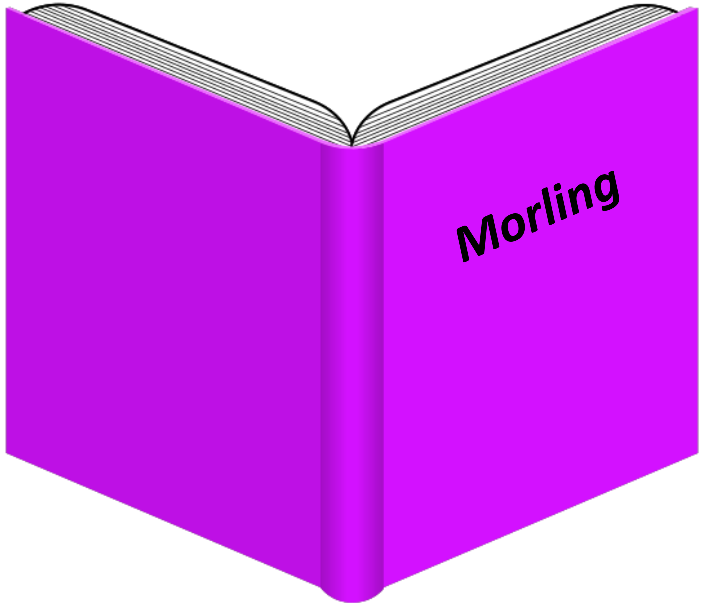


Research Methods and Statistics

Lecture 19: Analysis of Variance

Riet van Bork

Pictures source: pixabay.org

PROTECT OUR INTERNATIONAL EDUCATION

SAVE THE DATE

Strike & Protest at Dam Square, Dec 9th

PSYCHinAction:

AGAINST CUTS AND LANGUAGE BARRIERS

A Call To Protect International Education in Dutch Academia!

How to get involved?

- **Read, sign & share our open letter**
  - The UNL proposed to prohibit a number of programs from being taught in English, **including psychology!**

    Read here why we oppose this →
- **Join the Whatsapp community**

  & participate in workgroups
  - Share your ideas and work together with staff & students from all universities!
- **Make your voice heard on Dec 9th**

  at the strike & protest at Dam Square
  - It's time to stand together & **let's protect international education** in the Netherlands!

x 285,7 mm

PSYCHinAction:

AGAINST CUTS AND LANGUAGE BARRIERS

A Call To Protect International Education in Dutch Academia!

Strike and Protest? What is the difference?

- **Strike:**
  - **Stopping work** temporarily in response to workplace issues
  - The right to strike is **legally protected** in the Netherlands
  - There is no obligation to make up for missed work
  - Action cannot be taken against workers for striking
  - At university: ***staff members*** can strike, ***students*** cannot
- **Protest/Demonstration:**
  - Some kind of **physical demonstration** for/against a cause
  - E.g. going to Dam Square, marching, protest signs etc.
  - The right to legal protest is **also protected** in the Netherlands
  - At university: everyone can protest, both ***staff and students***

December 9th is both! **A staff strike was called by the unions**
And we will protest together on Dam Square at 12:00

PSYCHinAction:

AGAINST CUTS AND LANGUAGE BARRIERS

A Call To Protect International Education in Dutch Academia!

For more info check out →

- **PSYCHinAction:**

  https://linktr.ee/psychinaction_movement

- **WOinActie:**

  https://woinactie.blogspot.com

- **FNV:**

  https://www.fnv.nl/cao-sector/overheid/onderwijsonderzoek/kabinet-sloopt-hoger-onderwijs

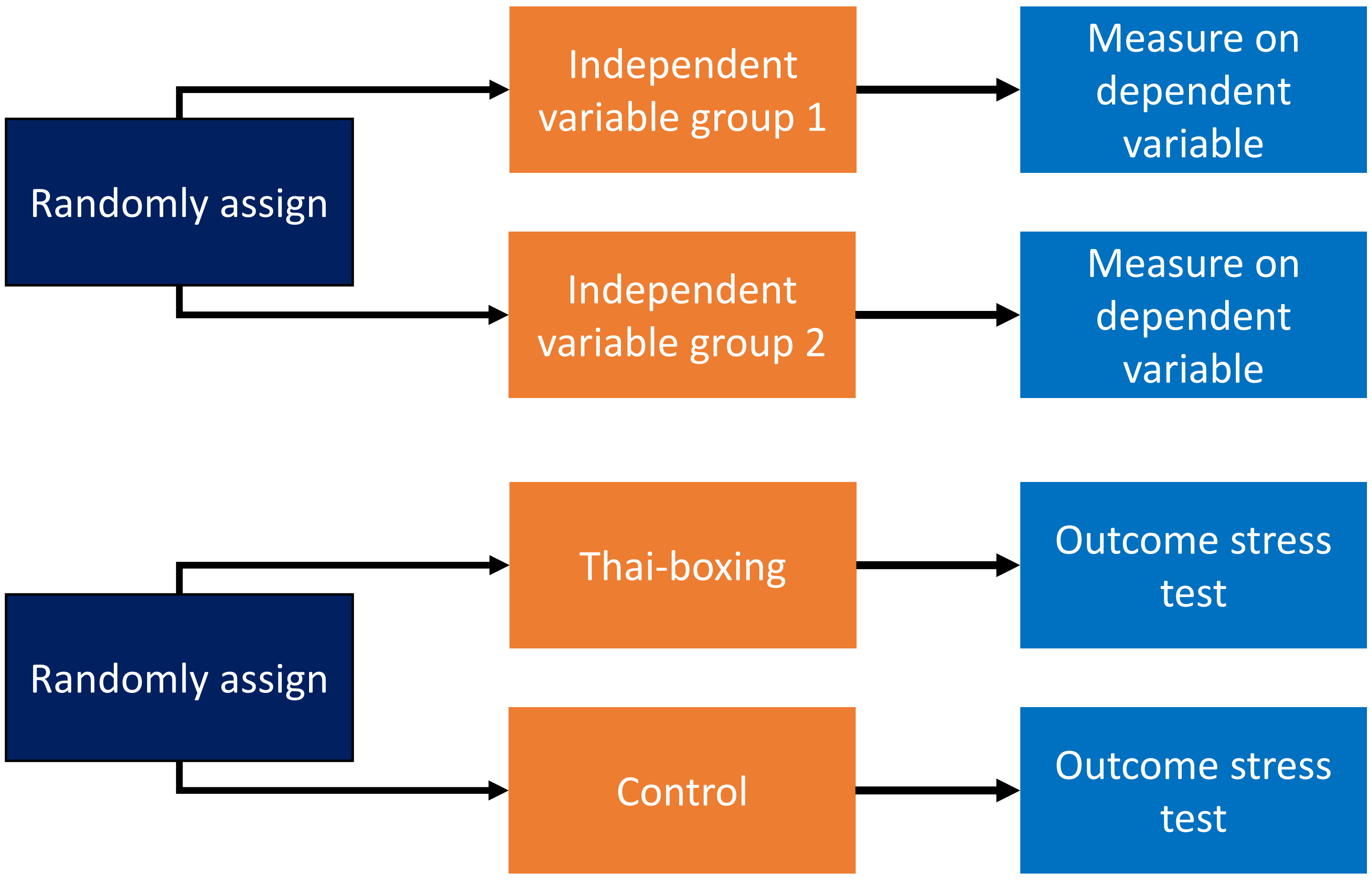
- **AOb:**

  https://www.aob.nl/actueel/artikelen/vakbonden-9-december-demonstratie-tegen-sloop-hogeronderwijs

Example: Strategies for stress relief

Picture source: pixabay.com

Morling Chapter 10: independent-groups design (posttest only):

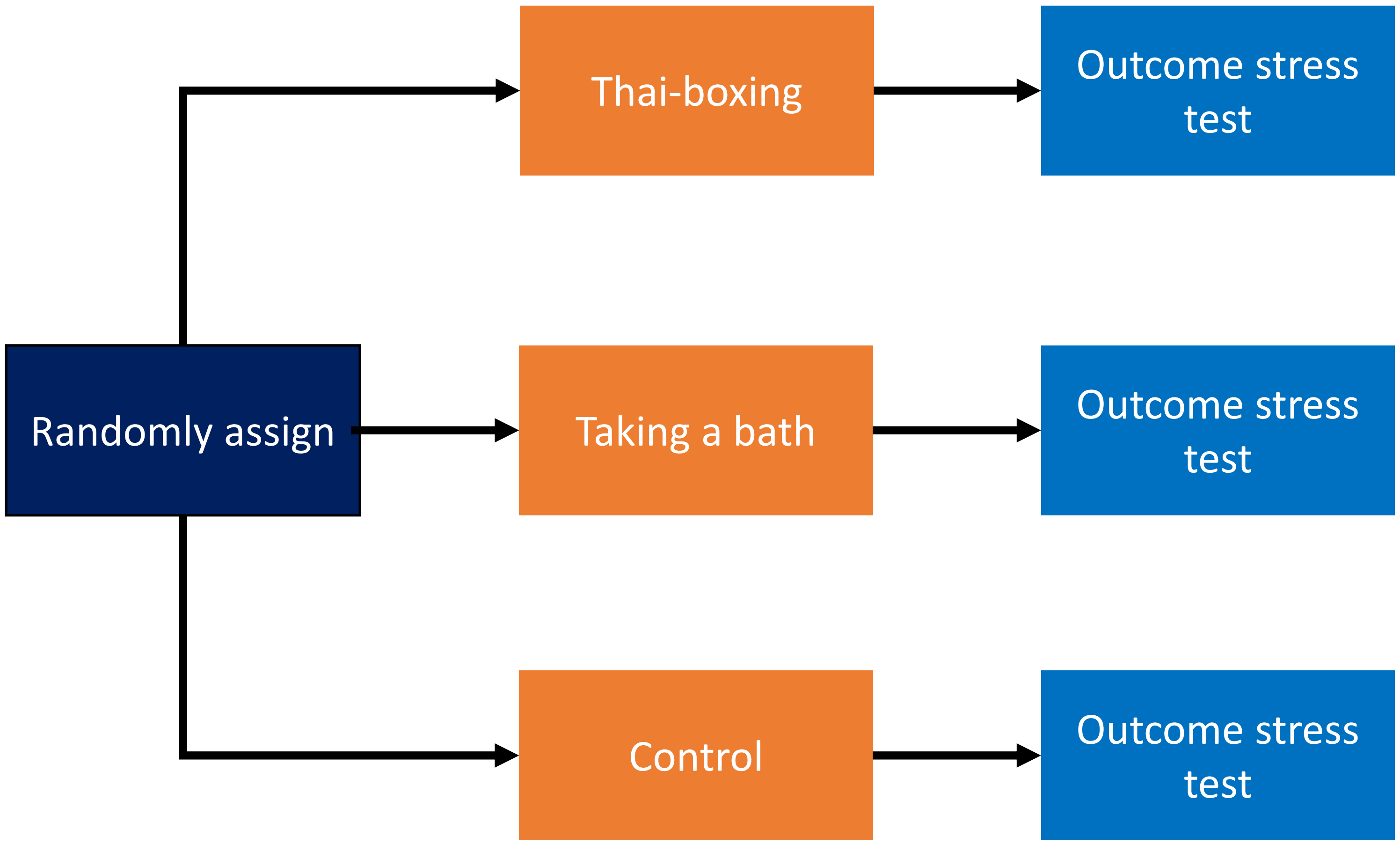


Explanatory variable/independent variable

Response variable/ dependent variable

	Quantitative	**Categorical**
Quantitative	Correlation Regression	t-test ANOVA
Categorical	Logistic regression	Contingency table

More than two groups:



Also more than two groups:

Strategy	**Age** elderly	 Young adults
Thai boxing	7.1	7.3
No sport	9.5	13.1

More than two groups: 3 x 2 x 3 factorial design

Age

elderly

Strategy

Lifestyle	**Thai boxing**	**bath**	**control**
Sporty	?	?	?
not sporty	?	?	?

young adults

Strategy

Lifestyle	**Thai boxing**	**bath**	**control**
Sporty	?	?	?
not sporty	?	?	?

children

Strategy

Lifestyle	**Thai boxing**	**bath**	**control**
Sporty	?	?	?
not sporty	?	?	?

