Exam - Statistics (WBMA009-05) 2026/2027

Date and time: October 29, 2026, 11.45-13.45h
Place: 5263.0202 Exam Hall 2

Rules to follow:

- This is a closed book exam. Consultation of books and notes is **not** permitted. You may use a simple (non-programmable) calculator.
- Write your name and student number on each sheet of paper.
- There are 4 exercises worth a total of 30 points.
- Your exam grade will be computed as follows:

$$\text{grade} = 1 + 0.3 \cdot \text{number of points}.$$

- ALWAYS include the relevant equation(s) and/or short derivations.
- Throughout the exam, you may assume that all required regularity conditions are satisfied, unless stated otherwise.
- **We wish you success in completing the exam!**

START OF EXAM

α	0.025	0.05	0.1	0.5	0.9	0.95	0.975
$\mathcal{N}(0,1)$	-1.96	-1.64	-1.28	0	1.28	1.64	1.96

Table 1: Quantiles q_α of the $\mathcal{N}(0,1)$ distribution (rounded to two decimal places).

1. **Sample from a Pareto distribution.** **14**
Consider a random sample from a Pareto distribution with parameter $\theta > 0$,

$$X_1, \ldots, X_n \sim \text{Pareto}(\theta)$$

and density

$$f_\theta(x) = \frac{\theta}{x^{\theta+1}} \quad (x \geq 1)$$

(a) 2 Show that $T = \sum_{i=1}^n \log(X_i)$ is a sufficient statistic for θ.

(b) 2 Show that the ML estimator is

$$\hat{\theta}_{ML} = \frac{n}{\sum_{i=1}^n \log X_i}.$$

(Confirm that the stationary point is a maximum by the 2nd derivative.)

(c) 2 Show that the expected Fisher information (for $n = 1$) is

$$I(\theta) = \frac{1}{\theta^2}.$$

(d) 2 Assume $n = 25$ and the realization

$$\frac{1}{25} \sum_{i=1}^{25} \log(X_i) = 0.4.$$

Provide an asymptotic two-sided 90% confidence interval for θ.

(e) 2 For $n = 25$, consider the test problem

$$H_0 : \theta = 2 \quad \text{vs.} \quad H_1 : \theta \neq 2.$$

Based on the asymptotic normality of the ML estimator, derive a rejection region for $\hat{\theta}_{ML}$ at significance level $\alpha = 0.05$.
Using the realization given in part (d), state whether H_0 is rejected.

For parts (f) and (g), you may use the fact that for $Y_i = \log(X_i)$, we have

$$E[Y_i] = \frac{1}{\theta}, \qquad \text{Var}(Y_i) = \frac{1}{\theta^2}.$$

(f) 2 Use a first-order Taylor expansion of $g(y) = 1/y$ around $y_0 = 1/\theta$ to show

$$\hat{\theta}_{ML} \approx \theta - \theta^2 \left(\bar{Y} - \frac{1}{\theta} \right).$$

(g) 2 Use the result from (f) and the central limit theorem to confirm that

$$\hat{\theta}_{ML} \approx \mathcal{N}\left(\theta, \frac{\theta^2}{n} \right).$$

HINT:
Quantiles of the Gaussian $\mathcal{N}(0, 1)$ distribution are provided on the 1st page.

2. **Confidence interval and p-value for mean difference.** $\boxed{4}$
Consider two independent Gaussian random samples of sizes $n = 12$ and $m = 8$ with known variance parameters $\sigma_X^2 = 1.5$ and $\sigma_Y^2 = 1$:

$$X_1, \ldots, X_{12} \sim \mathcal{N}(\mu_X, 1.5) \quad \text{and} \quad Y_1, \ldots, Y_8 \sim \mathcal{N}(\mu_Y, 1).$$

From observed data $x_1, \ldots, x_{12}$ and $y_1, \ldots, y_8$, the following empirical means were computed:

$$\bar{x} = 1.35 \quad \text{and} \quad \bar{y} = 0.71.$$

(a) $\boxed{2}$ Derive an exact two-sided 95% confidence interval for the mean difference

$$\mu_D := \mu_X - \mu_Y.$$

(b) $\boxed{2}$ Compute the p-value for the test problem

$$H_0 : \mu_D = 0 \quad \text{vs.} \quad H_1 : \mu_D \neq 0.$$

HINT: Quantiles of the Gaussian $\mathcal{N}(0, 1)$ distribution are provided on the 1st page.

