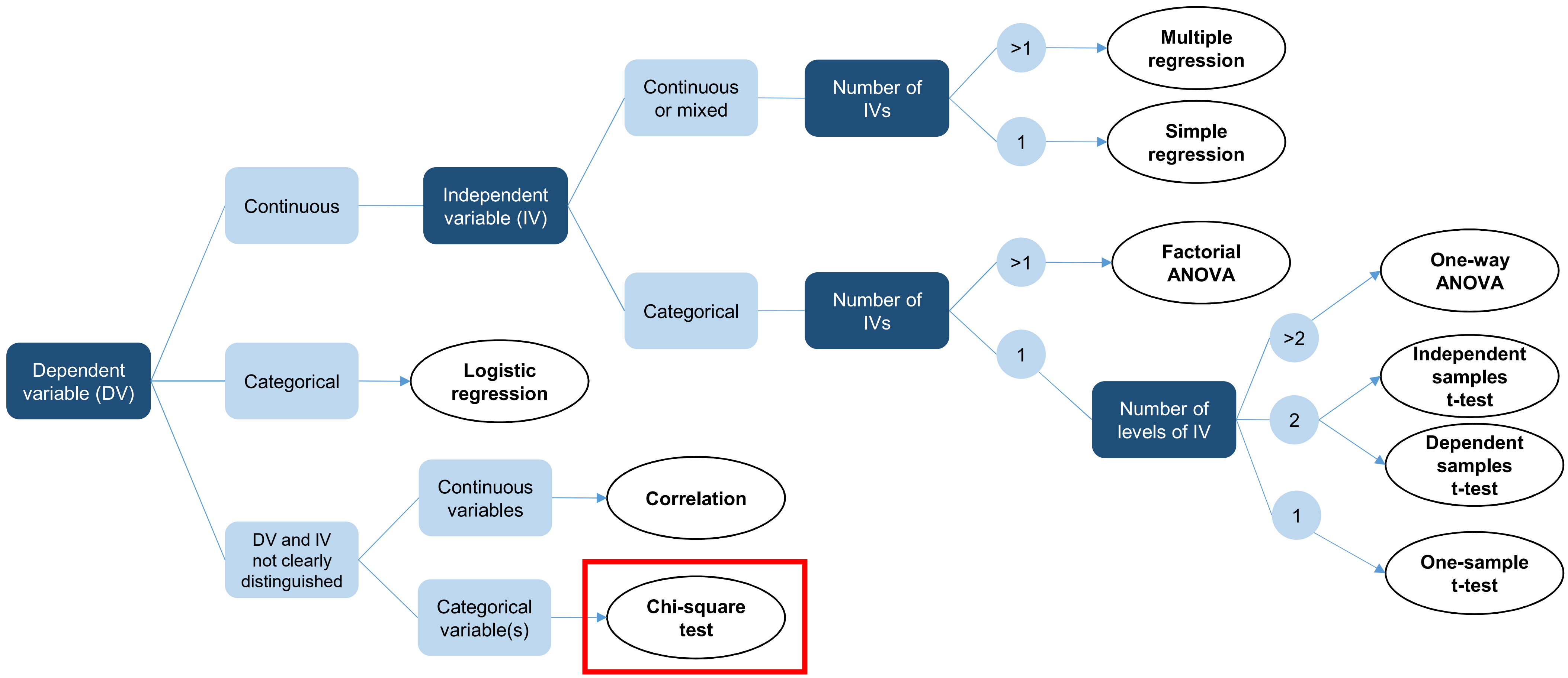


$$\eta^2_{Age} = \frac{575.8}{2767.4} = 0.208$$

$$\omega^2_{Age} = \frac{575.8 - 2 \times 69.8}{2767.4 + 69.8} = 0.154$$

Value of η^2 (ω^2)	Interpretation
.01	Small effect
.06	Medium effect
.14	Large effect

Kirk (1996)


