Paper 1, Section I

7H Statistics

Suppose that $X_1, \ldots, X_n$ are independent normally distributed random variables, each with mean μ and variance 1, and consider testing $H_0 : \mu = 0$ against $H_1 : \mu = 1$. Explain what is meant by the *critical region*, the *size* and the *power* of a test.

For $0 < \alpha < 1$, derive the test that is most powerful among all tests of size at most α. Obtain an expression for the power of your test in terms of the standard normal distribution function $\Phi(\cdot)$.

[Results from the course may be used without proof provided they are clearly stated.]

Paper 2, Section I

8H Statistics

Suppose that, given θ, the random variable X has $\mathbb{P}(X = k) = e^{-\theta}\theta^k/k!$, $k = 0, 1, 2, \ldots$. Suppose that the prior density of θ is $\pi(\theta) = \lambda e^{-\lambda\theta}$, $\theta > 0$, for some known λ (> 0). Derive the posterior density $\pi(\theta \mid x)$ of θ based on the observation $X = x$.

For a given loss function $L(\theta, a)$, a statistician wants to calculate the value of a that minimises the expected posterior loss

$$\int L(\theta, a)\pi(\theta \mid x) d\theta.$$

Suppose that $x = 0$. Find a in terms of λ in the following cases:

(a) $L(\theta, a) = (\theta - a)^2$;

(b) $L(\theta, a) = |\theta - a|$.

Paper 4, Section II

19H Statistics

Consider a linear model $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where $\mathbf{Y}$ is an $n \times 1$ vector of observations, X is a known $n \times p$ matrix, $\boldsymbol{\beta}$ is a $p \times 1$ $(p < n)$ vector of unknown parameters and $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of independent normally distributed random variables each with mean zero and unknown variance σ^2. Write down the log-likelihood and show that the maximum likelihood estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ of $\boldsymbol{\beta}$ and σ^2 respectively satisfy

$$X^T X\hat{\boldsymbol{\beta}} = X^T\mathbf{Y}, \qquad \frac{1}{\hat{\sigma}^4}(\mathbf{Y} - X\hat{\boldsymbol{\beta}})^T(\mathbf{Y} - X\hat{\boldsymbol{\beta}}) = \frac{n}{\hat{\sigma}^2}$$

(^{T} denotes the transpose). Assuming that X^TX is invertible, find the solutions $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ of these equations and write down their distributions.

Prove that $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ are independent.

Consider the model $Y_{ij} = \mu_i + \gamma x_{ij} + \varepsilon_{ij}$, $i = 1, 2, 3$ and $j = 1, 2, 3$. Suppose that, for all i, $x_{i1} = -1$, $x_{i2} = 0$ and $x_{i3} = 1$, and that ε_{ij}, $i, j = 1, 2, 3$, are independent $N(0, \sigma^2)$ random variables where σ^2 is unknown. Show how this model may be written as a linear model and write down $\mathbf{Y}$, X, $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$. Find the maximum likelihood estimators of μ_i $(i = 1, 2, 3)$, γ and σ^2 in terms of the Y_{ij}. Derive a $100(1 - \alpha)\%$ confidence interval for σ^2 and for $\mu_2 - \mu_1$.

[You may assume that, if $\mathbf{W} = (\mathbf{W_1}^T, \mathbf{W_2}^T)^T$ is multivariate normal with $\mathrm{cov}(\mathbf{W_1}, \mathbf{W_2}) = 0$, then $\mathbf{W_1}$ and $\mathbf{W_2}$ are independent.]

Paper 1, Section II

19H Statistics

Suppose $X_1, \ldots, X_n$ are independent identically distributed random variables each with probability mass function $\mathbb{P}(X_i = x_i) = p(x_i; \theta)$, where θ is an unknown parameter. State what is meant by a *sufficient statistic* for θ. State the factorisation criterion for a sufficient statistic. State and prove the Rao–Blackwell theorem.

Suppose that $X_1, \ldots, X_n$ are independent identically distributed random variables with

$$\mathbb{P}(X_i = x_i) = \binom{m}{x_i} \theta^{x_i}(1-\theta)^{m-x_i}, \quad x_i = 0, \ldots, m,$$

where m is a known positive integer and θ is unknown. Show that $\tilde{\theta} = X_1/m$ is unbiased for θ.

Show that $T = \sum_{i=1}^{n} X_i$ is sufficient for θ and use the Rao–Blackwell theorem to find another unbiased estimator $\hat{\theta}$ for θ, giving details of your derivation. Calculate the variance of $\hat{\theta}$ and compare it to the variance of $\tilde{\theta}$.

A statistician cannot remember the exact statement of the Rao–Blackwell theorem and calculates $\mathbb{E}(T \mid X_1)$ in an attempt to find an estimator of θ. Comment on the suitability or otherwise of this approach, giving your reasons.

[*Hint: If a and b are positive integers then, for $r = 0, 1, \ldots, a+b$,* $\binom{a+b}{r} = \sum_{j=0}^{r} \binom{a}{j}\binom{b}{r-j}$.]

Paper 3, Section II

20H Statistics

(a) Suppose that $X_1, \ldots, X_n$ are independent identically distributed random variables, each with density $f(x) = \theta \exp(-\theta x)$, $x > 0$ for some unknown $\theta > 0$. Use the generalised likelihood ratio to obtain a size α test of $H_0 : \theta = 1$ against $H_1 : \theta \neq 1$.

(b) A die is loaded so that, if p_i is the probability of face i, then $p_1 = p_2 = \theta_1$, $p_3 = p_4 = \theta_2$ and $p_5 = p_6 = \theta_3$. The die is thrown n times and face i is observed x_i times. Write down the likelihood function for $\theta = (\theta_1, \theta_2, \theta_3)$ and find the maximum likelihood estimate of θ.

Consider testing whether or not $\theta_1 = \theta_2 = \theta_3$ for this die. Find the generalised likelihood ratio statistic Λ and show that

$$2 \log_e \Lambda \approx T, \qquad \text{where } T = \sum_{i=1}^{3} \frac{(o_i - e_i)^2}{e_i},$$

where you should specify o_i and e_i in terms of $x_1, \ldots, x_6$. Explain how to obtain an approximate size 0.05 test using the value of T. Explain what you would conclude (and why) if $T = 2.03$.

