R introduction

Mandy Vogel

University Leipzig

September 10, 2015

Overview

Recap

Standard Errors

Parametric Frequentist Null Hypothesis Testing
- Understand Hypothesis Testing: Z-test
- Simulation Exercises

t-Tests
- One Sample t-test
- Two Sample t-test

Power t-test

Reading Excel Files

Table of Contents I

Tests

There are on one hand

- parametric tests
- non-parametric tests

On the other hand there are there are tests for

- central tendency
- proportions
- variability
- distribution functions
- associations

Parameter

To describe data we need a proper way to summarize them for easier understanding. Therefore we focus on three main areas:

- parameters of location: mean, trimmed means, median, min, max, quantiles, mode
- spread: variance, standard deviation, cœfficient of variation, IQR, range

Table of Contents I

Unreliability

- variance is i.a. used for measures of unreliability like confidence intervals
- Consider the properties that you would like a measure of unreliability to possess. As the variance of the data increases, what would happen to the unreliability of estimated parameters? Would it go up or down?

Unreliability

- variance is i.a. used for measures of unreliability like confidence intervals
- Consider the properties that you would like a measure of unreliability to possess. As the variance of the data increases, what would happen to the unreliability of estimated parameters? Would it go up or down?

Unreliability

Unreliability would go up as variance increased, so we would want to have the variance on the top (the numerator) of any divisions in our formula for unreliability

Unreliability

- What about sample size? Would you want your estimate of unreliability to go up or down as sample size, n, increased?

Unreliability

You would want unreliability to go down as sample size went up, so you would put sample size on the bottom of the formula for unreliability (i.e. in the denominator).

Standard Errors

- our measurement is a combination of variance and sample size
- and we enclose the term inside a square root to get something on the same scale as our measurements
- unreliability measures are called standard errors
- standard errors are used in the calculation of confidence intervals $CI = estimate \pm t_{\alpha/2} \times se$

Table of Contents I

Hypothesis Testing

- R. A. Fischer
- Karl Popper

The statistical null hypothesis testing can be regarded as a mathematical response to the philosophical need for falsification in hypothesis testing.

Hypothesis Testing

- the most common use of statistics by biologists is for null hypothesis tests
- biologists generally use null hypothesis tests based on parametric assumptions
- a large number of null hypothesis procedures have been developed to address particular applications and assumptions, but all take the same approach: They address the question:

How probable are the data if the null hypothesis is true?

Hypothesis Testing

A adherence to this framework can be traced to the following causes:

- there is a strong need for hypothesis testing in the biological science
- these methods generally produce plausible results
- this approach is strongly emphasized in biometric and introductory statistics

But: justification is seldom considered.

Hypothesis Testing

- you can never prove something
- we can only reject (modus tollens) or fail to reject

nothing in science is proven until it is disproved

The Null Hypothesis

- the null hypothesis is often a statement of no effect or no difference
- generally constructed to encompass all possible outcomes except an expected effect
- as a result the rejection of the null supports the expected effect (reductio ad absurdum)

The Null Hypothesis

We look at H_0 and not at the research hypothesis directly because

- as noted above we cannot prove that a hypothesis is true
- it is simply easier to consider statistical evidence from the perspective of H_0
  - the research hypothesis is often no more than an inexact supposition
  - the null hypothesis can often be expressed in exact mathematical terms

Significance Testing

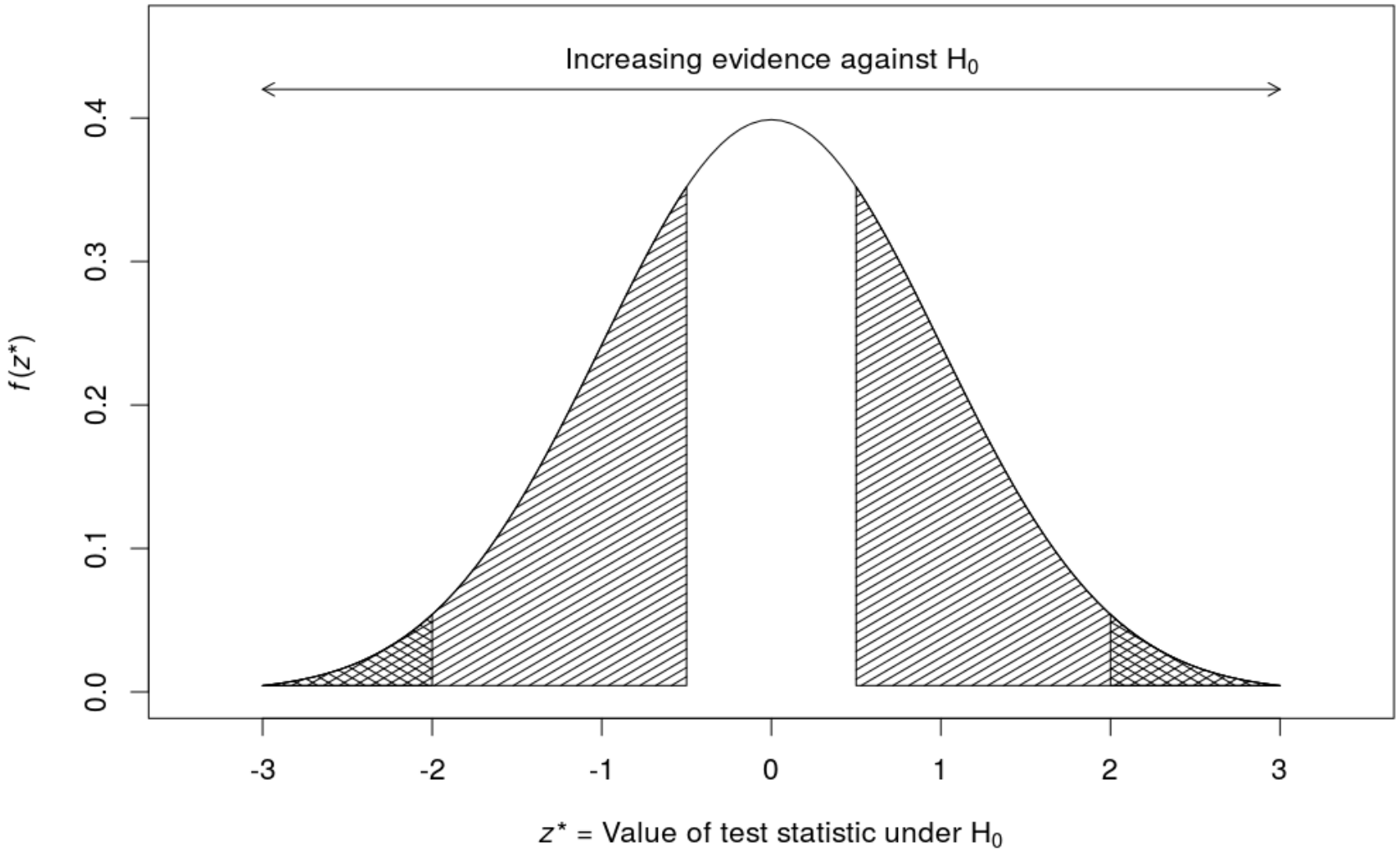
- we are limited to two possible decisions
  1. Reject H_0 or
  2. Fail to reject H_0

