

# Sample tasks that can appear in the final test

## Disclaimer

The aim of this document is to introduce the various type of tasks and questions that can appear in the final test. Please note, that this is not a complete list of possible questions and tasks that can appear in the final test. To prepare for the final test we recommend to be able to answer the questions asked during the exercises classes and in the homework assignments.

## Task 1

Suppose that a linear model was developed to model a glucose concentration in blood ( $Y$ ) after administering  $u$  units of a medication to a patient whose gender is  $g$  (either male or female) and weight is  $w$  in kilograms. Suppose that in the developed model, the effect of both the weight ( $w$ ) and the dose of the medication ( $u$ ) on the glucose concentration ( $Y$ ) appeared to be different for male and female. Further, it appeared that an increase of dose of the medication by one unit has, for persons of given gender and common weight, the same effect on the glucose concentration irrespective of the weight of those persons. Finally, suppose that all variables entered the model without any transformation.

- (i) Specify the regression function of the considered linear model. If you introduce new symbols, provide their correct mathematical definition.
- (ii) Suppose that the first two observations are based on the following two patients: 1. a man of 80 kg who was administered with 10 units of the medication, 2. a women of 60 kg who was administered with 8 units of the medication. Provide the first two rows of the model matrix.
- (iii) Consider symbols in which the model above is specified and write a null and an alternative hypotheses of the test which can be used to show that the effect of the weight ( $w$ ) on the glucose concentration is different for males and females.
- (iv) Consider symbols in which the model above is specified and define a parameter that expresses the mean difference in the glucose concentration of two male patients whose weighs are 80 kg (man #1) and 70 kg (man # 2), respectively, and who were administered with 10 (man #1) and 8 (man #2) units of the medication, respectively.

## Task 2

Suppose that a linear model (further referred to as model  $M_1$ ) was estimated using data of sample size  $n = 100$  to model the bribery amount in *thousands* CZK ( $Y$ ) that a politician of level  $lev$  (categorical factor with three levels: *municipal*, *regional*, *national*) obtains for assigning a contract amounting to  $x \times 100\,000$  CZK related to domain  $dom$  (categorical factor with two levels: *IT Service*, *constructions*). Suppose that the reference group reparameterizing pseudocontrasts with the first level as a reference (`contr.treatment` in R software) were used when fitting the model and the following least squares estimates of the regression coefficients were obtained (“:” denotes the interaction term):

Parameter	Estimate
$\beta_0$ - Intercept	5
$\beta_1$ - Level( <i>regional</i> )	10
$\beta_2$ - Level( <i>national</i> )	20
$\beta_3$ - Contract amount	1
$\beta_4$ - Domain( <i>constructions</i> )	-5
$\beta_5$ - Level( <i>regional</i> ) : Contract amount	5
$\beta_6$ - Level( <i>national</i> ) : Contract amount	-5
$\beta_7$ - Level( <i>regional</i> ) : Domain( <i>constructions</i> )	5
$\beta_8$ - Level( <i>national</i> ) : Domain( <i>constructions</i> )	30

In the following, let  $\beta = (\beta_0, \beta_1, \dots, \beta_8)^\top$ .

- (i) Suppose that the first two observations were based on data of the following characteristics: 1. contract of 500 000 CZK for *IT services* assigned on a *municipal* level, 2. contract of 10 000 000 CZK for *constructions* assigned on a *regional* level. Provide the first two rows of the model matrix of model  $M_1$ .
- (ii) Find a function of the regression coefficients  $\beta$  which defines parameter  $\theta_1$  that provides quantification of change of the expected bribe amount for contracts being decided on
  - (1) a *municipal* level ( $\theta_1$ ),
  - (2) *regional* level ( $\theta_2$ ),
  - (3) *national* level ( $\theta_3$ ),

when the contract amount increases by 100 000 CZK and the contract stays in the same domain (either *IT Services* or *constructions*). Calculate the estimates of  $\theta_1$ ,  $\theta_2$  and  $\theta_3$ .

