Statistics II - Test I

Coursecode: 202300026
Quartile: 1B 2023/24
Time: 22 January 2025
Time: 13:45 - 16:45
Teacher: Sophie Langer

Student's name:________________
Student ID:________________

General information:

- A regular scientific calculator is allowed, a programmable calculator ("GR") is not.
- All tables of probability distribution needed are attached.
- Other than a pen, no means are needed (or allowed) for answering the exam questions.
- Please write your name on every exam paper you hand in.
- Please write legibly. I cannot evaluate what I do not understand.

Distribution of the points:

Part A

Task	(a)	(b)	(c)	(d)	(e)	(f)	(g(i))	(g(ii))	(g(iii))	(g(iv))	Sum
	1	3	3	2	4	1	1	1	1	1	18

Part B

Task	1	2	3	4	5	6	7	8	9	Sum
	2	3	3	4	2+2	6+2	6	1+2+4	3	40

Achieved points:

Task	Part A	1	2	3	4	5	6	7	8	9	Sum

Part A: Basic concepts

(a) [1 Point] True or False? In a multivariate Gaussian distribution, if two variables are uncorrelated, they are guaranteed to be independent.

(b) [3 Points] In logistic regression, what is the role of the response variable, and why is a standard linear regression model unsuitable for this type of data? Describe the transformation applied to the linear regression to address this issue and provide the corresponding mathematical formula that defines the logistic regression model.

(c) [3 Points] True or False? Increasing the amount of explanatory variables in a linear regression model will always lead to an increase in the value of R^2. Justify your answer. Does the same hold for the adjusted R^2? Explain why or why not.

(d) [2 Points] What are the advantages and disadvantages of using a parametric test compared to a nonparametric test? Provide an example of a parametric and nonparametric test.

(e) [4 Points] In Bayesian statistics what are the three important components, what do they represent and how do they interact in Bayes' theorem?

(f) [1 Point] Describe the difference between white noise and iid noise.

(g) [4 Points] For a **strictly** stationary process $\{X_t\}$, indicate for each statement if it is true or false. Motivate your answer.

(i) for $t, s > 0$, then the bivariate vectors (X_s, X_t) and (X_t, X_s) have the same distribution

(ii) for $t, s > 0$, then $\text{Cov}(X_t, X_s) = \text{Cov}(X_s, X_t)$

(iii) for $t, s > 0$ then $\mathbf{E}(X_1 X_{1+s}) = \mathbf{E}(X_t X_{t+s})$

(iv) for $t, s, h > 0$, then the bivariate vectors (X_t, X_{s+h}) and (X_1, X_h) have the same distribution.

Part B: Theory

1. [2 Points] Suppose that $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}_X, \Sigma_X)$ and $\mathbf{Y} \sim \mathcal{N}_d(\boldsymbol{\mu}_Y, \Sigma_Y)$ and $\mathbf{X}$ and $\mathbf{Y}$ are independent. What is the distribution of $\mathbf{X} - C\mathbf{Y}$ where C is a $d \times d$ matrix of constants?

2. [3 Points] Consider the multiple linear regression model in matrix-vector notation $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 I)$ and assume that $X^T X$ is invertible. Let $\hat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \mathbf{Y}$ be the MLE and

$$\hat{\mathbf{Y}} = X\hat{\boldsymbol{\beta}}$$

the corresponding vector of predicted values. Show that the variance of the prediction is $\text{Var}(\hat{\mathbf{Y}}) = \sigma^2 X(X^T X)^{-1} X^T$.

3. [3 Points] At the beginning of a year a first grade class was randomly divided into two groups. One group was taught to read using a uniform method, while the other group was taught using an individual method. At the end of the year, each student was given a reading ability test. The results of two independent random sample of students from the two groups are

First group	:	227	176	252	149	16	55	234
Second group	:	202	14	165	171	292	271	

Use the Wald-Wolfowitz run test to determine whether the teaching method has an impact on students' reading ability. Evaluate the result at a significance level of $\alpha = 0.05$. Formally state the null and alternative hypotheses.

4. [4 Points] A group of alien enthusiasts believes they have discovered a way to improve brain function by wearing tinfoil hats. To test their theory, they recruit 10 volunteers and ask them to wear tinfoil hats for 8 hours a day over a four-week period. At the start and end of the experiment, the volunteers take a standard IQ test. The test scores are shown below:

Participant	First Test Score	Second Test Score
1	92	102
2	97	100
3	76	74
4	87	85
5	80	83
6	79	89
7	99	100
8	111	112
9	103	99
10	93	97

Using the Wilcoxon signed-rank test, determine whether there is evidence to support the claim that wearing tinfoil hats improves cognitive performance. Use the smaller of the two rank sums as the test statistic and evaluate the results at a 5% significance level. Formally state the null and alternative hypotheses.

5. [4 Points] Let $X_1, \dots, X_n \overset{i.i.d}{\sim} F$ and $\hat{F}_n$ be the empirical cumulative distribution function of F, and $T(X_1, \dots, X_n)$ an estimator. Consider the following bootstrap procedure.

   1. Draw $X_1^*, \dots, X_n^* \sim \hat{F}_n$
   2. $T^* = T(X_1^*, \dots, X_n^*)$
   3. Repeat R times step 1 and 2 to get $T_1^*, \dots, T_R^*$
   4. Compute the empirical bootstrap estimate $\theta^* := \frac{1}{R} \sum_{i=1}^{R} T_i^*$.

   Show that

   (a) [2 Points] Step 1 above can be replaced by

      1. Draw $X_1^*, \dots, X_n^*$ uniformly at random with replacement from $X_1, \dots, X_n$.

   (b) [2 Points] Find the empirical bootstrap estimate of the cumulative distribution function of T, $F_T(s) := \mathbf{P}(T \leq s)$.