Significance Testing – Example

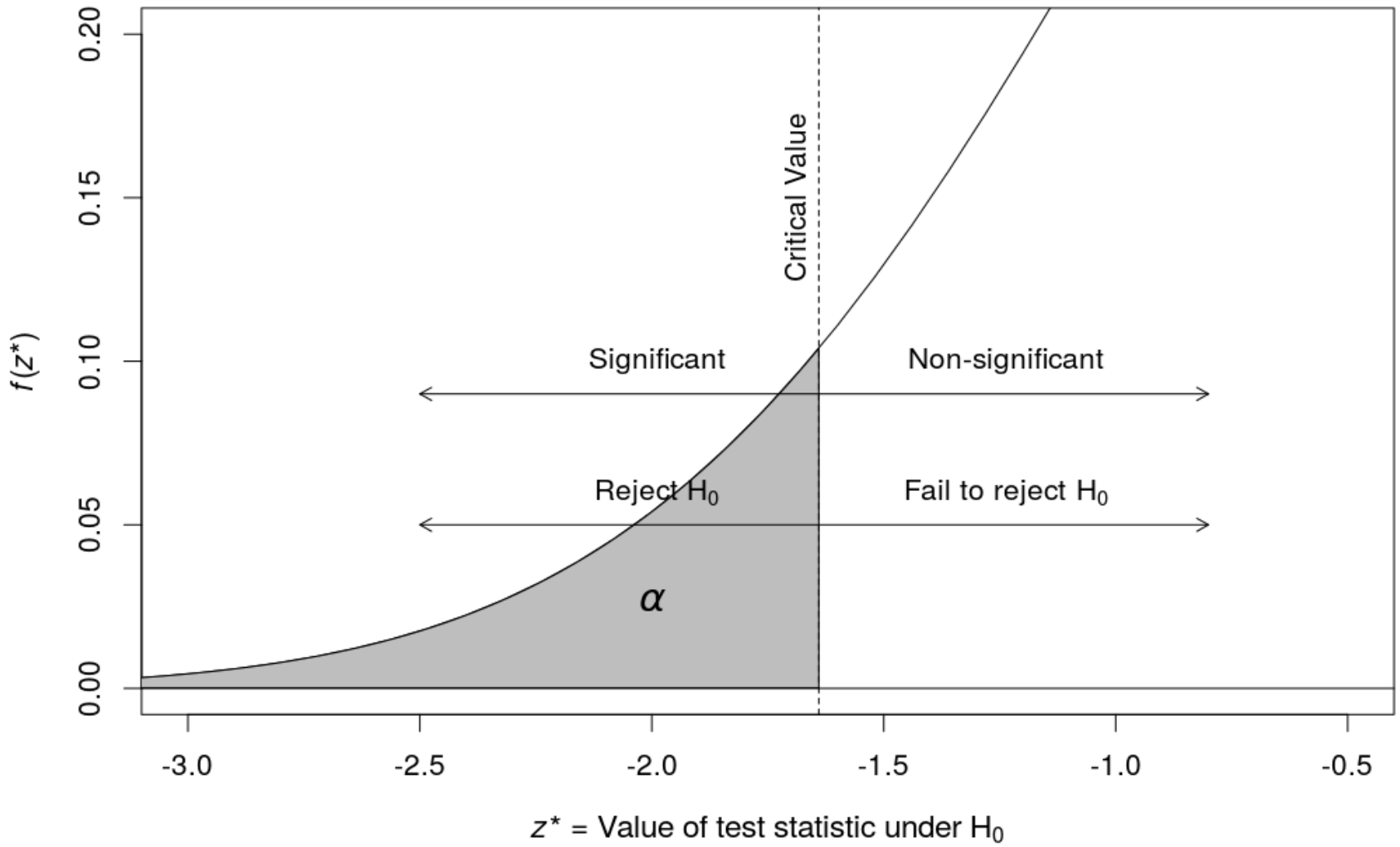
As null hypothesis we might predict that a parameter describing the difference between two populations is zero

- when H_0 is true we would expect an estimate of the parameter to take a value near zero
- an estimator called test statistic is used to quantify the difference between the parameter value quantified in H_0 and the estimated parameter based on the data
- the distribution of the test statistic under H_0 is known
- the p-value is the probability of seeing a test statistic as or more extreme than the test statistic observed if H_0 is true
- smaller p-values designate stronger evidence against H_0 (WHY?)

Significance Testing – p-val



Significance Testing – Critical Value



Significance Testing – Question

Question

If we perform a z-test to assess whether there is a difference in means between two groups a test statistic is calculated.
Why the difference not used directly?

Model of Null Hypothesis Testing (Fisher)

- State H_0
- Conduct an investigation that produces data concerning H_0
- Choose an appropriate test and test statistic, and calculate the test statistic
- Determine the p-value
- Reject H_0 if the p-value is small; otherwise, retain H_0

Fisher proposed that a threshold be established describing outcomes that would be extremely improbable if H_0 were true, allowing rejection of H_0. (α - level)

Model of Null Hypothesis Testing (Fisher)

- outcomes meeting this criterion are said to be statistically significant
- Fisher also proposed $\alpha = 0.05$
- later he recommended that fixed significance levels be too restrictive

Model of Null Hypothesis Testing (Neyman & Pearson)

Different in three important aspects:

1. the significance level should be chosen before beginning the data collection
2. incorporate explicitly the alternative hypothesis (Fisher opposed this)
3. introduction of type I and II errors

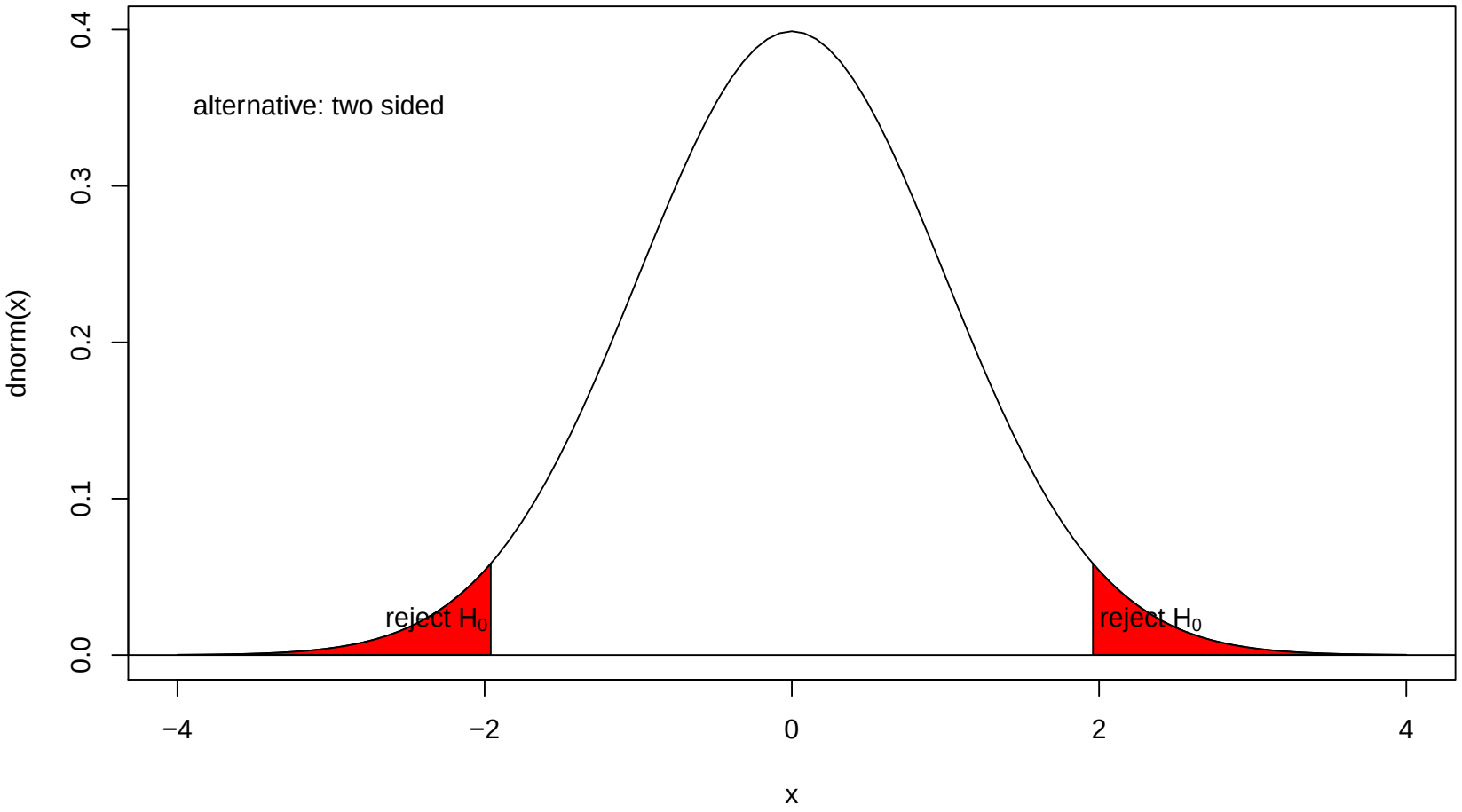
Alternatives

- it may be possible to anticipate directionality in the measured effect
- such scenarios can be accessed with upper- and lower-tailed tests
- H_0 stays $H_0 : X = Y$ or change to $H_0 : X \leq Y$ whereas $H_A : X > Y$