Explanatory variable/independent variable

Response variable/ dependent variable

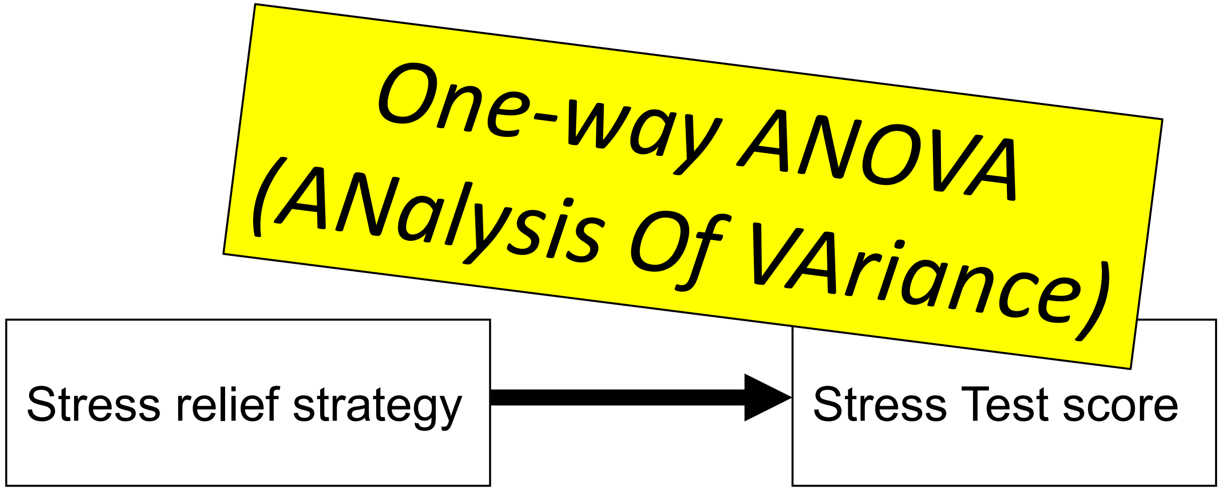
	Quantitative	Categorical
Quantitative	Correlation Regression	t-test ANOVA
Categorical	Logistic regression	Contingency table

Why not just a bunch of t-tests?

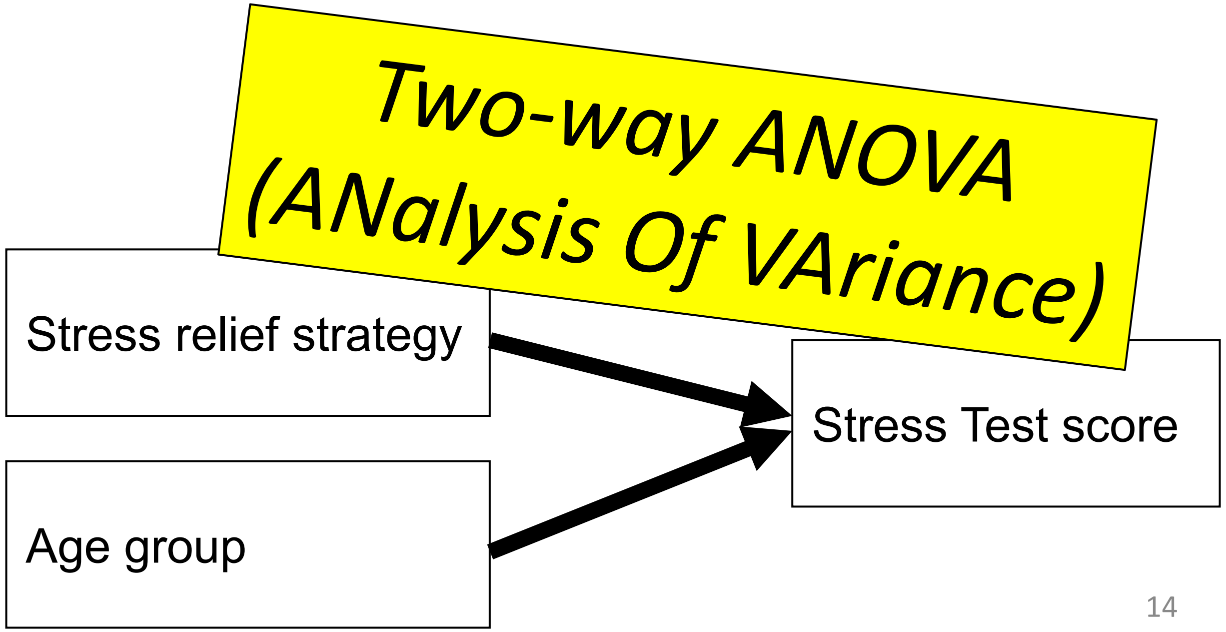
- With more than two groups you would need multiple t-tests to compare each pair of groups
- The number of tests increases fast with increasing numbers of groups
  - E.g, 4 groups = 6 tests, 5 groups = 10 tests, 10 groups = 45 tests
- If each test uses $\alpha = 0.05$ then the total probability of falsely rejecting H_0 in at least one of these tests will be much larger than 0.05!
- So, instead we do a two-step procedure:
  - Step 1: a single test (the ANOVA test) to test whether there is a difference between at least two groups
  - Step 2: if test in step 1 is significant, only then we check which groups differ.. (and use a correction)

More than two groups:

Strategy	
Thai boxing	Stress score
Bath	Stress score
No sport	Stress score



	Age	
Strategy	elderly	Young adults
Thai boxing	Stress score	Stress score
Bath	Stress score	Stress score
No sport	Stress score	Stress score



Three-way ANOVA (ANalysis Of VAriance)

Age

elderly

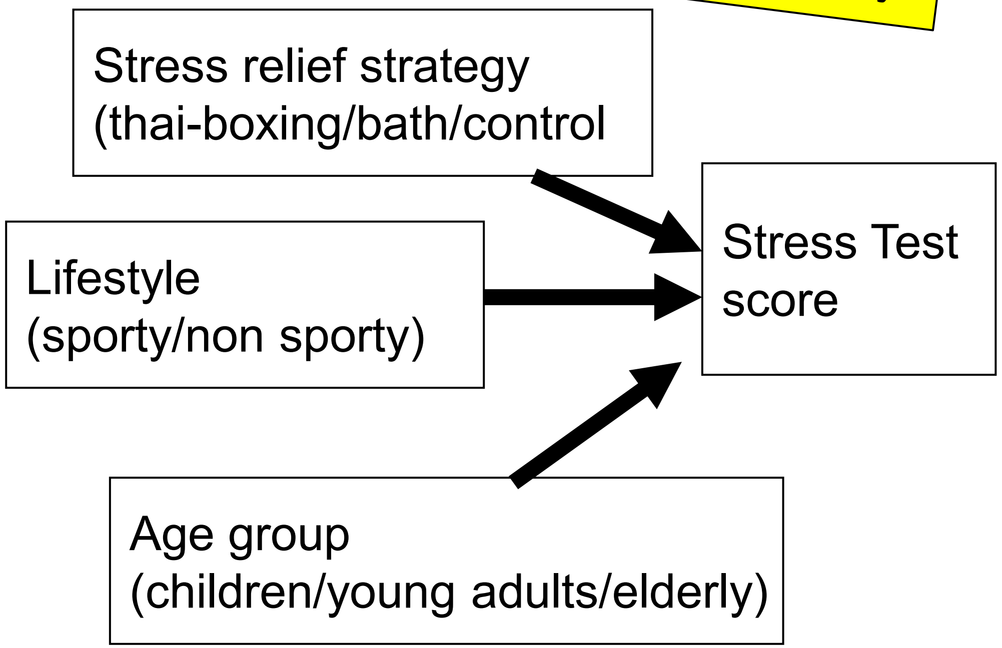
Lifestyle	Strategy: Thai boxing	bath	control
Sporty	?	?	?
not sporty	?	?	?

young adults

Lifestyle	Strategy: Thai boxing	bath	control
Sporty	?	?	?
not sporty	?	?	?

children

Lifestyle	Strategy: Thai boxing	bath	control
Sporty	?	?	?
not sporty	?	?	?



Example today

Explanatory categorical variables are referred to as *factors*

- 1 factor with three levels: thai-boxing vs bath vs control
- These are the 3 *independent groups*
- Quantitative dependent variable (or *response variable)*: stress level

- The statistical method to compare means of more than two groups is called:

Analysis of Variance (ANOVA)

