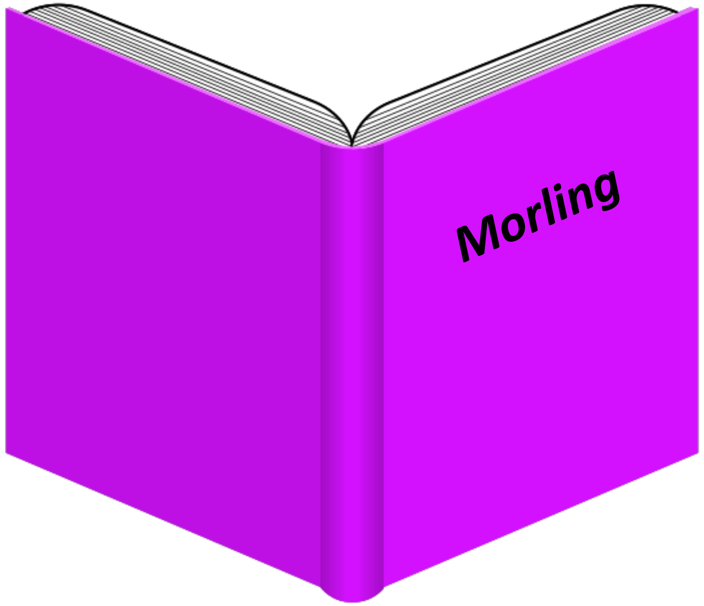


Research Methods and Statistics

Lecture 17: Comparing groups II

Riet van Bork

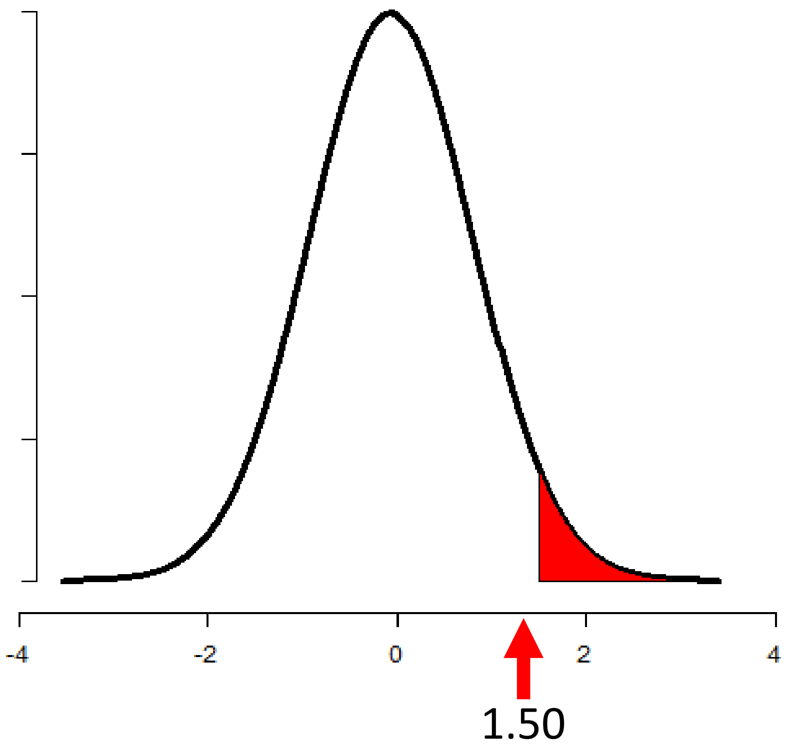
Pictures source: pixabay.org

Tests

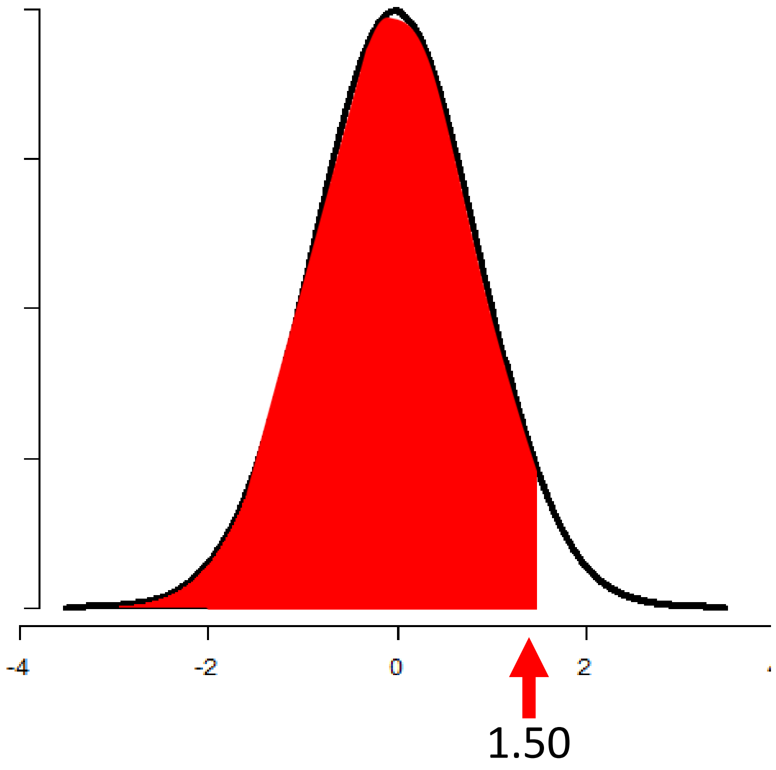
- This block we discuss many different tests. Which one to use in what case?
- Look for the relevant information:
    - Mean or proportion?
    - One sample or two samples?
    - Independent samples or dependent samples?
    - Confidence interval or significance test?
    - One-sided or two-sided test? (and in the first case, what direction?)

Suppose your test statistic is 1.50.
To compute the p-value first check your alternative hypothesis!

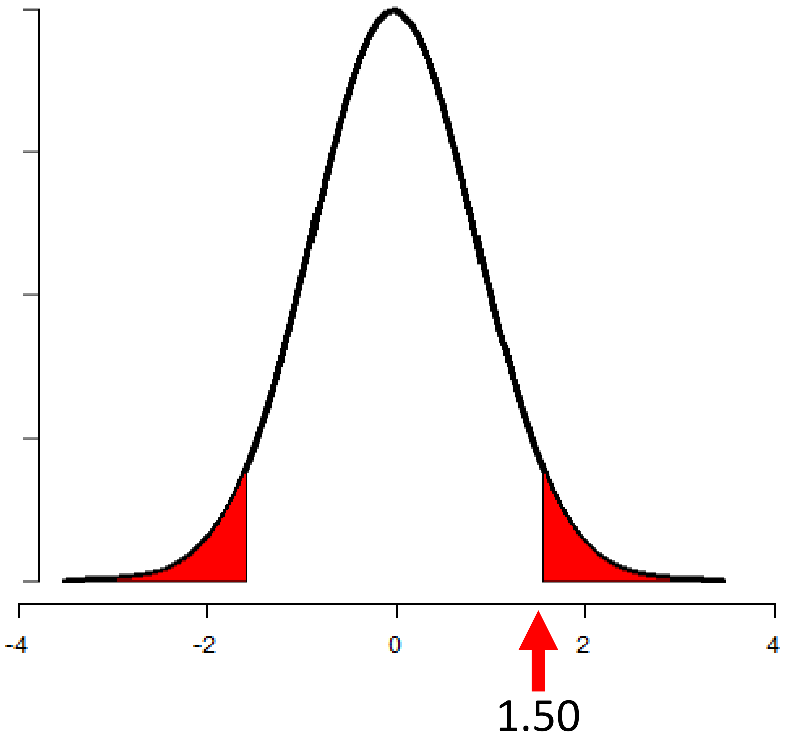
H_a: $p > p_0$ → calculate upper tail



H_a : $p < p_0$ → calculate lower tail



H_a: $p \neq p_0$ → calculate smallest tail and double it!



Today

- Independent vs dependent samples
- Comparing two proportions in independent samples
- Comparing two means in dependent samples
- Comparing two proportions in dependent samples

Today

- Independent vs dependent samples
- Comparing two proportions in independent samples
- Comparing two means in dependent samples
- Comparing two proportions in dependent samples

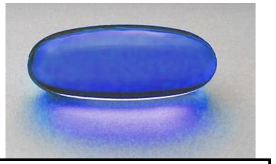
Example

Does caffeine increase concentration?

Coffee picture: https://commons.wikimedia.org/wiki/File:A_small_cup_of_coffee.JPG

Meditation photo: https://pixabay.com/nl/illustrations/bewustzijn-mindfulness-meditatie-4246466/

Does caffeine increase concentration?



versus

No caffeine
(placebo)

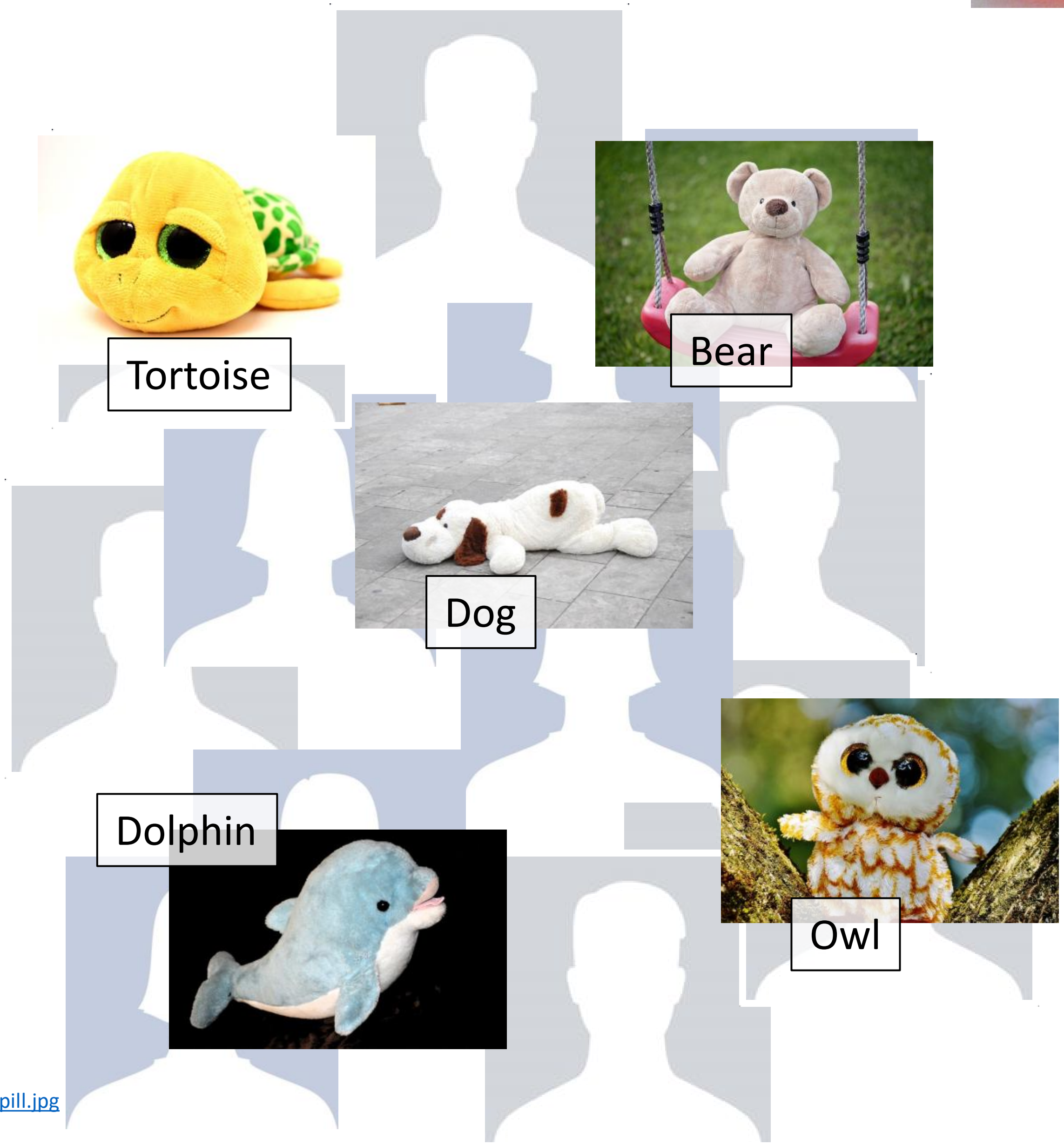
Caffeine

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Samples

- One or two samples
- Dependent or independent

One sample

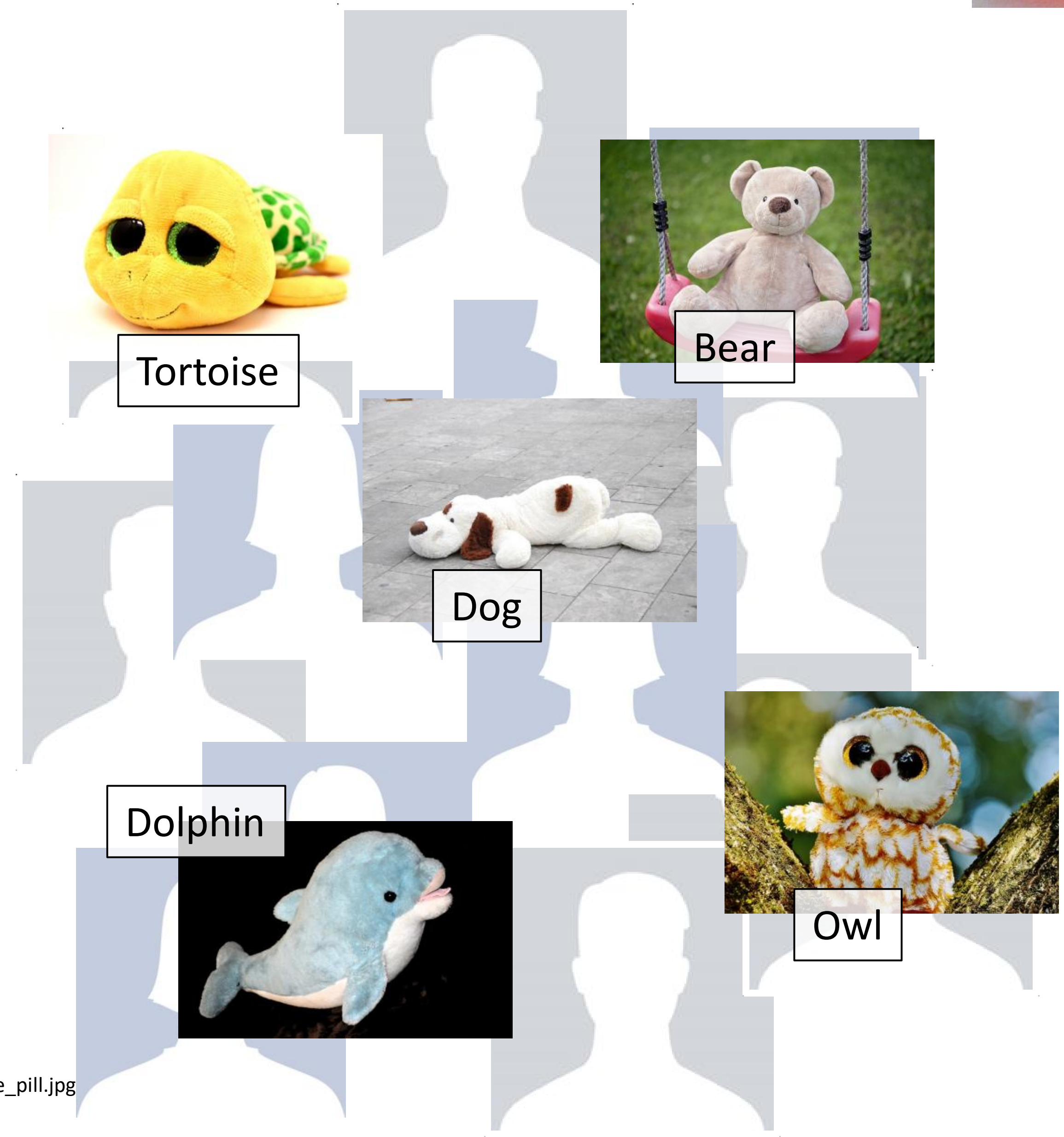


For example: Researcher John knows that in the population of stuffed animals the mean score on a concentration test is 10. He draws a sample of 25 animals and gives them a caffeine pill. In this sample he observes a mean concentration of 12 and an SD of 1.5. He wants to know whether animals with a caffeine pill score significantly different from 10.