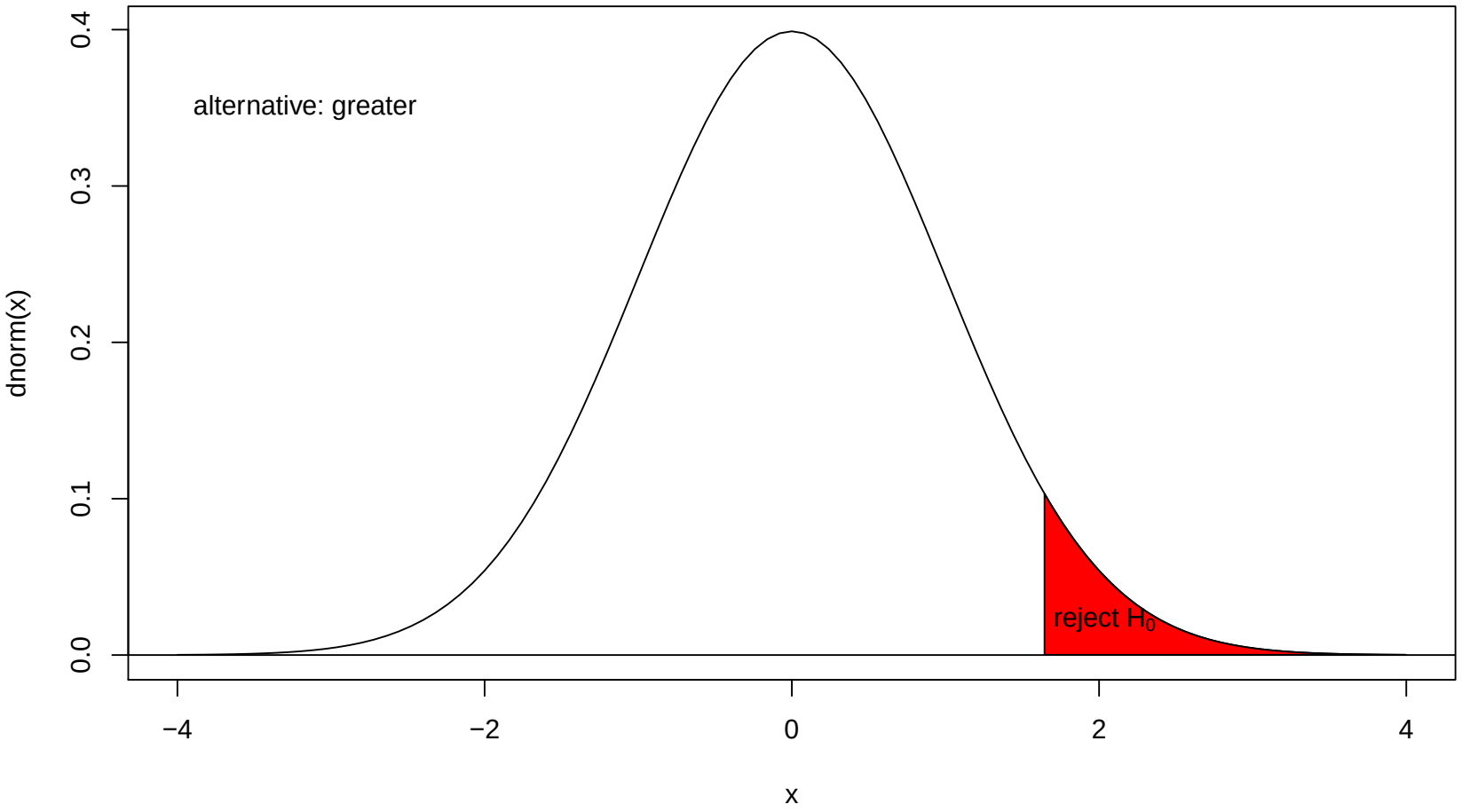
Question

Why – in a strict approach – does H_0 does not change to $H_0 : X \leq Y$. What could be the reason?

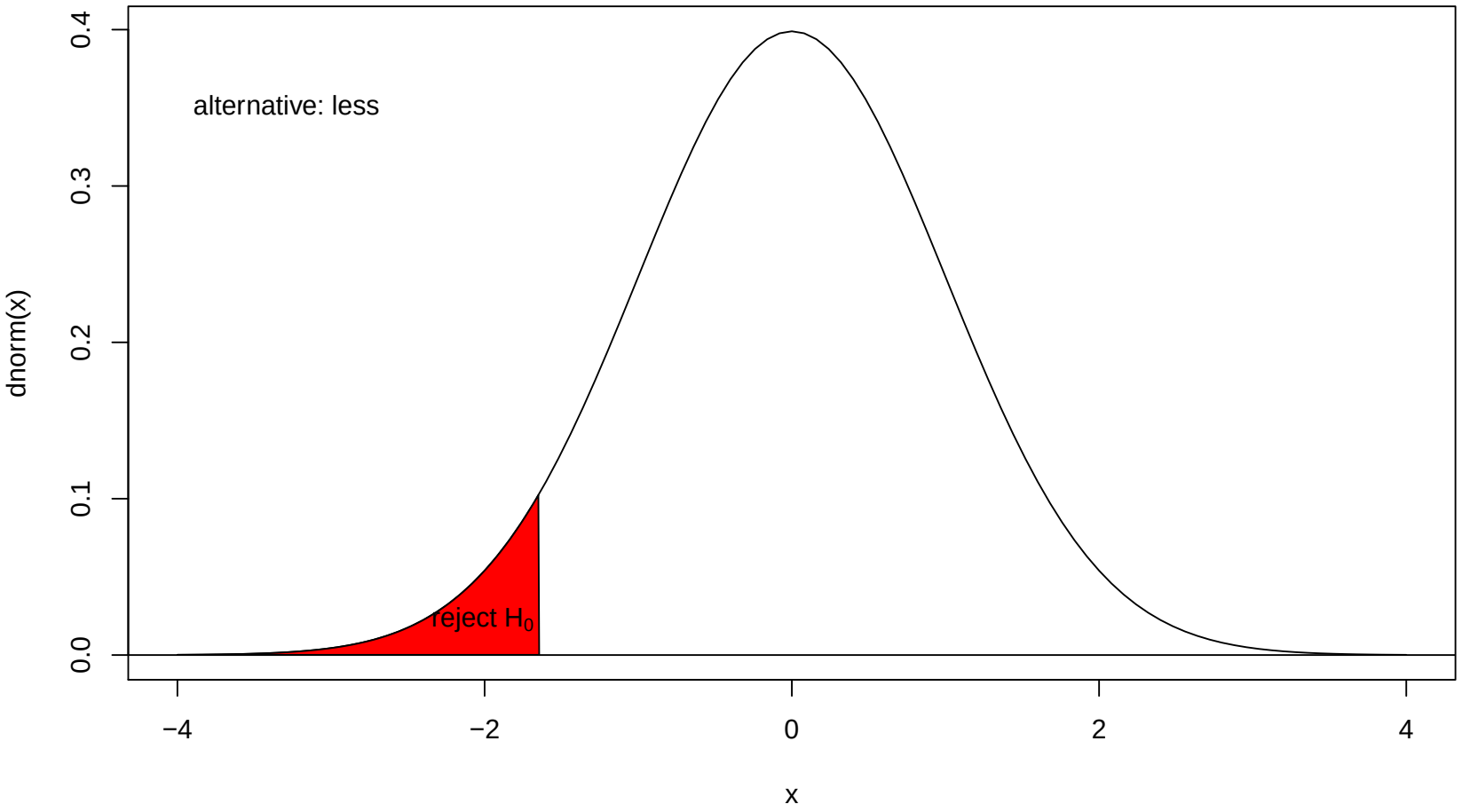
Alternatives



Alternatives



Alternatives



p-values

p-values

- are quantities that relate to the null hypothesis
  - you cannot have a p-value without a null hypothesis
  - the p-value measures how likely it is to see evidence as extreme or more extreme as that observed assuming the null hypothesis is true
  - small p-values are evidence against the null hypothesis; they are not the probability it is true!
  - Bayesian's use a different approach and typically end up with quantities that do have probabilistic interpretations

Non-significant Results

- failure to reject H_0 does not mean that H_0 is true
- different approaches in literature

Significant Results

- having sufficient evidence to reject H_0 dœs not mean H_0 is untrue
- strict Fisherian view: rejecting the null dœs not imply anything about H_A

Significance

- statistical significance should never be confused with scientific significance
- statistical significance tells us the surprise factor:
  - if all my assumptions are correct, and the null hypothesis is true, how surprised should I be by my data
  - at some level of surprise we choose to decide that our null hypothesis is unlikely to be true (usually we check to be sure our assumptions are reasonable)
- scientific significance is concerned with whether what we found is likely to have any relevance to our understanding of nature

Significance

- statistical significance is affected by sample size
- scientific significance is not
- getting more data often ensures statistical significance
  - new data technologies give us too much data, e.g. flow cytometry, sequencing
  - many things are scientifically uninteresting, but statistically significant

Warning

Now we implement a test on ourselves. In real life:

- while the quantities often seem simple
- NEVER IMPLEMENT THEM YOURSELF
- use good software that already exists (R, SAS, MatLab), numerical/scientific computing has many pitfalls for the unwary

Z-test for a population mean

The z-test is something like a t-test (it is like you would know almost everything about the perfect conditions (and therefore it has more power). It uses the normal distribution as distribution for its test statistic and is therefore a good example.

Objective

To investigate the significance of the difference between an assumed population mean μ_0 and a sample mean $\bar{x}$.

Limitations

1. It is necessary that the population variance σ^2 is known.
2. The test is accurate if the population is normally distributed. If the population is not normal, the test will still give an approximate guide.

Z-test for a population mean

Test statistic

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

1. Write a function which takes a vector, the population standard deviation and the population mean as arguments and which returns the Z score.
   - name the function `ztest`
   - set a default value for the population mean

You can always test your function using simulated values:
`rnorm(100,mean=0)` gives you a vector containing 100 normal distributed values with mean 0.