3. **Likelihood ratio test.** $\boxed{4}$
Consider a random sample

$$X_1, \ldots, X_{100} \sim \text{Poisson}(\lambda),$$

where $\lambda > 0$ is unknown. The density of the Poisson distribution is

$$f_\lambda(x) = \frac{\lambda^x}{x!} \exp(-\lambda), \qquad x = 0, 1, 2, \ldots$$

and the ML estimator is

$$\hat{\lambda}_{ML} = \frac{1}{100} \sum_{i=1}^{100} X_i.$$

We want to test

$$H_0 : \lambda = 1 \quad \text{vs.} \quad H_1 : \lambda \neq 1.$$

(a) $\boxed{2}$ Derive the likelihood ratio test statistic Λ.

(b) $\boxed{2}$ Under H_0, the likelihood ratio statistic asymptotically satisfies

$$-2 \log(\Lambda) \sim \chi_1^2.$$

Assume that the observed data satisfy

$$\sum_{i=1}^{100} x_i = 80.$$

Use the quantiles in Table 1 and the fact that $Z \sim \mathcal{N}(0, 1) \Rightarrow Z^2 \sim \chi_1^2$ to check whether the approximate p-value is smaller than 0.05.

4. **Neyman–Pearson Lemma.** $\boxed{8}$
Consider a random sample

$$X_1, X_2, X_3 \sim \text{Gamma}(2, \theta),$$

where $\theta > 0$ is an unknown rate parameter. The density is

$$f_\theta(x) = \theta^2 x \exp(-\theta x), \qquad x > 0.$$

We want to test the following hypotheses:

$$H_0 : \theta = 1 \quad \text{vs.} \quad H_1 : \theta = 2.$$

(a) $\boxed{2}$ Derive the Neyman–Pearson likelihood ratio

$$\Lambda = \frac{L(1)}{L(2)}$$

and show that it is strictly monotone in

$$S = X_1 + X_2 + X_3.$$

(b) $\boxed{2}$ Use the Neyman–Pearson lemma to obtain the most powerful test at significance level $\alpha = 0.05$. For which values of S is H_0 rejected?

For parts (c) and (d), let f_0 and f_1 denote the joint densities under H_0 and H_1, respectively, and let $\phi^* : \mathbb{R}^3 \to \{0, 1\}$ denote the Neyman–Pearson test derived in (b).

Let $k > 0$ be the likelihood-ratio threshold, so that

$$\phi^*(\mathbf{x}) = \begin{cases} 1 & \text{if } f_0(\mathbf{x}) < k f_1(\mathbf{x}), \\ 0 & \text{otherwise.} \end{cases}$$

Consider any other test function $\phi : \mathbb{R}^3 \to \{0, 1\}$ with

$$E_0[\phi(\mathbf{X})] \leq \alpha = 0.05.$$

(c) $\boxed{2}$ Show that

$$(\phi^*(\mathbf{x}) - \phi(\mathbf{x}))(k f_1(\mathbf{x}) - f_0(\mathbf{x})) \geq 0.$$

(d) $\boxed{2}$ Integrate the result from part (c) to prove that

$$E_1[\phi^*(\mathbf{X})] \geq E_1[\phi(\mathbf{X})],$$

i.e., that the Neyman–Pearson test is most powerful.

HINTS FOR PART (b):
(1) The sum of n iid Gamma(a, θ) random variables is Gamma$(n \cdot a, \theta)$ distributed.
(2) The 0.05-quantile of the Gamma$(6, 1)$ distribution is approximately 2.613.