Paper 1, Section I

7H Statistics

Consider an estimator $\hat{\theta}$ of an unknown parameter θ, and assume that $\mathbb{E}_\theta\big(\hat{\theta}^2\big) < \infty$ for all θ. Define the *bias* and *mean squared error* of $\hat{\theta}$.

Show that the mean squared error of $\hat{\theta}$ is the sum of its variance and the square of its bias.

Suppose that $X_1, \ldots, X_n$ are independent identically distributed random variables with mean θ and variance θ^2, and consider estimators of θ of the form $k\bar{X}$ where $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$.

(i) Find the value of k that gives an unbiased estimator, and show that the mean squared error of this unbiased estimator is θ^2/n.

(ii) Find the range of values of k for which the mean squared error of $k\bar{X}$ is smaller than θ^2/n.

Paper 2, Section I

8H Statistics

There are 100 patients taking part in a trial of a new surgical procedure for a particular medical condition. Of these, 50 patients are randomly selected to receive the new procedure and the remaining 50 receive the old procedure. Six months later, a doctor assesses whether or not each patient has fully recovered. The results are shown below:

	Fully recovered	Not fully recovered
Old procedure	25	25
New procedure	31	19

The doctor is interested in whether there is a difference in full recovery rates for patients receiving the two procedures. Carry out an appropriate 5% significance level test, stating your hypotheses carefully. [You do not need to derive the test.] What conclusion should be reported to the doctor?

[*Hint: Let* $\chi_k^2(\alpha)$ *denote the upper* 100α *percentage point of a* χ_k^2 *distribution. Then*

$$\chi_1^2(0.05) = 3.84,\ \chi_2^2(0.05) = 5.99,\ \chi_3^2(0.05) = 7.82,\ \chi_4^2(0.05) = 9.49.]$$

Paper 4, Section II

19H Statistics

Consider a linear model

$$\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \qquad (\dagger)$$

where X is a known $n \times p$ matrix, $\boldsymbol{\beta}$ is a $p \times 1$ $(p < n)$ vector of unknown parameters and $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of independent $N(0, \sigma^2)$ random variables with σ^2 unknown. Assume that X has full rank p. Find the least squares estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ and derive its distribution. Define the residual sum of squares RSS and write down an unbiased estimator $\hat{\sigma}^2$ of σ^2.

Suppose that $V_i = a + bu_i + \delta_i$ and $Z_i = c + dw_i + \eta_i$, for $i = 1, \ldots, m$, where u_i and w_i are known with $\sum_{i=1}^m u_i = \sum_{i=1}^m w_i = 0$, and $\delta_1, \ldots, \delta_m, \eta_1, \ldots, \eta_m$ are independent $N(0, \sigma^2)$ random variables. Assume that at least two of the u_i are distinct and at least two of the w_i are distinct. Show that $\mathbf{Y} = (V_1, \ldots, V_m, Z_1, \ldots, Z_m)^T$ (where T denotes transpose) may be written as in $(\dagger)$ and identify X and $\boldsymbol{\beta}$. Find $\hat{\boldsymbol{\beta}}$ in terms of the V_i, Z_i, u_i and w_i. Find the distribution of $\hat{b} - \hat{d}$ and derive a 95% confidence interval for $b - d$.

[*Hint: You may assume that* $\frac{RSS}{\sigma^2}$ *has a* χ^2_{n-p} *distribution, and that* $\hat{\beta}$ *and the residual sum of squares are independent. Properties of* χ^2 *distributions may be used without proof.*]

Paper 1, Section II

19H Statistics

Suppose that X_1, X_2, and X_3 are independent identically distributed Poisson random variables with expectation θ, so that

$$\mathbb{P}(X_i = x) = \frac{e^{-\theta}\theta^x}{x!} \quad x = 0, 1, \ldots,$$

and consider testing $H_0 : \theta = 1$ against $H_1 : \theta = \theta_1$, where θ_1 is a known value greater than 1. Show that the test with critical region $\{(x_1, x_2, x_3) : \sum_{i=1}^{3} x_i > 5\}$ is a likelihood ratio test of H_0 against H_1. What is the size of this test? Write down an expression for its power.

A scientist counts the number of bird territories in n randomly selected sections of a large park. Let Y_i be the number of bird territories in the ith section, and suppose that $Y_1, \ldots, Y_n$ are independent Poisson random variables with expectations $\theta_1, \ldots, \theta_n$ respectively. Let a_i be the area of the ith section. Suppose that $n = 2m$, $a_1 = \cdots = a_m = a(> 0)$ and $a_{m+1} = \cdots = a_{2m} = 2a$. Derive the generalised likelihood ratio Λ for testing

$$H_0 : \theta_i = \lambda a_i \text{ against } H_1 : \theta_i = \begin{cases} \lambda_1 & i = 1, \ldots, m \\ \lambda_2 & i = m+1, \ldots, 2m. \end{cases}$$

What should the scientist conclude about the number of bird territories if $2\log_e(\Lambda)$ is 15.67?

[*Hint: Let $F_\theta(x)$ be $\mathbb{P}(W \leqslant x)$ where W has a Poisson distribution with expectation θ. Then*

$$F_1(3) = 0.998, \quad F_3(5) = 0.916, \quad F_3(6) = 0.966, \quad F_5(3) = 0.433\,.]$$

Paper 3, Section II

20H Statistics

Suppose that $X_1, \ldots, X_n$ are independent identically distributed random variables with

$$\mathbb{P}(X_i = x) = \binom{k}{x} \theta^x (1-\theta)^{k-x}, \quad x = 0, \ldots, k,$$

where k is known and θ $(0 < \theta < 1)$ is an unknown parameter. Find the maximum likelihood estimator $\hat{\theta}$ of θ.

Statistician 1 has prior density for θ given by $\pi_1(\theta) = \alpha\theta^{\alpha-1}$, $0 < \theta < 1$, where $\alpha > 1$. Find the posterior distribution for θ after observing data $X_1 = x_1, \ldots, X_n = x_n$. Write down the posterior mean $\hat{\theta}_1^{(B)}$, and show that

$$\hat{\theta}_1^{(B)} = c\,\hat{\theta} + (1-c)\tilde{\theta}_1,$$

where $\tilde{\theta}_1$ depends only on the prior distribution and c is a constant in $(0,1)$ that is to be specified.

Statistician 2 has prior density for θ given by $\pi_2(\theta) = \alpha(1-\theta)^{\alpha-1}$, $0 < \theta < 1$. Briefly describe the prior beliefs that the two statisticians hold about θ. Find the posterior mean $\hat{\theta}_2^{(B)}$ and show that $\hat{\theta}_2^{(B)} < \hat{\theta}_1^{(B)}$.