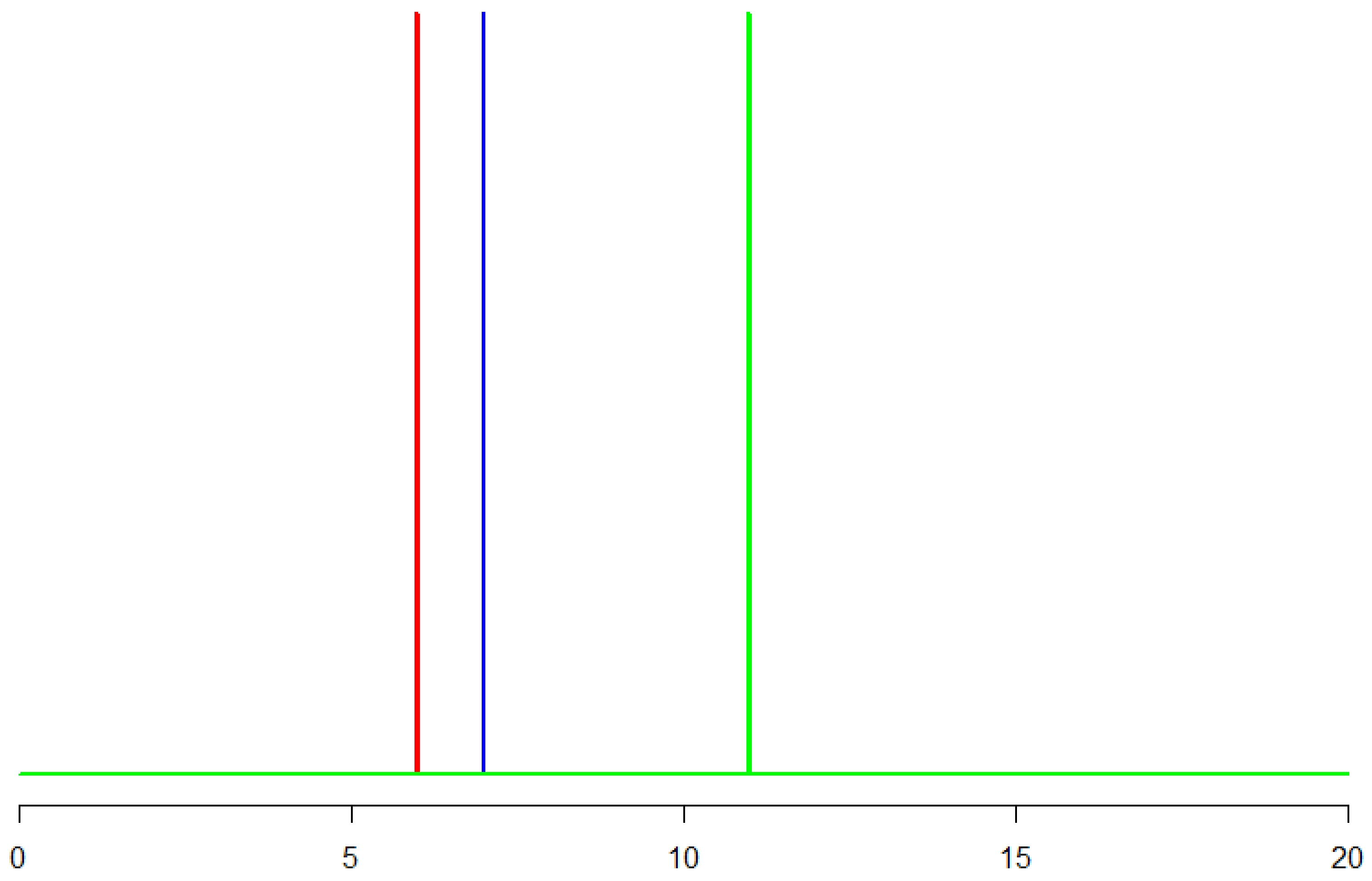
Today

1) Main idea underlying ANOVA

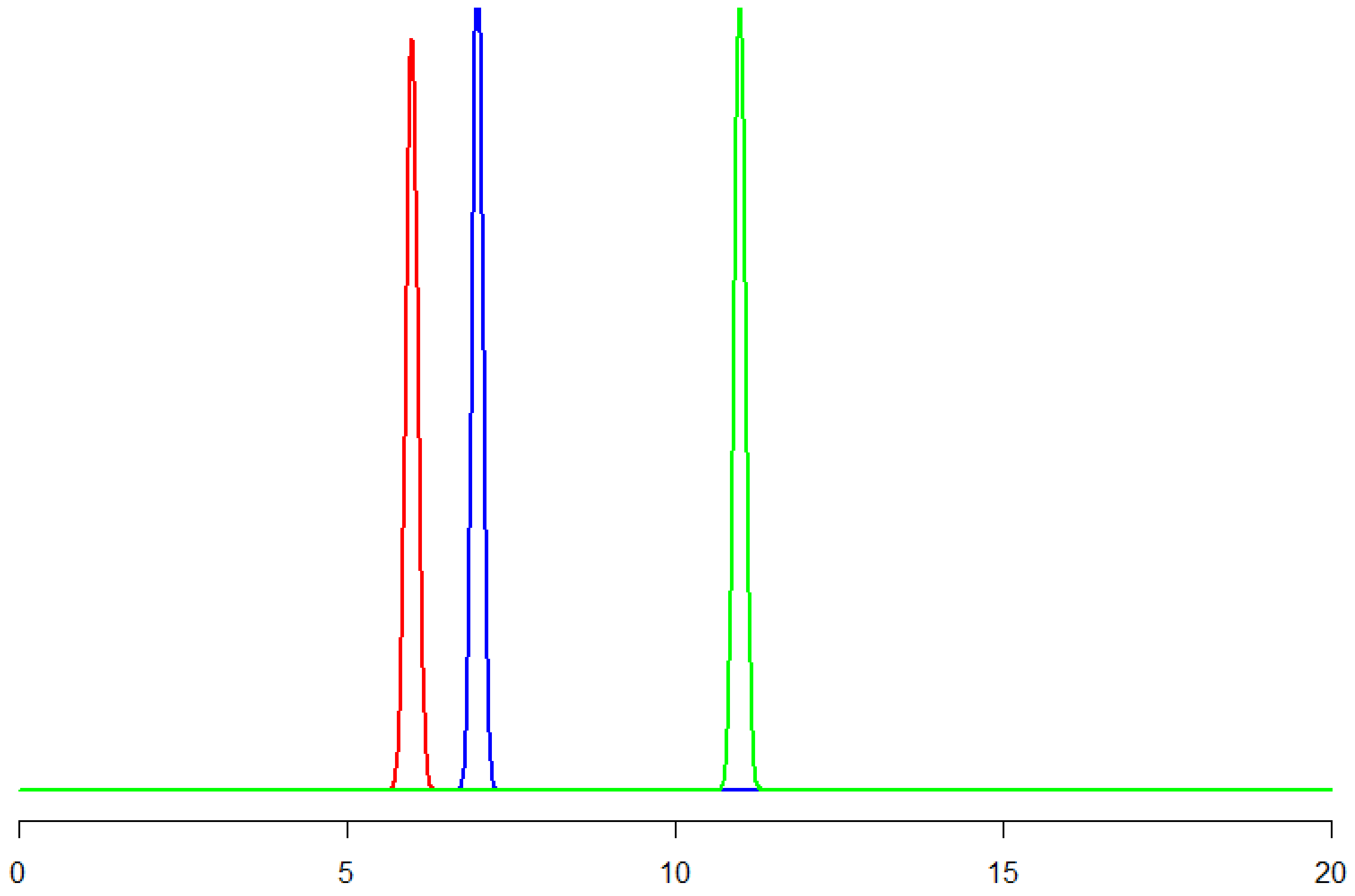
2) How to perform a one-way ANOVA

3) Estimating differences in groups

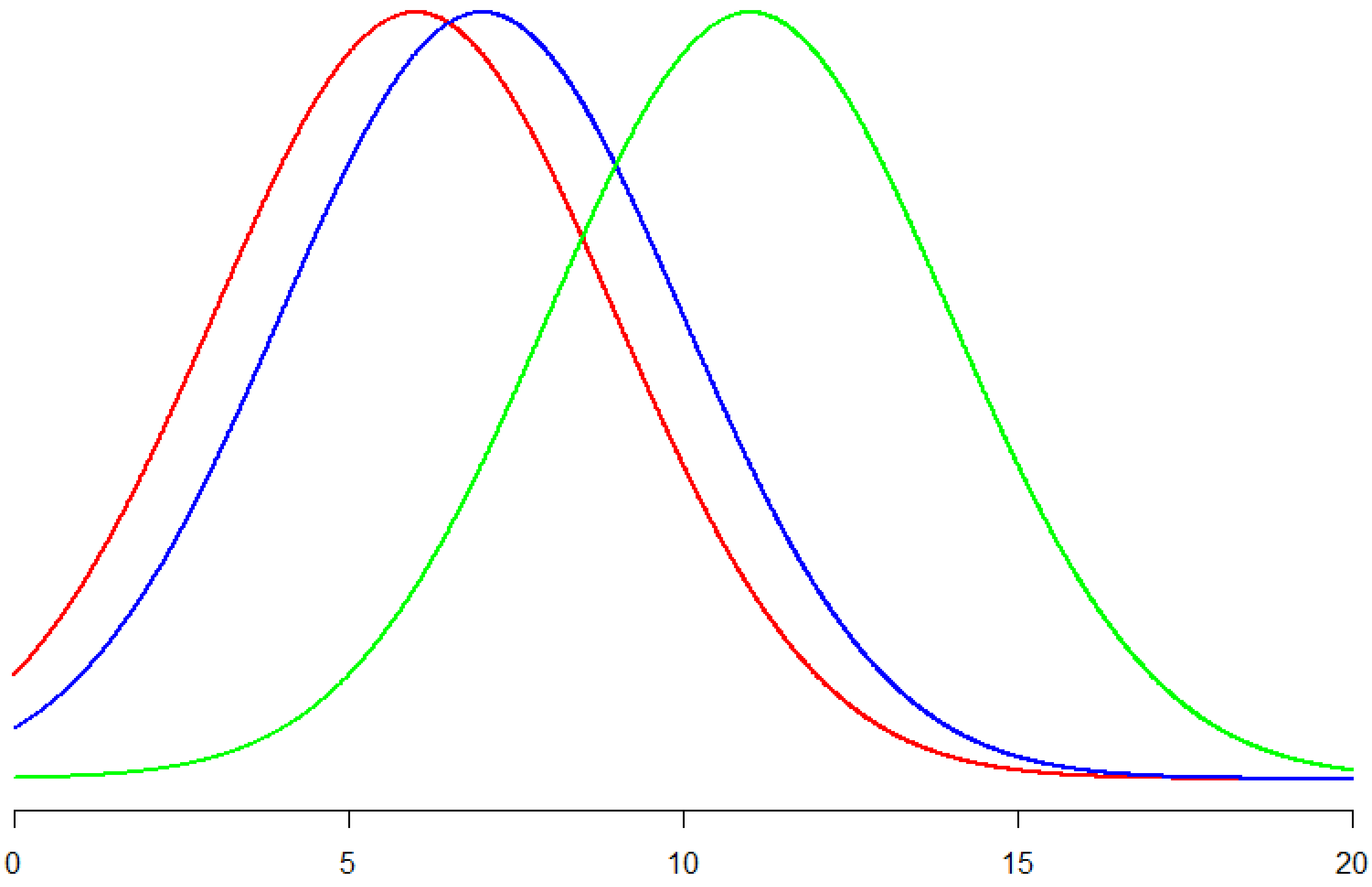
	Thai-boxing	Bath	Control
	6	7	11
	6	7	11
	6	7	11
	6	7	11
	6	7	11
	6	7	11
Mean	$\bar{y}_1$ = 6	$\bar{y}_2$ = 7	$\bar{y}_3$ = 11
Variance	$s^2_1 = 0$	$s^2_2 = 0$	$s^2_3 = 0$
Sample Size	$n_1 = 6$	$n_2 = 6$	$n_3 = 6$

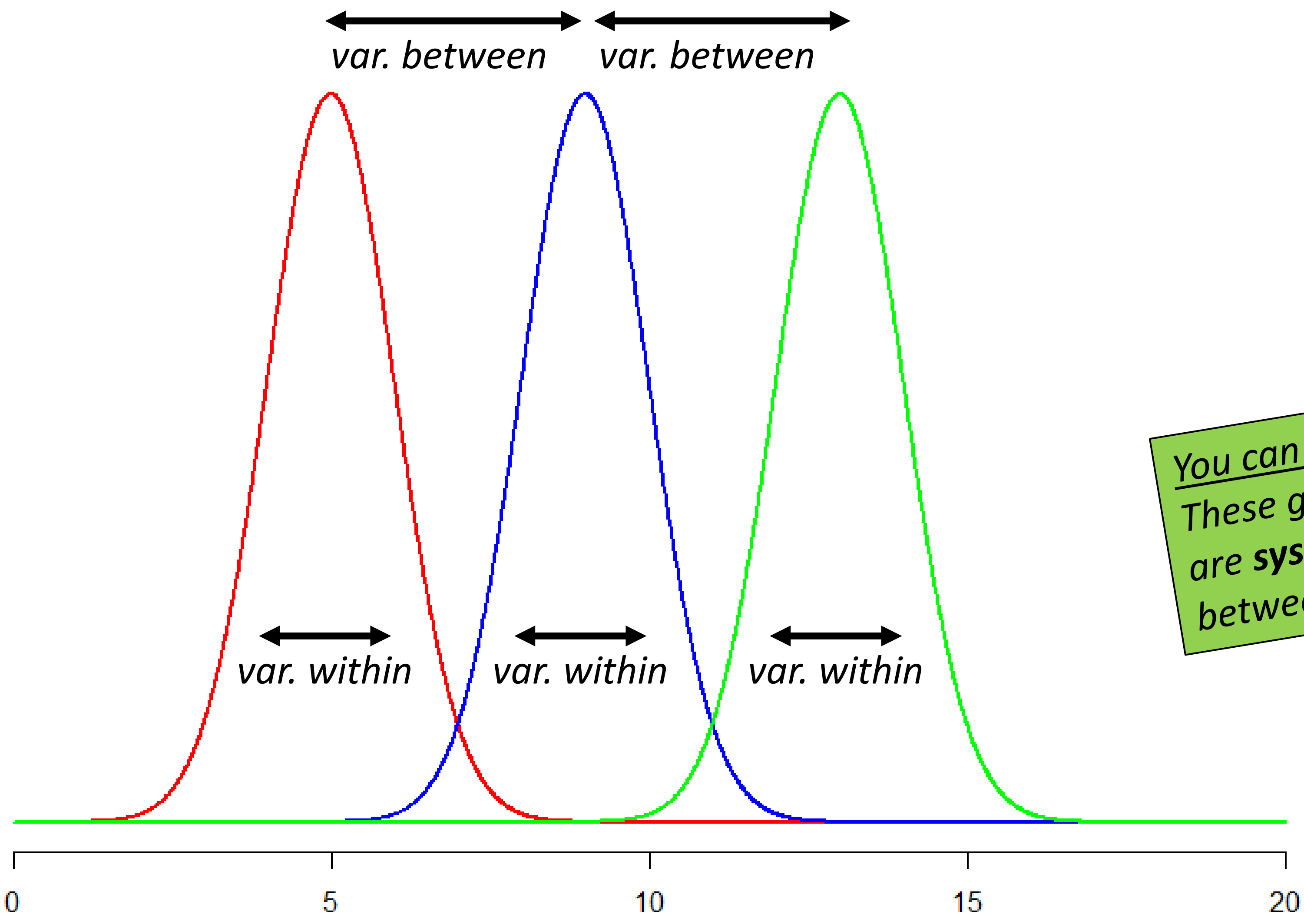

0
5
10
15
20

	Thai-boxing	Bath	Control
	6.1	7.1	11.1
	5.9	6.9	10.9
	6.0	7.0	10.9
	6.1	7.1	11.0
	6.0	6.9	11.1
	5.9	7.0	11.0
Mean	$\bar{y}_1$ = 6	$\bar{y}_2$ = 7	$\bar{y}_3$ = 11
Variance	$s^2{}_1$ = 0.008	$s^2{}_2$ = 0.008	$s^2{}_3$ = 0.008
Sample Size	n_1 = 6	n_2 = 6	n_3 = 6