- (iii) Find a function of the regression coefficients  $\beta$  which defines parameter  $\theta_4$  that provides quantification of the expected difference between the bribe amounts needed on *national* and *regional* level, respectively (difference “*national* minus *regional*”) for contracts on *constructions* amounting to 10 000 000 CZK. Calculate the estimate of  $\theta_4$ .
- (iv) Find the function of regression coefficients  $\beta$  which defines parameter  $\theta_5$  that provides quantification of the difference in the expected bribe amounts for the following two contracts. The first contract amounts to 1 000 000 CZK and it is for *constructions* on a *regional* level. The second contract amounts to 2 000 000 CZK and it is for *IT services* on a *national* level. Calculate the estimate of  $\theta_5$ .

### Task 3

We are interested in modeling the associate professor's salary (`salary.assoc`) given an information about the university type they are affiliated to (`type` - categorical variable with three levels), the number of professors (`n.prof`), the number of associate professors (`n.assoc`) and the number of assistant professors (`n.assist`).

Considering the given dataset the following model was fitted:

$$\text{salary.assoc} \sim (\text{type} + \text{n.prof} + \text{n.assoc} + \text{n.assist})^2$$

which we denote as `m1`.

In order to judge the effect of considered variables in the model we obtain the ANOVA table of type I

Analysis of Variance Table

Response: `salary.assoc`

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
<code>type</code>	2	1825984	912992	342.0239	< 2.2e-16
<code>n.prof</code>	1	129544	129544	48.5297	5.564e-12
<code>n.assoc</code>	1	1775	1775	0.6649	0.415009
<code>n.assist</code>	1	3563	3563	1.3349	0.248179
<code>type:n.prof</code>	2	513550	256775	96.1928	< 2.2e-16
<code>type:n.assoc</code>	2	223511	111756	41.8658	< 2.2e-16
<code>type:n.assist</code>	2	32569	16285	6.1006	0.002317
<code>n.prof:n.assoc</code>	1	23936	23936	8.9671	0.002810
<code>n.prof:n.assist</code>	1	9792	9792	3.6682	0.055717
<code>n.assoc:n.assist</code>	1	25785	25785	9.6597	0.001931
Residuals	1110	2963012	2669		

and also ANOVA tables of type II and III:

Anova Table (Type II tests)

Response: `salary.assoc`

	Sum Sq	Df	F value	Pr(>F)
<code>type</code>	122160	2	22.8818	1.828e-10
<code>n.prof</code>	92992	1	34.8365	4.759e-09
<code>n.assoc</code>	18968	1	7.1057	0.007795
<code>n.assist</code>	44045	1	16.4999	5.207e-05
<code>type:n.prof</code>	86497	2	16.2017	1.160e-07
<code>type:n.assoc</code>	37467	2	7.0179	0.000936
<code>type:n.assist</code>	57827	2	10.8316	2.194e-05
<code>n.prof:n.assoc</code>	50555	1	18.9387	1.474e-05
<code>n.prof:n.assist</code>	1553	1	0.5816	0.445830
<code>n.assoc:n.assist</code>	25785	1	9.6597	0.001931
Residuals	2963012	1110		

Anova Table (Type III tests)

Response: `salary.assoc`

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	4201754	1	1574.0561	< 2.2e-16
<code>type</code>	110598	2	20.7161	1.469e-09
<code>n.prof</code>	88772	1	33.2556	1.047e-08
<code>n.assoc</code>	2682	1	1.0048	0.316368
<code>n.assist</code>	54873	1	20.5566	6.418e-06
<code>type:n.prof</code>	86497	2	16.2017	1.160e-07
<code>type:n.assoc</code>	37467	2	7.0179	0.000936
<code>type:n.assist</code>	57827	2	10.8316	2.194e-05
<code>n.prof:n.assoc</code>	50555	1	18.9387	1.474e-05
<code>n.prof:n.assist</code>	1553	1	0.5816	0.445830
<code>n.assoc:n.assist</code>	25785	1	9.6597	0.001931
Residuals	2963012	1110		

Using the previous outputs make your decision regarding the following questions. Support your statements with the corresponding  $F$ -statistic value and  $P$ -value. If the corresponding values are not provided in the tables above, try to guess what could be an appropriate result using values given in the tables.