Suppose that α increases (but n, k and the x_i remain unchanged). How do the prior beliefs of the two statisticians change? How does c vary? Explain briefly what happens to $\hat{\theta}_1^{(B)}$ and $\hat{\theta}_2^{(B)}$.

[*Hint: The Beta*(α, β) $(\alpha > 0,\ \beta > 0)$ *distribution has density*

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1,$$

with expectation $\frac{\alpha}{\alpha+\beta}$ *and variance* $\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$. *Here,* $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$, $\alpha > 0$, *is the Gamma function.*]

Paper 1, Section I

7H Statistics

Let $x_1, \ldots, x_n$ be independent and identically distributed observations from a distribution with probability density function

$$f(x) = \begin{cases} \lambda e^{-\lambda(x-\mu)}, & x \geqslant \mu, \\ 0, & x < \mu, \end{cases}$$

where λ and μ are unknown positive parameters. Let $\beta = \mu + 1/\lambda$. Find the maximum likelihood estimators $\hat{\lambda}$, $\hat{\mu}$ and $\hat{\beta}$.

Determine for each of $\hat{\lambda}$, $\hat{\mu}$ and $\hat{\beta}$ whether or not it has a positive bias.

Paper 2, Section I

8H Statistics

State and prove the Rao–Blackwell theorem.

Individuals in a population are independently of three types $\{0, 1, 2\}$, with unknown probabilities p_0, p_1, p_2 where $p_0 + p_1 + p_2 = 1$. In a random sample of n people the ith person is found to be of type $x_i \in \{0, 1, 2\}$.

Show that an unbiased estimator of $\theta = p_0 p_1 p_2$ is

$$\hat{\theta} = \begin{cases} 1, & \text{if } (x_1, x_2, x_3) = (0, 1, 2), \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that n_i of the individuals are of type i. Find an unbiased estimator of θ, say θ^*, such that $\operatorname{var}(\theta^*) < \theta(1 - \theta)$.

Paper 4, Section II

19H Statistics

Explain the notion of a sufficient statistic.

Suppose X is a random variable with distribution F taking values in $\{1, \ldots, 6\}$, with $P(X = i) = p_i$. Let $x_1, \ldots, x_n$ be a sample from F. Suppose n_i is the number of these x_j that are equal to i. Use a factorization criterion to explain why $(n_1, \ldots, n_6)$ is sufficient for $\theta = (p_1, \ldots, p_6)$.

Let H_0 be the hypothesis that $p_i = 1/6$ for all i. Derive the statistic of the generalized likelihood ratio test of H_0 against the alternative that this is not a good fit.

Assuming that $n_i \approx n/6$ when H_0 is true and n is large, show that this test can be approximated by a chi-squared test using a test statistic

$$T = -n + \frac{6}{n} \sum_{i=1}^{6} n_i^2 .$$

Suppose $n = 100$ and $T = 8.12$. Would you reject H_0? Explain your answer.

Paper 1, Section II

19H Statistics

Consider the general linear model $Y = X\theta + \epsilon$ where X is a known $n \times p$ matrix, θ is an unknown $p \times 1$ vector of parameters, and ϵ is an $n \times 1$ vector of independent $N(0, \sigma^2)$ random variables with unknown variance σ^2. Assume the $p \times p$ matrix $X^T X$ is invertible. Let

$$\hat{\theta} = (X^T X)^{-1} X^T Y$$
$$\hat{\epsilon} = Y - X\hat{\theta}.$$

What are the distributions of $\hat{\theta}$ and $\hat{\epsilon}$? Show that $\hat{\theta}$ and $\hat{\epsilon}$ are uncorrelated.

Four apple trees stand in a 2×2 rectangular grid. The annual yield of the tree at coordinate (i, j) conforms to the model

$$y_{ij} = \alpha_i + \beta x_{ij} + \epsilon_{ij}, \quad i, j \in \{1, 2\},$$

where x_{ij} is the amount of fertilizer applied to tree (i, j), α_1, α_2 may differ because of varying soil across rows, and the ϵ_{ij} are $N(0, \sigma^2)$ random variables that are independent of one another and from year to year. The following two possible experiments are to be compared:

$$\text{I}: \ (x_{ij}) = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \quad \text{and} \quad \text{II}: \ (x_{ij}) = \begin{pmatrix} 0 & 2 \\ 3 & 1 \end{pmatrix}.$$

Represent these as general linear models, with $\theta = (\alpha_1, \alpha_2, \beta)$. Compare the variances of estimates of β under I and II.

With II the following yields are observed:

$$(y_{ij}) = \begin{pmatrix} 100 & 300 \\ 600 & 400 \end{pmatrix}.$$

Forecast the total yield that will be obtained next year if no fertilizer is used. What is the 95% predictive interval for this yield?

Paper 3, Section II
20H Statistics

Suppose x_1 is a single observation from a distribution with density f over $[0,1]$. It is desired to test $H_0 : f(x) = 1$ against $H_1 : f(x) = 2x$.

Let $\delta : [0,1] \to \{0,1\}$ define a test by $\delta(x_1) = i \iff$ 'accept H_i'. Let $\alpha_i(\delta) = P(\delta(x_1) = 1 - i \mid H_i)$. State the Neyman-Pearson lemma using this notation.

Let δ be the best test of size 0.10. Find δ and $\alpha_1(\delta)$.

Consider now $\delta : [0,1] \to \{0,1,\star\}$ where $\delta(x_1) = \star$ means 'declare the test to be inconclusive'. Let $\gamma_i(\delta) = P(\delta(x) = \star \mid H_i)$. Given prior probabilities π_0 for H_0 and $\pi_1 = 1 - \pi_0$ for H_1, and some w_0, w_1, let

$$\text{cost}(\delta) = \pi_0\big(w_0\alpha_0(\delta) + \gamma_0(\delta)\big) + \pi_1\big(w_1\alpha_1(\delta) + \gamma_1(\delta)\big).$$

Let $\delta^*(x_1) = i \iff x_1 \in A_i$, where $A_0 = [0, 0.5)$, $A_\star = [0.5, 0.6)$, $A_1 = [0.6, 1]$. Prove that for each value of $\pi_0 \in (0,1)$ there exist w_0, w_1 (depending on π_0) such that $\text{cost}(\delta^*) = \min_\delta \text{cost}(\delta)$. [*Hint*: $w_0 = 1 + 2(0.6)(\pi_1/\pi_0)$.]