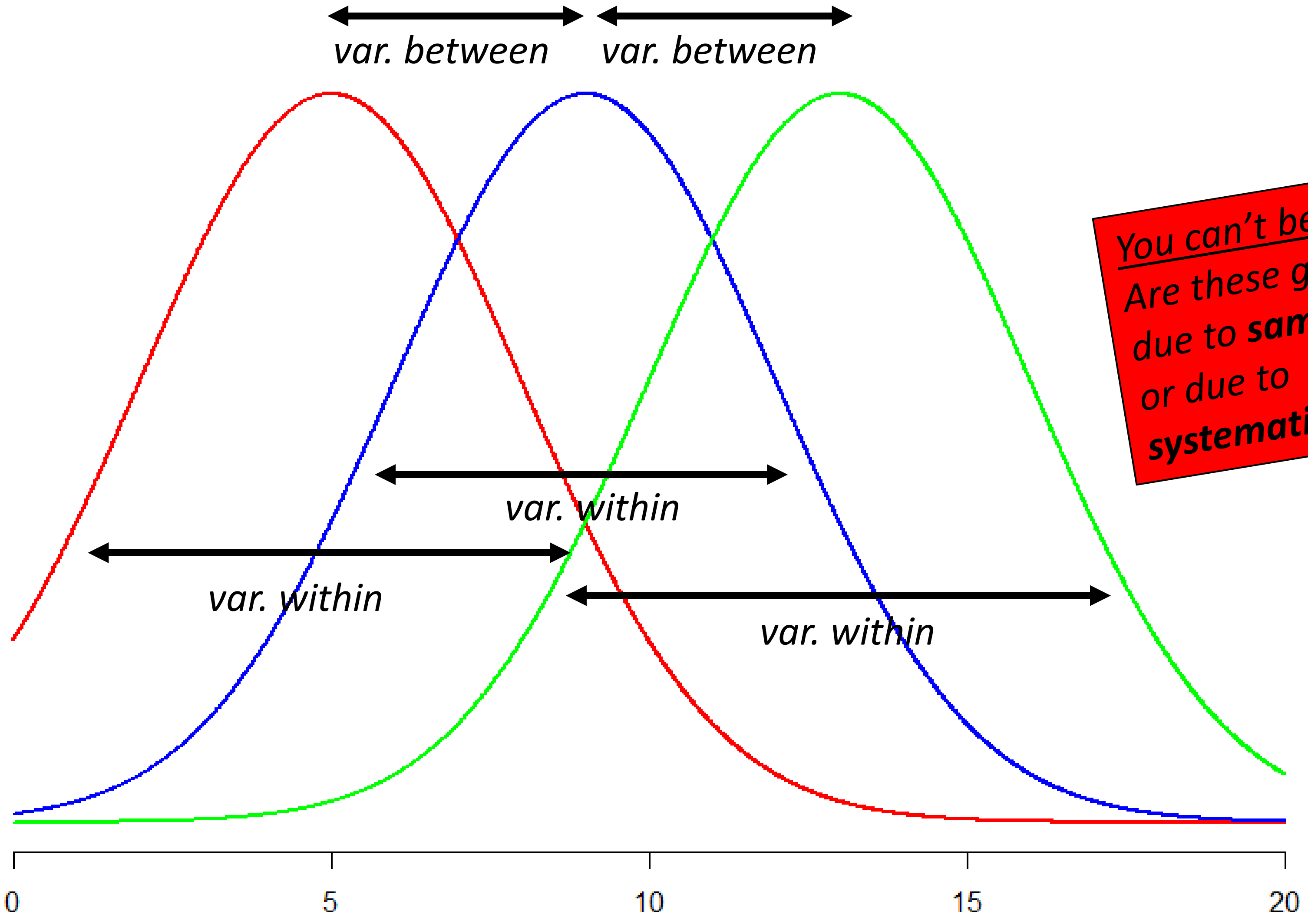

0
5
10
15
20

	Thai-boxing	Bath	Control
	8.2	6.6	13.4
	0.7	9.1	9.8
	4.4	4.0	6.1
	8.2	6.6	13.4
	6.3	4.0	9.8
	8.2	11.7	13.4
Mean	$\bar{y}_1 = 6$	$\bar{y}_2 = 7$	$\bar{y}_3 = 11$
Variance	$s^2_1 = 9.0$	$s^2_2 = 9.0$	$s^2_3 = 8.8$
Sample Size	$n_1 = 6$	$n_2 = 6$	$n_3 = 6$


0
5
10
15
20



<u>*You can be sure:*</u>
*These group differences are **systematic differences** between the groups*



<u>You can't be sure:</u>
Are these group differences due to **sampling variability** or due to **systematic differences?**

Main idea underlying ANOVA

If the between-variability is much larger than the within-variability you can be more sure of true group differences

Idea behind ANOVA

The idea: If there are no differences between groups (all groups come from the same population), then you expect the differences between people from different groups to bet the same as the differences between people from the same group!

- Focus on: $\frac{variability\ between\ groups}{variability\ within\ groups}$

$$\frac{variabilitybetween\ groups}{variabilitywithin\ groups} \approx 1$$

variability between and within are approx. equal
→ no evidence for group differences

$$\frac{variabilitybetween\ groups}{variabilitywithin\ groups} > 1$$

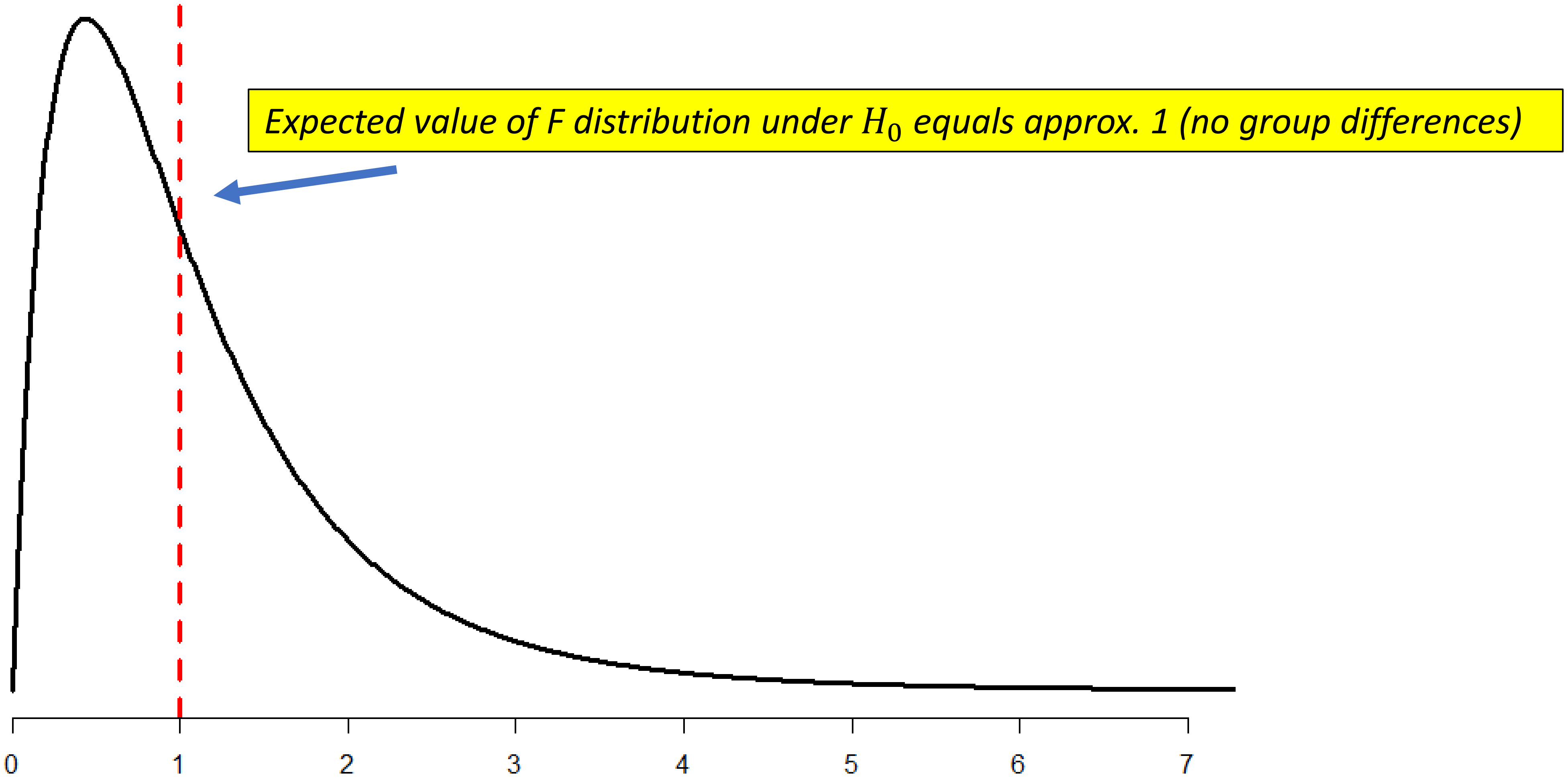
variability between groups is much larger than variability within groups ***→ group differences!***

Test statistic: $F = \frac{variability\ between\ groups}{variability\ within\ groups}$ $df_1 = g - 1$ $df_2 = N - g$

g = Number of groups

N = total sample size all groups together

F-distribution (F.DIST(..) in Excel)



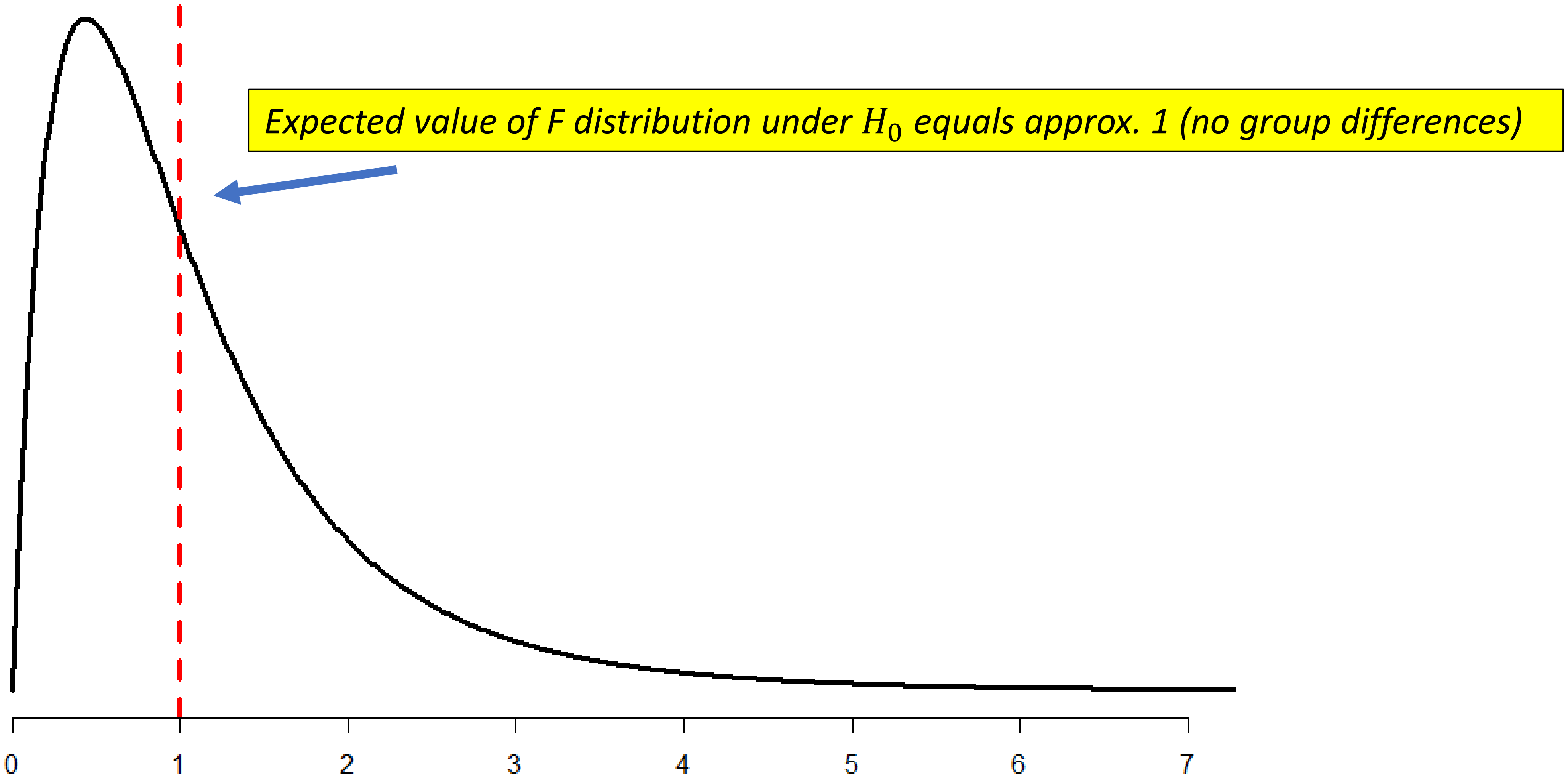
Advanced: why is expected value 1? (p.766)

- Why is the expected value of F under H_0 approx. 1?
- $F = \frac{between\ groups\ estimate\ of\ \sigma^2}{within\ groups\ estimate\ of\ \sigma^2}$
- Within groups estimate is based on variability within the groups, and between groups estimate is based on the differences between groups, both to estimate the variance in the population.
- Under H_0 , two unbiased estimates of σ^2:
- $within\ groups\ estimate\ of\ \sigma^2 = \frac{s_1^2+s_2^2+\cdots+s_g^2}{g}$
- $between\ groups\ estimate\ of\ \sigma^2 = \frac{n[(\bar{y}_1-\bar{y})^2+(\bar{y}_2-\bar{y})^2+\cdots+(\bar{y}_g-\bar{y})^2]}{g-1}$
- If both estimates are unbiased they have same expected value (hence ratio is 1!!). If H_0 is not true, then between groups estimate is biased as an estimate of σ^2 and will be bigger (and thus ratio > 1)