Z-test for a population mean

Write a function which takes a vector, the population standard deviation and the population mean as arguments and which gives the Z score as result.

```
> ztest <- function(x,x.sd,mu=0){
+     sqrt(length(x)) * (mean(x)-mu)/x.sd
+ }
> set.seed(1)
> ztest(rnorm(100),x.sd = 1)
[1] 1.088874
```

Z-test for a population mean

The function `pnorm(Z)` gives the probability of $x \leq Z$. Change your function so that it has the p-value (for a two sided test) as result.

```
> ztest <- function(x,x.sd,mu=0){
+     x <- x[!is.na(x)]
+     if(length(x) < 3) stop("too few values in x")
+     z <- sqrt(length(x)) * (mean(x)-mu)/x.sd
+     2*pnorm(-abs(z))
+ }
> set.seed(1)
> ztest(rnorm(100),x.sd = 1)
[1] 0.2762096
```

Z-test for a population mean

Now let the result be a named vector containing the estimated difference, Z, p and the n.

```
> ztest <- function(x,x.sd,mu=0){
+     x <- x[!is.na(x)]
+     if(length(x) < 3) stop("too few values in x")
+     est.diff <- mean(x)-mu
+     z <- sqrt(length(x)) * (est.diff)/x.sd
+     round(c(diff=est.diff,Z=z,pval=2*pnorm(-abs(z)),n=length(x)
+ }
> set.seed(1)
> ztest(rnorm(100),x.sd = 1)
    diff         Z      pval         n
  0.1089    1.0889    0.2762  100.0000
```

Z-test for a population mean

Variants

1. Z-test for two population means (variances known and equal)
2. Z-test for two population means (variances known and unequal)

To investigate the statistical significance of the difference between an assumed population mean μ_0 and a sample mean $\bar{x}$. There is a function `z.test()` in the `BSDA` package.

Limitations (again)

1. It is necessary that the population variance σ^2 is known.
2. The test is accurate if the population is normally distributed. If the population is not normal, the test will still give an approximate guide.

Simulation Exercises I

1. Now sample 100 values from a Normal distribution with mean 10 and standard deviation 2 and use a z-test to compare it against the population mean 10. What is the p-value?
2. Now we do the sampling and the testing 1000 times, what would be the number of statistically significant results? Use `replicate()` (which is a wrapper of `tapply()`) or a `for()` loop! Record at least the p-values and the estimated differences! Use `table()` to count the p-vals below 0.05. What type of error do you associate with it? What is the smallest absolute difference with a p-value below 0.05?
3. Repeat the simulation above, change the sample size to 1000,10,3 in each of the 1000 samples! How many p-values below 0.05? What are now the smallest absolute differences with a p-value below 0.05?

Simulation Exercises – Solutions

- Now sample 100 values from a Normal distribution with mean 10 and standard deviation 2 and use a z-test to compare it against the population mean 10. What is the p-value? What the estimated difference?

```
> ztest(rnorm(100,mean=10,sd=2),x.sd=2,mu=10)["pval"]
  pval
0.0441
> ztest(rnorm(100,mean=10,sd=2),x.sd=2,mu=10)["diff"]
   diff
-0.0655
> ztest(rnorm(100,mean=10,sd=2),x.sd=2,mu=10)[c("pval","
  pval   diff
0.4515 0.1506
```

Simulation Exercises, n = 100

- Now do the sampling and the testing 1000 times, what would be the number of statistically significant results? Use `replicate()` (which is a wrapper of `tapply()`) or a `for()` loop. Record at least the p-values and the estimated differences! Transform the result into a data frame.

using `replicate()`

```
> res100 <- replicate(1000, ztest(rnorm(100,mean=10,sd=2),x.sd=2,m
> res100 <- as.data.frame(t(res100))
> head(res100)
     diff       Z   pval   n
1  0.0643  0.3216 0.7478 100
2  0.0136  0.0681 0.9457 100
3 -0.0147 -0.0733 0.9416 100
4 -0.1114 -0.5570 0.5775 100
5  0.2567  1.2834 0.1994 100
6  0.0712  0.3559 0.7219 100
```

Simulation Exercises, n = 100

- Use `table()` to count the p-vals below 0.05. What type of error do you associate with it? What is the smallest absolute difference with a p-value below 0.05?

```
> table(res100$pval < 0.05)

FALSE  TRUE
  954    46
> tapply(abs(res100$diff),res100$pval < 0.05,summary)
$`FALSE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00010 0.05753 0.12140 0.13880 0.20170 0.39000

$`TRUE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.3937  0.4150  0.4444  0.4815  0.4970  0.6869

> min(abs(res100$diff[res100$pval<0.05]))
[1] 0.3937
```

Simulation Exercises, n = 1000

- Repeat the simulation above, change the sample size to 1000 in each of the 1000 samples! How many p-values below 0.05? What is now the smallest absolute difference with a p-value below 0.05?

```
> res1000 <- replicate(1000, ztest(rnorm(1000,mean=10,sd=2),
> res1000 <- as.data.frame(t(res1000))
> head(res1000)
     diff       Z   pval    n
1 -0.0237 -0.3743 0.7082 1000
2  0.0827  1.3071 0.1912 1000
3  0.0865  1.3678 0.1714 1000
4  0.0062  0.0987 0.9214 1000
5 -0.1036 -1.6387 0.1013 1000
6 -0.0653 -1.0326 0.3018 1000
```

Simulation Exercises, n = 1000

- Use `table()` to count the p-vals below 0.05. What type of error do you associate with it? What is the smallest absolute difference with a p-value below 0.05?

```
> table(res1000$pval < 0.05)

FALSE  TRUE
  953    47
> tapply(abs(res1000$diff),res1000$pval < 0.05,summary)
$`FALSE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.01870 0.03870 0.04468 0.06540 0.12210

$`TRUE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.1244  0.1326  0.1398  0.1483  0.1532  0.2193
```

Simulation Exercises, n = 10

- Repeat the simulation above, change the sample size to 10 in each of the 1000 samples! How many p-values below 0.05? What is now the smallest absolute difference with a p-value below 0.05?

```
> res10 <- replicate(1000, ztest(rnorm(10,mean=10,sd=2),x.sd
> res10 <- as.data.frame(t(res10))
> table(res10$pval < 0.05)

FALSE  TRUE
  952    48
```

Simulation Exercises, n = 10

```
> tapply(abs(res10$diff),res10$pval < 0.05,summary)
$`FALSE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.0017  0.2038  0.4074  0.4651  0.6783  1.2390

$`TRUE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.241   1.291   1.370   1.464   1.585   2.574
```

Simulation Exercises, n = 3

```
> res3 <- replicate(1000, ztest(rnorm(3,mean=10,sd=2),x.sd=2
> res3 <- as.data.frame(t(res3))
> table(res3$pval < 0.05)

FALSE  TRUE
  947    53
```

Simulation Exercises, n = 3

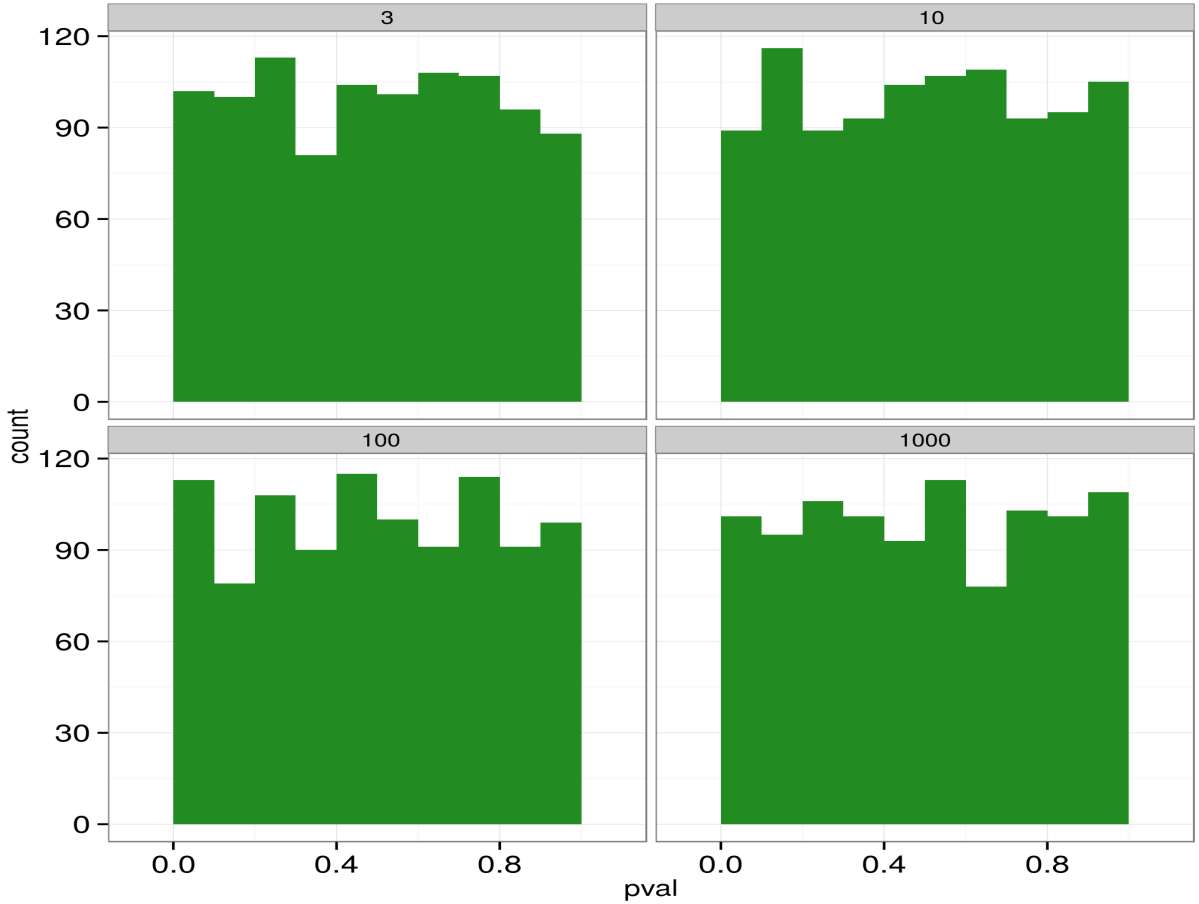
```
> tapply(abs(res3$diff),res3$pval < 0.05,summary)
$`FALSE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.0003  0.3641  0.7396  0.8202  1.2050  2.2620

$`TRUE`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  2.267   2.362   2.482   2.666   2.851   5.118
```