compare performance of caffeine group to some specific value

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg
Other pictures: pixabay.com

One sample



For example: Researcher John knows that in the population of stuffed animals the **mean** score on a concentration test is 10. He draws a sample of 25 animals and gives them a caffeine pill. In this sample he observes a mean concentration of 12 and an SD of 1.5. He wants to know whether animals with a caffeine pill score significantly **different** from **10**.

$$H_0: \mu = 10$$
$$H_a: \mu \neq 10$$

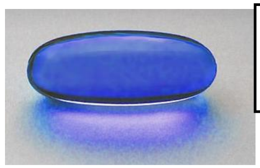
See chapter 9!

compare performance of caffeine group to some specific value

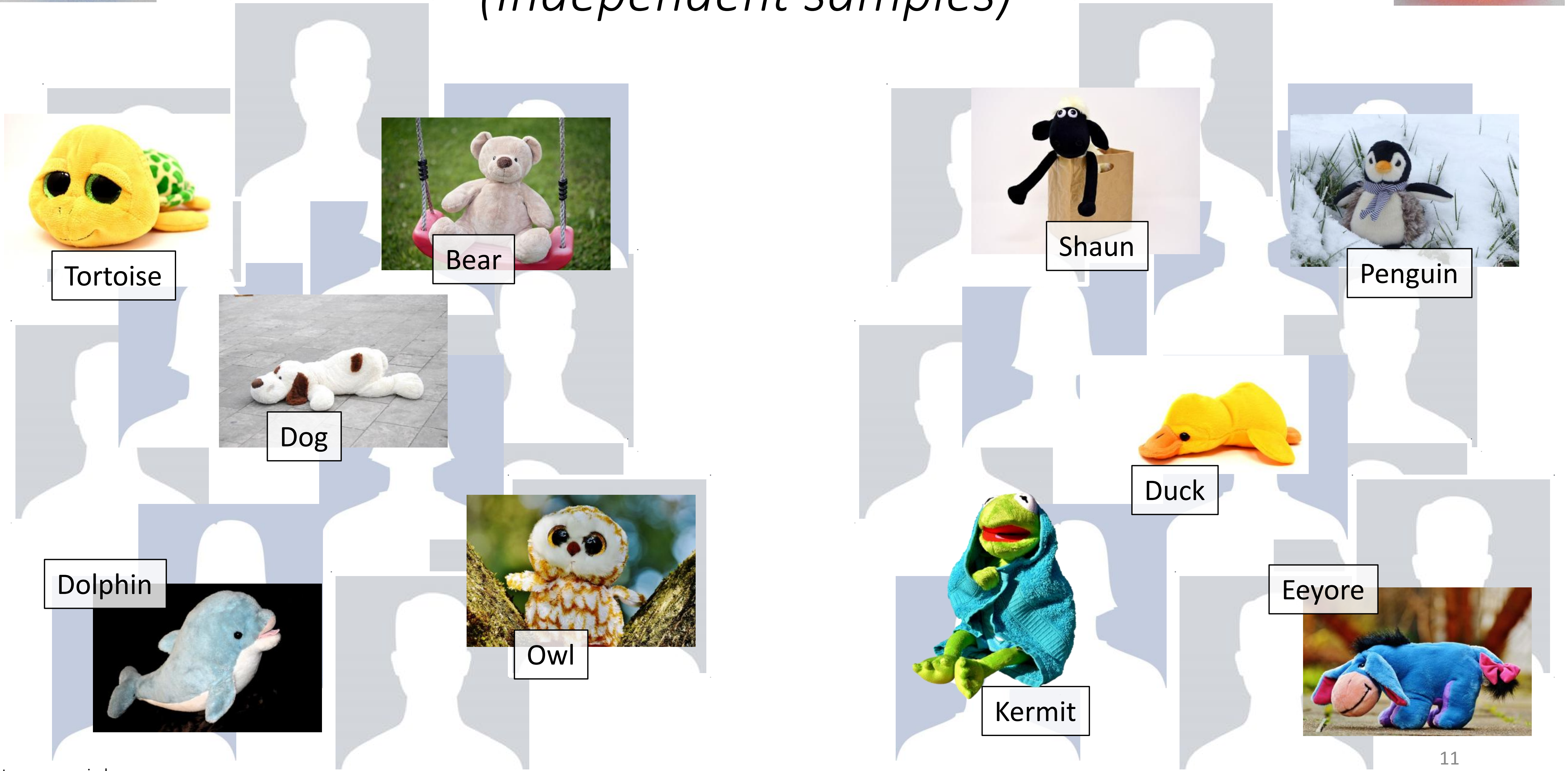
Photo Silhouette: facebook.com for non-commercial use only

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Picture source: pixabay.com



Between-groups design
(Independent samples)

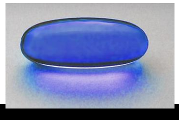


Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Photo Silhouette: facebook.com for non-commercial use only

Matched pairs design: dependent samples





Pair #1


Pair #2


Pair #3

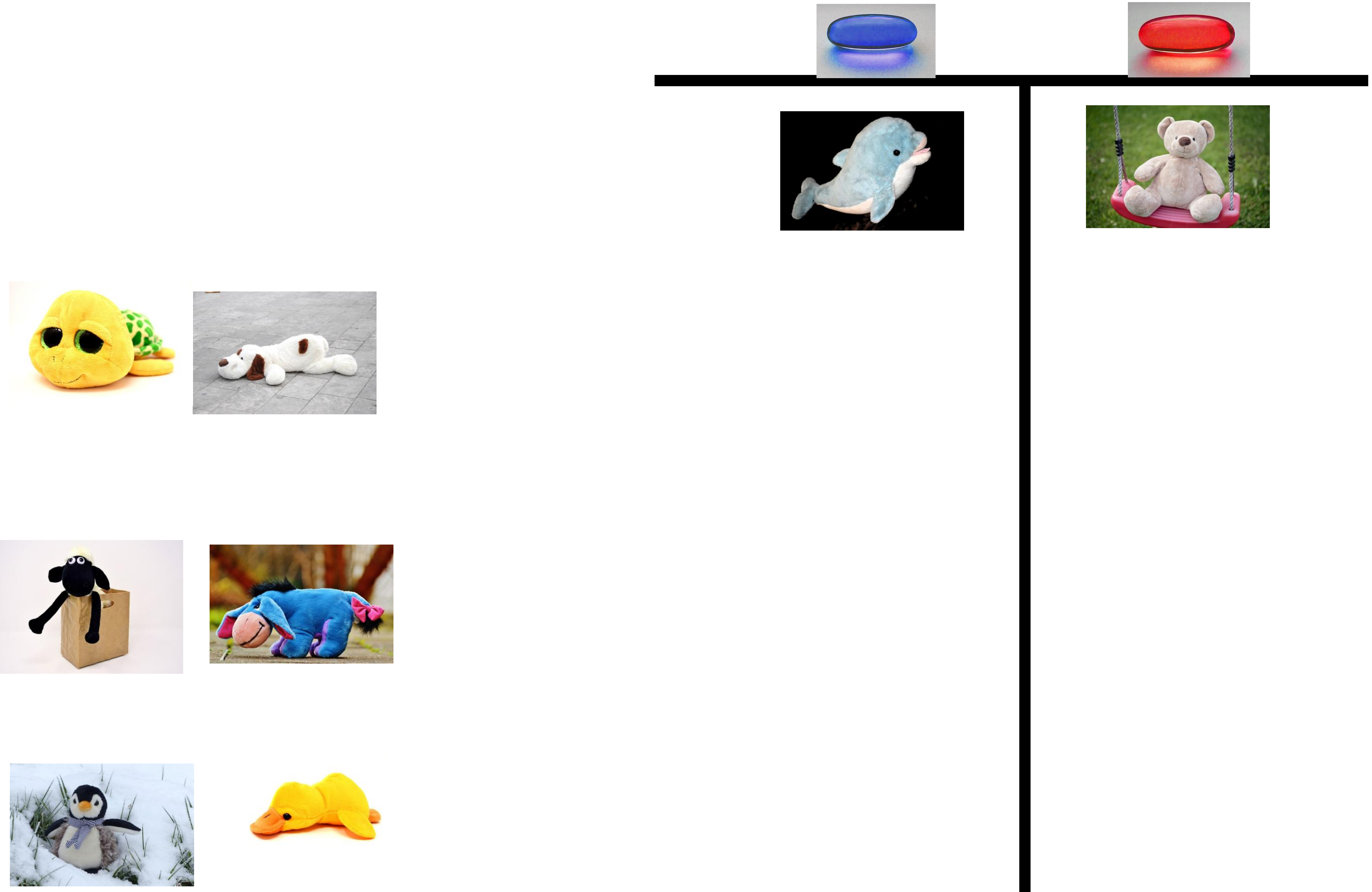

Pair #4

Make pairs on basis of an important background variable (e.g., age)

Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

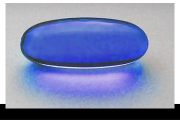
Matched pairs design: dependent samples



Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

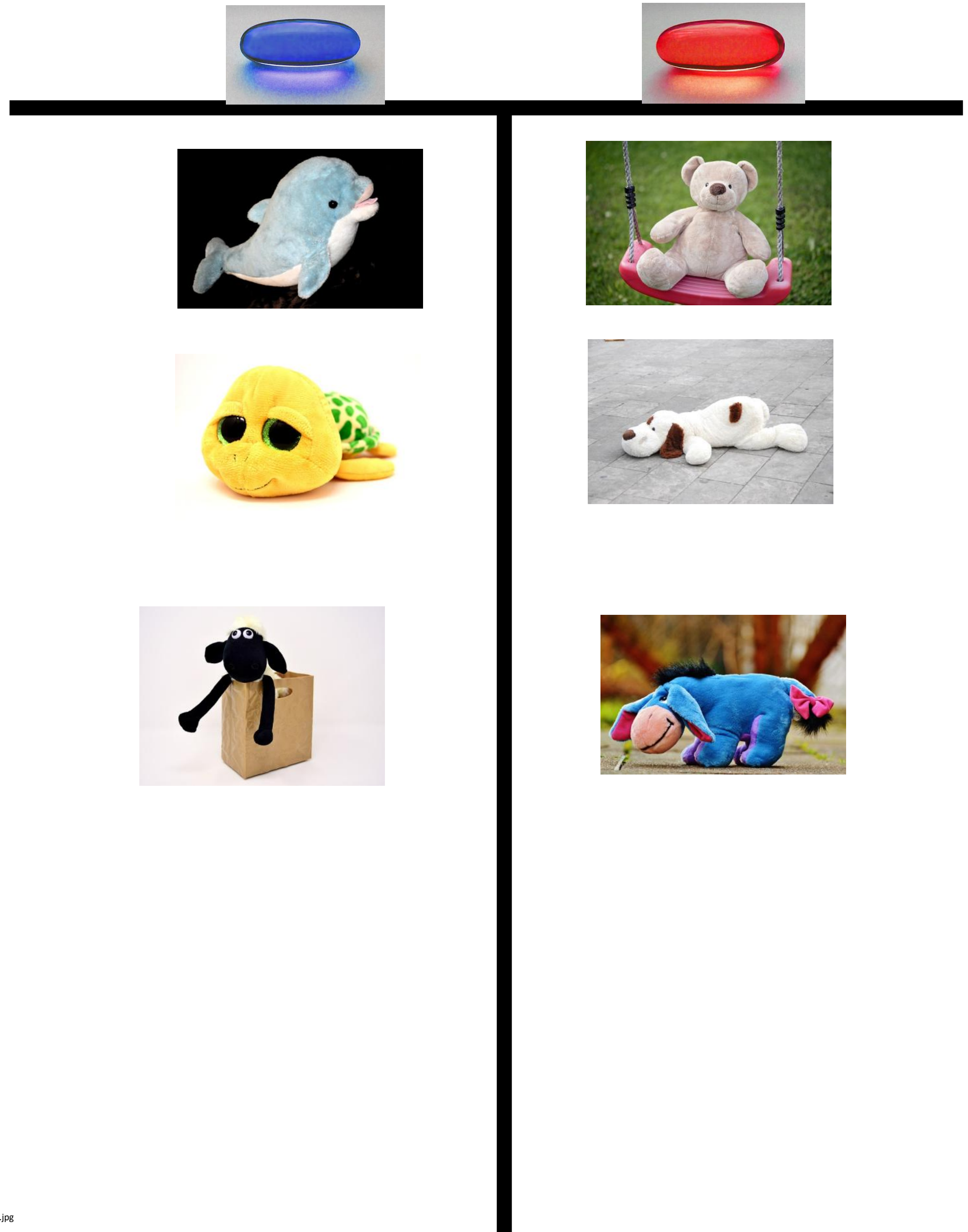
Matched pairs design: dependent samples



Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

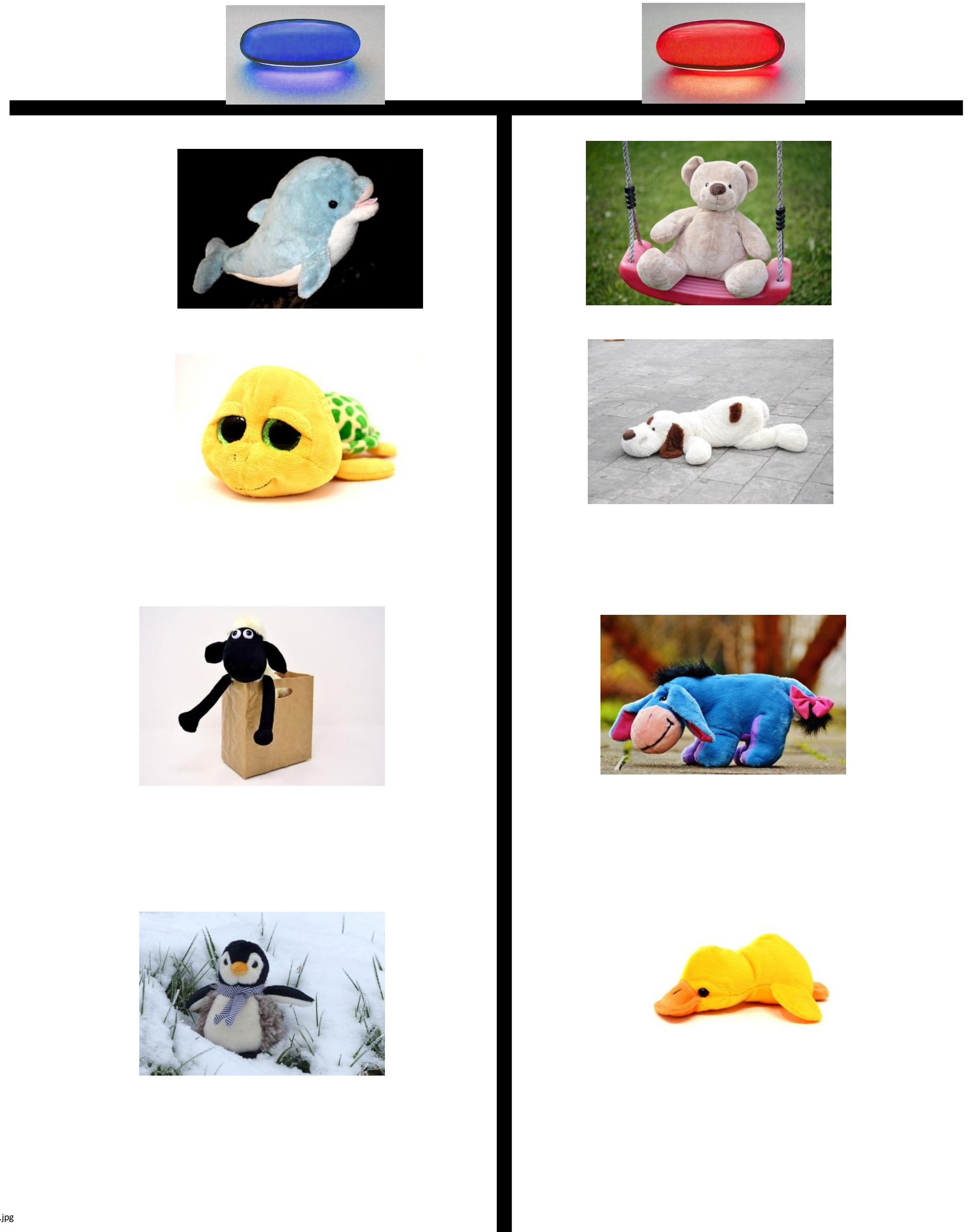
Matched pairs design: dependent samples



Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Matched pairs design: dependent samples



Matched on
background variable
Here: age

Picture source: pixabay.com

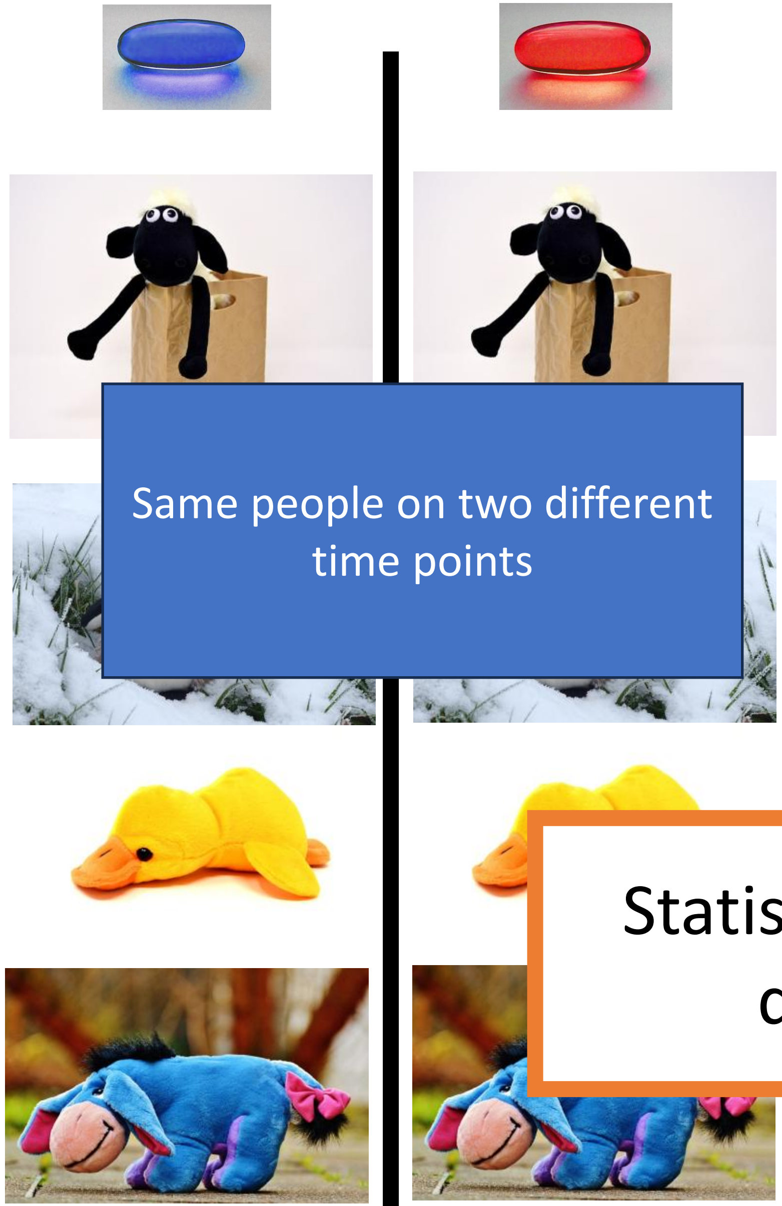
Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Within-groups design: Dependent samples
(specific form of matched pairs)

Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Two types of Dependent samples



Same people on two different time points

Different people but they form matched pairs, for example because they are matched on a background variable

Statistically you treat these designs the same!

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Picture source: pixabay.com

Parameter

- Proportion (p)

- Mean (μ)

Parameter	Sample(s)	CI or Sign test	Chapter
Proportion (categorical)	One sample	CI	Ch8
		Signif Test	Ch9
	Two independent samples	CI	Ch10b
		Signif Test	Ch10b
	Two dependent samples	CI	X
		Signif Test	Ch10b
Mean (quantitative)	One sample	CI	Ch8
		Signif Test	Ch9
	Two independent samples	CI	Ch10a
		Signif Test	Ch10a
	Two dependent samples	CI	Ch10b
		Signif Test	Ch10b

Green	Interim Exam II
Orange	Last week lecture
Yellow	Today

Parameter	Sample(s)	Statistic	Chapter
Proportion (categorical)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	CI	Ch10b ←
		Signif Test	Ch10b
	Two dependent samples	~~CI~~	~~X~~
		Signif Test	Ch10b
Mean (quantitative)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	~~CI~~	~~Ch10a~~
		~~Signif Test~~	~~Ch10a~~
	Two dependent samples	CI	Ch10b
		Signif Test	Ch10b

Green	Interim Exam II
Orange	Last week lectures
Yellow	Today

Today

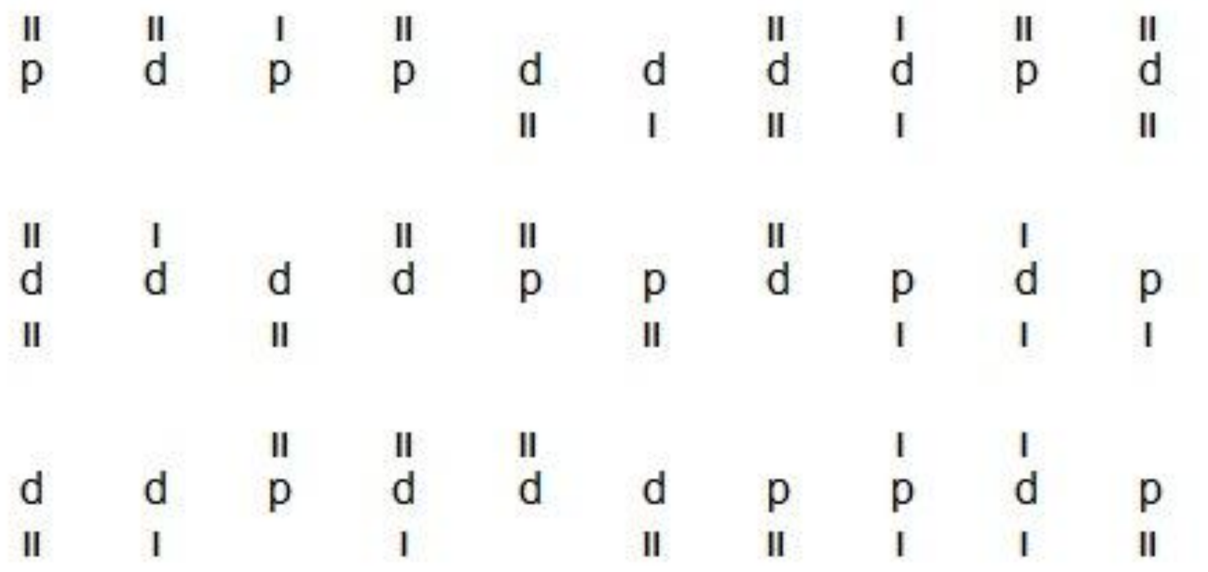
- Independent vs dependent samples
- Comparing two proportions in independent samples
- Comparing two means in dependent samples
- Comparing two proportions in dependent samples

Ada wants to know whether caffeine influences concentration. To study this she randomly assigns people to two conditions: an experimental condition in which people get a caffeine pill, and a control condition in which people get a placebo pill. She then has both groups participate in a concentration test. She uses the results to create a 95% confidence interval to assess whether the proportion of people who pass the concentration test differs between conditions.

Ada wants to know whether caffeine influences concentration. To study this she randomly **assigns people to either one of two conditions**: an experimental condition in which people get a caffeine pill, and a control condition in which people get a placebo pill. She then has both groups participate in a concentration test. She uses the results to create a **95% confidence interval** to assess whether the **proportion** of people who pass the concentration test **differs** between conditions.

Concentration measure

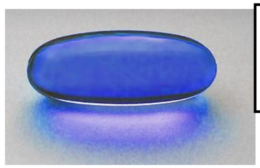
“cross out any letter "d" with two marks around, above it or below it, in any order”



