

## Mock Paper for Fundamental Theory of Statistical Inference, 2025/26.

You should answer *all* parts, which are equally weighted.

(1) Consider testing two simple hypotheses  $H_0 : \theta = \theta_0$  against  $H_1 : \theta = \theta_1$ , with action space  $\mathcal{A} = \{0, 1\}$ , where  $a = 0$  and  $a = 1$  correspond respectively to acceptance and rejection of  $H_0$ . Consider the loss function

$$L(\theta, a) = \begin{cases} K_0 a, & \theta = \theta_0, \\ K_1(1 - a), & \theta = \theta_1, \end{cases}$$

for constants  $K_0, K_1 > 0$ . Assume prior probabilities  $\pi_0 > 0$  and  $\pi_1 = 1 - \pi_0 > 0$  for  $H_0$  and  $H_1$  respectively.

(a) Derive the Bayes rule for this inference problem.

(b) Interpret the Bayes rule in terms of the frequentist approach to testing. Is it a likelihood ratio test? What is the critical value of the corresponding test statistic?

(c) Show that any likelihood ratio test which rejects  $H_0$  if  $f(y; \theta_1)/f(y; \theta_0) \geq C, C \geq 0$ , is admissible for the assumed form of loss function.

(d) Let  $\alpha$  and  $\beta$  be the Type 1 and Type 2 error probabilities of such a likelihood ratio test. Show that, if the critical value  $C$  of the test is fixed so that  $K_0\alpha = K_1\beta$ , then it is also a minimax test.

(2) Let  $Y_1, \dots, Y_p$  ( $p > 2$ ) be independent random variables such that  $Y_i \sim N(\theta_i, 1)$ . Write  $Y = (Y_1, \dots, Y_p)^T$  and  $\theta = (\theta_1, \dots, \theta_p)^T$ . Let  $\hat{\theta} \equiv \hat{\theta}(Y) = (\hat{\theta}_1(Y), \dots, \hat{\theta}_p(Y))^T$  be an estimator of  $\theta$ . Further, let  $g(Y) \equiv (g_1(Y), \dots, g_p(Y))^T = \hat{\theta} - Y$  and  $D_i(Y) = \partial g_i(Y)/\partial Y_i$ .

(a) Show that

$$\hat{R}(Y) = p + 2 \sum_{i=1}^p D_i(Y) + \sum_{i=1}^p \{g_i(Y)\}^2$$

is an unbiased estimator of the risk of  $\hat{\theta}$ , under squared error loss  $L(\theta, \hat{\theta}) = \sum_{i=1}^p (\hat{\theta}_i - \theta_i)^2$ .

(b) Suppose the estimator  $\hat{\theta}$  is of the form  $\hat{\theta} = bY$ , for  $b \in \mathbb{R}$ . Find the value  $b^*$  of  $b$  that minimises the unbiased risk estimator  $\hat{R}(Y)$ . Compare the estimator  $b^*Y$  with the James-Stein estimator.

CONTINUED

(c) The *soft threshold estimator* is defined by

$$\hat{\theta}_i = \begin{cases} Y_i + \lambda, & \text{if } Y_i < -\lambda \\ 0, & \text{if } -\lambda \leq Y_i \leq \lambda \\ Y_i - \lambda, & \text{if } Y_i > \lambda, \end{cases}$$

where  $\lambda > 0$  is a constant, to be specified. Show that for this estimator

$$\hat{R}(Y) = p + \sum_{i=1}^p \{-2I(|Y_i| \leq \lambda) + \min(Y_i^2, \lambda^2)\}.$$

[Here,  $I(A) = 1$  if  $A$  holds,  $= 0$  otherwise.]

It has been suggested that a suitable choice of the value of  $\lambda$  is that which minimises  $\hat{R}(Y)$ . Explain why determining this value only requires examining the value of  $\hat{R}(Y)$  at a finite number of values of  $\lambda$ .

(d) Find the form of the estimator  $\hat{\theta}^P$  that minimises the penalized sum of squares

$$\sum_{i=1}^p (Y_i - \theta_i)^2 + \lambda^2 J(\theta),$$

with  $J(\theta) = |\{\theta_i : \theta_i \neq 0\}|$ , the number of non-zero elements of  $\theta$ .

Calculate the risk of  $\hat{\theta}^P$  for the case  $p = 1$ . When is this estimator preferable to the unbiased estimator  $\hat{\theta} \equiv Y$ ? Quantify your answer for  $\lambda = 1, 2, 5$ .

**(3)** ‘There are serious difficulties in an approach to statistical inference based on the repeated sampling principle that does not take account of a conditionality principle’. Discuss, in no more than 300–400 words.