Simulation Visualization p-vals, effect sizes

1. Plotting the distributions of the pvals and the estimated difference per sample size.
2. What is the message?

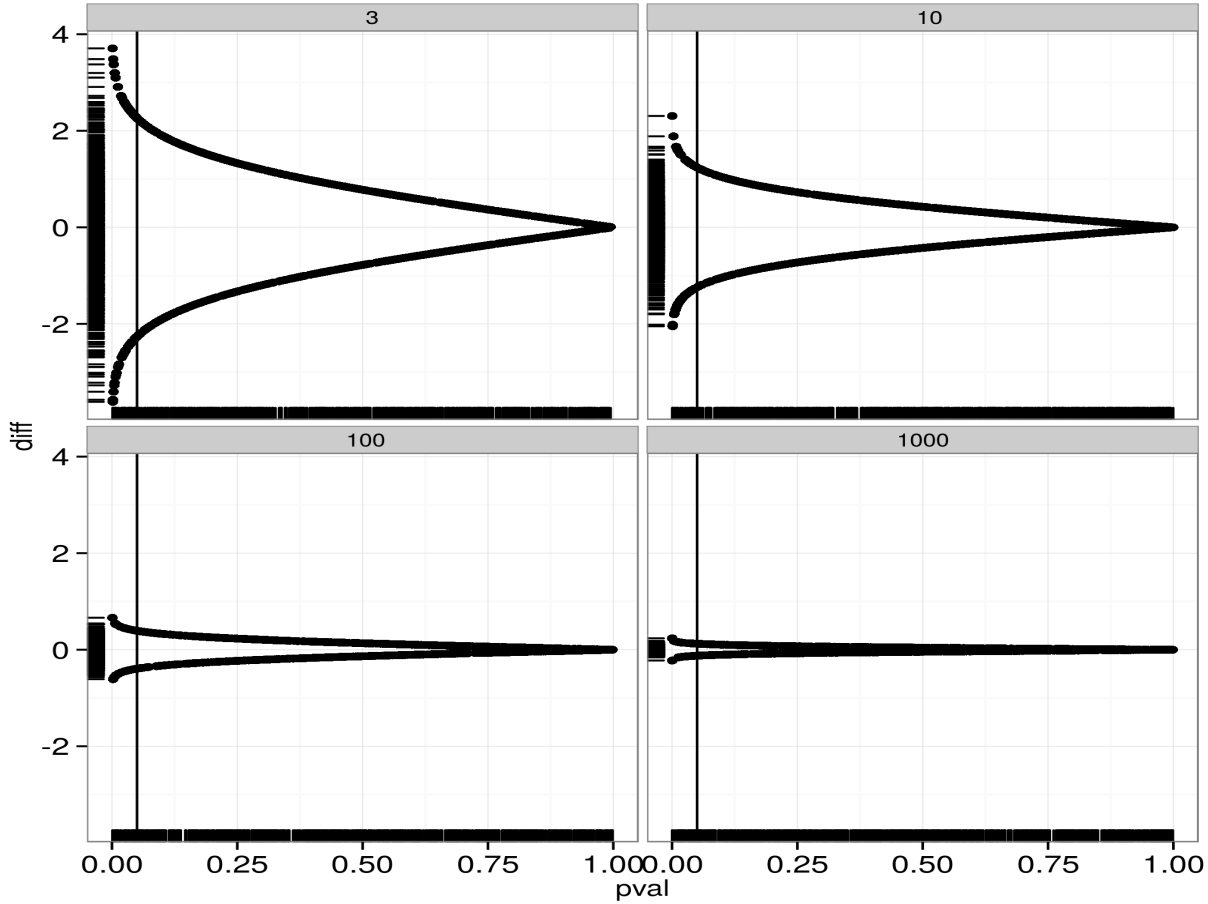
Simulation Visualization

Distributions of the pvals and the difference per sample size

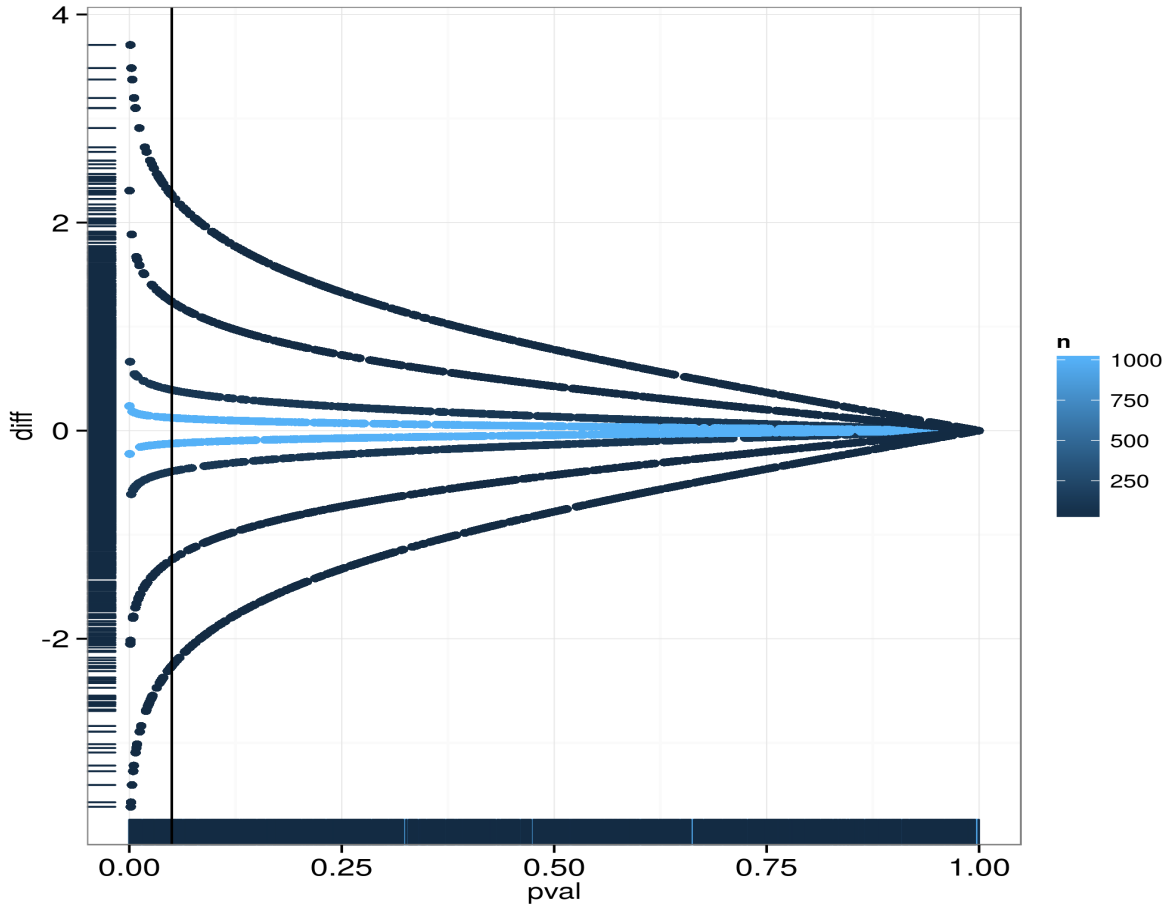


Simulation Visualization

Distributions of the pvals and the difference per sample size



Simulation Visualization



Example

- let the distribution under H_0 be $X \sim N(0, 1)$
- alternative:

$$H_0 : X \leq 0$$

- test statistic $Z = -1.2$
- what is the p-value?
- what is the p-value of the respective two tailed test?

Table of Contents I

t-tests

- if σ is unknown (this will be the general case) we must estimate it
- because of this step, the test statistic will no longer follow a normal distribution
- so e.g. in a one sample t-test: test (sample mean against a population mean)

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

where $\bar{x}$ is the sample mean, s is the sample standard deviation and n is the sample size. The degrees of freedom used in this test is $n - 1$

t-tests

A t-test is any statistical hypothesis test in which the test statistic follows a Student's t distribution if the null hypothesis is supported.

One Sample t-test

```
> set.seed(1)
> x <- rnorm(12)
> t.test(x,mu=0) ## population mean 0

One Sample t-test

data:  x
t = 1.1478, df = 11, p-value = 0.2754
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.2464740  0.7837494
sample estimates:
mean of x
0.2686377
```

One Sample t-test

```
> t.test(x,mu=1) ## population mean 1

One Sample t-test

data:  x
t = -3.125, df = 11, p-value = 0.009664
alternative hypothesis: true mean is not equal to 1
95 percent confidence interval:
 -0.2464740  0.7837494
sample estimates:
mean of x
0.2686377
```

Example/Exercise

- males, aged 45-54;
- population: systolic blood pressure normally distributed: $X \sim N(131, 12)$
- new sample: male n=85, mean=128
- what would be the appropriate test?
- state the hypothesis, calculate the statistic and the p-value!