Dependent variable (DV)
Continuous
Independent variable (IV)
Continuous or mixed
Number of IVs
>1
Multiple regression
1
Simple regression
Categorical
Number of IVs
>1
Factorial ANOVA
1
Number of levels of IV
>2
One-way ANOVA
2
Independent samples t-test
Dependent samples t-test
1
One-sample t-test
Categorical
Logistic regression
DV and IV not clearly distinguished
Continuous variables
Correlation
Categorical variable(s)
Chi-square test

Analyzing the association of two nominal variables

Observed frequencies

Previous experience with statistical software

Program	I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP	I have previously used other statistical software (e.g., SPSS, STATA).	**Total**	**Prop.**
Consumer Science					**5**	*0.03*
Management	74	26	2	11	**113**	*0.75*
Management & Technology	2	13	4	4	**23**	*0.15*
Other	7	3	0	1	**11**	*0.07*
Total	**86**	**43**	**6**	**17**	**Σ=152**	
Prop.	*0.57*	*0.28*	*0.04*	*0.11*		

Is there an association between the Master‘s program and previous experience with statistical software?

→ Comparison of the observed frequencies with the **expected frequencies under the H_0** (i.e., that there is **no** association)

Analyzing the association of two nominal variables

Expected frequencies under the H_0 (i.e., there is no association)

Previous experience with statistical software

Program		I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).	Total	Prop.
	Consumer Science	=.57*.03*152				**5**	*0.03*
	Management					**113**	*0.75*
	Management & Technology					**23**	*0.15*
	Other					**11**	*0.07*
	Total	**86**	**43**	**6**	**17**	**Σ=152**	
	Prop.	*0.57*	*0.28*	*0.04*	*0.11*		

Analyzing the association of two nominal variables

Expected frequencies under the H_0
(i.e., there is no association)

Previous experience with statistical software

		I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).	**Total**	**Prop.**
Program	Consumer Science	2.83				**5**	*0.03*
	Management	=.57*.75*152				**113**	*0.75*
	Management & Technology					**23**	*0.15*
	Other					**11**	*0.07*
	Total	**86**	**43**	**6**	**17**	**Σ=152**	
	Prop.	*0.57*	*0.28*	*0.04*	*0.11*		

Analyzing the association of two nominal variables

Expected frequencies under the H_0
(i.e., there is no association)

Previous experience with statistical software

Program		I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).	Total	Prop.
	Consumer Science	2.83	=.28*.03*152			**5**	*0.03*
	Management	63.93				**113**	*0.75*
	Management & Technology					**23**	*0.15*
	Other					**11**	*0.07*
	Total	**86**	**43**	**6**	**17**	**Σ=152**	
	Prop.	*0.57*	*0.28*	*0.04*	*0.11*		

Analyzing the association of two nominal variables

Expected frequencies under the H_0 (i.e., there is no association)

Previous experience with statistical software

Program		I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).	**Total**	**Prop.**
	Consumer Science	2.83	1.41	0.2	0.56	**5**	*0.03*
	Management	63.93	31.97	4.46	12.64	**113**	*0.75*
	Management & Technology	13.01	6.51	0.91	2.57	**23**	*0.15*
	Other	6.22	3.11	0.43	1.23	**11**	*0.07*
	Total	**86**	**43**	**6**	**17**	**Σ=152**	
	Prop.	*0.57*	*0.28*	*0.04*	*0.11*		

Chi-square test

Observed frequencies (*n*)

		Previous experience with statistical software			
		I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).
Program	Consumer Science	3	1	0	1
	Management	74	26	2	11
	Management & Technology	2	13	4	4
	Other	7	3	0	1

$$\chi^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(n_{i,j} - e_{i,j})^2}{e_{i,j}}$$

Expected frequencies under the H_0 (*e*)