6. [8 Points] In point counts for birds, you know how many birds you detected but not the number of birds present. Let Y be the number of birds detected and assume $Y \sim Bin(\eta, p)$, where the number of birds present η is unknown and the probability of success p is known.

(a) [6 Points] Assume the prior for η is a Poisson distribution $\eta \sim \text{Poisson}(m)$ with probability mass function

$$p(\eta|m) = \frac{m^\eta e^{-m}}{\eta!}.$$

Derive the posterior of η.

Hint: You might use that $\binom{a}{b} = \frac{a!}{b!(a-b)!}$ *for two integers* $a > b$.

(b) [2 Points] Using squared loss, what is the Bayes estimator of η?

Hint: You might use that if $X \sim \textit{Poisson}(m)$, $\mathbf{E}(X) = m$.

7. [6 Points] Consider the stochastic harmonic process

$$X_t = A\cos\left(\frac{2\pi t}{T} + B\right), \qquad t \in \mathbb{Z}, \tag{1}$$

with T a given integer. A and B are independent random variables with $\mathbf{E}(A) = 0$, $\text{Var}(A) = \sigma^2$, and B uniformly distributed on $[0, 2\pi]$ with pdf $f(x) = \frac{1}{2\pi}$ for $x \in [0, 2\pi]$. This process is stationary. Compute the autocovariance $\gamma_X(h) = \text{Cov}(X_t, X_{t+h})$ of the process and show that it is independent of t.

Hint: You might use the trigonometric identity

$$\cos(a)\cos(b) = \frac{1}{2}\cos(a-b) + \frac{1}{2}\cos(a+b)$$

and recall that $\int \cos(a + bx)dx = \frac{1}{b}\sin(a + bx)$.

8. [7 Points] Consider the time series

$$Y_t = \beta_0 + \beta_1 t + X_t,$$

where $\{X_t\}$ is weakly stationary with autocovariance $\gamma_X(h) = \text{Cov}(X_t, X_{t+h})$ and $\beta_1 \neq 0$ and the differenced series

$$Z_t = (1 - B)Y_t,$$

where B denotes the backward shift operator defined as $BW_t = W_{t-1}$ for any time series $\{W_t\}$.

(a) [1 Point] Is $\{Y_t\}$ a stationary time series? Motivate your answer.

(b) [2 Points] Show that $Z_t = \beta_1 + (1 - B)X_t$.

(c) [4 Points] Show that $\{Z_t\}$ is stationary and find its autocovariance function.

9. [3 Points] Show that the function $f(x) = \max\{0, x_1, x_2\}$ with $x_1, x_2 \in \mathbb{R}$ can be represented by a neural network with 2 hidden layers and ReLU activation function $\sigma(x) = \max\{x, 0\}$. Specify the number of neurons in each hidden layer.

TABLE G
Critical values of r in the runs test*

Given in the tables are various critical values of r for values of m and n less than or equal to 20. For the one-sample runs test, any observed value of r which is less than or equal to the smaller value, or is greater than or equal to the larger value in a pair is significant at the $\alpha = .05$ level.

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
											–	–	–	–	–	–	–	–	–
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
					–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
				9	9	–	–	–	–	–	–	–	–	–	–	–	–	–	–
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
			9	10	10	11	11	–	–	–	–	–	–	–	–	–	–	–	–
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
		–	9	10	11	12	12	13	13	13	13	–	–	–	–	–	–	–	–
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
		–	–	11	12	13	13	14	14	14	14	15	15	15	–	–	–	–	–
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
		–	–	11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
		–	–	–	13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
		–	–	–	13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
		–	–	–	13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
	–	–	–	–	13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
	–	–	–	–	–	15	16	17	18	19	19	20	20	21	21	22	22	23	23
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
	–	–	–	–	–	15	16	17	18	19	20	20	21	22	22	23	23	23	24
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
	–	–	–	–	–	15	16	18	18	19	20	21	22	22	23	23	24	24	25
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
	–	–	–	–	–	–	17	18	19	20	21	21	22	23	23	24	25	25	25
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
	–	–	–	–	–	–	17	18	19	20	21	22	23	23	24	25	25	26	26
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
	–	–	–	–	–	–	17	18	19	20	21	22	23	24	25	25	26	26	27
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
	–	–	–	–	–	–	17	18	20	21	22	23	23	24	25	26	26	27	27
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14
	–	–	–	–	–	–	17	18	20	21	22	23	24	25	25	26	27	27	28

* Adapted from Swed, and Eisenhart, C. (1943). Tables for testing randomness of grouping in a sequence of alternatives. *Annals of Mathematical Statistics*, **14**, 83–86, with the kind permission of the authors and publisher.

Wilcoxon Signed-Rank Test Critical Values Table

Reject H_0 if the test value is less than or equal to the value given in the table.

	One tailed, $\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
n	Two tailed, $\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$	$\alpha = 0.01$
5	1	--	--	--
6	2	1	--	--
7	4	2	0	--
8	6	4	2	1
9	8	6	3	2
10	11	8	5	3
11	14	11	7	5
12	17	14	10	7
13	21	17	13	10
14	26	21	16	13
15	30	25	20	16
16	36	30	24	19
17	41	35	28	23
18	47	40	33	28
19	54	46	38	32
20	60	52	43	37
21	68	59	49	43
22	75	66	56	49
23	83	73	62	55
24	92	81	69	61
25	101	90	77	68
26	110	98	85	76
27	120	107	93	84
28	130	117	102	92
29	141	127	111	100
30	152	137	120	109