- (i) Estimate the residual variance in model `m1`.
- (ii) Is it reasonable to assume instead of model `m1` its submodel where we only omit the interaction term between the number of professors and the number of associate professors (`n.prof:n.assoc`)?
- (iii) Is it reasonable to assume instead of model `m1` its submodel where we only omit the interaction term between the number of professors and the number of assistant professors (`n.prof:n.assist`)?
- (iv) Is it reasonable to assume only the following model `salary.assoc ~ type + n.prof + n.assoc` as a submodel of `salary.assoc ~ type + n.prof + n.assoc + n.assist`?

## Task 4

We are interested in the number of fires (`fire`) in some specific town district given the percentage of minority decreased by 25 (`minor25`) and the location of the district (covariate `fside` with two levels - North and South). Considering the given data we obtain the following model:

```
log(fire) ~ minor25 * fside.
```

The north district `fside = North` was considered to be a reference category (`contr.treatment` in R). The following summary table was obtained:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	2.218328	0.105162	21.094	< 2e-16	***
minor25	0.025328	0.003914	6.471	7.54e-08	***
fsideSouth	-0.191788	0.176365	-1.087	0.2829	
minor25:fsideSouth	-0.012590	0.005267	-2.390	0.0213	*

---

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

Residual standard error: 0.5224 on 43 degrees of freedom  
Multiple R-squared: 0.5686, Adjusted R-squared: 0.5385  
F-statistic: 18.89 on 3 and 43 DF, p-value: 5.768e-08

Using the output above answer the following questions.

- (i) Compare the number of fires in two different districts however, with the same percentage of minority where one district is located at the north side of the town and the other is located at the south part.
- (ii) Describe the effect of percentage of minority on the expected number of fires.
- (iii) Is the district locality (`fside`) statistically significant modifier of the effect of percentage of minority on the number of fires?
- (iv) Describe the effect of district location (`fside`) on the number of fires.
- (v) Interpret the absolute term in the model (Intercept).

## Task 5

We would like to model the logarithm of the number of fires (`fire`) in some town district given the logarithm of the percentage of minority and the locality where the district is situated (variable `fside` with two levels - North and South). Using the given dataset the following model was obtained:

```
log(fire) ~ log(minor) * fside.
```

For the modeling purposes we assigned the level North with value +1 and the level South with value -1 (`contr.sum` in R). The following summary table was obtained:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	1.15598	0.13491	8.568	7.54e-11	***
log(minor)	0.41486	0.04164	9.963	9.65e-13	***
fside1	-0.13982	0.13491	-1.036	0.3058	
log(minor):fside1	0.10395	0.04164	2.497	0.0164	*

---

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

Residual standard error: 0.4149 on 43 degrees of freedom  
Multiple R-squared: 0.7278, Adjusted R-squared: 0.7088  
F-statistic: 38.33 on 3 and 43 DF, p-value: 3.234e-12

Using the output above answer the following questions.

- (i) Estimate the average effect (an average over both possible district locations) of the logarithm of the percentage of minority on the logarithm of the number of fires.
- (ii) Describe the effect of the percentage of minority (`minor`) on the number of fires.
- (iii) Is the district location (`fside`) statistically significant modifier of the effect of the ratio of minors on the number of fires?

## Task 6

We model the amount of yield (yield) on the magnesium content (Mg) and calcium content (Ca). For this purpose we introduce the following covariates: lMg - logarithm of Mg, lMg2 - (logarithm of Mg) squared, lCa - logarithm of Ca. We build the following model

$$\text{yield} \sim \text{lMg} + \text{lMg}^2 + \text{lCa} ,$$

which was estimated from the collected data and we get the following results:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-16.9375	7.5351	-2.248	0.025186	*
lMg	18.5873	6.1000	3.047	0.002479	**
lMg2	-3.3544	1.2329	-2.721	0.006824	**
lCa	-0.8888	0.2581	-3.444	0.000641	***

---

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

Residual standard error: 1.045 on 364 degrees of freedom  
Multiple R-squared: 0.09675, Adjusted R-squared: 0.0893  
F-statistic: 13 on 3 and 364 DF, p-value: 4.408e-08

Answer the following questions with the help of the above output.