See e.g., https://en.wikipedia.org/wiki/D2_Test_of_Attention

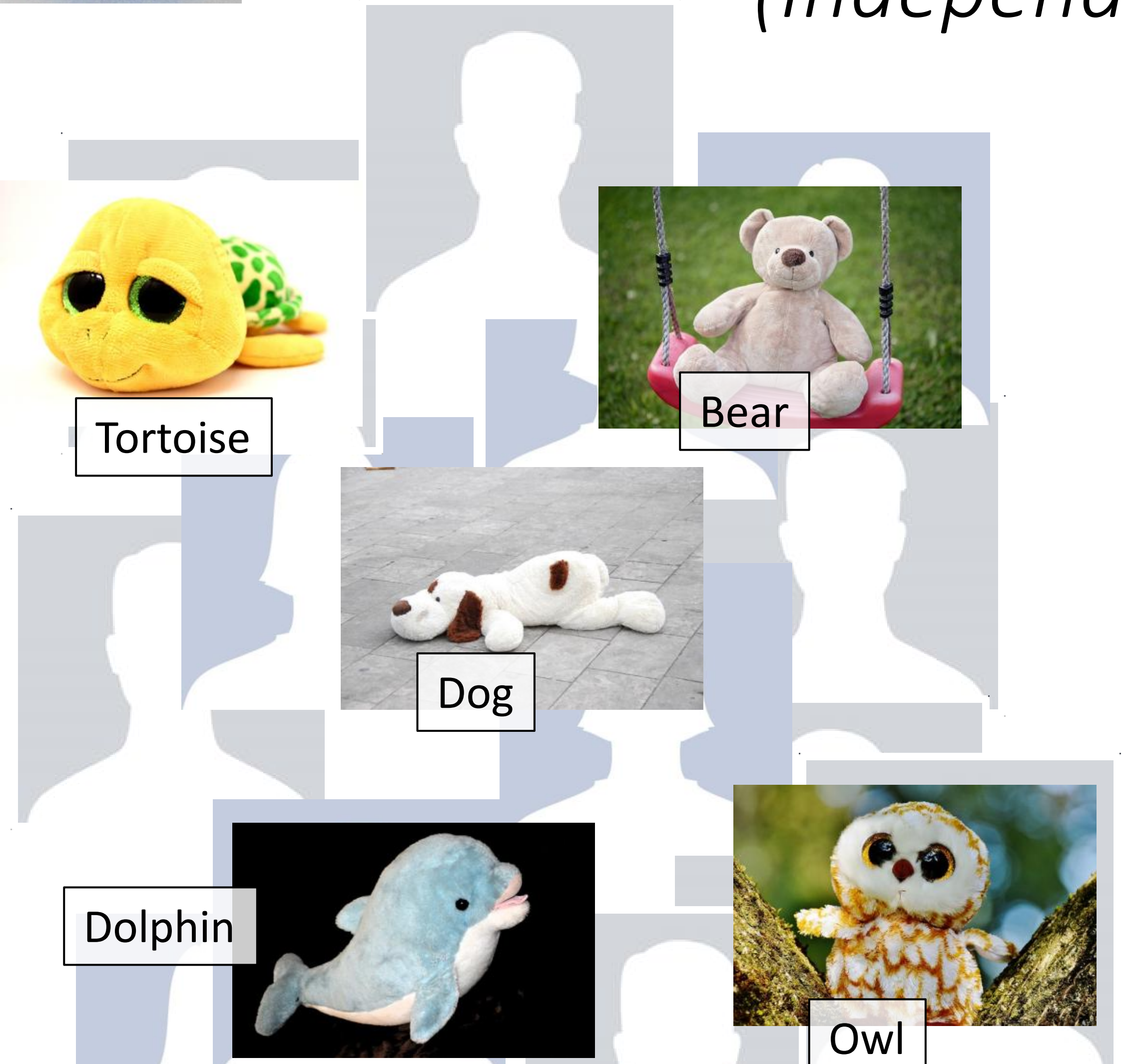
Between-groups design

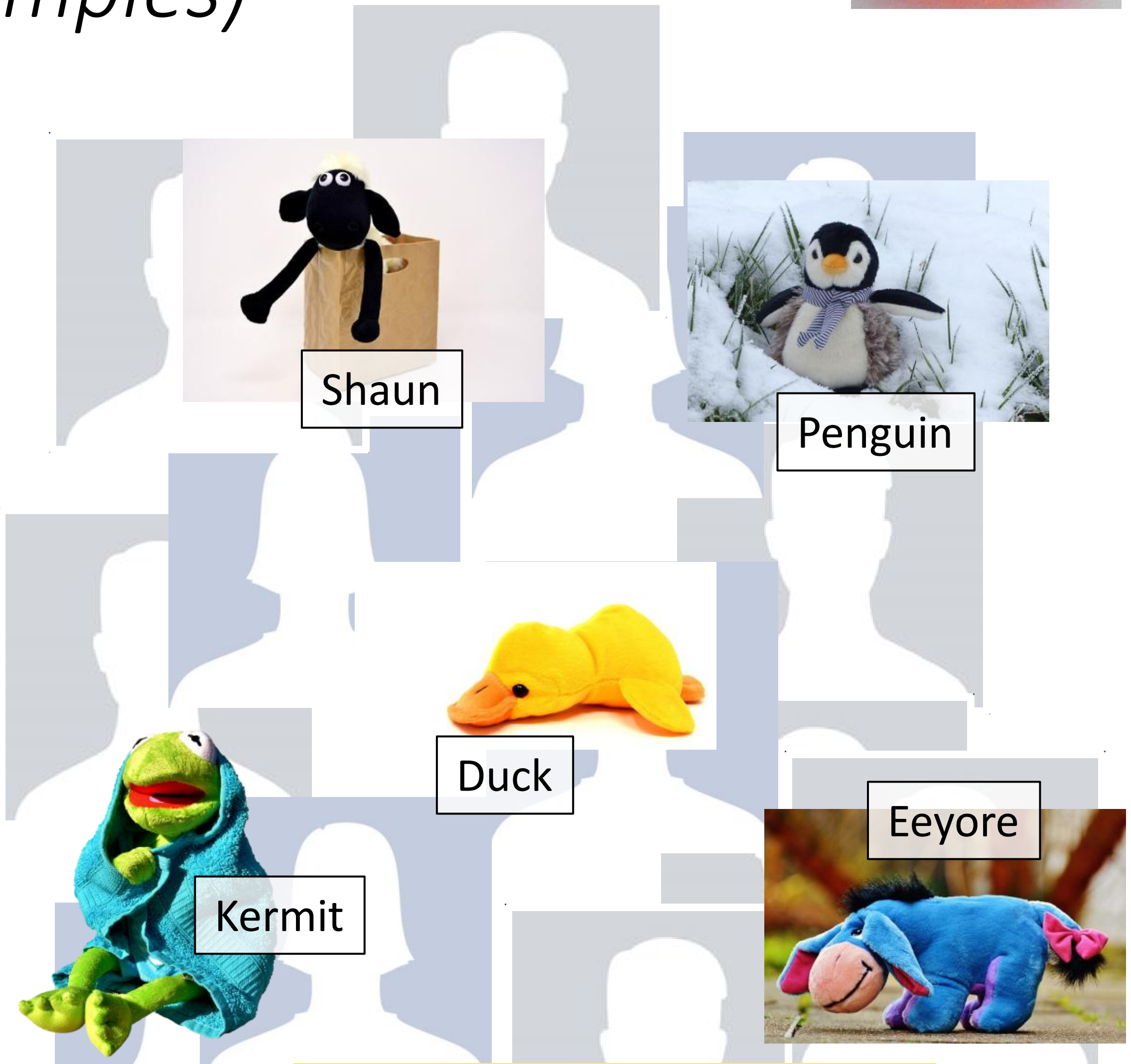
(Independent samples)



No caffeine (placebo)

Caffeine





50 passed; 18 failed

$n_1 = 68$

44 passed; 10 failed

$n_2 = 54$

Photo source pills: https://commons.wikime

Picture source: pixabay.com

Comparing two proportions: Confidence Interval

Confidence interval: Interval containing the most believable values for a parameter

- Are the proportions different in the population?
  - I.e., $p_1 = p_2$?

  → $p_1 - p_2 = 0$

- So we want to see whether $p_1 - p_2$ = 0 is a believable value or not
  - Is 0 included in the confidence interval of $p_1 - p_2$?
    - → yes? then the two proportions do not differ significantly
    - → no? then the two proportions differ significantly

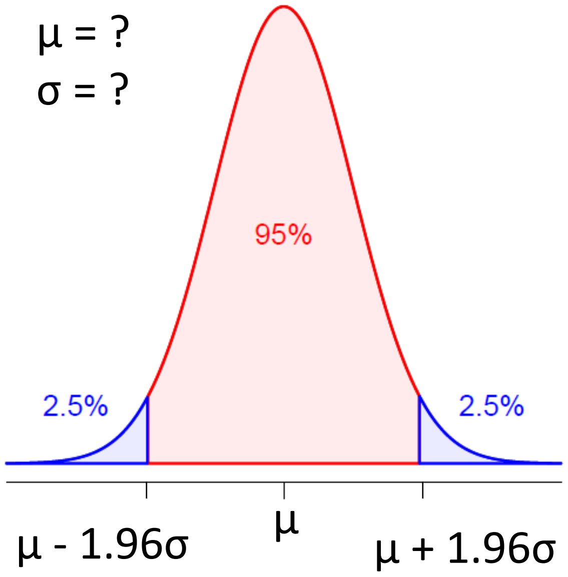
- Thus, we need the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Sampling distribution of $\hat{p}_1 - \hat{p}_2$

Is a normal distribution if $n_1 p_1 \geq$ **10** and $n_1(1-p_1) \geq$ **10** in group 1 and if $n_2 p_2 \geq$ **10** and $n_2(1-p_2) \geq$ **10** in group 2

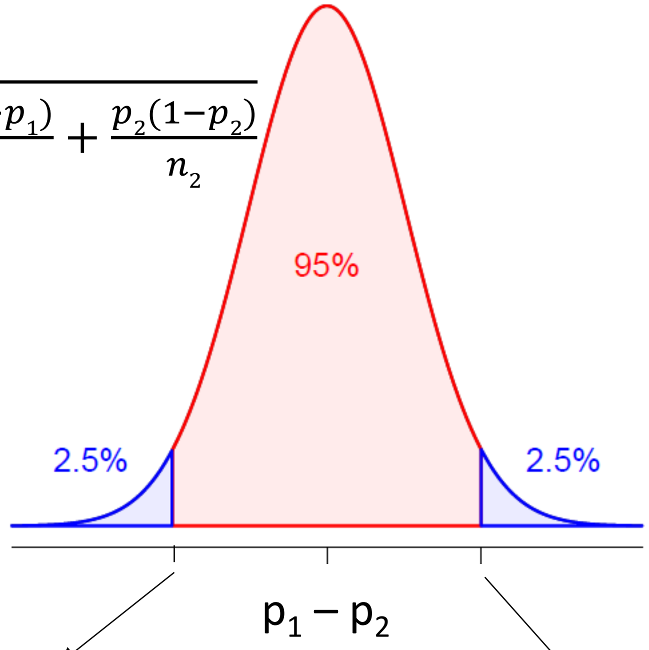
Smaller than for one proportion

μ = ?
σ = ?



$\mu = p_1 - p_2$

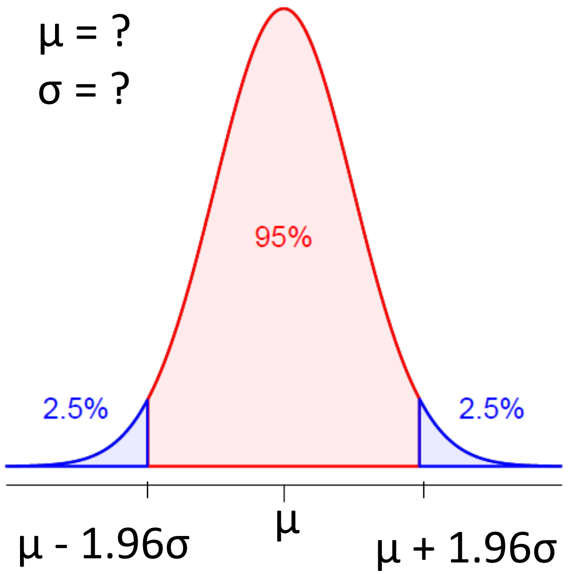
$$\sigma = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$



$$(p_1 - p_2) - 1.96\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

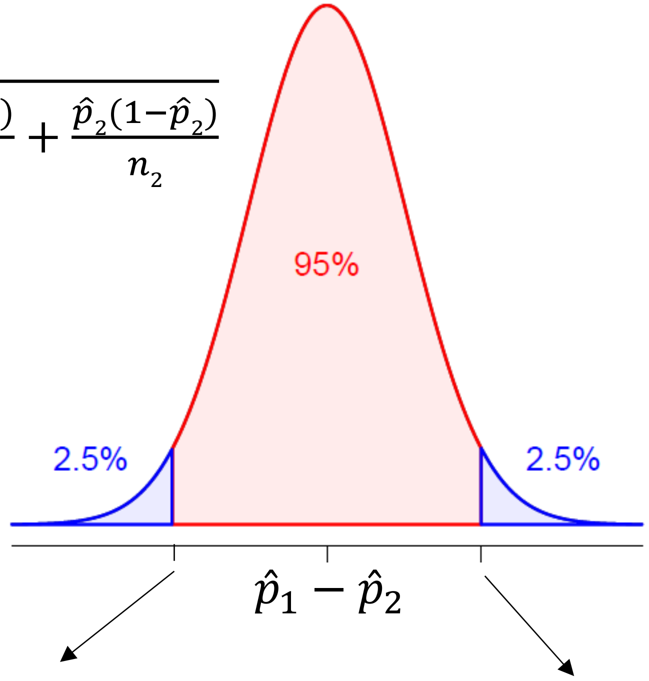
$$(p_1 - p_2) + 1.96\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Sampling distribution of $\hat{p}_1 - \hat{p}_2$



$\hat{\mu} = \hat{p}_1 - \hat{p}_2$

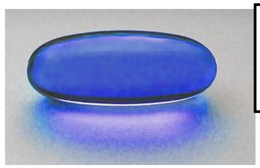
$\hat{\sigma}$ = se = $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$



$(\hat{p}_1 - \hat{p}_2) - 1.96\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ $(\hat{p}_1 - \hat{p}_2) + 1.96\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

general forumula for confidence interval of two proportion (the one on the formula sheet)

$\hat{p}_1 - \hat{p}_2 \pm$ z(se) with se= $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$



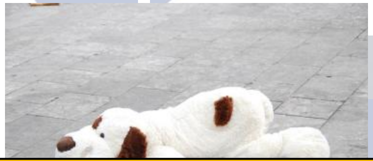
No caffeine (placebo)

Between-groups design
(Independent samples)

Caffeine

Tortoise

Bear



Shaun

Penguin

$$\hat{p}_1 = \frac{total\ passed\ group\ 1}{n_1} = \frac{50}{68} = 0.74$$

$$\hat{p}_2 = \frac{total\ passed\ group\ 2}{n_2} = \frac{44}{54} = 0.81$$

Dolphin

Owl

Kermit

Eeyore

50 passed; 18 failed

$n_1 = 68$

44 passed; 10 failed

$n_2 = 54$

Photo source pills: https://commons.wikime

Picture source: pixabay.com

Comparing two proportions: Confidence Interval

$\hat{p}_1 - \hat{p}_2 \pm$ z(se) with se= $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

In our case:

- $\hat{p}_1 = \frac{50}{68}$ =0.74, n_1 = 68
- $\hat{p}_2 = \frac{44}{54}$ =0.81, n_2 = 54

Thus:

- $\hat{p}_1 - \hat{p}_2 = -0.07$
- z = 1.96 (95% confidence)
- se= $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{0.74(1-0.74)}{68} + \frac{0.81(1-0.81)}{54}} = 0.075$

Filling in in the formula:

- upper bound: $\hat{p}_1 - \hat{p}_2$+ z(se) = -0.07 + 1.96*0.075 = 0.077
- lower bound: $\hat{p}_1 - \hat{p}_2$ – z(se) = -0.07 – 1.96*0.075 = -0.217

Conclusion:
95% confidence interval of
$\hat{p}_1 - \hat{p}_2$ equals (-0.217, 0.077)

As this intervals **does** contain 0, a difference of 0 (i.e., no difference) is a believable value

Alternative

Ada wants to know whether caffeine influences concentration. To study this she randomly **assigns people to either one of two conditions**: an experimental condition in which people get a caffeine pill, and a control condition in which people get a placebo pill. She then has both groups participate in a concentration test. She uses the results to do a **significance test with a significance level of 0.05** to assess whether the **proportion** of people who pass the concentration test **differs** between conditions.

Parameter	Sample(s)	Statistic	Chapter
Proportion (categorical)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	~~CI~~	~~Ch10b~~
		Signif Test	Ch10b ←
	Two dependent samples	~~CI~~	~~X~~
		Signif Test	Ch10b
Mean (quantitative)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	~~CI~~	~~Ch10a~~
		~~Signif Test~~	~~Ch10a~~
	Two dependent samples	CI	Ch10b
		Signif Test	Ch10b

(green)	Interim Exam II
(orange)	Last week lecture
(yellow)	Today

Comparing two proportions: Significance test

Is the difference we found, $\hat{p}_1 - \hat{p}_2 = -0.07$, due to sampling variability?
Or does it reflect a true difference in the population?

