Paper 4, Section II

27K Principles of Statistics

For $i = 1, \ldots, n$, the pairs (X_i, Y_i) have independent bivariate normal distributions, with $\mathbb{E}(X_i) = \mu_X$, $\mathbb{E}(Y_i) = \mu_Y$, $\mathrm{var}(X_i) = \mathrm{var}(Y_i) = \phi$, and $\mathrm{corr}(X_i, Y_i) = \rho$. The means μ_X, μ_Y are known; the parameters $\phi > 0$ and $\rho \in (-1, 1)$ are unknown.

Show that the joint distribution of all the variables belongs to an exponential family, and identify the natural sufficient statistic, natural parameter, and mean-value parameter. Hence or otherwise, find the maximum likelihood estimator $\hat{\rho}$ of ρ.

Let $U_i := X_i + Y_i$, $V_i := X_i - Y_i$. What is the joint distribution of (U_i, V_i)?

Show that the distribution of

$$\frac{(1+\hat{\rho})/(1-\hat{\rho})}{(1+\rho)/(1-\rho)}$$

is F_n^n. Hence describe a $(1-\alpha)$-level confidence interval for ρ. Briefly explain what would change if μ_X and μ_Y were also unknown.

[*Recall that the distribution* $F_{\nu_2}^{\nu_1}$ *is that of* $(W_1/\nu_1)/(W_2/\nu_2)$, *where, independently for* $j = 1$ *and* $j = 2$, W_j *has the chi-squared distribution with* ν_j *degrees of freedom.*]

Paper 3, Section II

27K Principles of Statistics

The parameter vector is $\mathbf{\Theta} \equiv (\Theta_1, \Theta_2, \Theta_3)$, with $\Theta_i > 0$, $\Theta_1 + \Theta_2 + \Theta_3 = 1$. Given $\mathbf{\Theta} = \boldsymbol{\theta} \equiv (\theta_1, \theta_2, \theta_3)$, the integer random vector $\boldsymbol{X} = (X_1, X_2, X_3)$ has a trinomial distribution, with probability mass function

$$p(\boldsymbol{x} \mid \boldsymbol{\theta}) = \frac{n!}{x_1!\, x_2!\, x_3!}\, \theta_1^{x_1}\, \theta_2^{x_2}\, \theta_3^{x_3}\,, \qquad \left(x_i \geqslant 0,\ \sum_{i=1}^{3} x_i = n \right). \tag{1}$$

Compute the score vector for the parameter $\mathbf{\Theta}^* := (\Theta_1, \Theta_2)$, and, quoting any relevant general result, use this to determine $\mathbb{E}(X_i)$ $(i = 1, 2, 3)$.

Considering (1) as an exponential family with mean-value parameter $\mathbf{\Theta}^*$, what is the corresponding natural parameter $\mathbf{\Phi} \equiv (\Phi_1, \Phi_2)$?

Compute the information matrix I for $\mathbf{\Theta}^*$, which has (i, j)-entry

$$I_{ij} = -\mathbb{E}\left(\frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \right) \quad (i, j = 1, 2)\,,$$

where l denotes the log-likelihood function, based on $\boldsymbol{X}$, expressed in terms of (θ_1, θ_2).

Show that the variance of $\log(X_1/X_3)$ is asymptotic to $n^{-1}(\theta_1^{-1} + \theta_3^{-1})$ as $n \to \infty$. [*Hint. The information matrix I_Φ for $\mathbf{\Phi}$ is I^{-1} and the dispersion matrix of the maximum likelihood estimator $\widehat{\mathbf{\Phi}}$ behaves, asymptotically (for $n \to \infty$) as I_Φ^{-1}.*]

Paper 2, Section II

28K Principles of Statistics

Carefully defining all italicised terms, show that, if a sufficiently general method of inference respects both the *Weak Sufficiency Principle* and the *Conditionality Principle*, then it respects the *Likelihood Principle*.

The position X_t of a particle at time $t > 0$ has the Normal distribution $\mathcal{N}(0, \phi t)$, where ϕ is the value of an unknown parameter Φ; and the time, T_x, at which the particle first reaches position $x \neq 0$ has probability density function

$$p_x(t) = \frac{|x|}{\sqrt{2\pi\phi t^3}} \exp\left(-\frac{x^2}{2\phi t}\right) \quad (t > 0)\,.$$

Experimenter E_1 observes X_τ, and experimenter E_2 observes T_ξ, where $\tau > 0$, $\xi \neq 0$ are fixed in advance. It turns out that $T_\xi = \tau$. What does the Likelihood Principle say about the inferences about Φ to be made by the two experimenters?

E_1 bases his inference about Φ on the distribution and observed value of X_τ^2/τ, while E_2 bases her inference on the distribution and observed value of ξ^2/T_ξ. Show that these choices respect the Likelihood Principle.

Paper 1, Section II

28K Principles of Statistics

Prove that, if T is complete sufficient for Θ, and S is a function of T, then S is the minimum variance unbiased estimator of $\mathbb{E}(S \mid \Theta)$.

When the parameter Θ takes a value $\theta > 0$, observables $(X_1, \ldots, X_n)$ arise independently from the exponential distribution $\mathcal{E}(\theta)$, having probability density function

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

Show that the family of distributions

$$\Theta \sim \text{Gamma}\ (\alpha, \beta) \quad (\alpha > 0, \beta > 0) , \tag{1}$$

with probability density function

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \quad (\theta > 0) ,$$

is a conjugate family for Bayesian inference about Θ (where $\Gamma(\alpha)$ is the Gamma function).

Show that the expectation of $\Lambda := \log \Theta$, under prior distribution (1), is $\psi(\alpha) - \log \beta$, where $\psi(\alpha) := (\mathrm{d}/\mathrm{d}\alpha) \log \Gamma(\alpha)$. What is the prior variance of Λ? Deduce the posterior expectation and variance of Λ, given $(X_1, \ldots, X_n)$.

Let $\widetilde{\Lambda}$ denote the limiting form of the posterior expectation of Λ as $\alpha, \beta \downarrow 0$. Show that $\widetilde{\Lambda}$ is the minimum variance unbiased estimator of Λ. What is its variance?