- (i) Describe the effect of the logarithm of calcium content on the yield. If possible estimate the partial effect of the logarithm of calcium content on the yield given the magnesium content.
- (ii) Describe the effect of the logarithm of magnesium content on the yield. If possible estimate the partial effect of the logarithm of magnesium content on the yield given the calcium content.
- (iii) Let  $U_i$  ( $i = 1, \dots, n$ ) stand for the the  $i$ -th residual (i.e.  $U_i = Y_i - \hat{Y}_i$ ), where  $Y_i$  is the yield for the  $i$ -th observation). Calculate the following quantities  $\sum_{i=1}^n U_i$  (i.e. sum of residuals) and  $\sum_{i=1}^n U_i^2$  (sum of residuals squared).

## Task 7

We model the amount of yield (`yield`) on the magnesium content (`Mg`) and calcium content (`Ca`). For this purpose we introduce the following covariates: `lMg` - logarithm of `Mg`, `lMg2` - (logarithm of `Mg`) squared, `lCa` - logarithm of `Ca`. We build the following model

```
yield ~ lMg*lCa ,
```

which was estimated from the collected data and we get the following results:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-23.653	11.878	-1.991	0.04719	*
lMg	13.051	4.841	2.696	0.00735	**
lCa	6.000	3.025	1.984	0.04802	*
lMg:lCa	-2.801	1.226	-2.285	0.02288	*

---

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

Residual standard error: 1.049 on 364 degrees of freedom

Multiple R-squared: 0.09141, Adjusted R-squared: 0.08392

F-statistic: 12.21 on 3 and 364 DF, p-value: 1.254e-07

Answer the following questions with the help of the above output.

- (i) Describe the effect of the logarithm of calcium content on the yield. If possible estimate the partial effect of the logarithm of calcium content on the yield given the magnesium content.
- (ii) Describe the effect of the logarithm of magnesium content on the yield. If possible estimate the partial effect of the logarithm of magnesium content on the yield given the calcium content.

## Task 8

We want to model the associate professor's salary (`salary.assoc`) given the number of professors (`n.prof`), the number of associate professors (`n.assoc`), the number of assistant professors (`n.assist`) and the university type (`ntype`). The `ntype` covariate has three levels (I, IIA, IIB) but these levels were in the corresponding regression matrix column replaced with numbers 0, 1, and 2.

The following linear regression model was considered:

```
salary.assoc ~ (n.prof + n.assoc + n.assist)*ntype ,
```

and considering the given dataset the following summary output was obtained:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	470.850049	6.936154	67.883	< 2e-16 ***
n.prof	0.192696	0.033174	5.809	8.2e-09 ***
n.assoc	0.008545	0.063948	0.134	0.893728
n.assist	-0.317322	0.085452	-3.713	0.000215 ***
ntype	-59.871714	3.799160	-15.759	< 2e-16 ***
n.prof:ntype	-0.060086	0.047579	-1.263	0.206902
n.assoc:ntype	0.449865	0.090156	4.990	7.0e-07 ***
n.assist:ntype	0.182415	0.089014	2.049	0.040667 *

---

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

Residual standard error: 54.61 on 1117 degrees of freedom  
Multiple R-squared: 0.421, Adjusted R-squared: 0.4174  
F-statistic: 116 on 7 and 1117 DF, p-value: < 2.2e-16

Using the output above answer the following questions:

- (i) Compare the associate professor's salary with respect to different university types while considering the same number of professors, associate professors and assistant professors.
- (ii) Describe the effect of the number of professors (`n.prof`) on the salary of the associate professors.



## Task 9

Again, we would like to model the associate professor's salary (`salary.assoc`) given the number of professors (`n.prof`) and the university type (`type`). The university type covariate has three levels (I, IIA, IIB) and for their parametrization we are using `contr.treatment` with the reference category I.

The following linear regression model was considered

```
salary.assoc ~ (type + n.prof)^2 ,
```

and using the given dataset we obtained the following model output:

```
Call:
lm(formula = salary.assoc ~ (type + n.prof)^2, data = Data)
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    470.6      7.70     61.1 < 0.001
typeIIA        -47.0      8.87     -5.3 < 0.001
typeIIB       -127.5      8.47    -15.0 < 0.001
n.prof           0.05      0.02      2.5  0.014
typeIIA:n.prof  0.12      0.04      2.8  0.004
typeIIB:n.prof  1.3       0.10     13.0 < 0.001
```

```
Residual standard error: 54.23 on 1119 degrees of freedom
Multiple R-squared:  0.428,    Adjusted R-squared:  0.426
F-statistic: 167.5 on 5 and 1119 DF,  p-value: < 2.2e-16
```

For some better interpretation purposes we decided that we use the number of professors decreased by 40 (`n.prof40`) instead of the original covariate (`n.prof`).