Hence prove that if δ is any test for which

$$\alpha_i(\delta) \leqslant \alpha_i(\delta^*), \quad i = 0, 1$$

then $\gamma_0(\delta) \geqslant \gamma_0(\delta^*)$ and $\gamma_1(\delta) \geqslant \gamma_1(\delta^*)$.

Paper 1, Section I

7H Statistics

Describe the generalised likelihood ratio test and the type of statistical question for which it is useful.

Suppose that $X_1, \ldots, X_n$ are independent and identically distributed random variables with the Gamma$(2, \lambda)$ distribution, having density function $\lambda^2 x \exp(-\lambda x)$, $x \geqslant 0$. Similarly, $Y_1, \ldots, Y_n$ are independent and identically distributed with the Gamma$(2, \mu)$ distribution. It is desired to test the hypothesis $H_0 : \lambda = \mu$ against $H_1 : \lambda \neq \mu$. Derive the generalised likelihood ratio test and express it in terms of $R = \sum_i X_i / \sum_i Y_i$.

Let $F^{(1-\alpha)}_{\nu_1,\nu_2}$ denote the value that a random variable having the F_{ν_1,ν_2} distribution exceeds with probability α. Explain how to decide the outcome of a size 0.05 test when $n = 5$ by knowing only the value of R and the value $F^{(1-\alpha)}_{\nu_1,\nu_2}$, for some ν_1, ν_2 and α, which you should specify.

[You may use the fact that the χ^2_k distribution is equivalent to the Gamma$(k/2, 1/2)$ distribution.]

Paper 2, Section I

8H Statistics

Let the sample $x = (x_1, \ldots, x_n)$ have likelihood function $f(x; \theta)$. What does it mean to say $T(x)$ is a sufficient statistic for θ?

Show that if a certain factorization criterion is satisfied then T is sufficient for θ.

Suppose that T is sufficient for θ and there exist two samples, x and y, for which $T(x) \neq T(y)$ and $f(x; \theta)/f(y; \theta)$ does not depend on θ. Let

$$T_1(z) = \begin{cases} T(z) & z \neq y \\ T(x) & z = y. \end{cases}$$

Show that T_1 is also sufficient for θ.

Explain why T is not minimally sufficient for θ.

Paper 4, Section II

19H Statistics

From each of 3 populations, n data points are sampled and these are believed to obey

$$y_{ij} = \alpha_i + \beta_i(x_{ij} - \bar{x}_i) + \epsilon_{ij}, \quad j \in \{1, \ldots, n\},\ i \in \{1, 2, 3\},$$

where $\bar{x}_i = (1/n)\sum_j x_{ij}$, the ϵ_{ij} are independent and identically distributed as $N(0, \sigma^2)$, and σ^2 is unknown. Let $\bar{y}_i = (1/n)\sum_j y_{ij}$.

(i) Find expressions for $\hat{\alpha}_i$ and $\hat{\beta}_i$, the least squares estimates of α_i and β_i.

(ii) What are the distributions of $\hat{\alpha}_i$ and $\hat{\beta}_i$?

(iii) Show that the residual sum of squares, R_1, is given by

$$R_1 = \sum_{i=1}^{3} \left[\sum_{j=1}^{n} (y_{ij} - \bar{y}_i)^2 - \hat{\beta}_i^2 \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 \right].$$

Calculate R_1 when $n = 9$, $\{\hat{\alpha}_i\}_{i=1}^3 = \{1.6, 0.6, 0.8\}$, $\{\hat{\beta}_i\}_{i=1}^3 = \{2, 1, 1\}$,

$$\left\{ \sum_{j=1}^{9} (y_{ij} - \bar{y}_i)^2 \right\}_{i=1}^{3} = \{138, 82, 63\}, \quad \left\{ \sum_{j=1}^{9} (x_{ij} - \bar{x}_i)^2 \right\}_{i=1}^{3} = \{30, 60, 40\}.$$

(iv) H_0 is the hypothesis that $\alpha_1 = \alpha_2 = \alpha_3$. Find an expression for the maximum likelihood estimator of α_1 under the assumption that H_0 is true. Calculate its value for the above data.

(v) Explain (stating without proof any relevant theory) the rationale for a statistic which can be referred to an F distribution to test H_0 against the alternative that it is not true. What should be the degrees of freedom of this F distribution? What would be the outcome of a size 0.05 test of H_0 with the above data?

Paper 1, Section II

19H Statistics

State and prove the Neyman-Pearson lemma.

A sample of two independent observations, (x_1, x_2), is taken from a distribution with density $f(x;\theta) = \theta x^{\theta-1}$, $0 \leqslant x \leqslant 1$. It is desired to test $H_0 : \theta = 1$ against $H_1 : \theta = 2$. Show that the best test of size α can be expressed using the number c such that

$$1 - c + c \log c = \alpha\,.$$

Is this the uniformly most powerful test of size α for testing H_0 against $H_1 : \theta > 1$?

Suppose that the prior distribution of θ is $P(\theta = 1) = 4\gamma/(1 + 4\gamma)$, $P(\theta = 2) = 1/(1+4\gamma)$, where $1 > \gamma > 0$. Find the test of H_0 against H_1 that minimizes the probability of error.

Let $w(\theta)$ denote the power function of this test at θ ($\geqslant 1$). Show that

$$w(\theta) = 1 - \gamma^\theta + \gamma^\theta \log \gamma^\theta.$$

Paper 3, Section II

20H Statistics

Suppose that X is a single observation drawn from the uniform distribution on the interval $\big[\theta-10, \theta+10\big]$, where θ is unknown and might be any real number. Given $\theta_0 \neq 20$ we wish to test $H_0 : \theta = \theta_0$ against $H_1 : \theta = 20$. Let $\phi(\theta_0)$ be the test which accepts H_0 if and only if $X \in A(\theta_0)$, where

$$A(\theta_0) = \begin{cases} \big[\theta_0 - 8, \infty\big), & \theta_0 > 20 \\ \big(-\infty, \theta_0 + 8\big], & \theta_0 < 20\,. \end{cases}$$

Show that this test has size $\alpha = 0.10$.

Now consider

$$C_1(X) = \{\theta : X \in A(\theta)\},$$
$$C_2(X) = \big\{\theta : X - 9 \leqslant \theta \leqslant X + 9\big\}\,.$$

Prove that both $C_1(X)$ and $C_2(X)$ specify 90% confidence intervals for θ. Find the confidence interval specified by $C_1(X)$ when $X = 0$.

