Paper 1, Section II
29J Principles of Statistics

Let $X_1, \ldots, X_n$ be random variables with joint probability density function in a statistical model $\{f_\theta : \theta \in \mathbb{R}\}$.

(a) Define the *Fisher information* $I_n(\theta)$. What do we mean when we say that the Fisher information *tensorises*?

(b) Derive the relationship between the Fisher information and the derivative of the score function in a regular model.

(c) Consider the model defined by $X_1 = \theta + \varepsilon_1$ and

$$X_i = \theta(1 - \sqrt{\gamma}) + \sqrt{\gamma}\, X_{i-1} + \sqrt{1-\gamma}\ \varepsilon_i \quad \text{for } i = 2, \ldots, n,$$

where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. $N(0,1)$ random variables, and $\gamma \in [0,1)$ is a known constant. Compute the Fisher information $I_n(\theta)$. For which values of γ does the Fisher information tensorise? State a lower bound on the variance of an unbiased estimator $\hat{\theta}$ in this model.

Paper 2, Section II
29J Principles of Statistics

Let $X_1, \ldots, X_n$ be i.i.d. random observations taking values in $[0,1]$ with a continuous distribution function F. Let $\hat{F}_n(x) = n^{-1} \sum_{i=1}^n \mathbf{1}_{\{X_i \leqslant x\}}$ for each $x \in [0,1]$.

(a) State the Kolmogorov–Smirnov theorem. Explain how this theorem may be used in a goodness-of-fit test for the null hypothesis $H_0 : F = F_0$, with F_0 continuous.

(b) Suppose you do not have access to the quantiles of the sampling distribution of the Kolmogorov–Smirnov test statistic. However, you are given i.i.d. samples $Z_1, \ldots, Z_{nm}$ with distribution function F_0. Describe a test of $H_0 : F = F_0$ with size exactly $1/(m+1)$.

(c) Now suppose that $X_1, \ldots, X_n$ are i.i.d. taking values in $[0,\infty)$ with probability density function f, with $\sup_{x \geqslant 0} \big(|f(x)| + |f'(x)|\big) < 1$. Define the density estimator

$$\hat{f}_n(x) = n^{-2/3} \sum_{i=1}^n \mathbf{1}_{\left\{X_i - \frac{1}{2n^{1/3}} \leqslant x \leqslant X_i + \frac{1}{2n^{1/3}}\right\}}, \quad x \geqslant 0.$$

Show that for all $x \geqslant 0$ and all $n \geqslant 1$,

$$\mathbb{E}\big[\big(\hat{f}_n(x) - f(x)\big)^2\big] \leqslant \frac{2}{n^{2/3}}.$$

Paper 3, Section II
28J Principles of Statistics

Let $X_1, \ldots, X_n \sim^{iid} \text{Gamma}(\alpha, \beta)$ for some known $\alpha > 0$ and some unknown $\beta > 0$. [The gamma distribution has probability density function

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad x > 0,$$

and its mean and variance are α/β and α/β^2, respectively.]

(a) Find the maximum likelihood estimator $\hat{\beta}$ for β and derive the distributional limit of $\sqrt{n}(\hat{\beta} - \beta)$. [You may not use the asymptotic normality of the maximum likelihood estimator proved in the course.]

(b) Construct an asymptotic $(1-\gamma)$-level confidence interval for β and show that it has the correct (asymptotic) coverage.

(c) Write down all the steps needed to construct a candidate to an asymptotic $(1-\gamma)$-level confidence interval for β using the nonparametric bootstrap.

Paper 4, Section II
28J Principles of Statistics

Suppose that $X \mid \theta \sim \text{Poisson}(\theta)$, $\theta > 0$, and suppose the prior π on θ is a gamma distribution with parameters $\alpha > 0$ and $\beta > 0$. [Recall that π has probability density function

$$f(z) = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\beta z}, \qquad z > 0,$$

and that its mean and variance are α/β and α/β^2, respectively.]

(a) Find the π-Bayes estimator for θ for the quadratic loss, and derive its quadratic risk function.

(b) Suppose we wish to estimate $\mu = e^{-\theta} = \mathbb{P}_\theta(X = 0)$. Find the π-Bayes estimator for μ for the quadratic loss, and derive its quadratic risk function. [*Hint: The moment generating function of a Poisson*(θ) *distribution is* $M(t) = \exp\big(\theta(e^t - 1)\big)$ *for* $t \in \mathbb{R}$*, and that of a Gamma*(α, β) *distribution is* $M(t) = (1 - t/\beta)^{-\alpha}$ *for* $t < \beta$.]

(c) State a sufficient condition for an admissible estimator to be minimax, and give a proof of this fact.

(d) For each of the estimators in parts (a) and (b), is it possible to deduce using the condition in (c) that the estimator is minimax for some value of α and β? Justify your answer.