→ Significance test

Step 1: Assumptions:

- Variable is categorical
- Data are obtained using random sampling
- For one sided test:
  $n_1p_1 \geq 10$ and $n_1(1-p_1) \geq 10$ in group 1 and
  $n_2p_2 \geq 10$ and $n_2(1-p_2) \geq 10$ in group 2
- Two sided test is more robust against violations of that condition and so a bit more lenient (see p. 532):
  $n_1p_1 \geq 5$ and $n_1(1-p_1) \geq 5$ in group 1 and
  $n_2p_2 \geq 5$ and $n_2(1-p_2) \geq 5$ in group 2

Comparing two proportions: Significance test

Step 2: Hypothesis

- H_0: $p_1 - p_2 = 0$
- H_a: $p_1 - p_2 \neq 0$

(i.e., in this case two sided, but you can also make it one sided if e.g., the researcher expects the caffeine group to have better concentration)

Comparing two proportions: Significance test

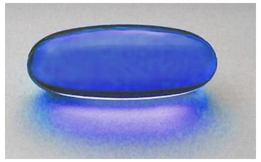
Step 3: Test statistic

$$z = \frac{estimate - value\ under\ H_0}{standard\ error\ under\ H_0} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{se_0}$$

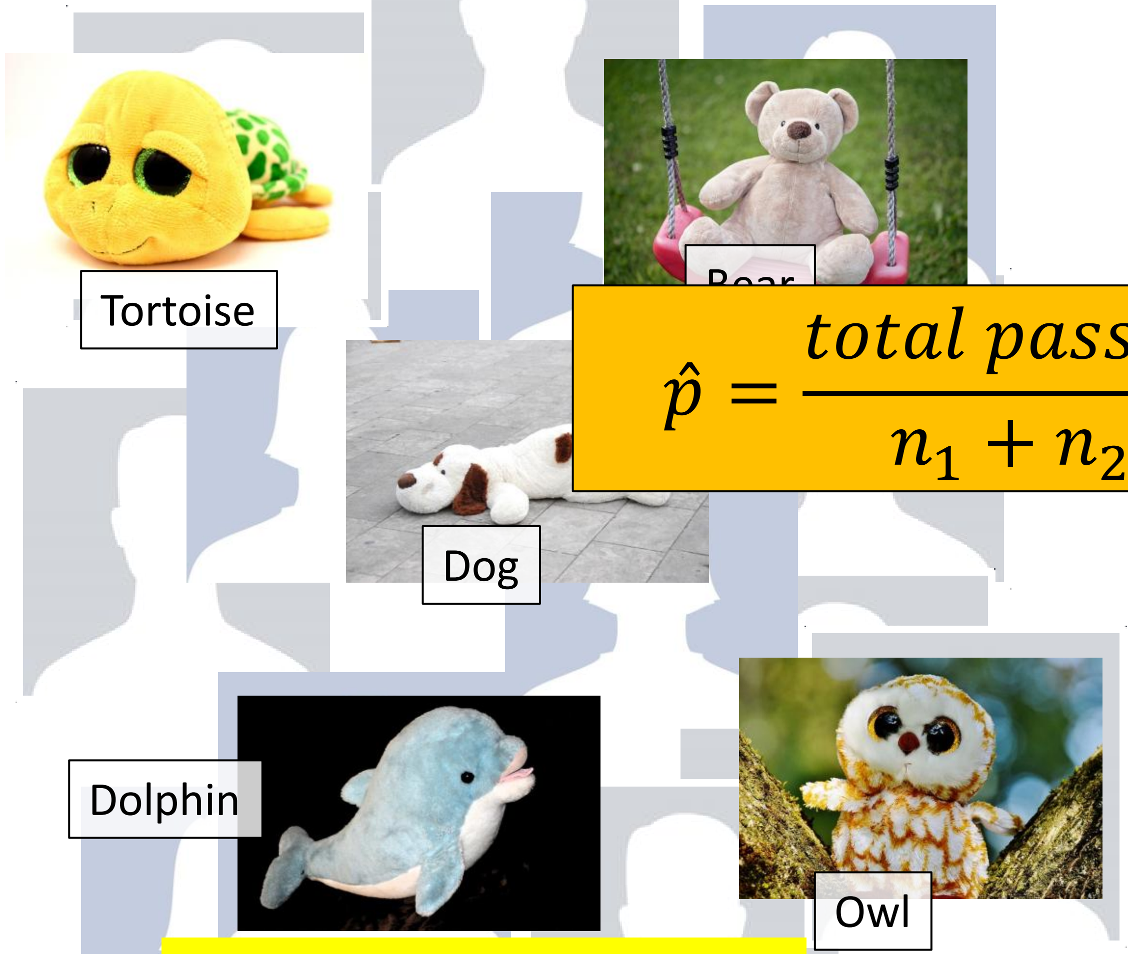
with $se_0 = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n_1} + \frac{\hat{p}\,(1-\hat{p})}{n_2}} = \sqrt{\hat{p}\,(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

$\hat{p}$ = Pooled estimate = overall proportion if you take both groups together

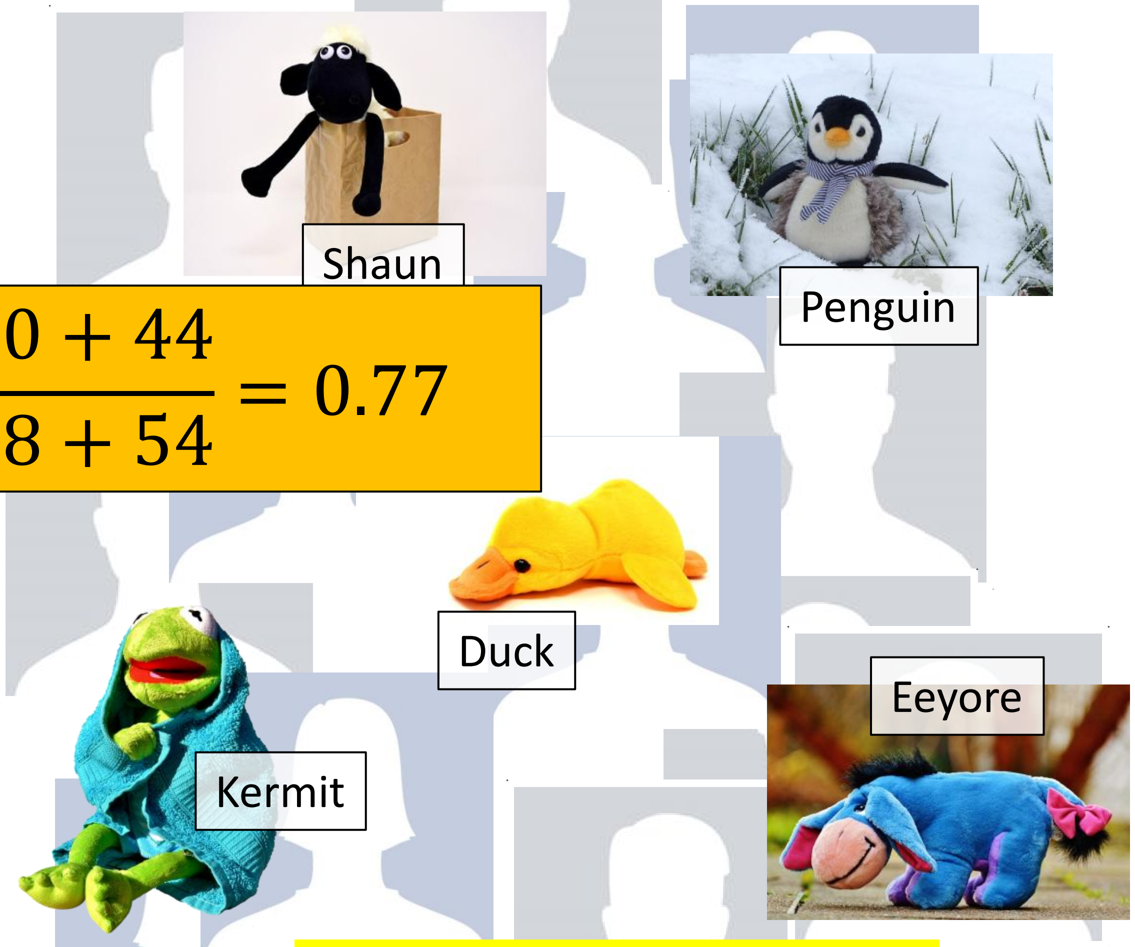
We do this because under H_0 the proportions are the same. And so now we have two estimates of that proportion (from two samples) and so we want to combine them



Between-groups design
(Independent samples)



$$\hat{p} = \frac{total\ passed}{n_1 + n_2} = \frac{50 + 44}{68 + 54} = 0.77$$



50 passed; 18 failed
$n_1 = 68$

44 passed; 10 failed
$n_2 = 54$

Photo source pills: https://commons.wikime

Picture source: pixabay.com

Comparing two proportions: Significance test

Step 3: Test statistic

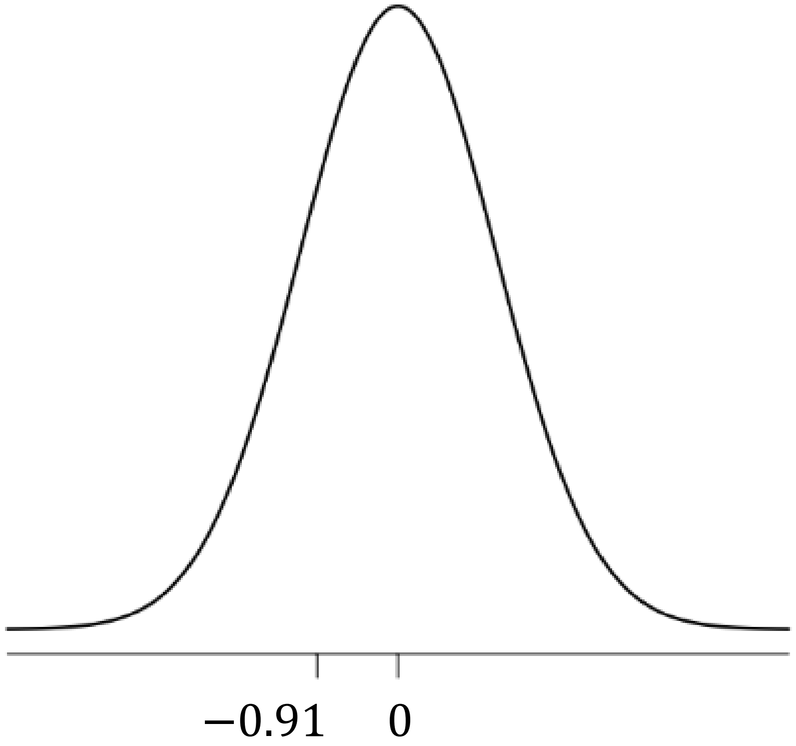
$$\mathrm{se}_0 = \sqrt{\hat{p}\,(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{0.77\,(1-0.77)\left(\frac{1}{68}+\frac{1}{54}\right)} = 0.077$$

$$z = \frac{estimate - null\ hypothesis\ value}{standard\ error} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{se_0} = \frac{-0.07}{0.077} = -0.91$$

Comparing two proportions: Significance test

Step 4: P-value

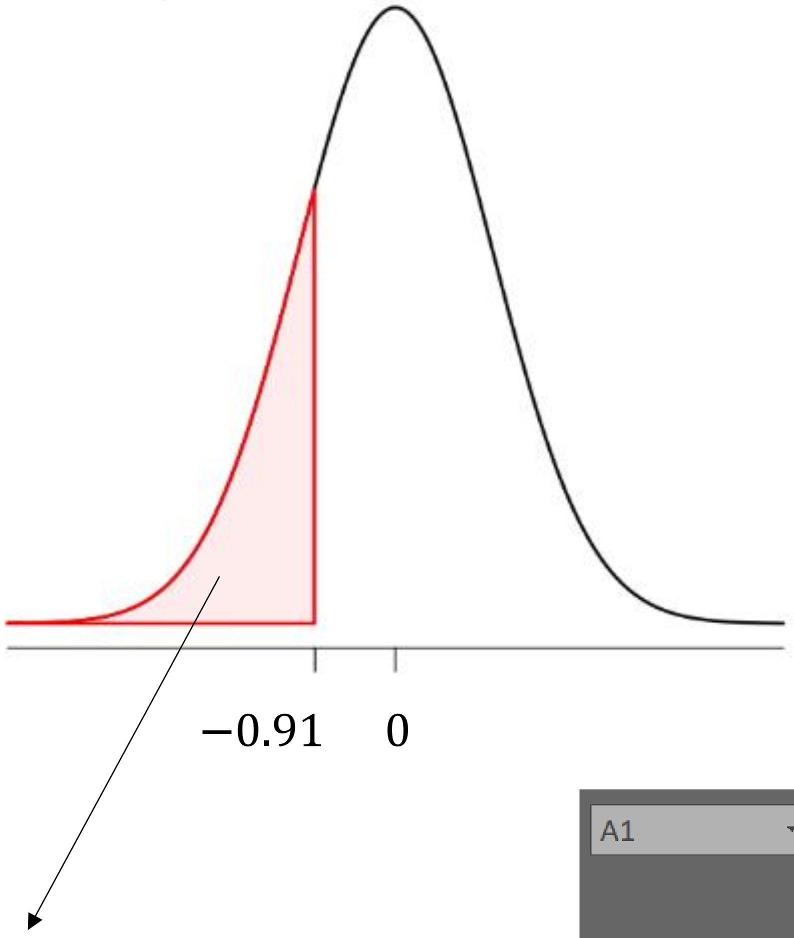
$z = -0.91$



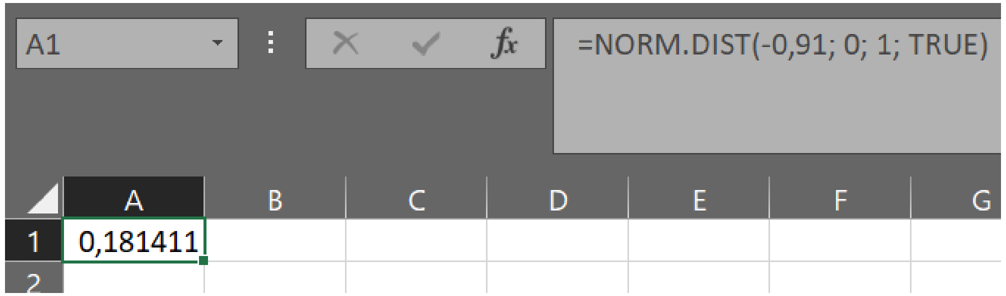
Comparing two proportions: Significance test