Let $L_i(X)$ be the length of the confidence interval specified by $C_i(X)$. Let $\beta(\theta_0)$ be the probability of the Type II error of $\phi(\theta_0)$. Show that

$$E[L_1(X) \mid \theta = 20] = E\left[\int_{-\infty}^{\infty} 1_{\{\theta_0 \in C_1(X)\}} d\theta_0 \;\middle|\; \theta = 20\right] = \int_{-\infty}^{\infty} \beta(\theta_0)\, d\theta_0.$$

Here $1_{\{B\}}$ is an indicator variable for event B. The expectation is over X. [Orders of integration and expectation can be interchanged.]

Use what you know about constructing best tests to explain which of the two confidence intervals has the smaller expected length when $\theta = 20$.

Paper 1, Section I

7H Statistics

Consider the experiment of tossing a coin n times. Assume that the tosses are independent and the coin is biased, with unknown probability p of heads and $1-p$ of tails. A total of X heads is observed.

(i) What is the maximum likelihood estimator $\widehat{p}$ of p?

Now suppose that a Bayesian statistician has the Beta(M, N) prior distribution for p.

(ii) What is the posterior distribution for p?

(iii) Assuming the loss function is $L(p, a) = (p-a)^2$, show that the statistician's point estimate for p is given by

$$\frac{M+X}{M+N+n}.$$

[The Beta(M, N) distribution has density $\dfrac{\Gamma(M+N)}{\Gamma(M)\Gamma(N)} x^{M-1}(1-x)^{N-1}$ for $0<x<1$ and mean $\dfrac{M}{M+N}$.]

Paper 2, Section I

8H Statistics

Let $X_1, \ldots, X_n$ be random variables with joint density function $f(x_1, \ldots, x_n; \theta)$, where θ is an unknown parameter. The null hypothesis $H_0 : \theta = \theta_0$ is to be tested against the alternative hypothesis $H_1 : \theta = \theta_1$.

(i) Define the following terms: critical region, Type I error, Type II error, size, power.

(ii) State and prove the Neyman–Pearson lemma.

Paper 1, Section II

19H Statistics

Let $X_1, \ldots, X_n$ be independent random variables with probability mass function $f(x;\theta)$, where θ is an unknown parameter.

(i) What does it mean to say that T is a sufficient statistic for θ? State, but do not prove, the factorisation criterion for sufficiency.

(ii) State and prove the Rao–Blackwell theorem.

Now consider the case where $f(x;\theta) = \dfrac{1}{x!}(-\log\theta)^x\theta$ for non-negative integer x and $0 < \theta < 1$.

(iii) Find a one-dimensional sufficient statistic T for θ.

(iv) Show that $\widetilde{\theta} = \mathbb{1}_{\{X_1=0\}}$ is an unbiased estimator of θ.

(v) Find another unbiased estimator $\widehat{\theta}$ which is a function of the sufficient statistic T and that has smaller variance than $\widetilde{\theta}$. You may use the following fact without proof: $X_1 + \cdots + X_n$ has the Poisson distribution with parameter $-n\log\theta$.

Paper 3, Section II

20H Statistics

Consider the general linear model

$$Y = X\beta + \epsilon$$

where X is a known $n \times p$ matrix, β is an unknown $p \times 1$ vector of parameters, and ϵ is an $n \times 1$ vector of independent $N(0, \sigma^2)$ random variables with unknown variance σ^2. Assume the $p \times p$ matrix $X^T X$ is invertible.

(i) Derive the least squares estimator $\widehat{\beta}$ of β.

(ii) Derive the distribution of $\widehat{\beta}$. Is $\widehat{\beta}$ an unbiased estimator of β?

(iii) Show that $\frac{1}{\sigma^2}\|Y - X\widehat{\beta}\|^2$ has the χ^2 distribution with k degrees of freedom, where k is to be determined.

(iv) Let $\widetilde{\beta}$ be an unbiased estimator of β of the form $\widetilde{\beta} = CY$ for some $p \times n$ matrix C. By considering the matrix $\mathbb{E}[(\widehat{\beta} - \widetilde{\beta})(\widehat{\beta} - \beta)^T]$ or otherwise, show that $\widehat{\beta}$ and $\widehat{\beta} - \widetilde{\beta}$ are independent.

[You may use standard facts about the multivariate normal distribution as well as results from linear algebra, including the fact that $I - X(X^T X)^{-1} X^T$ is a projection matrix of rank $n - p$, as long as they are carefully stated.]

Paper 4, Section II

19H Statistics

Consider independent random variables $X_1, \ldots, X_n$ with the $N(\mu_X, \sigma_X^2)$ distribution and $Y_1, \ldots, Y_n$ with the $N(\mu_Y, \sigma_Y^2)$ distribution, where the means μ_X, μ_Y and variances σ_X^2, σ_Y^2 are unknown. Derive the generalised likelihood ratio test of size α of the null hypothesis $H_0 : \sigma_X^2 = \sigma_Y^2$ against the alternative $H_1 : \sigma_X^2 \neq \sigma_Y^2$. Express the critical region in terms of the statistic $T = \dfrac{S_{XX}}{S_{XX} + S_{YY}}$ and the quantiles of a beta distribution, where

$$S_{XX} = \sum_{i=1}^n X_i^2 - \frac{1}{n}\left(\sum_{i=1}^n X_i\right)^2 \text{ and } S_{YY} = \sum_{i=1}^n Y_i^2 - \frac{1}{n}\left(\sum_{i=1}^n Y_i\right)^2 .$$

[You may use the following fact: if $U \sim \Gamma(a, \lambda)$ and $V \sim \Gamma(b, \lambda)$ are independent, then $\dfrac{U}{U+V} \sim \text{Beta}(a, b)$.]

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^2_1	χ^2_2	χ^2_3	χ^2_4	χ^2_5	χ^2_6
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 $\mathbb{E}[(Z-a)^4] = 3b^4$.]

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.

Paper 1, Section I

7H Statistics

What does it mean to say that an estimator $\hat{\theta}$ of a parameter θ is *unbiased*?

An n-vector Y of observations is believed to be explained by the model

$$Y = X\beta + \varepsilon,$$

where X is a known $n \times p$ matrix, β is an unknown p-vector of parameters, $p < n$, and ε is an n-vector of independent $N(0, \sigma^2)$ random variables. Find the maximum-likelihood estimator $\hat{\beta}$ of β, and show that it is unbiased.