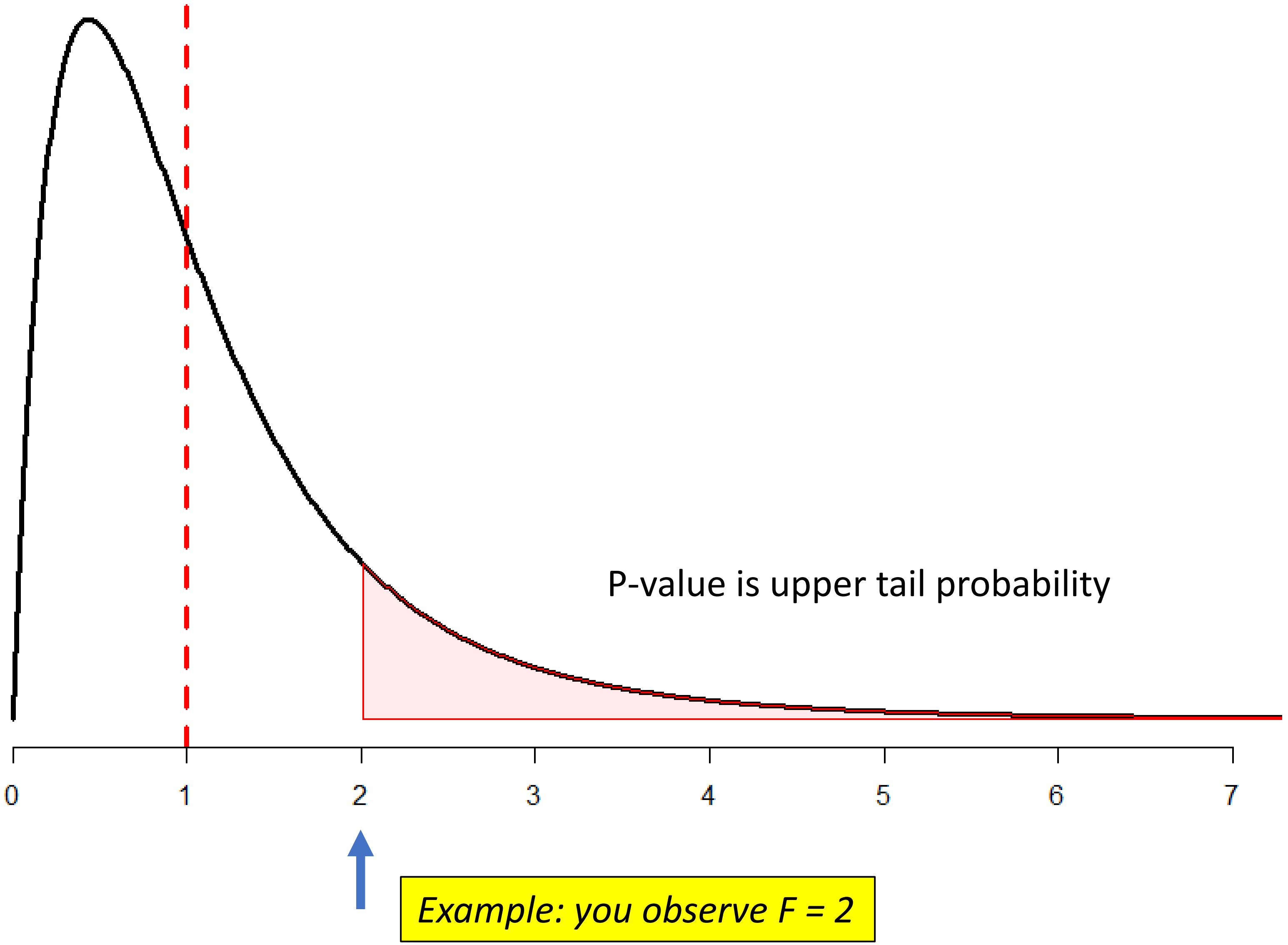
Remember: sample variance: $s^2 = \frac{\sum(y_i-\bar{y})^2}{n-1}$

Remember: variance of sampling distribution = $\frac{\sigma^2}{n}$

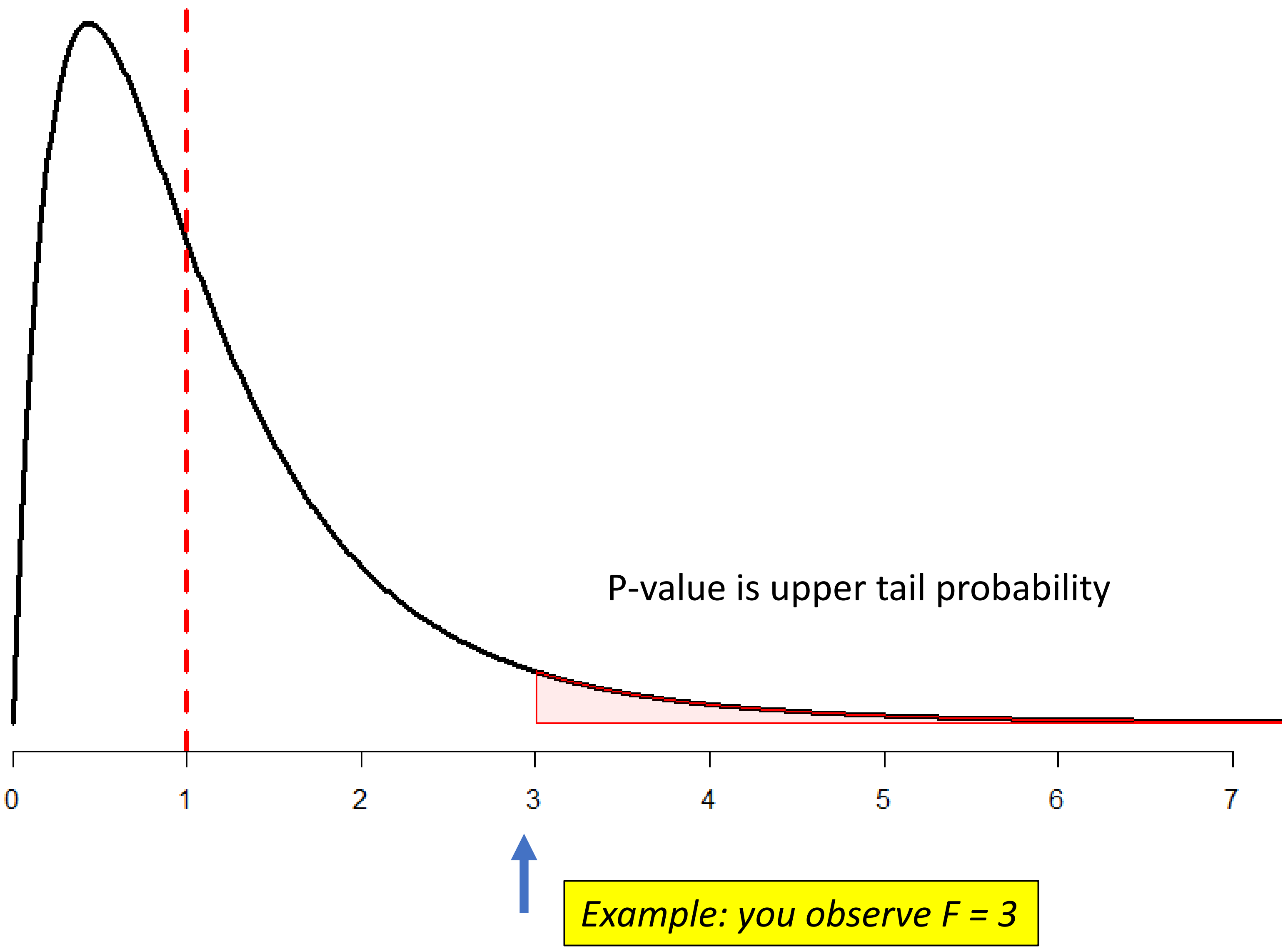
F-distribution (F.DIST(..) in Excel)



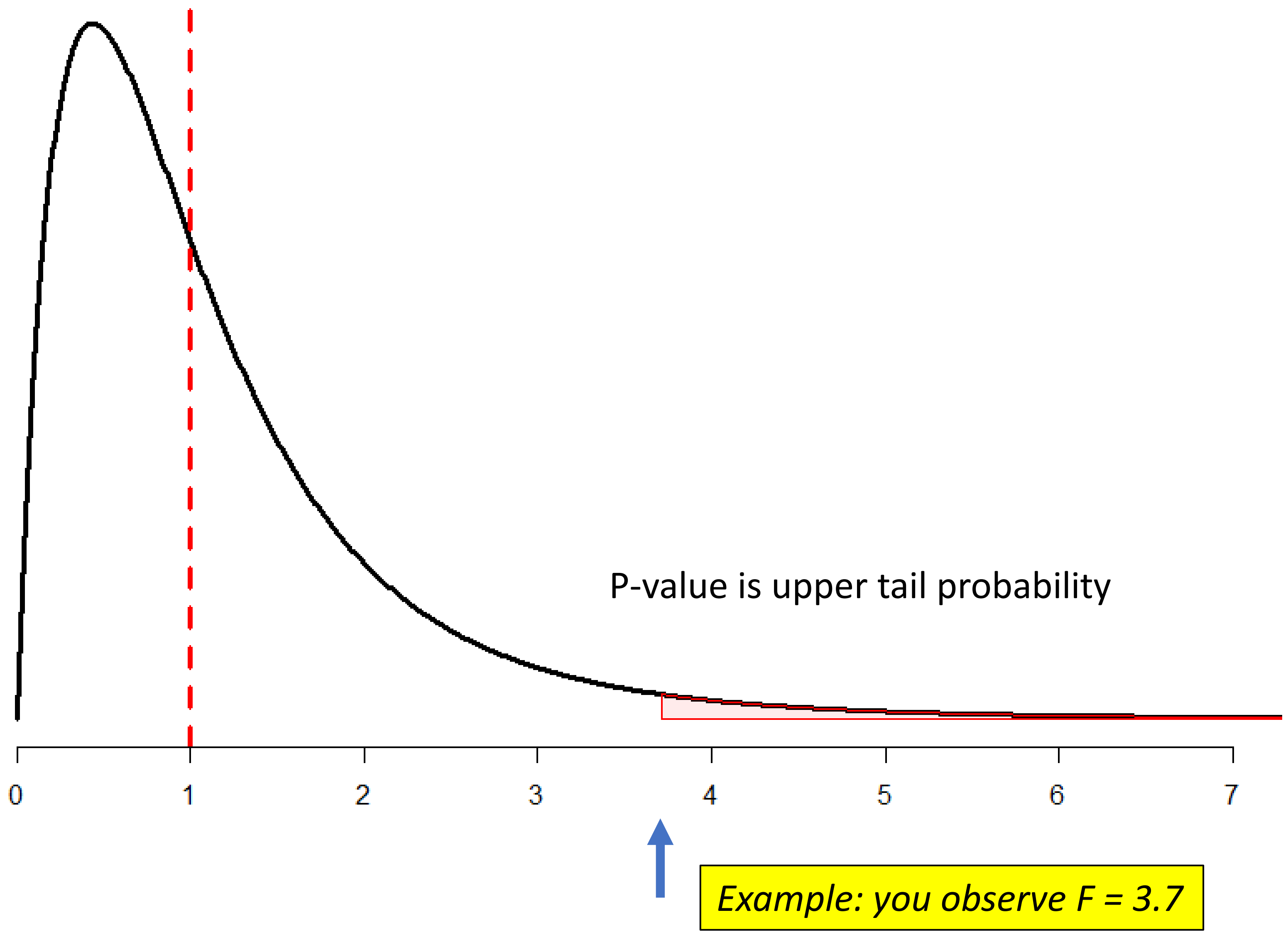
F-distribution (F.DIST(..) in Excel)



F-distribution (F.DIST(..) in Excel)



F-distribution (F.DIST(..) in Excel)



Today

1) Main idea underlying ANOVA

2) **How to perform an ANOVA**

3) Estimating differences in groups

	Thai-boxing	Bath	Control
	8	10	16
	9	9	12
	5	8	13
	4	7	10
	4	5	9
	6	3	6
Mean	$\bar{y}_1 = 6$	$\bar{y}_2 = 7$	$\bar{y}_3 = 11$
Variance	$s^2_1 = 4.4$	$s^2_2 = 6.8$	$s^2_3 = 12$
Sample Size	$n_1 = 6$	$n_2 = 6$	$n_3 = 6$

Step 1: Assumptions

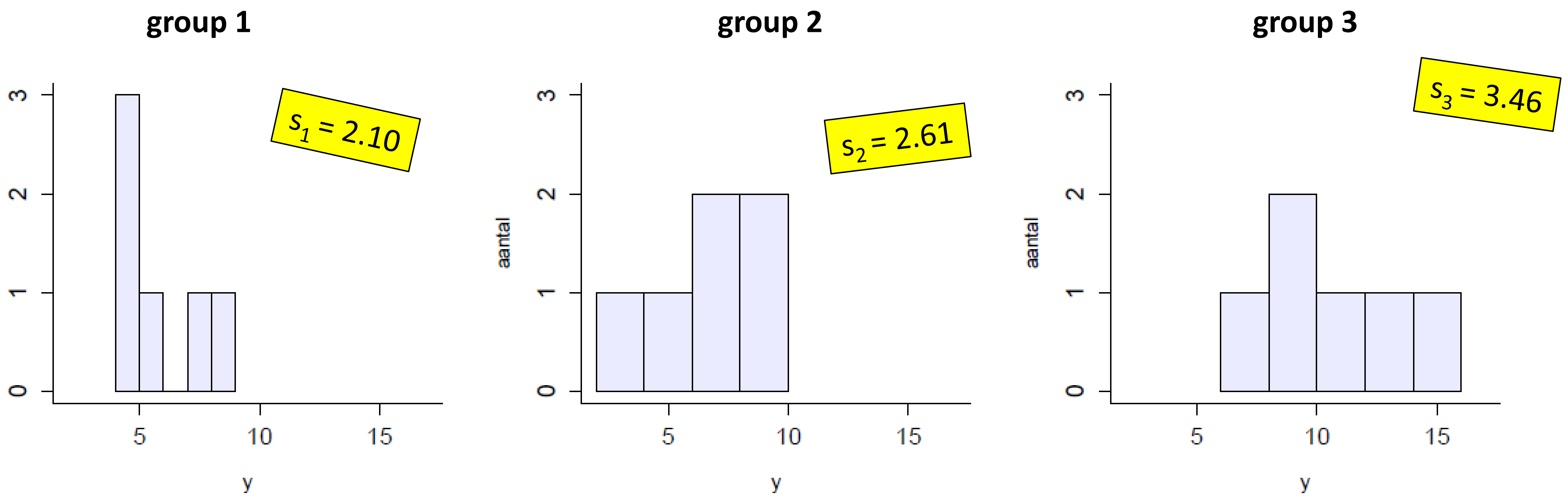
- A quantitative variable in more than two groups