Step 4: P-value

$$z = -0.91$$



P(z < - 0.91) = 0.18



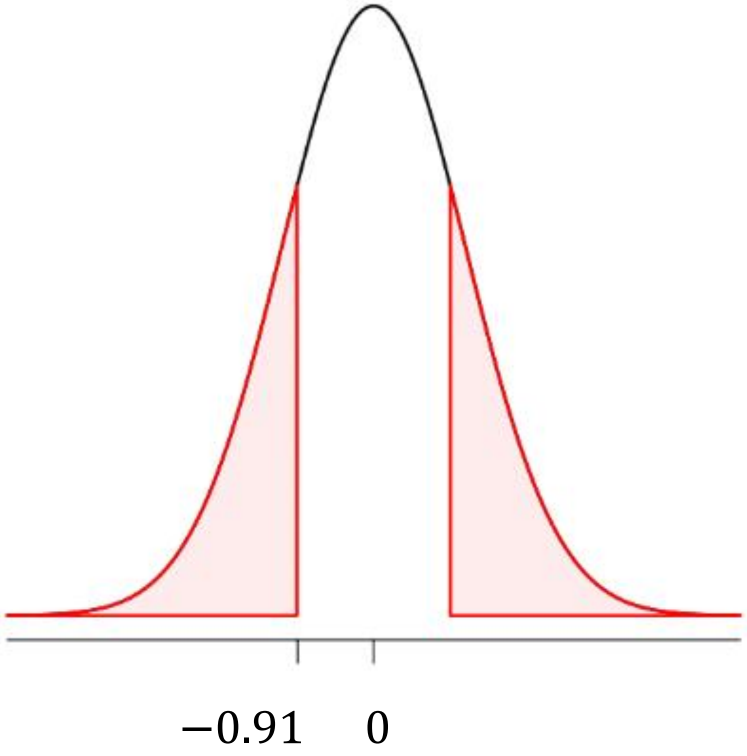
Comparing two proportions: Significance test

Step 4: P-value

$z = -0.91$

P-value = 2 *0.18

= 0.36



Because it's a two-sided test

Comparing two proportions: Significance test

Step 5: Conclusion

Suppose we choose $\alpha = 0.05$

P-value = 0.36, which is larger than α

→not reject H_0

→Sample proportions do not differ significantly from each other

Parameter	Sample(s)	Statistic	Chapter
Proportion (categorical)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	~~CI~~	~~Ch10b~~
		~~Signif Test~~	~~Ch10b~~
	Two dependent samples	~~CI~~	~~X~~
		Signif Test	Ch10b
Mean (quantitative)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	~~CI~~	~~Ch10a~~
		~~Signif Test~~	~~Ch10a~~
	Two dependent samples	CI	Ch10b
		Signif Test	Ch10b

Interim Exam II

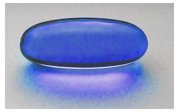
Last week lecture

Today

Today

- Independent vs dependent samples
- Comparing two proportions in independent samples
- Comparing two means in dependent samples
- Comparing two proportions in dependent samples

Within-groups design
(Dependent samples)







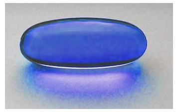



Timepoint 1
Mean: $\bar{x}_1$; sd: s_1
n = 4

Timepoint 2
Mean: $\bar{x}_2$; sd: s_2
n = 4

Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg



Matched pairs design

Matched on, for example, age

<u>Pair members A</u>
Mean: $\bar{x}_1$; sd: s_1
n = 4

<u>Pair members B</u>
Mean: $\bar{x}_2$; sd: s_2
n = 4

Picture source: pixabay.com

Photo source pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Another example of dependent samples is when you sample n couples (e.g., spouses or siblings)

Matched pairs design

First-born: family a | Second-born: family a

First-born: family b | Second-born: family b

First-born: family c | Second-born: family c

First-born: family d | Second-born: family d

Comparing the means of first-borns to that of second-borns

First borns: Mean: $\bar{x}_1$; sd: s_1, n = 4

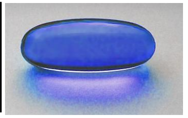
Second borns: Mean: $\bar{x}_2$; sd: s_2, n = 4

Picture source: pixabay.com

Comparing two means: Dependent samples

- Both a within groups design, or matching pairs of people result in what is called *matched pairs.* And in both these cases the samples are said to be *dependent.*
- The statistical tests in both of these cases is the same!
- If your samples are dependent you can subtract the scores of the one sample from the other and consider the difference scores as a single sample.

Matched pairs design

No caffeine (placebo) 

 Caffeine

x_d


18.3


14.5

18.3 – 14.5 = **2.8**

4 differences -> n = 4


13.7


14.9

13.7 – 14.9 = **-1.2**


8.9


12.7

8.9 – 12.7 = **-3.8**

$\bar{x}_d = 0.63$
$s_d = 3.84$
$n = 4$


16.1


11.4

16.1 – 11.4 = **4.7**

Note that for the calculations it does not matter whether you have repeated measures or paired samples

Picture source: pixabay.com

Comparing two means: Dependent samples

- To see whether caffeine changes concentration in the case of dependent samples, we want to know whether $\mu_d = \mu_1 - \mu_2$ differs from 0

→ 95%-confidence interval

(does the 95%-confidence interval of μ_d contain 0)

  - **Same procedure as confidence interval for a single mean (Chapter 8.3)**

- → Significance test

(Can we reject $H_{0:}\ \mu_d = 0$)

  - **Same procedure as significance test for a single mean (Chapter 9.3)**

Comparing two means: Dependent samples

- To see whether caffeine changes concentration in the case of dependent samples, we want to know whether $\mu_d = \mu_1 - \mu_2$ differs from 0

→ 95%-confidence interval **Same procedure as confidence interval for a single mean (Chapter 8.3)**

$\bar{x}_d \pm t_{0.025}(se)$ with $se = s_d/\sqrt{n}$

$df = n - 1$

n is the number of pairs here!

- → Significance test **Same procedure as significance test for a single mean (Chapter 9.3)**

$t = \frac{\bar{x}_d - 0}{se}$ with $se = s_d/\sqrt{n}$

$df = n - 1$

n is the number of pairs here!

Confidence interval

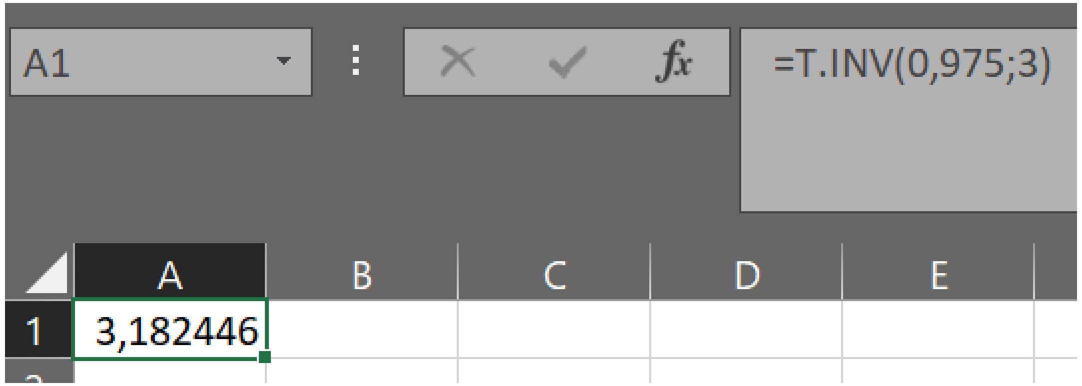
n=4 (see previous slide) ← $\bar{x}_d = 0.63$, $s_d = 3.84$

95% confidence interval to test if $\mu_d = \mu_1 - \mu_2$ differs from 0

Formula: $\bar{x}_d \pm t_{0.025}(se)$ with $se = s_d/\sqrt{n}$

$t_{0.025}$ (df = 4-1 = 3) = 3.182

Calculate critical value →



$$\text{se} = \frac{\text{s}_\text{d}}{\sqrt{n}} = \frac{3.84}{\sqrt{4}} = 1.92$$

→Lower bound: $0.63 - 3.182se = 0.63 - 3.182 \times 1.92 = -5.48$

→Upper bound: $0.63 + 3.182se = 0.63 + 3.182 \times 1.92 = 6.74$

Conclusion: the 95% confidence interval ranges (-5.48, 6.74), which means that 0 is in the interval and thus a difference of 0 is a believable value!

Significance test

n=4 (see previous slide) ← $\bar{x}_d = 0.63$, $s_d = 3.84$

Significance test with $\alpha = 0.05$ to test if $\mu_d = \mu_1 - \mu_2$ differs from 0. → $H_0\!: \mu_d = 0$, $H_a\!: \mu_d \neq 0$ (two-sided)

Step 1:

See lecture 15

Step 2:

H_0: $\mu_d = \mu_1 - \mu_2 = 0$

H_a: $\mu_d = \mu_1 - \mu_2 \neq 0$

Step 3:

$t = \frac{\bar{x}_d - 0}{se}$ with $se = s_d/\sqrt{n}$

$se = \frac{s_d}{\sqrt{n}} = \frac{3.84}{\sqrt{4}} = 1.92$

$t = \frac{0.63 - 0}{1.92} = 0.33$

Step 4:

Positive t-value, so smallest tail is upper tail and thus we want 1- T.DIST(..)

df = 4 – 1 = 3

From excel: (1-T.DIST(0,33; 3; TRUE)) = 0.3816

→probability of upper tail is 0.382

→But we need the two sided P-value: 2*0.3816 = 0.763

Step 5:

- P-value larger than level of significance (α)
- Not reject H_0
- Mean concentration scores are not significantly different between conditions

In practice you would use more subjects than n = 4

Parameter	Sample(s)	Statistic	Chapter
Proportion (categorical)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	~~CI~~	~~Ch10b~~
		~~Signif Test~~	~~Ch10b~~
	Two dependent samples	~~CI~~	~~X~~
		Signif Test	Ch10b
Mean (quantitative)	One sample	~~CI~~	~~Ch8~~
		~~Signif Test~~	~~Ch9~~
	Two independent samples	~~CI~~	~~Ch10a~~
		~~Signif Test~~	~~Ch10a~~
	Two dependent samples	~~CI~~	~~Ch10b~~
		~~Signif Test~~	~~Ch10b~~

Interim Exam II

Last week lecture

Today

Today

- Independent vs dependent samples
- Comparing two proportions in independent samples
- Comparing two means in dependent samples
- Comparing two proportions in dependent samples

English | Login | Help | Admin Center | Contact Us

APA PsycNET® | American Psychological Association

SEARCH | BROWSE | APA THESAURUS | Recent Searches | My List | My PsycNET

Transcendental Meditation and concentration ability.

EXPORT | Add To My List | Database: PsycINFO | Journal Article

Sabel, Bernhard A.

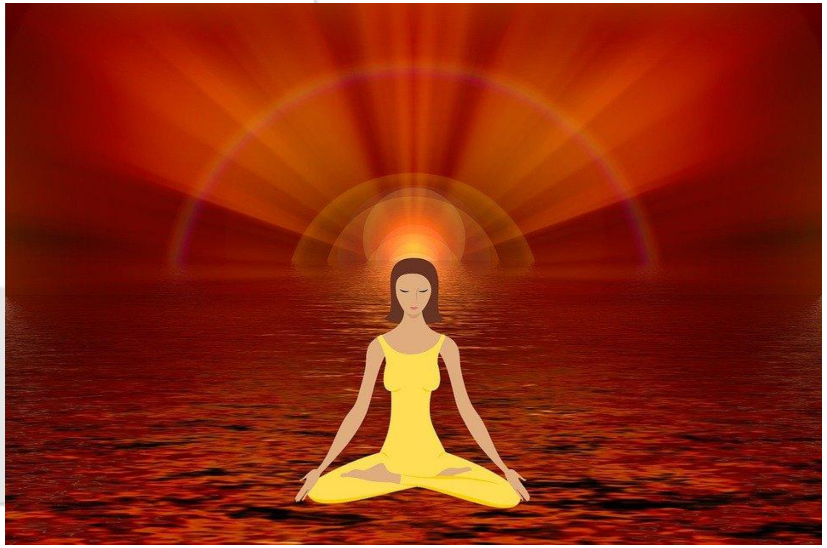
Full text from publisher

Citation

Sabel, B. A. (1980). Transcendental Meditation and concentration ability. *Perceptual and Motor Skills, 50*(3, Pt 1), 799-802.
http://dx.doi.org/10.2466/pms.1980.50.3.799

Abstract

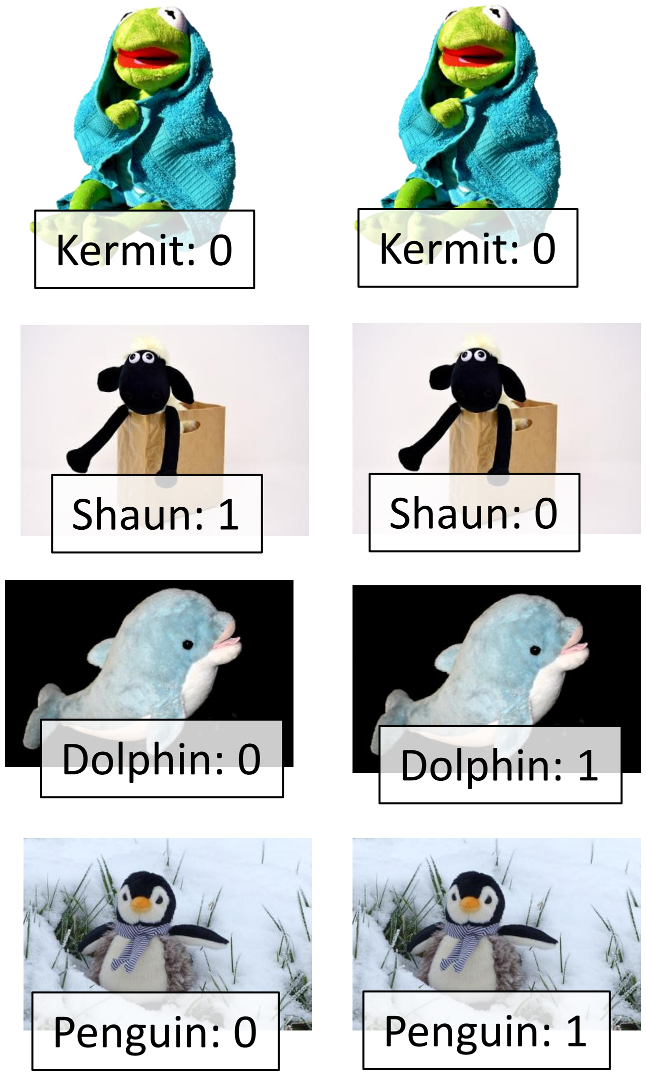
Assigned 60 practitioners (mean age 28 yrs) of Transcendental Meditation to 2 treatment groups. One group meditated for 20 min while the other read a text quietly. Both groups were tested before and after treatment to measure their concentration ability. Meditation had no measurable short-term effects on concentration and the Ss' experience of meditation was not correlated with the concentration score. (12 ref) (PsycINFO Database Record (c) 2016 APA, all rights reserved)



Picture source: pixabay.com

Within-groups design
(Dependent samples)