Example/Exercise

Assume now σ is unknown and s is calculated from the sample data and found to be 12 mm Hg. Follow the same hypothesis testing process as we did in before:

- state H_0, H_A, and α
- calculate the test statistic
- calculate the p-value
- conclusion?
- compare with the result from the z-test above.

Values

In null hypothesis testing one should report

- the distribution under H_0 (often implicitly)
- the test statistic
- the p-value
- the effect size

Confidence Intervals and Hypothesis Testing

- there is a fundamental connection between p-values and confidence intervals
- if we can reject H_0 for a two sided hypothesis test using significance level α, μ_0 will not be contained in a confidence interval for μ using a confidence level $1 - \alpha$
- the confidence interval is calculated:
  $CI = estimate \pm t_{\alpha/2} \times se$

Confidence Intervals and Hypothesis Testing

- the confidence interval is calculated:
  $CI = estimate \pm t_{\alpha/2} \times se$
- calculate the confidence interval for the mean in the exercise above!

Inference for Two Population Means

- comparing two means (more: ANOVA)
- in null hypothesis testing the null hypothesis would be:

$$H_0 : \mu_x - \mu_y = D_0 (= 0)$$

- the two-tailed alternative would be:

$$H_A : \mu_x - \mu_y \neq D_0 (= 0)$$

- the lower-tailed alternative would be:

$$H_A : \mu_x - \mu_y < D_0 (= 0)$$

Inference for Two Population Means

There are three possible cases

1. paired samples
2. population variances are equal
3. population variances are unequal

Two Sample t-tests

There are two ways to perform a two sample t-test in R:

- given two vectors `x` and `y` containing the measurement values from the respective groups `t.test(x,y)`
- given one vector `x` containing all the measurement values and one vector `g` containing the group membership $t.test(x \sim g)$ (read: x dependend on g)

Two Sample t-tests: two vector syntax

```
> set.seed(1)
> x <- rnorm(12)
> y <- rnorm(12)
> g <- sample(c("A","B"),12,replace = T)
> t.test(x,y)

Welch Two Sample t-test

data:  x and y
t = 0.5939, df = 20.012, p-value = 0.5592
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5966988  1.0717822
sample estimates:
 mean of x  mean of y
0.26863768 0.03109602
```

Two Sample t-tests: formula syntax

```
> t.test(x ~ g)

Welch Two Sample t-test

data:  x by g
t = -0.6644, df = 6.352, p-value = 0.5298
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.6136329  0.9171702
sample estimates:
mean in group A mean in group B
      0.1235413       0.4717726
```

Paired t-test

- appropriate for experimental designs in which pairs of identical or highly similar experimental units are assigned to distinct treatments or in before - after cases
- matched pairing helps to account for confounding
- test statistic

$$t = \frac{\bar{x_D} - D_0}{S_D/\sqrt{n}}$$

- so we see it reduces to a one sample test of the differences against a population difference

Example/Exercise

- load the `asbio` package and load the data set `sc.twins` by typing:

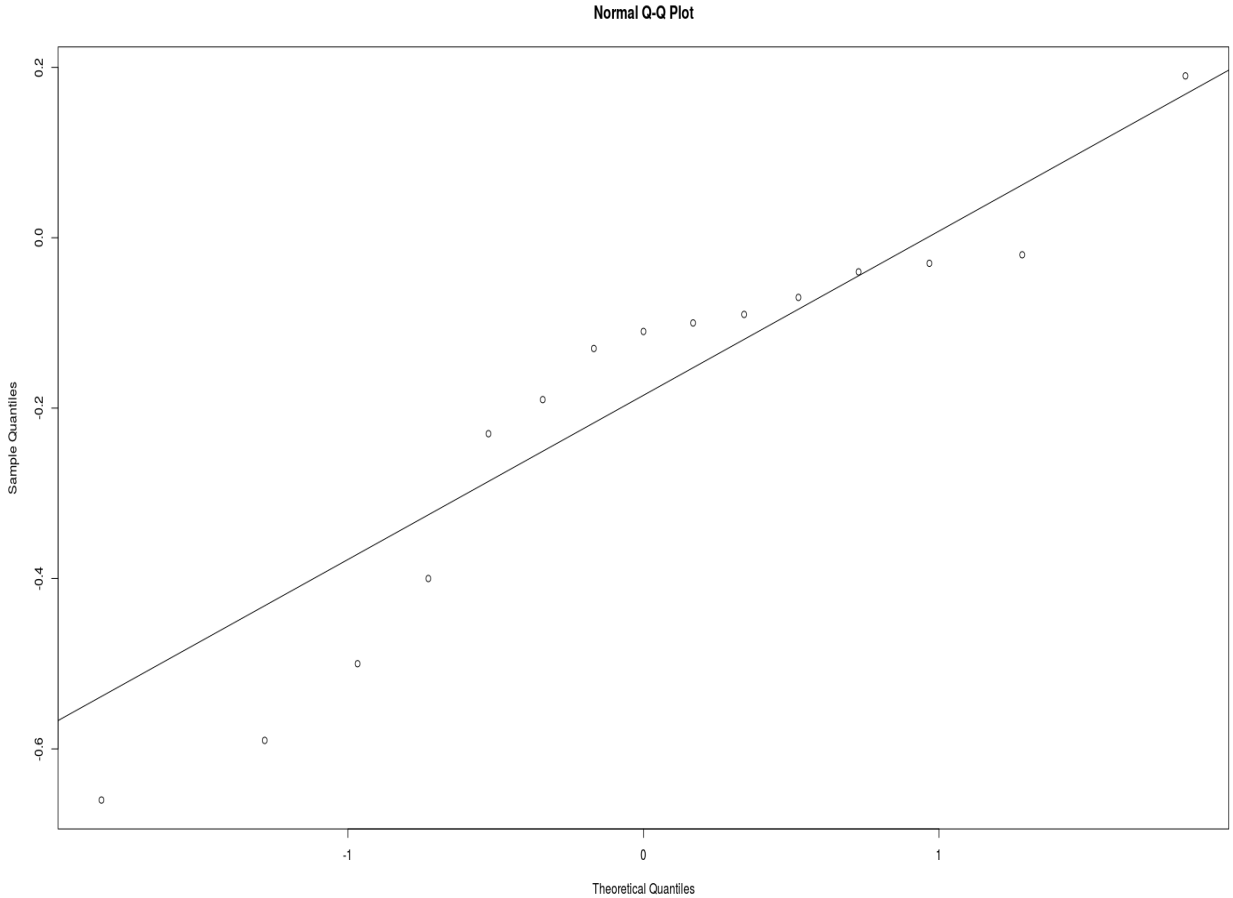
  `> data(sc.twin)`
- look at the help of the data and t.test (?sc.twin, ?t.test)
- perform a paired t-test
- remember the assumptions of a t-test... what I have to check?

Example/Exercise

- checking for normality visually:

```
> sc.twin$diff <- sc.twin$affected - sc.twin$unaffected
> qqnorm(sc.twin$diff)
> qqline(sc.twin$diff)
```

Example/Exercise



Example/Exercise

- checking for normality by a test:
- Shapiro-Wilk test often used:

```
> shapiro.test(sc.twin$diff)

Shapiro-Wilk normality test

data:  sc.twin$diff
W = 0.9041, p-value = 0.1099
```

- the test is not significant but the graphic and the test in combination indicate deviation from normality, so a non-parametric test may be more appropriate (we will run this test in the non-parametric tests section)

Pooled Variance t-test

- null hypothesis: $H_0 : \mu_1 = \mu_2$
- two-tailed alternative: $H_0 : \mu_1 \neq \mu_2$
- upper-tailed: $H_0 : \mu_1 > \mu_2$
- assumptions:
  1. parent distributions are normally distributed
  2. observations are independent
  3. the parent distributions underlying the treatments have equal variances (there are tests, e.q. Fligner-Killeen `fligner.test()`)
- the degrees of freedom are $n_1 + n_2 - 2$

Welch Approximate t-Test

- corrects the degrees of freedom (that is why there are often noninteger degrees of freedom)
- assumptions:
  1. parent distributions are normally distributed
  2. observations are independent

Student/Welch in R

- if not stated otherwise `t.test()` will not assume that the variances in the both groups are equal
- if one knows that both populations have the same variance set the `var.equal` argument to TRUE to perform a student's t-test

Student's t-test

```
> t.test(x, y, var.equal = T)

Two Sample t-test

data:  x and y
t = 0.5939, df = 22, p-value = 0.5586
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5918964  1.0669797
sample estimates:
 mean of x  mean of y
0.26863768 0.03109602
```

t-test

- the t-test, especially the Welch test is appropriate whenever the underlying values are normally distributed
- it is also recommended for group sizes ≥ 30 (robust against deviation from normality)
- assumes independence of observations