Paper 3, Section I

8H Statistics

In a demographic study, researchers gather data on the gender of children in families with more than two children. For each of the four possible outcomes GG, GB, BG, BB of the first two children in the family, they find 50 families which started with that pair, and record the gender of the third child of the family. This produces the following table of counts:

First two children	Third child B	Third child G
GG	16	34
GB	28	22
BG	25	25
BB	31	19

In view of this, is the hypothesis that the gender of the third child is independent of the genders of the first two children rejected at the 5% level?

[*Hint: the* 95% *point of a* χ_3^2 *distribution is* 7.8147, *and the* 95% *point of a* χ_4^2 *distribution is* 9.4877.]

Paper 1, Section II

18H Statistics

What is the *critical region* C of a test of the null hypothesis $H_0 : \theta \in \Theta_0$ against the alternative $H_1 : \theta \in \Theta_1$? What is the *size* of a test with critical region C? What is the *power function* of a test with critical region C?

State and prove the Neyman–Pearson Lemma.

If $X_1, \ldots, X_n$ are independent with common $\text{Exp}(\lambda)$ distribution, and $0 < \lambda_0 < \lambda_1$, find the form of the most powerful size-α test of $H_0 : \lambda = \lambda_0$ against $H_1 : \lambda = \lambda_1$. Find the power function as explicitly as you can, and prove that it is increasing in λ. Deduce that the test you have constructed is a size-α test of $H_0 : \lambda \leqslant \lambda_0$ against $H_1 : \lambda = \lambda_1$.

Paper 2, Section II

19H Statistics

What does it mean to say that the random d-vector X has a *multivariate normal* distribution with mean μ and covariance matrix Σ?

Suppose that $X \sim N_d(0, \sigma^2 I_d)$, and that for each $j = 1, \ldots, J$, A_j is a $d_j \times d$ matrix. Suppose further that

$$A_j A_i^T = 0$$

for $j \neq i$. Prove that the random vectors $Y_j \equiv A_j X$ are independent, and that $Y \equiv (Y_1^T, \ldots, Y_J^T)^T$ has a multivariate normal distribution.

[*Hint: Random vectors are independent if their joint MGF is the product of their individual MGFs.*]

If $Z_1, \ldots, Z_n$ is an independent sample from a univariate $N(\mu, \sigma^2)$ distribution, prove that the sample variance $S_{ZZ} \equiv (n-1)^{-1} \sum_{i=1}^n (Z_i - \bar{Z})^2$ and the sample mean $\bar{Z} \equiv n^{-1} \sum_{i=1}^n Z_i$ are independent.

Paper 4, Section II

19H Statistics

What is a *sufficient statistic*? State the factorization criterion for a statistic to be sufficient.

Suppose that $X_1, \ldots, X_n$ are independent random variables uniformly distributed over $[a, b]$, where the parameters $a < b$ are not known, and $n \geqslant 2$. Find a sufficient statistic for the parameter $\theta \equiv (a, b)$ based on the sample $X_1, \ldots, X_n$. Based on your sufficient statistic, derive an unbiased estimator of θ.

1/I/7H **Statistics**

A Bayesian statistician observes a random sample $X_1, \ldots, X_n$ drawn from a $N(\mu, \tau^{-1})$ distribution. He has a prior density for the unknown parameters μ, τ of the form

$$\pi_0(\mu, \tau) \propto \tau^{\alpha_0 - 1} \exp\left(-\tfrac{1}{2} K_0 \tau (\mu - \mu_0)^2 - \beta_0 \tau\right) \sqrt{\tau},$$

where α_0, β_0, μ_0 and K_0 are constants which he chooses. Show that after observing $X_1, \ldots, X_n$ his posterior density $\pi_n(\mu, \tau)$ is again of the form

$$\pi_n(\mu, \tau) \propto \tau^{\alpha_n - 1} \exp\left(-\tfrac{1}{2} K_n \tau (\mu - \mu_n)^2 - \beta_n \tau\right) \sqrt{\tau},$$

where you should find explicitly the form of α_n, β_n, μ_n and K_n.

1/II/18H **Statistics**

Suppose that $X_1, \ldots, X_n$ is a sample of size n with common $N(\mu_X, 1)$ distribution, and $Y_1, \ldots, Y_n$ is an independent sample of size n from a $N(\mu_Y, 1)$ distribution.

(i) Find (with careful justification) the form of the size-α likelihood–ratio test of the null hypothesis $H_0 : \mu_Y = 0$ against alternative $H_1 : (\mu_X, \mu_Y)$ unrestricted.

(ii) Find the form of the size-α likelihood–ratio test of the hypothesis

$$H_0 : \mu_X \geqslant A, \mu_Y = 0,$$

against $H_1 : (\mu_X, \mu_Y)$ unrestricted, where A is a given constant.

Compare the critical regions you obtain in (i) and (ii) and comment briefly.

2/II/19H **Statistics**

Suppose that the joint distribution of random variables X, Y taking values in $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ is given by the joint probability generating function

$$\varphi(s,t) \equiv E[s^X t^Y] = \frac{1-\alpha-\beta}{1-\alpha s-\beta t},$$

where the unknown parameters α and β are positive, and satisfy the inequality $\alpha+\beta<1$. Find $E(X)$. Prove that the probability mass function of (X, Y) is

$$f(x, y \,|\, \alpha, \beta) = (1-\alpha-\beta)\binom{x+y}{x}\alpha^x\beta^y \qquad (x, y \in \mathbb{Z}^+),$$

and prove that the maximum-likelihood estimators of α and β based on a sample of size n drawn from the distribution are

$$\hat{\alpha} = \frac{\overline{X}}{1+\overline{X}+\overline{Y}}, \qquad \hat{\beta} = \frac{\overline{Y}}{1+\overline{X}+\overline{Y}},$$

where $\overline{X}$ (respectively, $\overline{Y}$) is the sample mean of $X_1, \dots, X_n$ (respectively, $Y_1, \dots, Y_n$).

By considering $\hat{\alpha}+\hat{\beta}$ or otherwise, prove that the maximum-likelihood estimator is biased. Stating clearly any results to which you appeal, prove that as $n \to \infty$, $\hat{\alpha} \to \alpha$, making clear the sense in which this convergence happens.