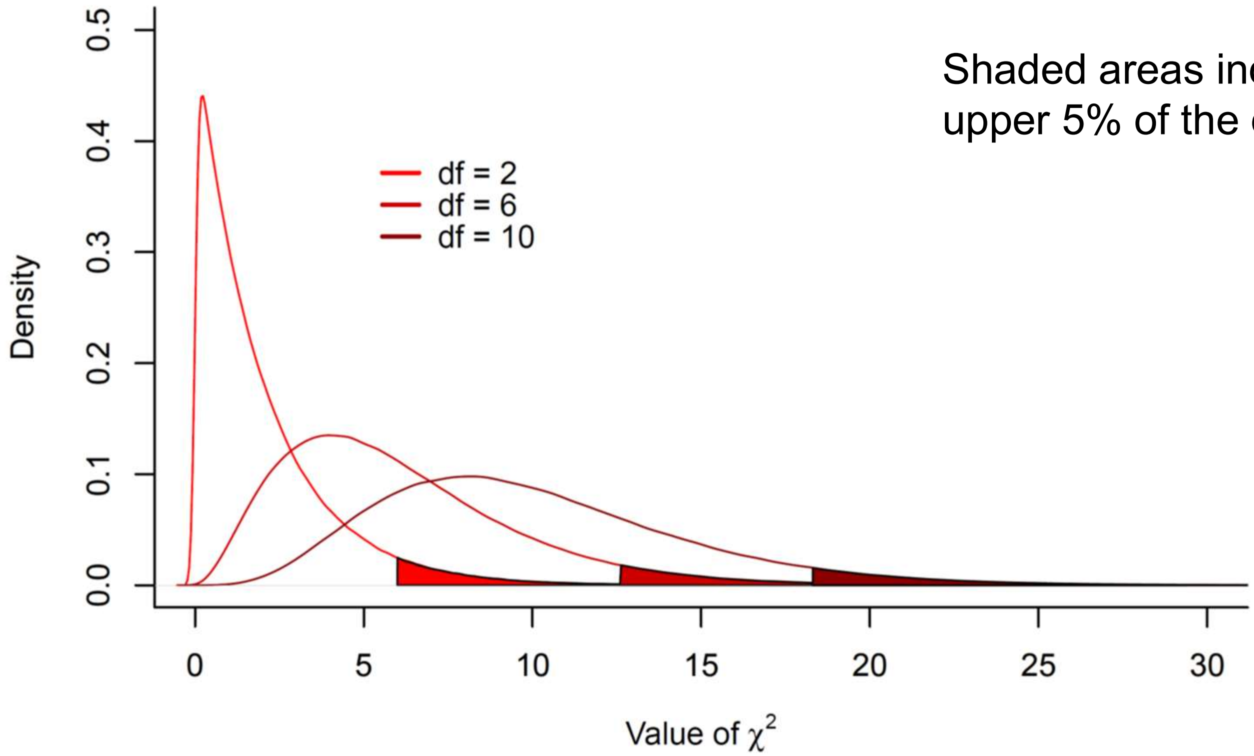
		Previous experience with statistical software			
		I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).
Program	Consumer Science	2.83	1.41	0.2	0.56
	Management	63.93	31.97	4.46	12.64
	Management & Technology	13.01	6.51	0.91	2.57
	Other	6.22	3.11	0.43	1.23

		Previous experience with statistical software			
		I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).
Program	Consumer Science	0.01	0.12	0.2	0.35
	Management	1.58	1.11	1.36	0.21
	Management & Technology	9.32	6.48	10.53	0.79
	Other	0.1	0	0.43	0.04

Difference between observed and expected frequencies

$$\chi^2 = 32.64$$

Chi-square distribution and degrees of freedom



Shaded areas indicate the upper 5% of the distribution

Chi-square test

Degrees of freedom (df)

χ^2(df = (r-1)×(c-1)