What we've learned

- if we talk about testing you should know the definition of the following terms, (a) for a one sample (b) for a two sample t-test
  - Null hypothesis
  - Alternative hypothesis
  - Test statistic
  - Significance level
  - Critical value
  - Decision rule
  - Type I error
  - Type II error

Table of Contents I

four possible situations

		Situation	
		H_0 is true	H_0 is false
Conclusion	H_0 is not rejected	Correct decision	Type II error (β)
	H_0 is rejected	Type I error (α)	Correct decision

Power analysis t-test

- in a power analysis desired experimental characteristics (effect size, sample size) are considered in context of α and β
- post hoc power analysis are generally considered inappropriate
- Power $= 1 - \beta \propto \frac{E\alpha\sqrt{n}}{\sigma}$ where E is the true effect size

Power analysis Example

Logging and insects

- research hypothesis: logging decreases species richness by at least 5 percent
- $\alpha = 0.05$
- $1 - \beta = 0.8$
- σ is assumed to be 10 percent, also normality
- What happens when s is unknown

Power analysis Example

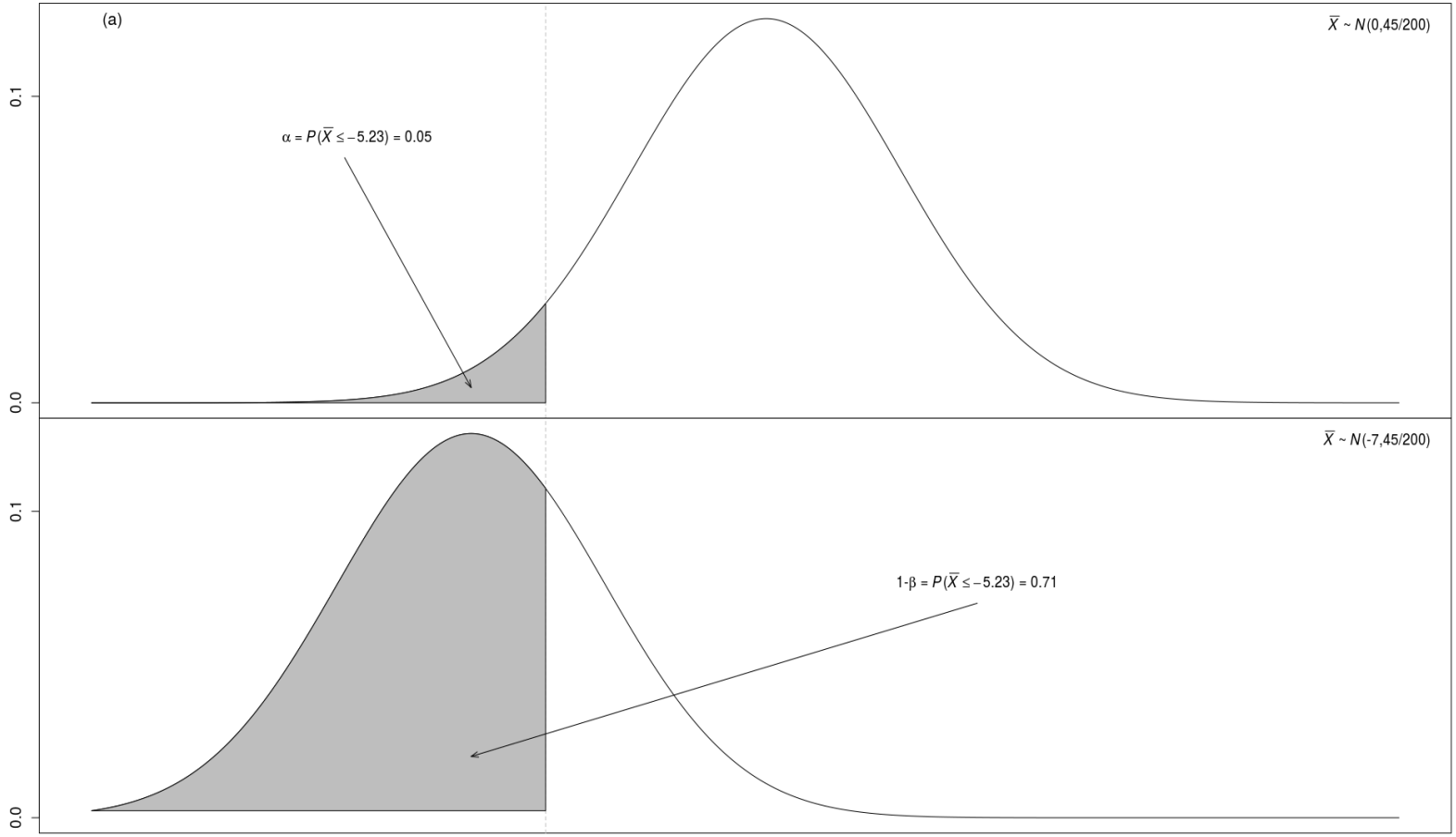
Logging and insects

```
> power.t.test(sd = 10,power = 0.8,  delta = 5)

     Two-sample t test power calculation

              n = 63.76576
          delta = 5
             sd = 10
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

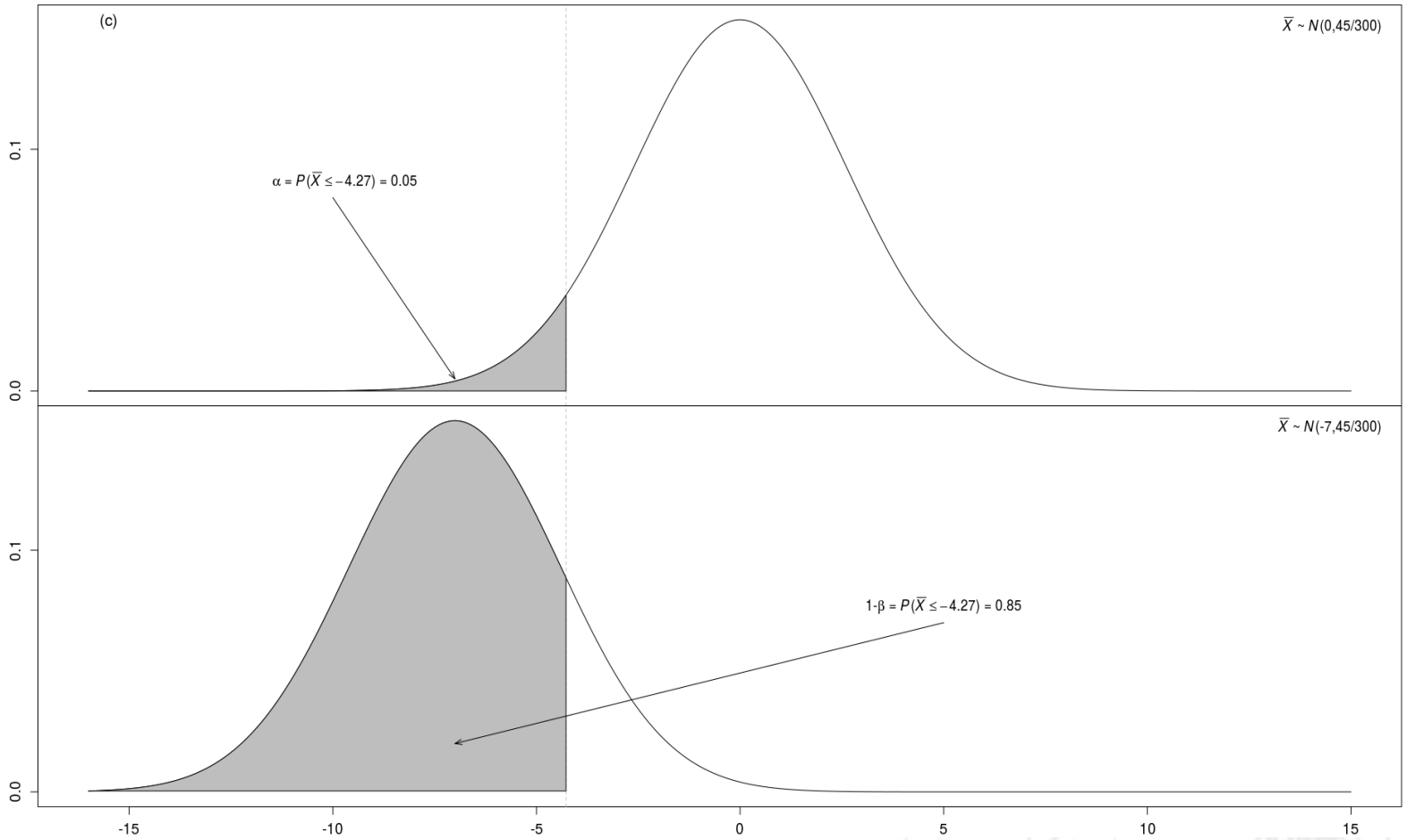
Power Illustration (normal distributed example)

- research hypothesis: smoking reduces Alzheimer by 7
- $\alpha = 0.05$
- $1 - \beta = 0.8$
- assume σ to be 45
- assume normality
- are $n = 200$ sufficient?

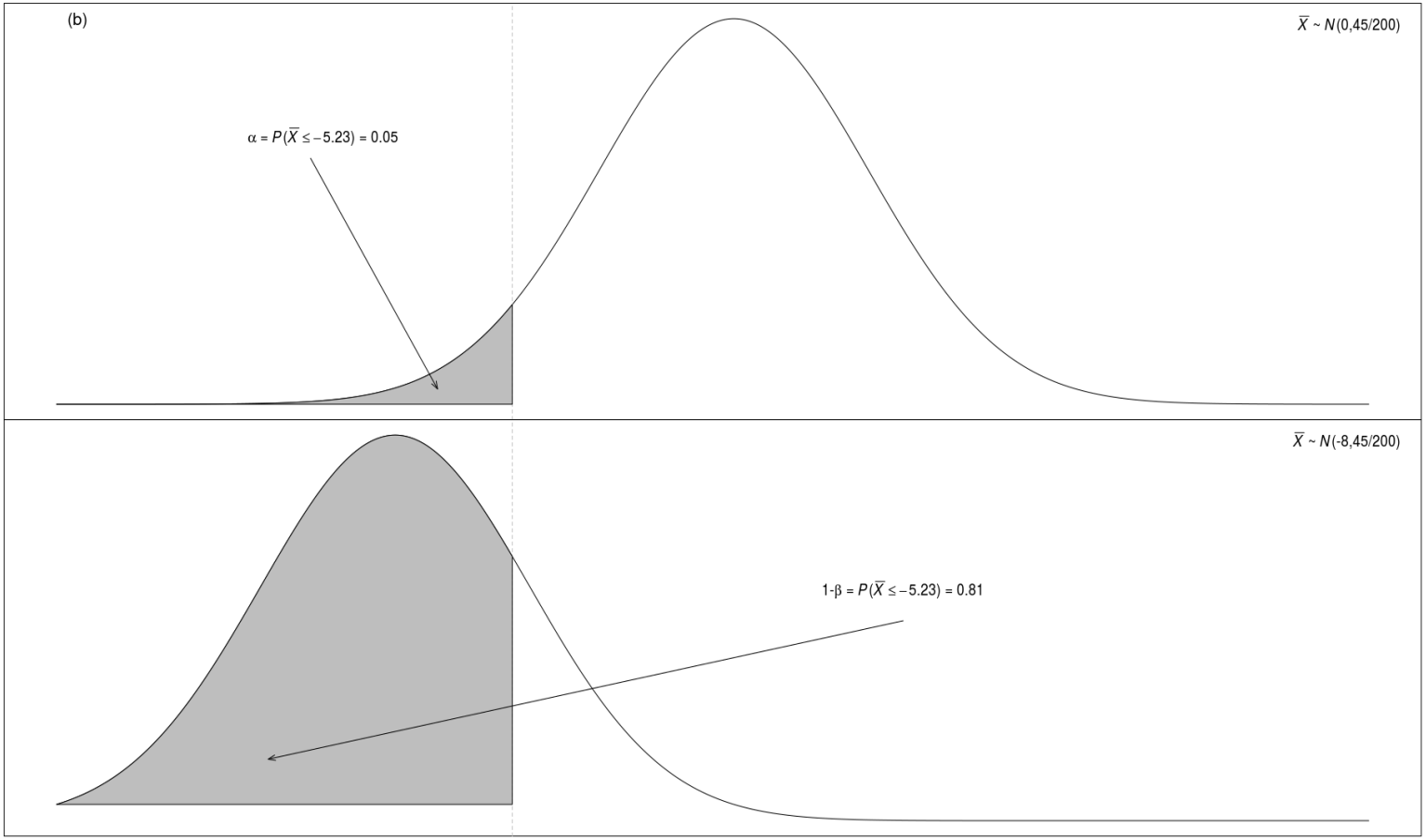
Power Visualization



Power Visualization: increase n



Power Visualization: increase effect size



Power Visualization: increase α

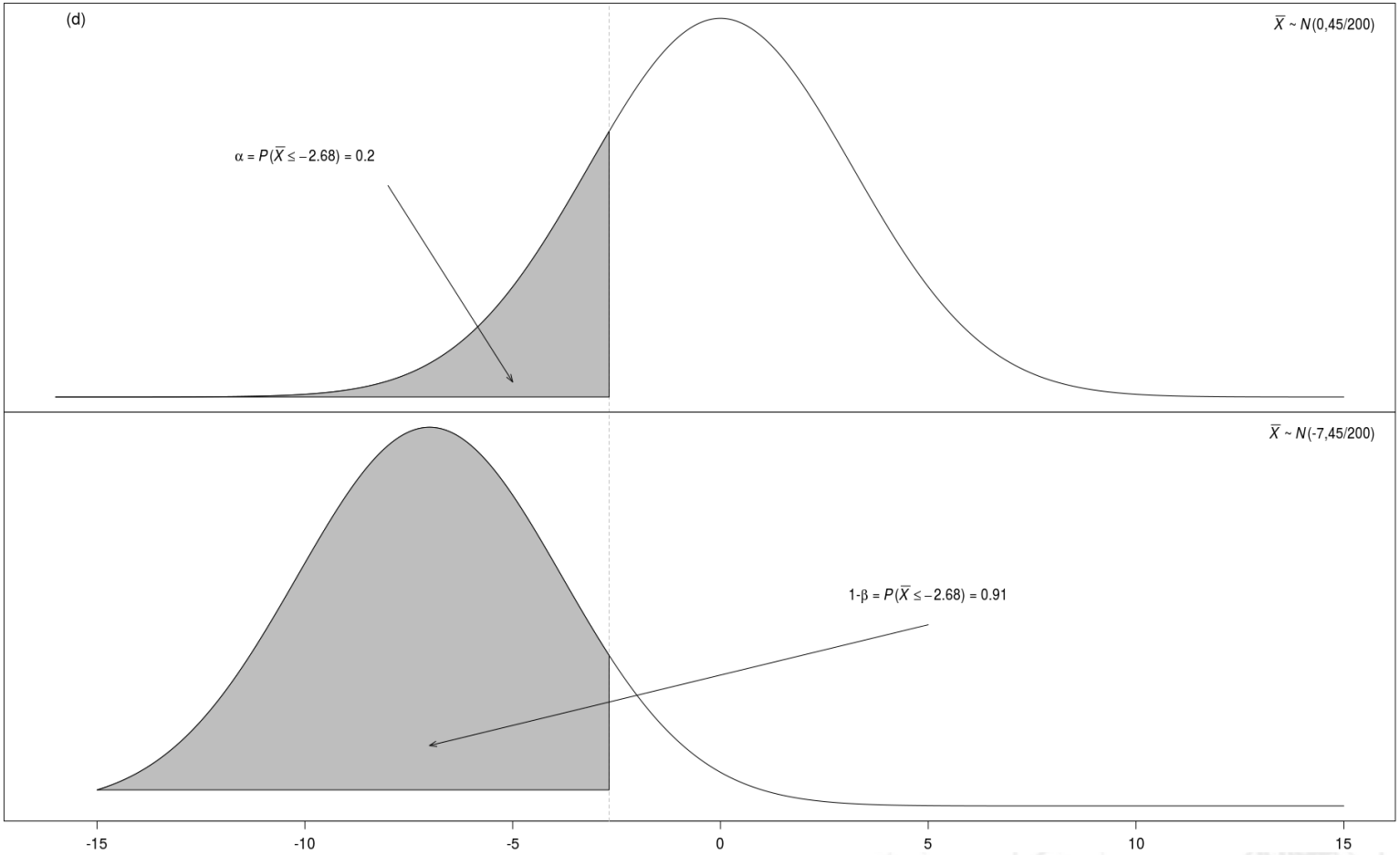


Table of Contents I

Excel Files

- there is a package for reading excel files: `readxl`
- the command is `read_excel()`

Exercise:

1. there is again a data directory which contains the data, download it (if you have not done it already)
2. use something like