- Independent random samples

- Normal distribution in the population with equal standard deviations

In this course we focus on a specific simpler case where group sizes are equal

- Group sizes are equal
  - In that case the formulas are simpler and so you do not need statistical software other than Excel



- Normal?
  - Moderate violations are not problematic, and even less so if sample size is large
- Equal standard deviations (SD)?
  - If sample size is equal over groups → violations are not problematic
  - Unequal sample sizes? If largest SD is not more than 2 x smallest SD → violations are not problematic

Step 2: Hypotheses

H_0: $\mu_1 = \mu_2 = \cdots = \mu_g$

(i.e., population means of all groups are equal)

H_a: At least two of the population means are different

Step 3: Statistic

Test statistic: $$F = \frac{between\ groups\ estimate\ of\ \sigma^2}{within\ groups\ estimate\ of\ \sigma^2} = \frac{variability\ between\ groups}{variability\ within\ groups}$$

df_1 = g - 1

df_2 = N - g

Step 3: Statistic

$$F = \frac{between\ groups\ estimate\ of\ \sigma^2}{\boldsymbol{within\ groups\ estimate\ of\ \sigma^2}}$$

- Variability within groups:

$$\text{within} - \text{groups variance estimate of } \sigma^2 = \frac{s_1^2 + s_2^2 + \cdots + s_g^2}{g}$$

- → In our case:

$$\text{within} - \text{groups variance estimate of } \sigma^2 = \frac{s_1^2 + s_2^2 + \cdots + s_g^2}{g}$$

$$= \frac{4.4+6.8+12}{3} = \frac{23.2}{3} = 7.73$$

	Thai-boxing	**Bath**	**Control**
	8	10	16
	9	9	12
	5	8	13
	4	7	10
	4	5	9
	6	3	6
Mean	$\bar{y}_1$ = 6	$\bar{y}_2$ = 7	$\bar{y}_3$ = 11
Variance	$s^2{}_1$ = 4.4	$s^2{}_2$ = 6.8	$s^2{}_3$ = 12
Sample Size	n_1 = 6	n_2 = 6	n_3 = 6

Step 3: Statistic

$$F = \frac{\boldsymbol{between\ groups\ estimate\ of\ \sigma^2}}{within\ groups\ estimate\ of\ \sigma^2}$$

- Variability between groups:

$$\text{between} - \text{groups variance estimate of } \sigma^2 = \\ = \frac{n\left[(\bar{y}_1 - \bar{y})^2 + (\bar{y}_2 - \bar{y})^2 + \cdots + \left(\bar{y}_g - \bar{y}\right)^2\right]}{g-1}$$

  - n: sample size in each group (equal sample sizes)
  - $\bar{y}$: overall mean of all sample data

- → In our case:

for equal sample sizes you can calculate: $\bar{y}$ = (6+7+11)/3 = 8

$$\text{between} - \text{groups variance estimate of } \sigma^2 \\ = \frac{n\left[(\bar{y}_1-\bar{y})^2+(\bar{y}_2-\bar{y})^2+\cdots+\left(\bar{y}_g-\bar{y}\right)^2\right]}{g-1} = \frac{6[(6-8)^2+(7-8)^2+(11-8)^2]}{2} = \frac{84}{2} = 42$$

	Thai-boxing	Bath	Control
	8	10	16
	9	9	12
	5	8	13
	4	7	10
	4	5	9
	6	3	6
Mean	$\bar{y}_1$ = 6	$\bar{y}_2$ = 7	$\bar{y}_3$ = 11
Variance	s^2_1 = 4.4	s^2_2 = 6.8	s^2_3 = 12
Sample Size	n_1 = 6	n_2 = 6	n_3 = 6

Step 3: Statistic

- Thus:

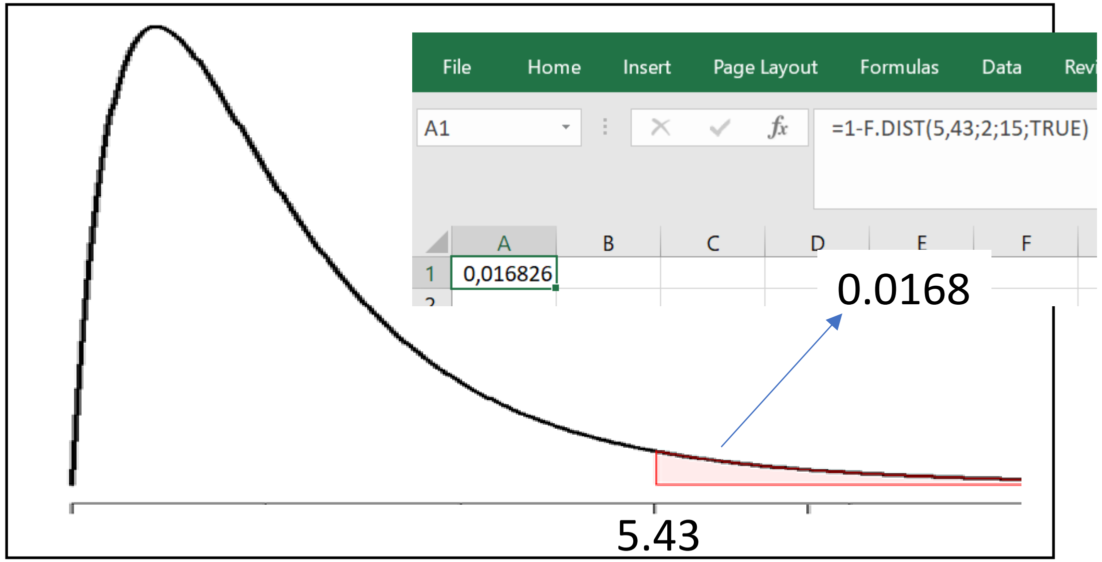
Between-groups variance estimate of σ^2 = 42

Within-groups variance estimate of σ^2 = 7.73

→ $$F = \frac{between\ groups\ estimate\ of\ \sigma^2}{within\ groups\ estimate\ of\ \sigma^2} = \frac{42}{7.73} = 5.43$$

Step 4: P-value

- So F = 5.43, what is the corresponding P-value?



- Excel
  - $df_1 = g - 1 = 3 - 1 = 2$
  - $df_2 = N - g = 18 - 3 = 15$
  - g: number of groups
  - N total number of subjects: $n_1 + n_2 + _{\dots} n_g$

Thus: P < 0.05

Because we want a P-value in the F-distribution

= **1 - F.DIST**(5,43; 2; 15; TRUE)

Because you want upper tail (always! For F-test)

Our calculated F-value

df1

df2

Always TRUE if you want to calculate a P-value

Step 5: Conclusion

- Suppose $\alpha = 0.05$
- $p < 0.05$

- Reject H_0

- There is a significant main effect of stress relief strategy!

- At least two groups differ

The ANOVA table

$MS_{group} = MS_{between}$ = Between-groups variance estimate of σ^2

Source	DF	SS	MS	F	P
Group	?	?	42	42/7.73=5.43	0.017
Error	?	?	7.73		
Total	?				

$MS_{error} = MS_{within}$ = Within-groups variance estimate of σ^2

The ANOVA table

$df_{group} = df_1 = g\text{-}1$

Source	DF	SS	MS	F	P
Group	?	?	42	5.43	0.017
Error	?	?	7.73		
Total	?				

$df_{error} = df_2 = N - g$

$df_{total} = df_1 + df_2 = df_{group} + df_{error}$

The ANOVA table

$df_{group} = df_1 = g\text{-}1$

Source	DF	SS	MS	F	P
Group	2	?	42	5.43	0.017
Error	15	?	7.73		
Total	17				

$df_{error} = df_2 = N - g$

$df_{total} = df_1 + df_2 = df_{group} + df_{error} = N - 1$

The ANOVA table

SS_{group} = Sum of Squares Between = (Between-groups variance estimate of σ^2) × (df_{group})

Source	DF	SS	MS	F	P
Group	2	?	42	5.43	0.017
Error	15	?	7.73		
Total	17				

SS_{within} = Sum of Squares Within = (Within-groups variance estimate of σ^2) × (df_{error})

The ANOVA table

SS_{group} = Sum of Squares Between = (Between-groups variance estimate of σ^2) $\times$ (df_{group})

Source	DF	SS	MS	F	P
Group	2	84	42	5.43	0.017
Error	15	115.95	7.73		
Total	17				

SS_{within} = Sum of Squares Within = (Within-groups variance estimate of σ^2) $\times$ (df_{error})

Advanced: what are these sum of squares?

- For the between variability it is easier to see:

$$\text{between} - \text{groups estimate of } \sigma^2 =$$

$$= \frac{\overbrace{n\left[(\bar{y}_1 - \bar{y})^2 + (\bar{y}_2 - \bar{y})^2 + \cdots + \left(\bar{y}_g - \bar{y}\right)^2\right]}^{\textit{sum of squares between}}}{g - 1}$$

$g - 1$= degrees of freedom between = df_1

Advanced: what are these sum of squares?

- $\text{within} - \text{groups estimate of } \sigma^2 = \frac{s_1^2+s_2^2+\cdots+s_g^2}{g}$
- $s_l^2 = \frac{\sum(y_i-\bar{y}_l)^2}{n-1}, l = 1,2,..g$

(this is just the formula for the variance you have learned in Chapter 2)

- Plugging this into the within groups formula:

$$\frac{\frac{\sum(y_i-\bar{y}_1)^2}{n-1}+\frac{\sum(y_i-\bar{y}_2)^2}{n-1}+\cdots+\frac{\sum(y_i-\bar{y}_g)^2}{n-1}}{g} = \frac{\overbrace{\sum(y_i-\bar{y}_1)^2+\sum(y_i-\bar{y}_2)^2+\cdots+\sum(y_i-\bar{y}_g)^2}^{\text{sum of squares within}}}{(n-1)g} = \frac{SS_{within}}{N-g}$$

$(n-1)g = (n*g) - g = N - g$
= degrees of freedom within $= df_2$

The ANOVA table

Source	DF	SS	MS	F	P
Group	$g - 1$	SS_{group} $= MS_{group} \times (g - 1)$	Between-groups variance estimate of σ^2 $MS_{group} = SS_{group} / (g - 1)$	MS_{group} / MS_{error}	From excel, use F.dist(..)
Error	$N - g$	SS_{error} $= MS_{error} \times (N - g)$	Within-groups variance estimate of σ^2 $MS_{error} = SS_{error} / (N - g)$	-	-
Total	$N - 1$	-	-	-	-