3/I/8H **Statistics**

If $X_1, \dots, X_n$ is a sample from a density $f(\cdot|\theta)$ with θ unknown, what is a 95% confidence set for θ?

In the case where the X_i are independent $N(\mu, \sigma^2)$ random variables with σ^2 known, μ unknown, find (in terms of σ^2) how large the size n of the sample must be in order for there to exist a 95% confidence interval for μ of length no more than some given $\varepsilon > 0$.

[*Hint: If* $Z \sim N(0,1)$ *then* $P(Z > 1.960) = 0.025$.]

4/II/19H **Statistics**

(i) Consider the linear model

$$Y_i = \alpha + \beta x_i + \varepsilon_i \,,$$

where observations Y_i, $i = 1, \ldots, n$, depend on known explanatory variables x_i, $i = 1, \ldots, n$, and independent $N(0, \sigma^2)$ random variables ε_i, $i = 1, \ldots, n$.

Derive the maximum-likelihood estimators of α, β and σ^2.

Stating clearly any results you require about the distribution of the maximum-likelihood estimators of α, β and σ^2, explain how to construct a test of the hypothesis that $\alpha = 0$ against an unrestricted alternative.

(ii) A simple ballistic theory predicts that the range of a gun fired at angle of elevation θ should be given by the formula

$$Y = \frac{V^2}{g} \sin 2\theta \,,$$

where V is the muzzle velocity, and g is the gravitational acceleration. Shells are fired at 9 different elevations, and the ranges observed are as follows:

θ (degrees)	5	15	25	35	45	55	65	75	85
$\sin 2\theta$	0.1736	0.5	0.7660	0.9397	1	0.9397	0.7660	0.5	0.1736
Y (m)	4322	11898	17485	20664	21296	19491	15572	10027	3458

The model

$$Y_i = \alpha + \beta \sin 2\theta_i + \varepsilon_i \qquad (*)$$

is proposed. Using the theory of part (i) above, find expressions for the maximum-likelihood estimators of α and β.

The t-test of the null hypothesis that $\alpha = 0$ against an unrestricted alternative does not reject the null hypothesis. Would you be willing to accept the model $(*)$? Briefly explain your answer.

[*You may need the following summary statistics of the data. If* $x_i = \sin 2\theta_i$, *then* $\bar{x} \equiv n^{-1} \sum x_i = 0.63986$, $\bar{Y} = 13802$, $S_{xx} \equiv \sum (x_i - \bar{x})^2 = 0.81517$, $S_{xy} = \sum Y_i (x_i - \bar{x}) = 17186$.]

1/I/7C **Statistics**

Let $X_1, \ldots, X_n$ be independent, identically distributed random variables from the $N\left(\mu, \sigma^2\right)$ distribution where μ and σ^2 are unknown. Use the generalized likelihood-ratio test to derive the form of a test of the hypothesis $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$.

Explain carefully how the test should be implemented.

1/II/18C **Statistics**

Let $X_1, \ldots, X_n$ be independent, identically distributed random variables with

$$\mathbb{P}\left(X_i = 1\right) = \theta = 1 - \mathbb{P}\left(X_i = 0\right) ,$$

where θ is an unknown parameter, $0 < \theta < 1$, and $n \geqslant 2$. It is desired to estimate the quantity $\phi = \theta(1-\theta) = n\mathbb{V}\mathrm{ar}\left(\left(X_1 + \cdots + X_n\right)/n\right)$.

(i) Find the maximum-likelihood estimate, $\hat{\phi}$, of ϕ.

(ii) Show that $\hat{\phi}_1 = X_1\left(1 - X_2\right)$ is an unbiased estimate of ϕ and hence, or otherwise, obtain an unbiased estimate of ϕ which has smaller variance than $\hat{\phi}_1$ and which is a function of $\hat{\phi}$.

(iii) Now suppose that a Bayesian approach is adopted and that the prior distribution for θ, $\pi(\theta)$, is taken to be the uniform distribution on $(0,1)$. Compute the Bayes point estimate of ϕ when the loss function is $L(\phi, a) = (\phi - a)^2$.

[*You may use that fact that when r, s are non-negative integers,*

$$\int_0^1 x^r (1-x)^s dx = r!s!/(r+s+1)! \quad]$$

2/II/19C **Statistics**

State and prove the Neyman–Pearson lemma.

Suppose that X is a random variable drawn from the probability density function

$$f(x \mid \theta) = \tfrac{1}{2} |\, x \,|^{\theta-1} e^{-|x|} / \Gamma(\theta), \quad -\infty < x < \infty\,,$$

where $\Gamma(\theta) = \int\limits_0^\infty y^{\theta-1} e^{-y} dy$ and $\theta \geqslant 1$ is unknown. Find the most powerful test of size α, $0 < \alpha < 1$, of the hypothesis $H_0 : \theta = 1$ against the alternative $H_1 : \theta = 2$. Express the power of the test as a function of α.

Is your test uniformly most powerful for testing $H_0 : \theta = 1$ against $H_1 : \theta > 1$? Explain your answer carefully.

3/I/8C **Statistics**

Light bulbs are sold in packets of 3 but some of the bulbs are defective. A sample of 256 packets yields the following figures for the number of defectives in a packet:

No. of defectives	0	1	2	3
No. of packets	116	94	40	6

Test the hypothesis that each bulb has a constant (but unknown) probability θ of being defective independently of all other bulbs.

[*Hint: You may wish to use some of the following percentage points:*

Distribution	χ_1^2	χ_2^2	χ_3^2	χ_4^2	t_1	t_2	t_3	t_4
90% *percentile*	2·71	4·61	6·25	7·78	3·08	1·89	1·64	1·53
95% *percentile*	3·84	5·99	7·81	9·49	6·31	2·92	2·35	2·13

]

4/II/19C **Statistics**

Consider the linear regression model

$$Y_i = \alpha + \beta x_i + \epsilon_i, \qquad 1 \leqslant i \leqslant n\,,$$

where $\epsilon_1, \ldots, \epsilon_n$ are independent, identically distributed $N(0, \sigma^2)$, $x_1, \ldots, x_n$ are known real numbers with $\sum_{i=1}^n x_i = 0$ and α, β and σ^2 are unknown.

(i) Find the least-squares estimates $\widehat{\alpha}$ and $\widehat{\beta}$ of α and β, respectively, and explain why in this case they are the same as the maximum-likelihood estimates.