```
> xx <- read_excel("data/datencaroline.xlsx")
```

Exercises I

1. use a t-test to compare `speichermodul` according to `katalysator`, visualize it.
2. Interpret the result.
3. What is the problem?

Exercises I

1. use the following code to do the test on every subset `konzentration` and `sample`, try to figure what is happening to understand the result:

```
data.l <- split(xx,list(xx$konzentration,xx$sample),drop=T)
tmp.l <- lapply(data.l,function(x) {
    tob <- t.test(x$speichermodul ~ x$katalysator)
    tmp <- data.frame(
        konz = unique(x$konzentration),
        samp = unique(x$sample),
        mean.group.1 = tob$estimate[1],
        mean.group.2 = tob$estimate[2],
        name.test.stat = tob$statistic,
        conf.lower = tob$conf.int[1],
        conf.upper = tob$conf.int[2],
        pval = tob$p.value,
        alternative = tob$alternative,
        tob$method)})
res <- Reduce(rbind,tmp.l)
```

2. make plots to visualize the results.
3. how many tests have a statistically significant result? How many did you expect? Is there a tendency?

Exercises - Solutions

- how many tests have an statistically significant result?
- How many did you expect?

```
> table(res$pval < 0.05)

FALSE  TRUE
    7     3
> nrow(res) * 0.05
[1] 0.5
> nrow(res)
[1] 10
```