Today

1) Main idea underlying ANOVA

2) How to perform an ANOVA

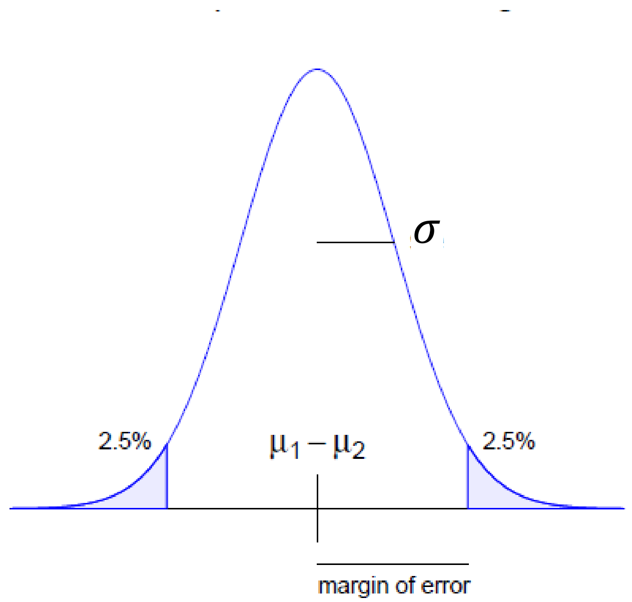
3) **Estimating differences in groups**

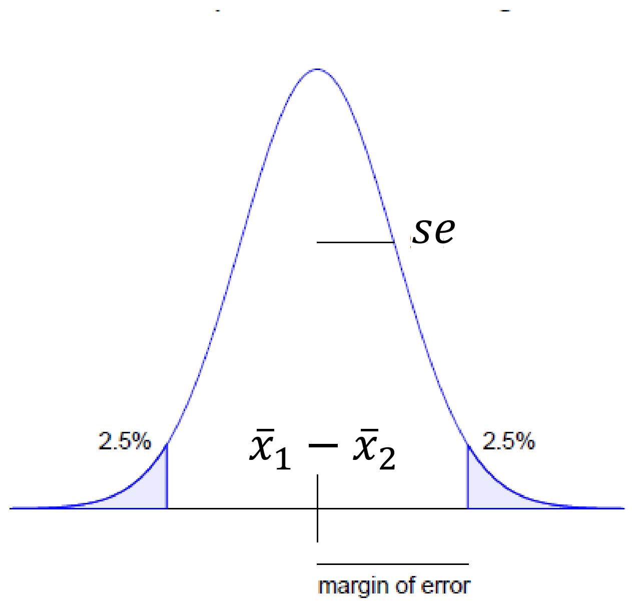
Remember from Ch 10 (Two means *in total*)

- Confidence interval for two population means
  - Sampling distribution $\bar{x}_1 - \bar{x}_2$
  - Estimate sampling distribution:
    - Mean = $\bar{x}_1 - \bar{x}_2$
    - Std. dev. = se = $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

95% confidence interval for $\mu_1 - \mu_2$:

- $\bar{x}_1 - \bar{x}_2 \pm t_{0.025}(\text{se})$ with se = $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$





ANOVA based confidence interval for two means

95% confidence interval for $\mu_1 - \mu_2$

- $\bar{x}_1 - \bar{x}_2 \pm$ margin of error
- $\bar{x}_1 - \bar{x}_2 \pm \mathrm{t}_{0.025}(\mathrm{se})$

The same for Ch 10 and ANOVA

- But different standard error! We assume equal standard deviations, so:
  - se $= \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}} =$ s $\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
  - s = $\sqrt{\frac{s_1^2 + s_2^2 + \cdots + s_g^2}{g}}$ (= square root of "within variability")
- What df to use for $\mathrm{t}_{0.025}$?
  - df = df_2 = N - g

"Fisher method"

Different for ANOVA as you have more information

Key difference with t-test (Ch 10) is that in ANOVA (Ch 14), you have more information because you have more groups to estimate the within group variability σ

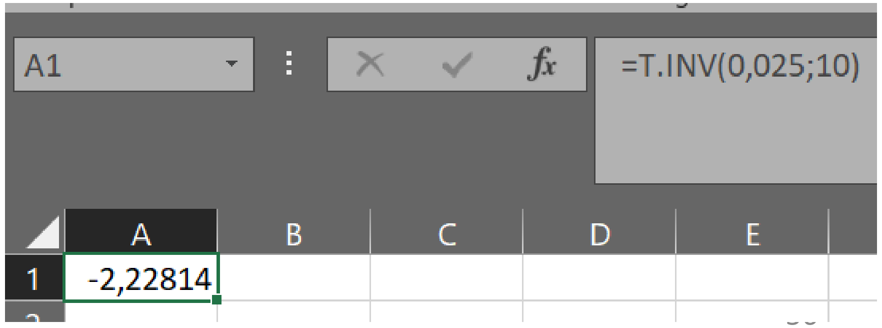
"Regular" (Chapter 10) confidence interval

	Thai-boxing	Bath	Control
Mean	$\bar{y}_1 = 6$	$\bar{y}_2 = 7$	$\bar{y}_3 = 11$
Variance	$s^2_1 = 4.4$	$s^2_2 = 6.8$	$s^2_3 = 12$
Sample Size	$n_1 = 6$	$n_2 = 6$	$n_3 = 6$

95% confidence interval

- $(\bar{y}_1 - \bar{y}_3) \pm t_{0.025}(\text{se})$
  - se $= \sqrt{\frac{s_1^2}{n_1} + \frac{s_3^2}{n_3}} = \sqrt{\frac{4.4}{6} + \frac{12}{6}} = 1.65$
  - $t_{0.025}$= 2.228 (from Excel, df = 6 + 6 – 2 = 10)
- Lower: -5 – 2.228 × 1.65 = -8.68
- Upper: -5 + 2.228 × 1.65 = -1.32

These will change in the Fisher method!

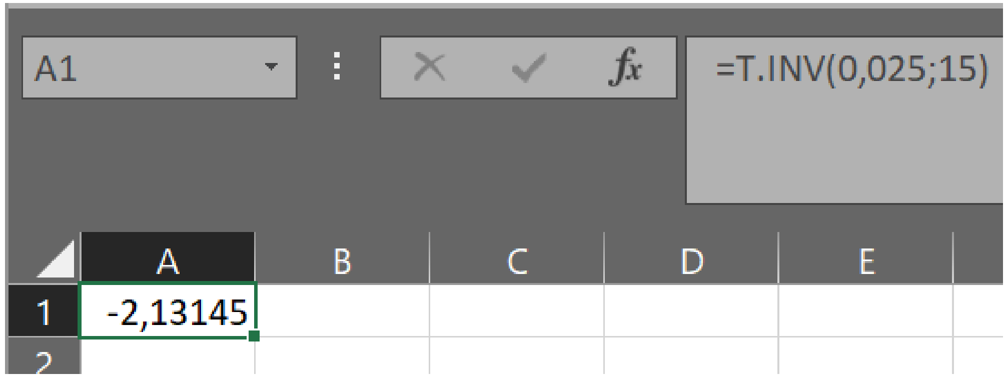


Fisher Method confidence interval (i.e., based on ANOVA, Ch 14)

	Thai-boxing	**Bath**	**Control**
Mean	$\bar{y}_1 = 6$	$\bar{y}_2 = 7$	$\bar{y}_3 = 11$
Variance	$s^2_1 = 4.4$	$s^2_2 = 6.8$	$s^2_3 = 12$
Sample Size	$n_1 = 6$	$n_2 = 6$	$n_3 = 6$

95% confidence interval

- $(\bar{y}_1 - \bar{y}_3) \pm t_{0\ 025}(se)$
  - $s = \sqrt{\frac{s_1^2+s_2^2+s_3^2}{3}} = \sqrt{\frac{4.4+6.8+12}{3}} = 2.78$
  - $se = s\sqrt{\frac{1}{n_1}+\frac{1}{n_3}} = 2.78 \times \sqrt{\frac{1}{6}+\frac{1}{6}} = 1.61$
  - $t_{0.025}$= 2.131 (from Excel, df = N – g = 6 + 6 + 6 – 3 = 15)
- Lower: -5 – 2.131 × 1.61 = -8.43
- Upper: -5 + 2.131 × 1.61 = -1.57

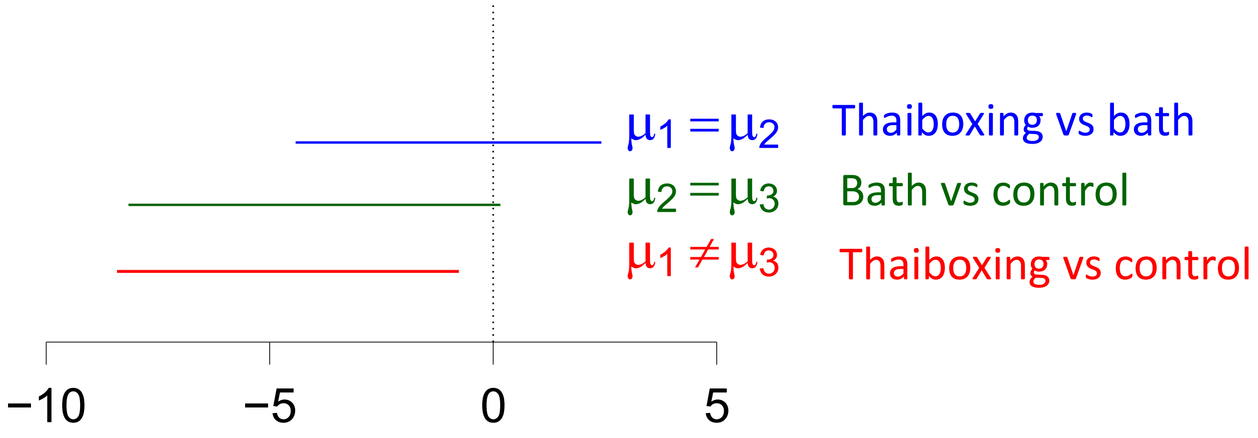


`Regular´ confidence interval

$(-4.42, 2.42)$

$(-8.16, 0.16)$

$(-8.68, -1.32)$

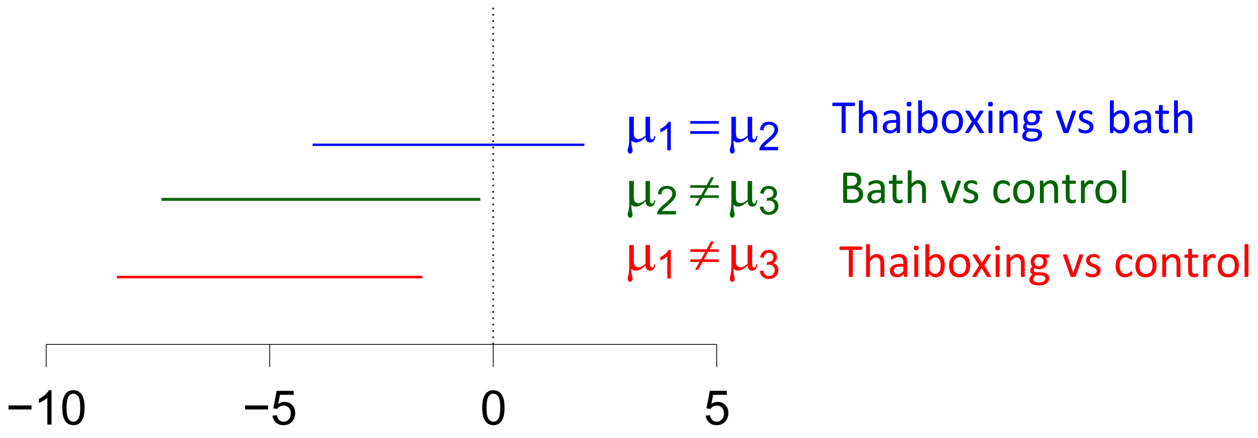


Fisher method

$(-4.04, 2.04)$

$(-7.42, -0.58)$

$(-8.43, -1.57)$



For anova, use Fisher method!

Why all the hassle?

- Why first an F-test and then in a second step additional comparisons?
- Why not just a bunch of t-tests to begin with?

- This all has to do with *multiple comparisons*
- If you have multiple groups to compare, you can end up with quite some comparisons to make..
- And this might lead to wrong conclusions because you *capitalize on chance*

Multiple comparisons

Journal of Personality and Social Psychology
2011, Vol. 100, No. 3, 407–425

0022-3514/11/$12.00 DOI: 10.1037/a0021524

Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect

Daryl J. Bem
Cornell University

The term *psi* denotes anomalous processes of information or energy transfer that are currently unexplained in terms of known physical or biological mechanisms. Two variants of psi are *precognition*
awareness) and *premonition* (affective apprehension) of a future event that could not
pated through any known inferential process. Precognition and premonition are
cases of a more general phenomenon: the anomalous retroactive influence of some
individual's current responses, whether those responses are conscious or noncon-
affective. This article reports 9 experiments, involving more than 1,000 participants,
ive influence by "time-reversing" well-established psychological effects so that the
es are obtained before the putatively causal stimulus events occur. Data are presented
effects: precognitive approach to erotic stimuli and precognitive avoidance of
roactive priming; retroactive habituation; and retroactive facilitation of recall. The
) in psi performance across all 9 experiments was 0.22, and all but one of the
statistically significant results. The individual-difference variable of stimulus seek-
of extraversion, was significantly correlated with psi performance in 5 of the
articipants who scored above the midpoint on a scale of stimulus seeking achieving
of 0.43. Skepticism about psi, issues of replication, and theories of psi are also

Table 7

Psi Performance in All Nine Experiments: Probability Levels (p), Effect Sizes (d), and Correlations (r) With Stimulus Seeking (SS)

Phenomenon tested and experiment	p full sample	d full sample	Correlation with SS	p high SS	d high SS	p low SS	d low SS
Precognitive approach/avoidance							
1. Detection of Erotic Stimuli	.01	0.25	.18*	.00002	0.71	.524	−0.01
2. Avoidance of Negative Stimuli[a]	.009	0.20	.17**	.001	0.45	.215	0.08
Retroactive priming							
3. Retroactive Priming I[a]	.007	0.26	−.05	.148	0.17	.036	0.24
4. Retroactive Priming II[a]	.014	0.23	−.07	.059	0.27	.035	0.23
Retroactive habituation							
5. Retroactive Habituation I							
Negative trials[b]	.014	0.22					
6. Retroactive Habituation II							
Negative trials[b]	.037	0.15					
Erotic trials	.039	0.14	0.24***	.002	0.57	.219	−0.09
7. Retroactive Induction of Boredom[a]	.096	0.09	.16**	.018	0.22	.483	0.00
Retroactive facilitation of recall							
8. Facilitation of Recall I	.029	0.19	.22**	.0003	0.57	.525	−0.08
9. Facilitation of Recall II	.002	0.42	−.10	.049	0.44	.013	0.40
Mean effect size (d)		0.22			0.43		0.10

[a] Probabilities and effect sizes in this row are based on the mean of the t values across the variations of the data analysis. [b] The Stimulus-Seeking Scale was not administered in this experiment.