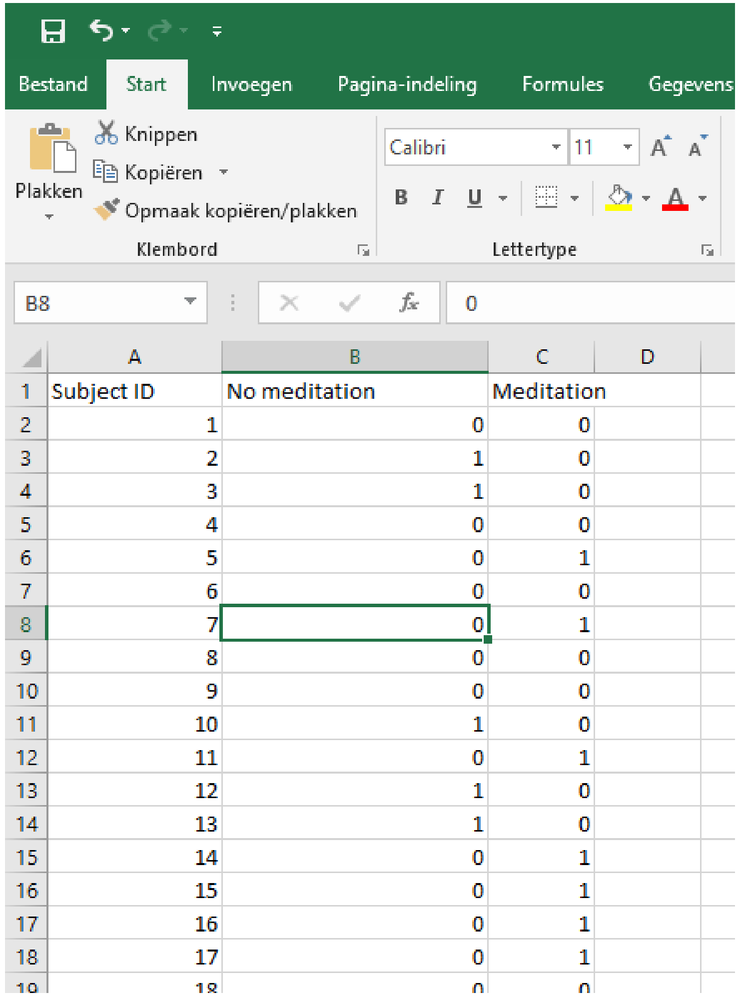
Meditation | **No Meditation**



1: passed concentration task
0: did not pass concentration task

Picture source: pixabay.com

Dependent samples: McNemar’s test



	A	B	C
1	Subject ID	No meditation	Meditation
2	1	0	0
3	2	1	0
4	3	1	0
5	4	0	0
6	5	0	1
7	6	0	0
8	7	0	1
9	8	0	0
10	9	0	0
11	10	1	0
12	11	0	1
13	12	1	0
14	13	1	0
15	14	0	1
16	15	0	1
17	16	0	1
18	17	0	1
19	18	0	0

(0: failed; 1: passed)

Step 1: assumptions

- The variable is categorical
- The sum of the frequencies (0 to 1; or 1 to 0) should be at least 30

	No meditation	
Meditation	Pass	Fail
Pass	400	31
Fail	19	50

the sum of these frequencies should be at least 30 (which is the case now as 31+19 = 50)

Step 2: hypotheses

H_0: $p_{meditation} - p_{no\ meditation} = 0$

H_A: $p_{meditation} - p_{no\ meditation} > 0$

Dependent samples: McNemar's test

Step 3: test statistic

a	*b*
c	*d*

	No meditation	
Meditation	Pass	Fail
Pass	400	31
Fail	19	50

$$z = \frac{b - c}{\sqrt{b + c}} = \frac{31 - 19}{\sqrt{31 + 19}} = \frac{12}{\sqrt{50}} = \frac{12}{7.07} = 1.70$$

Alternatively, you can also do:

Meditation	**No meditation**	**d**	**f (frequency)**
1	1	0	400
1	0	1	31
0	1	-1	19
0	0	0	50

$$z = \frac{f_{10} - f_{01}}{\sqrt{f_{10} + f_{01}}} = \frac{31 - 19}{\sqrt{31 + 19}} = \frac{12}{\sqrt{50}} = \frac{12}{7.07} = 1.70$$

(0: failed; 1: passed)

Dependent samples: McNemar's test

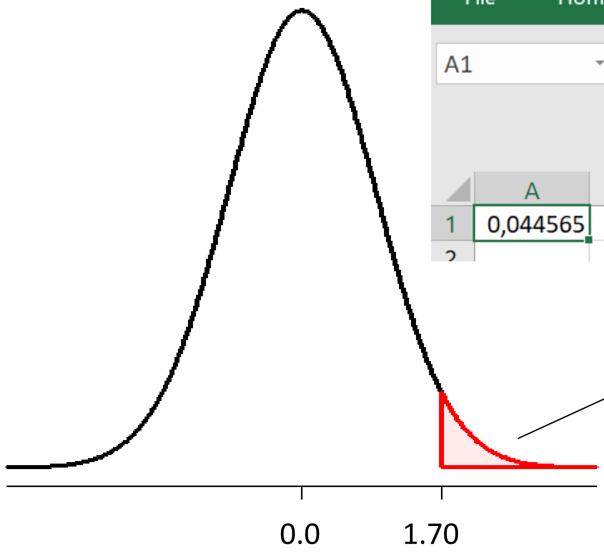
Step 3: test statistic

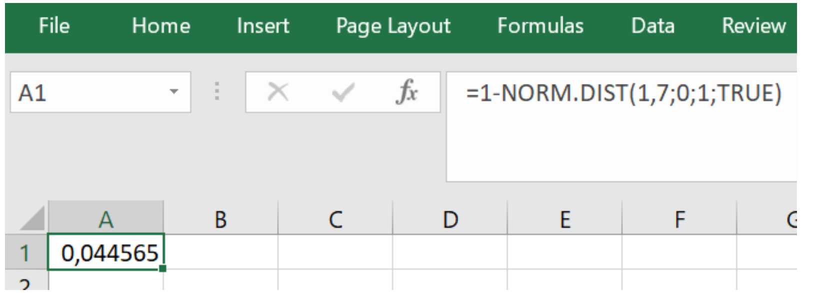
$$z = \frac{b - c}{\sqrt{b + c}} = \frac{31 - 19}{\sqrt{31 + 19}} = \frac{12}{\sqrt{50}} = \frac{12}{7.07} = 1.70$$

Step 4: P-value

Since H_A: $p_{meditation} - p_{no\ meditation} > 0$, we want *P(z > 1.70)*

P-value = 0.045 (note: if H_A *were undirected, you would calculate smallest tail and double it!)*





P(z > 1.70) = 0.0446

Step 5: Conclusion

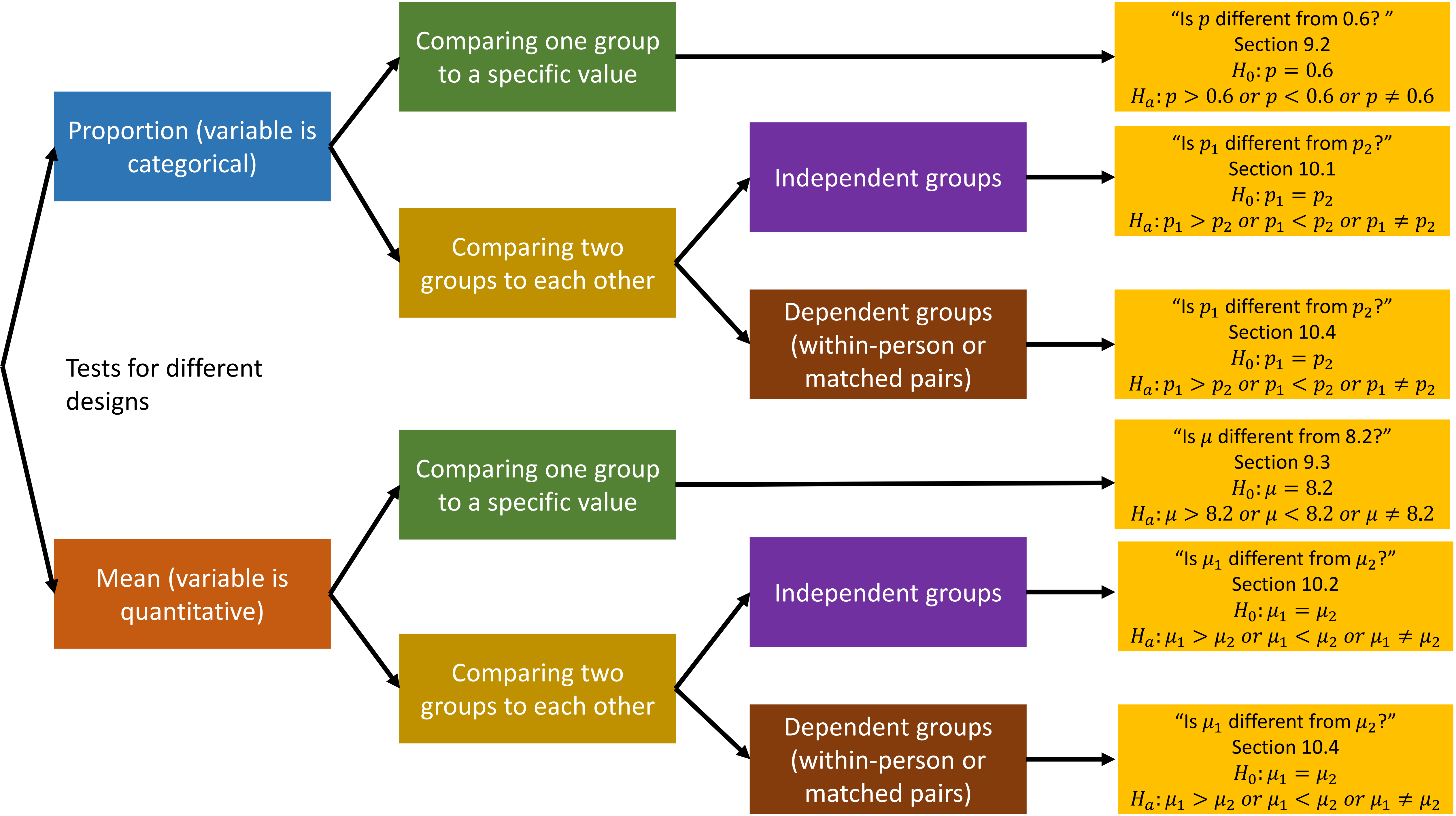
For an $\alpha = 0.05$, the P-value < α

We reject H_0 in favor of H_a. Significantly larger proportion of passing concentration test in meditation condition

Parameter	Sample(s)	Statistic	Chapter
Proportion (categorical)	One sample	CI	Ch8
		Signif Test	Ch9
	Two independent samples	CI	Ch10b
		Signif Test	Ch10b
	Two dependent samples	CI	X
		Signif Test	Ch10b
Mean (quantitative)	One sample	CI	Ch8
		Signif Test	Ch9
	Two independent samples	CI	Ch10a
		Signif Test	Ch10a
	Two dependent samples	CI	Ch10b
		Signif Test	Ch10b

(green)	Interim Exam II
(yellow)	Interim Exam III

Thursday, you learn one more test (chi-square test), and that is it for Interim Exam III


Tests for different designs
Proportion (variable is categorical)
Comparing one group to a specific value
Comparing two groups to each other
Independent groups
Dependent groups (within-person or matched pairs)
"Is p different from 0.6? "
Section 9.2
$H_0: p = 0.6$
$H_a: p > 0.6$ or $p < 0.6$ or $p \neq 0.6$
"Is p_1 different from p_2?"
Section 10.1
$H_0: p_1 = p_2$
$H_a: p_1 > p_2$ or $p_1 < p_2$ or $p_1 \neq p_2$
"Is p_1 different from p_2?"
Section 10.4
$H_0: p_1 = p_2$
$H_a: p_1 > p_2$ or $p_1 < p_2$ or $p_1 \neq p_2$
Mean (variable is quantitative)
Comparing one group to a specific value
Comparing two groups to each other
Independent groups
Dependent groups (within-person or matched pairs)
"Is μ different from 8.2?"
Section 9.3
$H_0: \mu = 8.2$
$H_a: \mu > 8.2$ or $\mu < 8.2$ or $\mu \neq 8.2$
"Is μ_1 different from μ_2?"
Section 10.2
$H_0: \mu_1 = \mu_2$
$H_a: \mu_1 > \mu_2$ or $\mu_1 < \mu_2$ or $\mu_1 \neq \mu_2$
"Is μ_1 different from μ_2?"
Section 10.4
$H_0: \mu_1 = \mu_2$
$H_a: \mu_1 > \mu_2$ or $\mu_1 < \mu_2$ or $\mu_1 \neq \mu_2$

Comparing one group to a specific proportion → "Is p different from 0.6?" → $p_0 = 0.6$ → Test-statistic, Use z-distribution: $z = \frac{\hat{p} - p_0}{se_0}, se_0 = \sqrt{\frac{p_0(1-p_0)}{n}}$

Comparing independent groups proportions → "Is p_1 different from p_2?" → Test-statistic, Use z-distribution: $z = \frac{\widehat{p_1} - \widehat{p_2} - 0}{se_0}, se_0 = \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}},$ $\hat{p}$=pooled estimate, see p.532

Comparing dependent groups proportions → "Is p_1 different from p_2?" → Test-statistic, use z-distribution: McNemar test (pp576-577): $z = \frac{f_{10} - f_{01}}{\sqrt{f_{10} + f_{01}}}$, f_{01} = frequency of pairs scoring 0,1; f_{10}=freq. of pairs that score 1,0 (see slides lecture 17)

Comparing one group to a specific mean → "Is μ different from 8.2?" → $\mu_0 = 8.2$ → Test-statistic, Use t-distribution: $t = \frac{\bar{x} - \mu_0}{se}, se = \frac{s}{\sqrt{n}}$, with df=n-1

Comparing independent groups means → "Is μ_1 different from μ_2?" → Test-statistic, Use t-distribution: $t = \frac{\overline{x_1} - \overline{x_2} - 0}{se}, se = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$ df will be given (don't need to calculate)

Comparing dependent groups means → "Is μ_1 different from μ_2?" → Test-statistic, Use t-distribution: $t = \frac{\overline{x_d} - 0}{se}, se = \frac{s_d}{\sqrt{n}}$, with df=n-1. $\overline{x_d}$ is the mean of the difference scores, See p.570-571

These give you a test-statistic. The next step to get a p-value is to calculate a tail probability using:

Z-value, Left tail probability:
=norm.dist(z-value;0;1;true)
or
Z-value, Right tail probability:
=1-norm.dist(z-value;0;1;true)

T-value, Left tail probability:
=t.dist(t-value;df;true)
or
T-value, Right tail probability:
=1-t.dist(t-value;df;true)

If H_a is directed (one-sided test): Calculate left or right tail depending on direction of H_a

If H_a is undirected (two-sided test): Calculate smaller tail (that is left tail for negative z-value or t-value, or right tail for positive z-value or t-value) and then double it!

