

TMA4267 Linear statistical models

Spring 2026

Thea Bjørnland

Part 2, Chapters 6, 7, 8, 9, 10

Tentative plan

Week	Tuesday	Thursday	Friday
6		Chapter 6 and R	Assignments 1, 2, 3, Install R and RStudio
7	Chapter 7 (7.3, 7.6, 7.7)	Assignment 4 Room VE22 - SMASH	
8	Chapter 7/8	Chapter 7/8	Cancelled*
9	Chapter 7/8	Assignment 5	Cancelled*
10	Chapter 9	Chapter 9	Cancelled*
11	Chapter 10	Assignment 6 <i>Main building (Hovedbygningen) 326</i>	Cancelled*

* Due to no-show, contact us if you need extra help

Repetition

$$Y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \varepsilon_i \quad Y = X\beta + \varepsilon$$

$$H_0 : C\beta = 0$$

$$H_1 : C\beta \neq 0$$

Call:

```
lm(formula = y ~ x1 + x2)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.9807	-5.1992	-0.0145	3.4825	9.9923

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.8972	3.5586	0.252	0.803971
x1	1.9854	0.4060	4.890	0.000138 ***
x2	-3.9486	2.0082	-1.966	0.065821 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.938 on 17 degrees of freedom

Multiple R-squared: 0.6491, Adjusted R-squared: 0.6078

F-statistic: 15.72 on 2 and 17 DF, p-value: 0.0001363

$$C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Repetition

$$Y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \varepsilon_i \quad Y = X\beta + \varepsilon$$

$$H_0 : C\beta = 0$$

$$H_1 : C\beta \neq 0$$

Call:

```
lm(formula = y ~ x1 + x2)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.9807	-5.1992	-0.0145	3.4825	9.9923

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.8972	3.5586	0.252	0.803971
x1	1.9854	0.4060	4.890	0.000138 ***
x2	-3.9486	2.0082	-1.966	0.065821 .

$$C = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.938 on 17 degrees of freedom

Multiple R-squared: 0.6491, Adjusted R-squared: 0.6078

F-statistic: 15.72 on 2 and 17 DF, p-value: 0.0001363

8.7 LIKELIHOOD RATIO TESTS

The tests in Sections 8.1, 8.2, and 8.4 were derived using informal methods based on finding sums of squares that have chi-square distributions and are independent. These same tests can be obtained more formally by the likelihood ratio approach. Likelihood ratio tests have some good properties and sometimes have optimal properties.

Theorem 8.7b. If $\mathbf{y}$ is $N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, then the F test for $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ in Theorem 8.4b is equivalent to the likelihood ratio test.

Hypothesis testing: some basics

Performing a frequentist hypothesis test in practice [\[edit \]](#)

The typical steps involved in performing a frequentist hypothesis test in practice are:

1. Define a hypothesis (claim which is testable using data).
2. Select a relevant statistical test with associated [test statistic](#) T .
3. Derive the distribution of the test statistic under the null hypothesis from the assumptions.
In standard cases this will be a well-known result. For example, the test statistic might follow a [Student's t distribution](#) with known degrees of freedom, or a [normal distribution](#) with known mean and variance.
4. Select a significance level (α), the maximum acceptable [false positive rate](#). Common values are 5% and 1%.
5. Compute from the observations the observed value t_{obs} of the test statistic T .
6. Decide to either reject the null hypothesis in favor of the alternative or not reject it. The [Neyman-Pearson](#) decision rule is to reject the null hypothesis H_0 if the observed value t_{obs} is in the critical region, and not to reject the null hypothesis otherwise.^{[\[31\]](#)}

Multiple testing

Data dredging (also known as **data snooping** or ***p*-hacking**)^{[1][a]} is the misuse of **data analysis** to find patterns in data that can be presented as **statistically significant**, thus dramatically increasing and understating the risk of **false positives**. This is done by performing many **statistical tests** on the data and only reporting those that come back with **significant** results.^[2]

