Paper 1, Section I

7E Statistics

Suppose $X_1, \ldots, X_n$ are independent $N(0, \sigma^2)$ random variables, where σ^2 is an unknown parameter. Explain carefully how to construct the uniformly most powerful test of size α for the hypothesis $H_0 : \sigma^2 = 1$ versus the alternative $H_1 : \sigma^2 > 1$.

Paper 2, Section I

8E Statistics

A washing powder manufacturer wants to determine the effectiveness of a television advertisement. Before the advertisement is shown, a pollster asks 100 randomly chosen people which of the three most popular washing powders, labelled A, B and C, they prefer. After the advertisement is shown, another 100 randomly chosen people (not the same as before) are asked the same question. The results are summarized below.

	A	B	C
before	36	47	17
after	44	33	23

Derive and carry out an appropriate test at the 5% significance level of the hypothesis that the advertisement has had no effect on people's preferences.

[*You may find the following table helpful:*

	χ_1^2	χ_2^2	χ_3^2	χ_4^2	χ_5^2	χ_6^2
95 percentile	3.84	5.99	7.82	9.49	11.07	12.59

.]

Paper 1, Section II
19E Statistics

Consider the the linear regression model

$$Y_i = \beta x_i + \epsilon_i,$$

where the numbers $x_1, \ldots, x_n$ are known, the independent random variables $\epsilon_1, \ldots, \epsilon_n$ have the $N(0, \sigma^2)$ distribution, and the parameters β and σ^2 are unknown. Find the maximum likelihood estimator for β.

State and prove the Gauss–Markov theorem in the context of this model.

Write down the distribution of an arbitrary linear estimator for β. Hence show that there exists a linear, unbiased estimator $\widehat{\beta}$ for β such that

$$\mathbb{E}_{\beta,\sigma^2}[(\widehat{\beta} - \beta)^4] \leqslant \mathbb{E}_{\beta,\sigma^2}[(\widetilde{\beta} - \beta)^4]$$

for all linear, unbiased estimators $\widetilde{\beta}$.

[Hint: If $Z \sim N(a, b^2)$ then $\mathbb{E}[(Z-a)^4] = 3\,b^4$.]

Paper 3, Section II
20E Statistics

Let $X_1, \ldots, X_n$ be independent $\text{Exp}(\theta)$ random variables with unknown parameter θ. Find the maximum likelihood estimator $\hat{\theta}$ of θ, and state the distribution of $n/\hat{\theta}$. Show that $\theta/\hat{\theta}$ has the $\Gamma(n,n)$ distribution. Find the $100\,(1-\alpha)\%$ confidence interval for θ of the form $[0, C\,\hat{\theta}]$ for a constant $C > 0$ depending on α.

Now, taking a Bayesian point of view, suppose your prior distribution for the parameter θ is $\Gamma(k, \lambda)$. Show that your Bayesian point estimator $\hat{\theta}_B$ of θ for the loss function $L(\theta, a) = (\theta - a)^2$ is given by

$$\hat{\theta}_B = \frac{n+k}{\lambda + \sum_i X_i}.$$

Find a constant $C_B > 0$ depending on α such that the posterior probability that $\theta \leqslant C_B\,\hat{\theta}_B$ is equal to $1-\alpha$.

[*The density of the* $\Gamma(k,\lambda)$ *distribution is* $f(x;\,k,\lambda) = \lambda^k x^{k-1} e^{-\lambda x}/\Gamma(k)$, *for* $x > 0$.]

Paper 4, Section II

19E Statistics

Consider a collection $X_1, \ldots, X_n$ of independent random variables with common density function $f(x;\theta)$ depending on a real parameter θ. What does it mean to say T is a sufficient statistic for θ? Prove that if the joint density of $X_1, \ldots, X_n$ satisfies the factorisation criterion for a statistic T, then T is sufficient for θ.

Let each X_i be uniformly distributed on $[-\sqrt{\theta}, \sqrt{\theta}\,]$. Find a two-dimensional sufficient statistic $T = (T_1, T_2)$. Using the fact that $\hat{\theta} = 3X_1^2$ is an unbiased estimator of θ, or otherwise, find an unbiased estimator of θ which is a function of T and has smaller variance than $\hat{\theta}$. Clearly state any results you use.