Example question

Selma wants to know whether the proportion of people who believe in heaven is larger than the proportion of people who believe in hell. To test this, she asks 1314 people two questions: “do you believe in hell?” (yes/no) and “do you believe in heaven?” (yes/no). The results she finds are as follows:

	Belief in Hell	
Belief in heaven	Yes	No
Yes	955	162
No	9	188

Selma chooses a significance level of 0.01. Calculate a p-value. Based on this p-value she should

a) accept the null-hypothesis.

b) reject the null-hypothesis in favor of the alternative hypothesis that the proportion believing in heaven is larger than the proportion of people who believe in hell.

c) not reject the null hypothesis.

Example question

Dependent sample: same people asked two questions

One-sided test
$p_{heaven} > p_{hell}$

Selma wants to know whether the proportion of people who believe in heaven is larger than the proportion of people who believe in hell. To test this, she asks 1314 people two questions: "do you believe in hell?" (yes/no) and "do you believe in heaven?" (yes/no). The results she finds are as follows:

Categorical, proportions instead of means

	Belief in Hell	
Belief in heaven	Yes	No
Yes	955	162
No	9	188

Beware: α = 0.01

Selma chooses a significance level of 0.01. Calculate a p-value. Based on this p-value she should

a) accept the null-hypothesis.

b) reject the null-hypothesis in favor of the alternative hypothesis that the proportion believing in heaven is larger than the proportion of people who believe in hell.

c) not reject the null hypothesis.

This means we're comparing dependent groups proportions

$$z = \frac{b - c}{\sqrt{b + c}} = \frac{162 - 9}{\sqrt{162 + 9}} = 11.7$$

a	b
c	d

	Belief in Hell	
Belief in heaven	Yes	No
Yes	955	162
No	9	188

heaven	hell	d (x1-x2)	freq
1	1	0	955
1	0	1	162
0	1	-1	9
0	0	0	188

$f_{10} > f_{01}$ is in line with H_a

f_{10}: cases that build to heaven larger proportion

f_{01}: cases that build to hell larger proportion

H_A: $p_{heaven} - p_{hell} > 0$

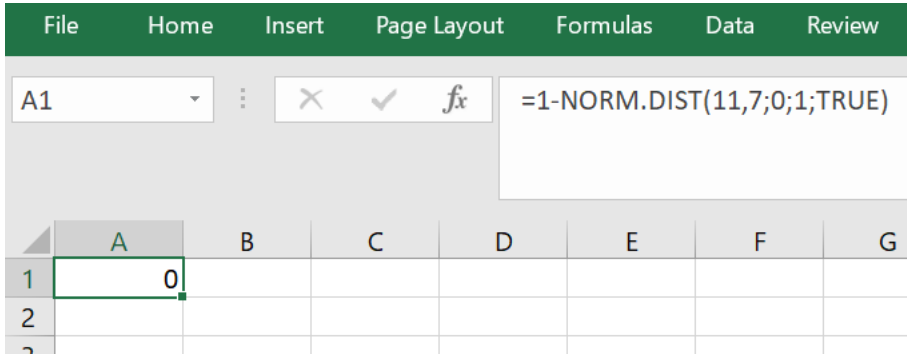
From H_a you expect f_{10}- $f_{01} > 0$

$$f_{10} = 162, f_{01} = 9, z = \frac{f_{10} - f_{01}}{\sqrt{f_{10} + f_{01}}} = \frac{162 - 9}{\sqrt{162 + 9}} = 11.7$$

Solution

To compare dependent groups proportions, we do a McNemar test

$$z = \frac{f_{10} - f_{01}}{\sqrt{f_{10} + f_{01}}} = \frac{162 - 9}{\sqrt{162 + 9}} = 11.7$$



H_A: $p_{heaven} - p_{hell} > 0$

Therefore, we want an upper tail probability: 1-NORM.DIST(..)

The p-value is extremely small (so small that excel gives 0)

This is smaller than 0,01 (significance level α) and thus the null hypothesis can be rejected in favor of H_a: $p_{heaven} > p_{hell}$

(**Note**: If H_a were undirected, the p-value would have to be doubled!)

Example question

Selma wants to know whether the proportion of people who believe in heaven is larger than the proportion of people who believe in hell. To test this, she asks 1314 people two questions: “do you believe in hell?” (yes/no) and “do you believe in heaven?” (yes/no). The results she finds are as follows:

	Belief in Hell	
Belief in heaven	Yes	No
Yes	955	162
No	9	188

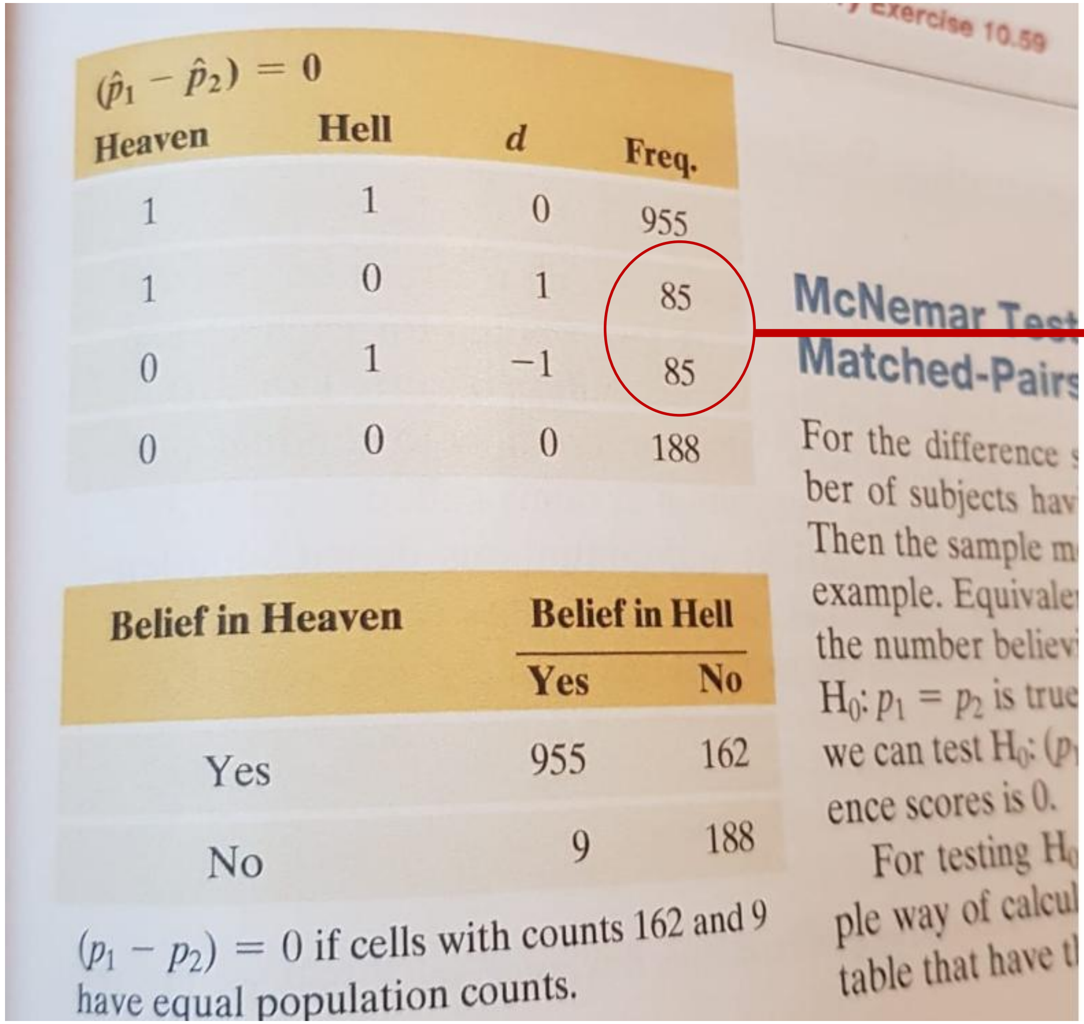
Selma chooses a significance level of 0.01. Calculate a p-value. Based on this p-value she should

a) accept the null-hypothesis.

b) reject the null-hypothesis in favor of the alternative hypothesis that the proportion believing in heaven is larger than the proportion of people who believe in hell.

c) not reject the null hypothesis.

Error in previous edition book, page 503:
The current edition no longer has this mistake.



Exercise 10.59

$(\hat{p}_1 - \hat{p}_2) = 0$

Heaven	Hell	d	Freq.
1	1	0	955
1	0	1	85
0	1	−1	85
0	0	0	188

Belief in Heaven	Belief in Hell	
	Yes	No
Yes	955	162
No	9	188

$(p_1 - p_2) = 0$ if cells with counts 162 and 9 have equal population counts.

McNemar Test
Matched-Pairs

For the difference
ber of subjects hav
Then the sample m
example. Equivale
the number believ
H_0: $p_1 = p_2$ is true
we can test H_0: $(p_1$
ence scores is 0.

For testing H_0
ple way of calcul
table that have t

Should be 162 and 9!!

Excercise from old edition of book 10.12

10.12 TV watching A researcher predicts that the percentage of people who do not watch TV is higher now than before the advent of the Internet. Let p_1 denote the population proportion of American adults in 1975 who reported watching no TV. Let p_2 denote the corresponding population proportion in 2008.

a. Set up null and alternative hypotheses to test the researcher's prediction.

b. According to General Social Surveys, 57 of the 1483 subjects sampled in 1975 and 87 of the 1324 subjects sampled in 2008 reported watching no TV. Find the sample estimates of p_1 and p_2.

c. Show steps of a significance test. Explain whether the results support the reseacher's claim.

Solution

$a)\ \ H_0: p_1 - p_2 = 0$

$H_a: p_1 - p_2 < 0$

b) $\hat{p}_1 = \frac{57}{1483} = 0.038$

$\hat{p}_2 = \frac{87}{1324} = 0.066$

c) Step 1: Assumptions:

- Variable is categorical
- Data are obtained using random sampling

<u>For one sided test</u>

$n_1p_1 \geq 10$ and $n_1(1-p_1) \geq 10$ in group 1 and

$n_2p_2 \geq 10$ and $n_2(1-p_2) \geq 10$ in group 2

Solution

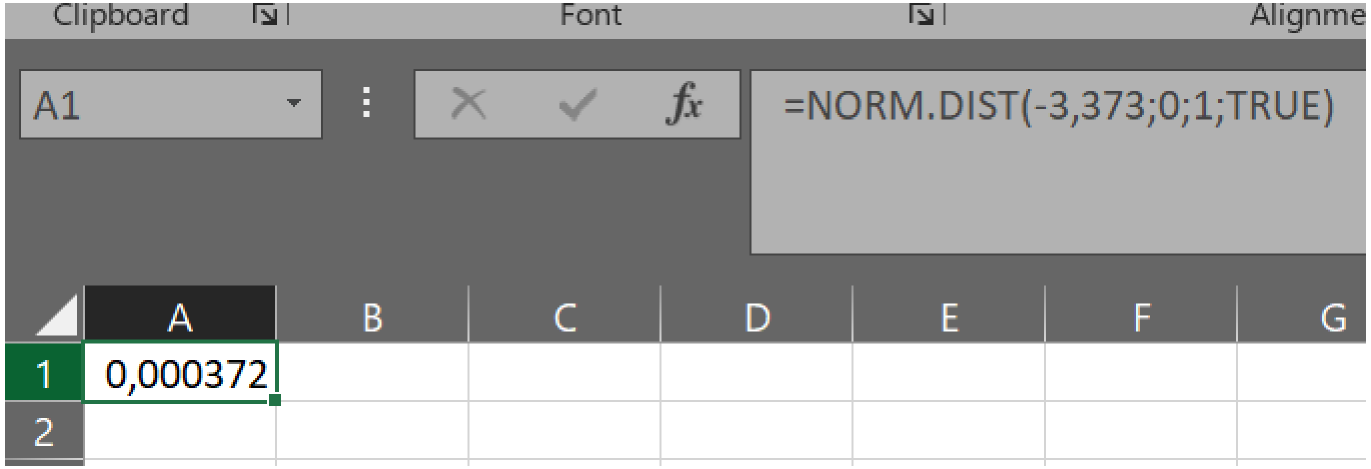
Step 2: hypotheses, see part (a)
Step 3: calculate test-statistic:

- $\hat{p}(pooled\ estimate) = \frac{144}{2807} = 0.051$

$$\mathrm{se}_0 = \sqrt{\hat{p}\,(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{(0.051 * 0.9487)\left(\frac{1}{1483}+\frac{1}{1324}\right)} = 0.0083$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{se_0} = \frac{0.038 - 0.066}{0.0083} = -3.373$$

Solution



Step 4 calculate p-value:

One sided p-value, lower tail probability because of the alternative hypothesis

P-value = NORM.DIST(-3,373; 0; 1; TRUE) = 0.0004

Step 5: conclusion:

For an alpha level of 0.05 the p-value is significant which means that the null hypothesis should be rejected in favor of the alternative hypothesis that the proportion p_1 (1957) is smaller than p_2 (2008)

Excercise 10.22

10.22 Chelation useless? Chelation is an alternative therapy for heart disease that uses repeated intravenous administration of a human-made amino acid in combination with oral vitamins and minerals. Practitioners believe it removes calcium deposits from buildup in arteries. However, the evidence for a positive effect is anecdotal or comes from nonrandomized, uncontrolled studies. A double-blind randomized clinical trial comparing chelation to placebo used a treadmill test in which the response was the length of time until a subject experienced ischemia (lack of blood flow and oxygen to the heart muscle).

a. After 27 weeks of treatment, the sample mean time for the chelation group was 652 seconds. A 95% confidence interval for the population mean for chelation minus the population mean for placebo was −53 to 36 seconds. Explain how to interpret the confidence interval.

b. A significance test comparing the means had P-value = 0.69. Specify the hypotheses for this test, which was two-sided.

c. The authors concluded from the test, "There is no evidence to support a beneficial effect of chelation therapy" (*Source:* M. Knudtson et al., (2002), *JAMA*, vol. 287, p. 481, 2002.). Explain how this conclusion agrees with inference based on the values in the confidence interval.

Solution 10.22

a) Anything between -53 and 36 seconds can be considered believable values for the difference between the chelation group and placebo group. As such a difference of 0 (i.e., no difference) is also a believable value.

$b)$ $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 \neq 0$

c) A two-sided hypothesis test with alpha =0.05 corresponds to a 95% confidence interval. In this case the 95% confidence interval included 0 which means that a difference of 0 is believable (null hypothesis should not be rejected). Correspondingly, the P-value of 0.69 is also larger than 0.05 which means that also in a significance test with alpha=0.05, the null hypothesis should not be rejected.

Highlighted excercises in book

- Some other excercises to focus on: 10.3, 10.6, 10.7, 10.8, 10.49, 10.50, 10.56, 10.57, 10.60, 10.62, 10.63