Instead of fitting a new model with the new covariate is it possible to simply plug-in some values from the previous output to the table below? Replace the question marks with appropriate values where possible.

```
Call:
lm(formula = salary.assoc ~ (type + n.prof40)^2, data = Data)
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)           ?           ?           ?           ?
typeIIA               ?           ?           ?           ?
typeIIB               ?           ?           ?           ?
n.prof40              ?           ?           ?           ?
typeIIA:n.prof40     ?           ?           ?           ?
typeIIB:n.prof40     ?           ?           ?           ?
```

```
Residual standard error: ? on ? degrees of freedom
Multiple R-squared:  ?,    Adjusted R-squared:  ?
F-statistic: ? on ? and ? DF,  p-value: ?
```

## Task 10

We are analyzing the angle of the occipital bone (*oca*) on the back of a human's head and we are searching for a possible effect caused by a gender (*fgender*) and population (*fpopul*) to which the subject belongs. Gender is coded as a factor with two levels *female* and *male* and the population information is a factor with three levels ordered as *AUSTR*, *BERG* and *BURIAT*. We used a parametrization given by *contr.sum*. The following model was obtained:

```
oca ~ fgender * fpopul,
```

and considering the data at hand we obtained the following summary output:

```
lm(formula = oca ~ fgender * fpopul, data = Howells,
    contrasts = list(fpopul = "contr.sum", fgender = "contr.sum"))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	115.6667	0.3211	360.246	<2e-16 ***
fgender1	0.6167	0.3211	1.921	0.0560 .
fpopul1	-0.7542	0.4541	-1.661	0.0981 .
fpopul2	1.0958	0.4541	2.413	0.0166 *
fgender1:fpopul1	-0.7292	0.4541	-1.606	0.1097
fgender1:fpopul2	-0.5292	0.4541	-1.165	0.2451

---

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

Residual standard error: 4.974 on 234 degrees of freedom  
Multiple R-squared: 0.0697, Adjusted R-squared: 0.04982  
F-statistic: 3.507 on 5 and 234 DF, p-value: 0.004473

To remind, the contrast matrix under the assumed parametrization takes the following form:

	fgender1		fpopul1	fpopul2
female	1		AUSTR 1	0
male	-1	,	BERG 0	1
			BURIAT -1	-1

- What is an expected occipital bone angle for a men from AUSTR population?
- What is an expected occipital bone angle for a female from BURIAT population?
- Compare the occipital bone angles for a men and female considering each population separately.

## Task II

We would like to model a relationship between the occipital angle (*oca*) and two covariates: a gender (*fgender*) and population (*fpopul*) to which the subject belongs. The gender covariate is a factor with two levels *female* and *male*. The population covariate has three levels labelled as *AUSTR*, *BERG*, *BURIAT*. Using the available data the following model was fitted:

```
oca ~ fpopul * fgender
```

For the fitted model the Anova of type II table was calculated:

Anova Table (Type II tests)

Response: *oca*

	Sum Sq	Df	F value	Pr(>F)
<i>fpopul</i>	330.0	?	5.50	0.005
<i>fgender</i>	120.0	?	?	0.020
<i>fpopul:fgender</i>	160.0	?	?	0.072
Residuals	6000.0	200		

- (i) Give an explanation on how the  $p$ -value in the first line of of the Anova table is obtained (the line which corresponds to *fpopul* with  $p$ -value 0.005).
- (ii) Instead of the question marks in the Anova table above fill in the appropriate numbers.  
(2 points)
- (iii) Using the Anova of type II table above, can we conclude that the differences in mean occipital angles for males and females are statistically significantly different across different populations (using a significance level of  $\alpha = 0.05$ )? Provide the corresponding  $p$ -value and the test statistic value.

## Maximum likelihood theory (overview)

Suppose we have a random sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$  from the distribution with a density  $f(\mathbf{x}; \boldsymbol{\theta})$  with respect to a  $\sigma$ -finite measure  $\mu$  and that the density is known up to unknown  $p$ -dimensional parameter  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^\top \in \Theta$ . Let  $\boldsymbol{\theta}_X = (\theta_{X1}, \dots, \theta_{Xp})^\top$  be the true value of the parameter.

Define the *likelihood function* as

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^n f(\mathbf{X}_i; \boldsymbol{\theta})$$

and the *log-likelihood function* as

$$\ell_n(\boldsymbol{\theta}) = \log L_n(\boldsymbol{\theta}) = \sum_{i=1}^n \log f(\mathbf{X}_i; \boldsymbol{\theta}).$$

The *maximum likelihood estimator* of parameter  $\boldsymbol{\theta}_X$  is defined as

$$\hat{\boldsymbol{\theta}}_n = \arg \max_{\boldsymbol{\theta} \in \Theta} L_n(\boldsymbol{\theta}).$$

Usually we search for the maximum likelihood estimator  $\hat{\boldsymbol{\theta}}_n$  as a solution of the system of likelihood equations  $\mathbf{U}_n(\hat{\boldsymbol{\theta}}_n) \stackrel{!}{=} \mathbf{0}$ , where the random vector

$$\mathbf{U}_n(\boldsymbol{\theta}) = \sum_{i=1}^n \mathbf{U}(\mathbf{X}_i; \boldsymbol{\theta}) = \sum_{i=1}^n \frac{\partial \log f(\mathbf{X}_i; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$$

is called *the score statistic*.

Under appropriate regularity assumptions

$$\sqrt{n} (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_X) \xrightarrow[n \rightarrow \infty]{d} \mathbf{N}_p(\mathbf{0}, I^{-1}(\boldsymbol{\theta}_X)),$$

where

$$I(\boldsymbol{\theta}_X) = -\mathbf{E} \frac{\partial^2 \log f(\mathbf{X}_1; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_X}.$$

is the Fisher information matrix.

To make an inference about  $\boldsymbol{\theta}_X$  usually one needs to estimate the information matrix  $I(\boldsymbol{\theta}_X)$ . In regression context we usually use *the observed information matrix* defined at  $\hat{\boldsymbol{\theta}}_n$  which is defined as

$$\hat{I}_n = -\frac{1}{n} \frac{\partial \mathbf{U}_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top} = -\frac{1}{n} \sum_{i=1}^n \frac{\partial^2 \log f(\mathbf{X}_i; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_n}.$$

## Inference about the vector parameter $\boldsymbol{\theta}$

Suppose we want to test the null hypothesis  $H_0 : \boldsymbol{\theta}_X = \boldsymbol{\theta}_0$  against the alternative  $H_1 : \boldsymbol{\theta}_X \neq \boldsymbol{\theta}_0$ . One of the possible test is *Wald test* and it is based on the following test statistic

$$W_n = n (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)^\top \hat{I}_n (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0).$$

It can be shown that under the null hypothesis  $W_n$  converges in distribution to a  $\chi^2$ -distribution with  $p$  degrees of freedom.

The (asymptotic) confidence set for  $\theta_X$  is then constructed as

$$\{\theta; n(\widehat{\theta}_n - \theta)^\top \widehat{I}_n (\widehat{\theta}_n - \theta) \leq \chi_p^2(1 - \alpha)\},$$