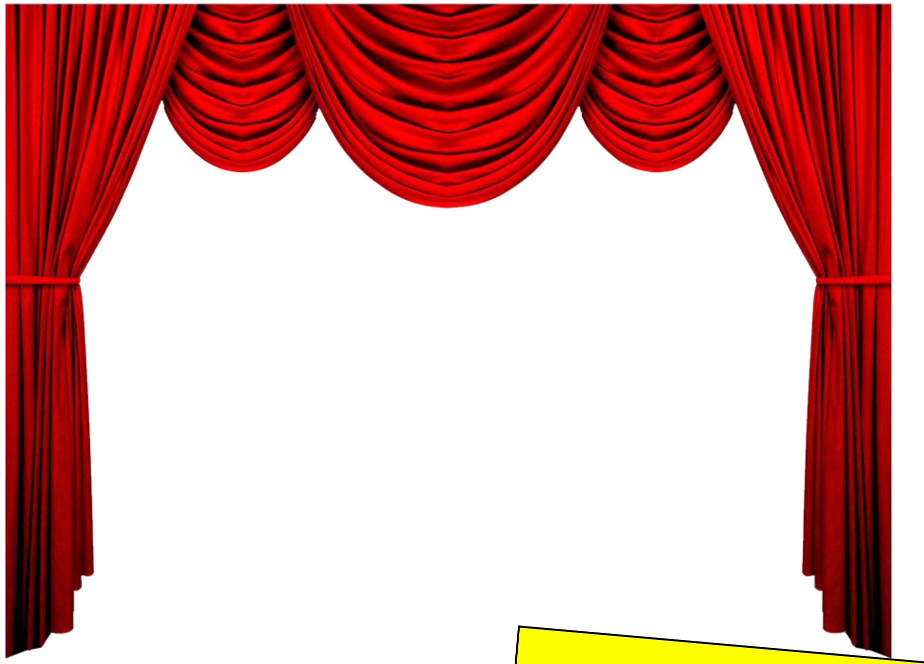
$^{*}p < .05$. $^{**}p < .02$. $^{***}p < .01$.

From Bem (2011)

Study 1

Photo source: http://pngimg.com/download/32324

Study 1



Result: 53.1%

[z = 2.30, *p*=.011]

Photo source: https://en.wikipedia.org/wiki/Lesbian_erotica

Curtain picture source: http://pngimg.com/download/32292

Study 1

- Material
    - Erotic pictures
    - Positive pictures
    - Romantic pictures
    - Negative pictures

- High-low impact versions

Study 1

- There are 8 tests possible!
  - 4 (kind of pictures) x 2 (high/low impact)

- H_0: no “feeling the future”
  - With regular $\alpha = 0.05$, probability of rejecting H_0 *while it is true* = .05
    - Type I error

- With regular $\alpha = 0.05$, probability of rejecting H_0 *while it is true* for at least one of 8 tests is:
  - $1 - 0.95^8 = 0.34$ *(you don’t have to understand this calculation)*

= capitalization on chance

Bonferroni method

$$\alpha^* = \frac{\alpha}{\text{Number of tests}}$$

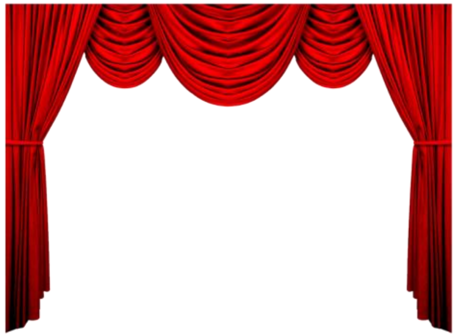
α^*: Significance level that you need to use per test

α: Significance level that you actually wanted
(= Type I error), mostly 0.05.

Photo source: https://en.wikipedia.org/wiki/Carlo_Emilio_Bonferroni

Bonferroni method

- E.g.,: The paper by Bem (2011):
  - 8 kinds of pictures = 8 tests!
  - Bem found for the high-erotic pictures: p = .01



- Bonferroni correction:

$$\boxed{\alpha^* = \frac{\alpha}{\text{Number of tests}} = \frac{0.05}{8} = 0.006}$$

Curtain picture source: http://pngimg.com/download/32292

Photo source: https://en.wikipedia.org/wiki/Lesbian_erotica

Multiple comparison

Danger of doing multiple statistical tests

If H_0 is true, there is still a 0.05 probability to reject H_0 **on each test you do**

In ANOVA, there may be many groups under consideration:

- E.g., 10 groups
  - g(g-1)/2 = (10*9)/2 = 90/2 = 45 possible comparisons
    - 1 versus 2, 1 versus 3, …, 2 versus 3, 2 versus 4, …., 9 versus 10 etc.
  - of those comparisons 45*.05 = 2.25 are expected to be significant if H_0 is true

Multiple comparison

- One solution is the “Bonferroni method”

→If you plan 5 tests, use an error probability of 0.05 / 5 = 0.01 and make 99% confidence intervals

  - Overall confidence level will be approx. 95%

- Another solution is the “Tukey Method”

→ complex, no formula

Summary

- When comparing more than two groups (categorical independent variable) on a qauntitative dependent variable, one can do an ANOVA test.
- This is a two-step procedure:
  - First one does an ANOVA test (F-test) to test whether there is any difference between groups
  - Second, if the ANOVA test is significant (in favor of the alternative hypothesis that there is at least one difference), one can use Fisher method confidence intervals to test which groups differ.
- The reasoning behind the F-test is that one compares the variability between people to the variability within people, and in case the variability between people is much larger than within, this suggests group differences.
- The reason for first doing one F-test is to protect the testing procedure for capitalization on chance: only in step 2 one will do "multiple comparisons"
- Another protection against capitalization on chance is that one can choose to do a correction on alpha, such as the Bonferroni correction or Tukey method (the Tukey method is not part of the exam material).

Example exam question 1

Matteo did an experiment to test whether there is a difference in what color increases appetite. He creates four conditions: red, green, yellow and blue, and assigns a random sample of 10 people to each of the conditions. He summarizes his findings in an ANOVA table. What value should be filled in at the question mark?

Source	DF	SS	MS	F	P
Group		126	42	**?**	
Error		255.6	7.1		
Total					

a) 5.92
b) 0.49
c) 0.169

Example exam question 1

Matteo did an experiment to test whether there is a difference in what color increases appetite. He creates four conditions: red, green, yellow and blue, and assigns a random sample of 10 people to each of the conditions. He summarizes his findings in an ANOVA table. What value should be filled in at the question mark?

Source	DF	SS	MS	F	P
Group		126	42	**?**	
Error		255.6	7.1		
Total					

a) **5.92 (42/7.1)**
b) 0.49
c) 0.169

Example exam question 2

- What is the between-groups variance?

Group 1	Group 2	Group 3
26	38	28
28	38	28
34	34	36
38	32	30
24	28	38
$\bar{y}_1 = 30$	$\bar{y}_2 = 34$	$\bar{y}_3 = 32$
$s^2 = 34$	$s^2 = 18$	$s^2 = 22$

A) 4

B) 20

C) 24.67

Example exam question 2

- What is the between-groups variance?

Group 1	Group 2	Group3
26	38	28
28	38	28
34	34	36
38	32	30
24	28	38
$\bar{y}_1$= 30	$\bar{y}_2$= 34	$\bar{y}_3$= 32
s_1^2= 34	s_2^2= 18	s_3^2= 22

A)4

B) 20

C) 24.67

Practice question

Matteo did an experiment to test whether there is a difference in what color increases appetite. He creates four conditions: red, green, yellow and blue, and assigns a random sample of 10 people to each of the conditions.

Excercise: finish the ANOVA table

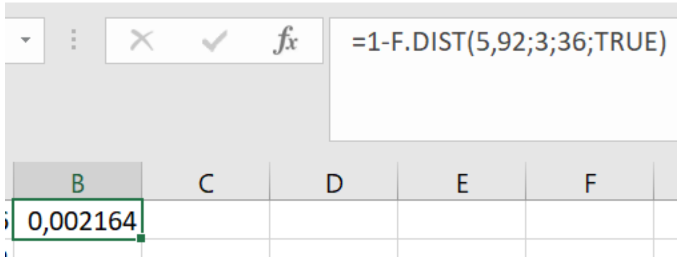
Source	**DF**	**SS**	**MS**	**F**	**P**
Group	**?**	126	42	**?**	**?**
Error	**?**	255.6	7.1		
Total	**?**				

Practice question

Matteo did an experiment to test whether there is a difference in what color increases appetite. He creates four conditions: red, green, yellow and blue, and assigns a random sample of 10 people to each of the conditions.

Excercise: finish the ANOVA table

Source	**DF**	**SS**	**MS**	**F**	**P**
Group	**3**	126	42	**5.92**	**0.002**
Error	**36**	255.6	7.1		
Total	**39**				



Excercise 14.2

14.2 Satisfaction with banking A bank conducts a survey in which it randomly samples 400 of its customers. The survey asks the customers which way they use the bank the most: (1) interacting with a teller at the bank, (2) using ATMs, or (3) using the bank's online banking service. It also asks their level of satisfaction with the service they most often use (on a scale of 0 to 10 with 0 = very poor and 10 = excellent). Does mean satisfaction differ according to how they most use the bank?

TRY

a. Identifying notation, state the null and alternative hypotheses for conducting an ANOVA with data from the survey.

b. Report the *df* values for this ANOVA. Above what F test statistic values give a P-value below 0.05?

c. For the data, $F = 0.46$ and the P-value equals 0.63. What can you conclude?

d. What were the assumptions on which the ANOVA was based? Which assumption is the most important?

Excercise 14.2

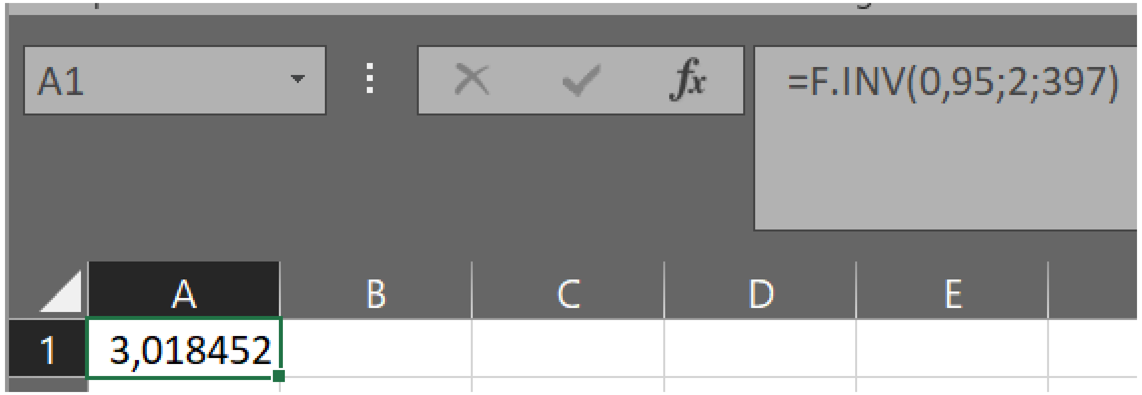
a) $H_0: \mu_1 = \mu_2 = \mu_3$

$H_a: \mu_1 \neq \mu_2$ or $\mu_1 \neq \mu_3$ or $\mu_2 \neq \mu$_3 (at least two population means are different)

b) df1 = 3-1=2

df2= 400-3

Critical f statistic = F.INV(0,95; df1; df2) = 3.018



c) The null hypothesis cannot be rejected

d)

- quantitative response variable for more than two groups
- Independent random samples
- Normal population distribution with equal standard deviation
- First assumption is clearly true, and the second is also true since the question states that they selected a random sample. Although the ANOVA test is quite robust against violations of the third assumption, in this case it is not clear whether the sample sizes are equal across the groups and so you do want to check whether there are no strong deviations from normality, and whether the largest SD is not more than 2 x smallest SD. So this will be most important to check.

Other excercises to focus on

- 14.1, 14.3, 14.4, 14.5, 14.7, 14.8, 14.12, 14.13, 14.14

If you have questions on how to answer these excercises: use the discussion board and we will help you!