= 3 × 3

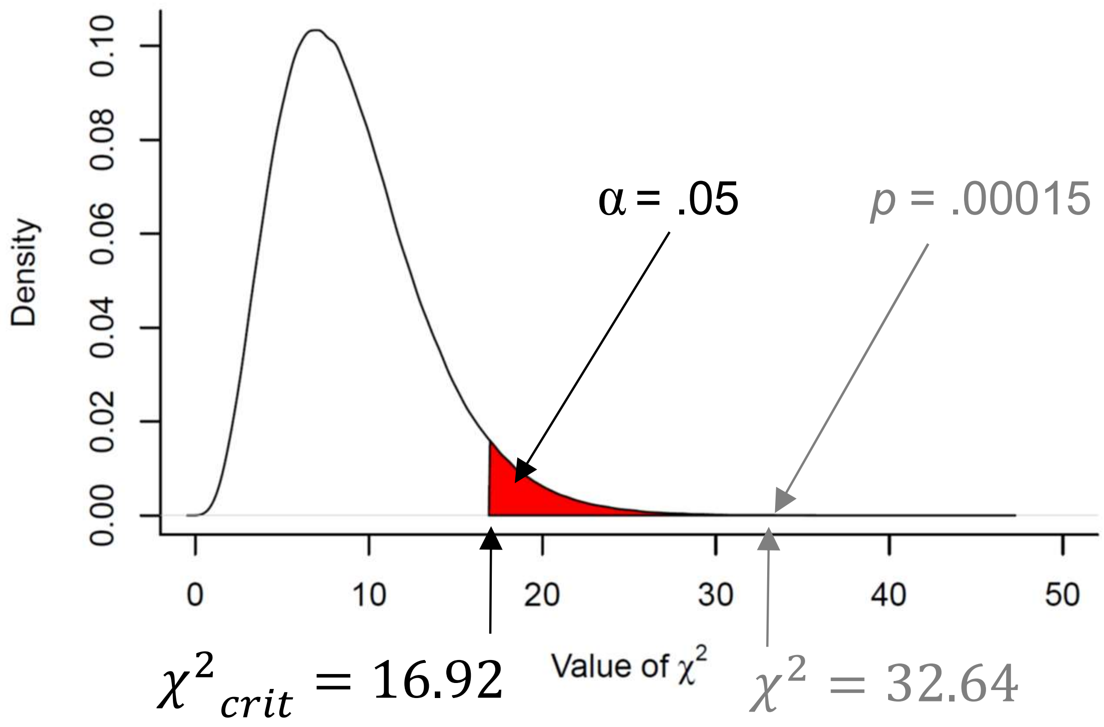
= 9)

c = 4

r = 4

Previous experience with statistical software

Program	I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).
Consumer Science	3	1	0	1
Management	74	26	2	11
Management & Technology	2	13	4	4
Other	7	3	0	1



Effect size for chi-square test

$$\text{Cramer's V} = \sqrt{\frac{\chi^2}{n \times \min(r-1, c-1)}}$$

r: number of rows
c: number of columns
n: total number of observations

Alternative effect size measure

$$\omega = V\sqrt{\min(r-1, c-1)}$$

$$\text{Cramer's V} = \sqrt{\frac{32.64}{152 \times 3}} = .27$$

Previous experience with statistical software

Program	I have no previous experience with statistical software.	I have previously used R.	I have previously used both R and JASP.	I have previously used other statistical software (e.g., SPSS, STATA).	**Total**	**Prop.**
Consumer Science	3	1	0	1	**5**	*0.03*
Management	74	26	2	11	**113**	*0.75*
Management & Technology	2	13	4	4	**23**	*0.15*
Other	7	3	0	1	**11**	*0.07*
Total	**86**	**43**	**6**	**17**	**Σ=152**	
Prop.	*0.57*	*0.28*	*0.04*	*0.11*		

Effect size for chi-square test

$$\text{Cramer's V} = \sqrt{\frac{\chi^2}{n \times \min(r-1, c-1)}}$$

r: number of rows
c: number of columns
n: total number of observations

$\min(r-1, c-1)$	Small effect	Medium effect	Large effect
1	0.10	0.30	0.50
2	0.07	0.21	0.35
3	0.06	0.17	0.29
4	0.05	0.15	0.25
5	0.04	0.13	0.22

Interpretation

Value of *V*

Self-quiz questions

- Which statistical test is indicated for each of the following situations? For each test also give the test statistic that is used to compute a p-value
  - Comparing the means of three or more groups across one or several factors
  - Association of two nominal-level variables
  - Comparing the means of two independent groups
- Give effect size measures for each of the tests
- Imagine that in a factorial ANOVA, you obtain a p-value of .02 for the interaction between two factors. How do you interpret this result?

Background reading for next week

Howell, D. C. (2017). Regression. In: D. C. Howell, *Fundamental statistics for the behavioral sciences* (9th ed.) (p. 226–264). Wadsworth Cengage Learning, Belmont.