where  $\chi_p^2(1 - \alpha)$  is the  $1 - \alpha$  quantile of  $\chi^2$ -distribution with  $p$  degrees of freedom.

### Inference about $\theta_{Xk}$ (the $k$ -th coordinate of $\theta_X$ )

Suppose we want to test the null hypothesis  $H_0 : \theta_{Xk} = \theta_0$  against the alternative  $H_1 : \theta_{Xk} \neq \theta_0$ . One of the possible test is *Wald test* and it is based on the following test statistic

$$T_n = \frac{\sqrt{n}(\widehat{\theta}_{nk} - \theta_0)}{\sqrt{i_n^{kk}}},$$

where  $\widehat{\theta}_{nk}$  is the  $k$ -th element of  $\widehat{\theta}_n$  and  $i_n^{kk}$  is the  $k$ -th diagonal element of  $\widehat{I}_n^{-1}$  (i.e. the **inverse** of the matrix  $\widehat{I}_n$ ). The test statistic  $T_n$  under the null hypothesis converges to a standard normal distribution  $N(0, 1)$ .

The (asymptotic) confidence interval for  $\theta_{Xk}$  is given by

$$\left(\widehat{\theta}_{nk} - \frac{u_{1-\alpha/2}\sqrt{i_n^{kk}}}{\sqrt{n}}, \widehat{\theta}_{nk} + \frac{u_{1-\alpha/2}\sqrt{i_n^{kk}}}{\sqrt{n}}\right). \quad (1)$$

### Task 12

Let  $(X_1, Y_1)^\top, \dots, (X_n, Y_n)^\top$  be independent identically distributed random vectors. Suppose that the conditional density of  $Y_1$  given  $X_1$  is

$$f_{Y|X}(y|x; \beta) = \beta x e^{-\beta xy} \mathbb{I}\{y > 0\},$$

where  $\beta > 0$  is an unknown parameter. Further suppose that the distribution of  $X_1$  does not depend on  $\beta$ .

- (i) Find the maximum likelihood estimator of  $\beta$ .
- (ii) Construct a test of the null hypothesis  $H_0 : \beta = \beta_0$  against the alternative  $H_1 : \beta \neq \beta_0$ .
- (iii) Construct a confidence interval for  $\beta$ .

### Task 13

Suppose that you observe independent random vectors  $(X_1, Y_1)^\top, \dots, (X_n, Y_n)^\top$ , such such that

$$P(Y_1 = k | X_1) = \frac{[\lambda(X_1)]^k e^{-\lambda(X_1)}}{k!}, \quad k = 0, 1, 2, \dots,$$

where  $\lambda(x) = \exp\{\beta_0 + \beta_1 x\}$  and the distribution of  $X_1$  does not depend on the unknown parameter  $\beta = (\beta_0, \beta_1)^\top$ .

- (i) Find the maximum likelihood estimator of  $\beta = (\beta_0, \beta_1)$ .
- (ii) Derive a test for the null hypothesis  $H_0 : (\beta_0, \beta_1)^\top = (0, 0)^\top$  against the alternative that  $H_1 : (\beta_0, \beta_1)^\top \neq (0, 0)^\top$ .
- (iii) Derive a test for the null hypothesis  $H_0 : \beta_1 = 0$  against the alternative that  $H_1 : \beta_1 \neq 0$ .
- (iv) Find the confidence set for the vector parameter  $\beta = (\beta_0, \beta_1)^\top$ .
- (v) Find the confidence interval for the parameter  $\beta_1$ .

#### Task 14

Suppose that you observe independent identically distributed random vectors  $(X_1, Y_1)^\top, \dots, (X_n, Y_n)^\top$  such that the conditional distribution  $Y_i | X_i$  is  $N(\beta_1 X_i + \beta_2 X_i^2, 1)$ .

- (i) Find the maximum likelihood estimator of  $\beta = (\beta_0, \beta_1)$ .
- (ii) Derive a test for the null hypothesis  $H_0 : (\beta_1, \beta_2)^\top = (0, 0)^\top$  against the alternative that  $H_1 : (\beta_1, \beta_2)^\top \neq (0, 0)^\top$ .
- (iii) Derive a test for the null hypothesis  $H_0 : \beta_1 = 0$  against the alternative that  $H_1 : \beta_1 \neq 0$ .
- (iv) Find the confidence set for the vector parameter  $\beta = (\beta_1, \beta_2)^\top$ .
- (v) Find the confidence interval for the parameter  $\beta_1$ .

#### Task 15

Suppose that you observe independent identically distributed random vectors  $(X_1, Y_1)^\top, \dots, (X_n, Y_n)^\top$  such that

$$P(Y_1 = 1 | X_1) = \frac{\exp\{\alpha + \beta X_1\}}{1 + \exp\{\alpha + \beta X_1\}}, \quad P(Y_1 = 0 | X_1) = \frac{1}{1 + \exp\{\alpha + \beta X_1\}},$$

where the distribution of  $X_1$  does not depend on the unknown parameters  $\alpha$  and  $\beta$ .

- (i) Derive a test for the null hypothesis  $H_0 : \beta = 0$  against the alternative that  $H_1 : \beta \neq 0$ .
- (ii) Find the confidence interval for the parameter  $\beta$ .