Paper 1, Section II

28K Principles of Statistics

Define *admissible*, *Bayes*, *minimax* decision rules.

A random vector $X = (X_1, X_2, X_3)^{\mathrm{T}}$ has independent components, where X_i has the normal distribution $\mathcal{N}(\theta_i, 1)$ when the parameter vector Θ takes the value $\theta = (\theta_1, \theta_2, \theta_3)^{\mathrm{T}}$. It is required to estimate Θ by a point $a \in \mathbb{R}^3$, with loss function $L(\theta, a) = \|a - \theta\|^2$. What is the risk function of the maximum-likelihood estimator $\widehat{\Theta} := X$? Show that $\widehat{\Theta}$ is dominated by the estimator $\widetilde{\Theta} := (1 - \|X\|^{-2})\, X$.

Paper 2, Section II

28K Principles of Statistics

Random variables $X_1, \ldots, X_n$ are independent and identically distributed from the normal distribution with unknown mean M and unknown precision (inverse variance) H. Show that the likelihood function, for data $X_1 = x_1, \ldots, X_n = x_n$, is

$$L_n(\mu, h) \propto h^{n/2} \exp\left(-\tfrac{1}{2}h\left\{n\,(\overline{x} - \mu)^2 + S\right\}\right),$$

where $\overline{x} := n^{-1}\sum_i x_i$ and $S := \sum_i (x_i - \overline{x})^2$.

A bivariate prior distribution for (M, H) is specified, in terms of hyperparameters $(\alpha_0, \beta_0, m_0, \lambda_0)$, as follows. The marginal distribution of H is $\Gamma(\alpha_0, \beta_0)$, with density

$$\pi(h) \propto h^{\alpha_0 - 1} e^{-\beta_0 h} \quad (h > 0),$$

and the conditional distribution of M, given $H = h$, is normal with mean m_0 and precision $\lambda_0 h$.

Show that the conditional prior distribution of H, given $\mathrm{M} = \mu$, is

$$H \mid \mathrm{M} = \mu \quad \sim \quad \Gamma\left(\alpha_0 + \tfrac{1}{2}, \beta_0 + \tfrac{1}{2}\lambda_0\,(\mu - m_0)^2\right).$$

Show that the posterior joint distribution of (M, H), given $X_1 = x_1, \ldots, X_n = x_n$, has the same form as the prior, with updated hyperparameters $(\alpha_n, \beta_n, m_n, \lambda_n)$ which you should express in terms of the prior hyperparameters and the data.

[You may use the identity

$$p(t-a)^2 + q(t-b)^2 = (t-\delta)^2 + pq(a-b)^2,$$

where $p + q = 1$ and $\delta = pa + qb$.]

Explain how you could implement Gibbs sampling to generate a random sample from the posterior joint distribution.

Paper 3, Section II

27K Principles of Statistics

Random variables $X_1, X_2, \ldots$ are independent and identically distributed from the exponential distribution $\mathcal{E}(\theta)$, with density function

$$p_X(x \mid \theta) = \theta e^{-\theta x} \quad (x > 0)\,,$$

when the parameter Θ takes value $\theta > 0$. The following experiment is performed. First X_1 is observed. Thereafter, if $X_1 = x_1, \ldots, X_i = x_i$ have been observed ($i \geqslant 1$), a coin having probability $\alpha(x_i)$ of landing heads is tossed, where $\alpha : \mathbb{R} \to (0,1)$ is a known function and the coin toss is independent of the X's and previous tosses. If it lands heads, no further observations are made; if tails, X_{i+1} is observed.

Let N be the total number of X's observed, and $\mathbf{X} := (X_1, \ldots, X_N)$. Write down the likelihood function for Θ based on data $\mathbf{X} = (x_1, \ldots, x_n)$, and identify a minimal sufficient statistic. What does the likelihood principle have to say about inference from this experiment?

Now consider the experiment that only records $Y := X_N$. Show that the density function of Y has the form

$$p_Y(y \mid \theta) = \exp\{a(y) - k(\theta) - \theta y\}\,.$$

Assuming the function $a(\cdot)$ is twice differentiable and that both $p_Y(y \mid \theta)$ and $\partial p_Y(y \mid \theta)/\partial y$ vanish at 0 and ∞, show that $a'(Y)$ is an unbiased estimator of Θ, and find its variance.

Stating clearly any general results you use, deduce that

$$-k''(\theta)\, \mathbb{E}_\theta\{a''(Y)\} \geqslant 1\,.$$

Paper 4, Section II

27K Principles of Statistics

What does it mean to say that a $(1 \times p)$ random vector ξ has a *multivariate normal distribution*?

Suppose $\xi = (X, Y)$ has the bivariate normal distribution with mean vector $\mu = (\mu_X, \mu_Y)$, and dispersion matrix

$$\Sigma = \begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{pmatrix}.$$

Show that, with $\beta := \sigma_{XY}/\sigma_{XX}$, $Y - \beta X$ is independent of X, and thus that the conditional distribution of Y given X is normal with mean $\mu_Y + \beta(X - \mu_X)$ and variance $\sigma_{YY\cdot X} := \sigma_{YY} - \sigma_{XY}^2/\sigma_{XX}$.

For $i = 1, \ldots, n$, $\xi_i = (X_i, Y_i)$ are independent and identically distributed with the above distribution, where all elements of μ and Σ are unknown. Let

$$S = \begin{pmatrix} S_{XX} & S_{XY} \\ S_{XY} & S_{YY} \end{pmatrix} := \sum_{i=1}^{n} (\xi_i - \overline{\xi})^{\mathrm{T}} (\xi_i - \overline{\xi}),$$

where $\overline{\xi} := n^{-1} \sum_{i=1}^{n} \xi_i$.

The *sample correlation coefficient* is $r := S_{XY}/\sqrt{S_{XX}S_{YY}}$. Show that the distribution of r depends only on the population correlation coefficient $\rho := \sigma_{XY}/\sqrt{\sigma_{XX}\sigma_{YY}}$.

Student's t-statistic (on $n-2$ degrees of freedom) for testing the null hypothesis $H_0 : \beta = 0$ is

$$t := \frac{\widehat{\beta}}{\sqrt{S_{YY\cdot X}/(n-2)S_{XX}}},$$

where $\widehat{\beta} := S_{XY}/S_{XX}$ and $S_{YY\cdot X} := S_{YY} - S_{XY}^2/S_{XX}$. Its density when H_0 is true is

$$p(t) = C\left(1 + \frac{t^2}{n-2}\right)^{-\frac{1}{2}(n-1)},$$

where C is a constant that need not be specified.

Express t in terms of r, and hence derive the density of r when $\rho = 0$.

How could you use the sample correlation r to test the hypothesis $\rho = 0$?