(ii) Determine the maximum-likelihood estimate $\widehat{\sigma}^2$ of σ^2 and find a multiple of it which is an unbiased estimate of σ^2.

(iii) Determine the joint distribution of $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma}^2$.

(iv) Explain carefully how you would test the hypothesis $H_0 : \alpha = \alpha_0$ against the alternative $H_1 : \alpha \neq \alpha_0$.

1/I/7C **Statistics**

A random sample $X_1, \ldots, X_n$ is taken from a normal distribution having unknown mean θ and variance 1. Find the maximum likelihood estimate $\hat{\theta}_M$ for θ based on $X_1, \ldots, X_n$.

Suppose that we now take a Bayesian point of view and regard θ itself as a normal random variable of known mean μ and variance τ^{-1}. Find the Bayes' estimate $\hat{\theta}_B$ for θ based on $X_1, \ldots, X_n$, corresponding to the quadratic loss function $(\theta - a)^2$.

1/II/18C **Statistics**

Let X be a random variable whose distribution depends on an unknown parameter θ. Explain what is meant by a sufficient statistic $T(X)$ for θ.

In the case where X is discrete, with probability mass function $f(x|\theta)$, explain, with justification, how a sufficient statistic may be found.

Assume now that $X = (X_1, \ldots, X_n)$, where $X_1, \ldots, X_n$ are independent non-negative random variables with common density function

$$f(x|\theta) = \begin{cases} \lambda e^{-\lambda(x-\theta)} & \text{if } x \geqslant \theta, \\ 0 & \text{otherwise.} \end{cases}$$

Here $\theta \geq 0$ is unknown and λ is a known positive parameter. Find a sufficient statistic for θ and hence obtain an unbiased estimator $\hat{\theta}$ for θ of variance $(n\lambda)^{-2}$.

[*You may use without proof the following facts: for independent exponential random variables X and Y, having parameters λ and μ respectively, X has mean λ^{-1} and variance λ^{-2} and min$\{X, Y\}$ has exponential distribution of parameter $\lambda + \mu$.*]

2/II/19C **Statistics**

Suppose that $X_1, \ldots, X_n$ are independent normal random variables of unknown mean θ and variance 1. It is desired to test the hypothesis $H_0 : \theta \leq 0$ against the alternative $H_1 : \theta > 0$. Show that there is a uniformly most powerful test of size $\alpha = 1/20$ and identify a critical region for such a test in the case $n = 9$. If you appeal to any theoretical result from the course you should also prove it.

[*The 95th percentile of the standard normal distribution is 1.65.*]

3/I/8C **Statistics**

One hundred children were asked whether they preferred crisps, fruit or chocolate. Of the boys, 12 stated a preference for crisps, 11 for fruit, and 17 for chocolate. Of the girls, 13 stated a preference for crisps, 14 for fruit, and 33 for chocolate. Answer each of the following questions by carrying out an appropriate statistical test.

(a) Are the data consistent with the hypothesis that girls find all three types of snack equally attractive?

(b) Are the data consistent with the hypothesis that boys and girls show the same distribution of preferences?

4/II/19C **Statistics**

Two series of experiments are performed, the first resulting in observations $X_1, \ldots, X_m$, the second resulting in observations $Y_1, \ldots, Y_n$. We assume that all observations are independent and normally distributed, with unknown means μ_X in the first series and μ_Y in the second series. We assume further that the variances of the observations are unknown but are all equal.

Write down the distributions of the sample mean $\bar{X} = m^{-1} \sum_{i=1}^{m} X_i$ and sum of squares $S_{XX} = \sum_{i=1}^{m} (X_i - \bar{X})^2$.

Hence obtain a statistic $T(X, Y)$ to test the hypothesis $H_0 : \mu_X = \mu_Y$ against $H_1 : \mu_X > \mu_Y$ and derive its distribution under H_0. Explain how you would carry out a test of size $\alpha = 1/100$.

1/I/7D **Statistics**

The fast-food chain McGonagles have three sizes of their takeaway haggis, Large, Jumbo and Soopersize. A survey of 300 randomly selected customers at one branch choose 92 Large, 89 Jumbo and 119 Soopersize haggises.

Is there sufficient evidence to reject the hypothesis that all three sizes are equally popular? Explain your answer carefully.

Distribution	t_1	t_2	t_3	χ^2_1	χ^2_2	χ^2_3	$F_{1,2}$	$F_{2,3}$
95% percentile	6·31	2·92	2·35	3·84	5·99	7·82	18·51	9·55

1/II/18D **Statistics**

In the context of hypothesis testing define the following terms: (i) simple hypothesis; (ii) critical region; (iii) size; (iv) power; and (v) type II error probability.

State, without proof, the Neyman–Pearson lemma.

Let X be a single observation from a probability density function f. It is desired to test the hypothesis

$$H_0 : f = f_0 \qquad \text{against} \qquad H_1 : f = f_1,$$

with $f_0(x) = \frac{1}{2}|x|\, e^{-x^2/2}$ and $f_1(x) = \Phi'(x)$, $-\infty < x < \infty$, where $\Phi(x)$ is the distribution function of the standard normal, $N(0,1)$.

Determine the best test of size α, where $0 < \alpha < 1$, and express its power in terms of Φ and α.

Find the size of the test that minimizes the sum of the error probabilities. Explain your reasoning carefully.

2/II/19D **Statistics**

Let $X_1, \ldots, X_n$ be a random sample from a probability density function $f(x \mid \theta)$, where θ is an unknown real-valued parameter which is assumed to have a prior density $\pi(\theta)$. Determine the optimal Bayes point estimate $a(X_1, \ldots, X_n)$ of θ, in terms of the posterior distribution of θ given $X_1, \ldots, X_n$, when the loss function is

$$L(\theta, a) = \begin{cases} \gamma(\theta - a) & \text{when } \theta \geqslant a, \\ \delta(a - \theta) & \text{when } \theta \leqslant a, \end{cases}$$

where γ and δ are given positive constants.

Calculate the estimate explicitly in the case when $f(x \mid \theta)$ is the density of the uniform distribution on $(0, \theta)$ and $\pi(\theta) = e^{-\theta}\theta^n/n!$, $\theta > 0$.

3/I/8D **Statistics**

Let $X_1, \ldots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2, where μ and σ^2 are unknown. Derive the form of the size-α generalized likelihood-ratio test of the hypothesis $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$, and show that it is equivalent to the standard t-test of size α.

[*You should state, but need not derive, the distribution of the test statistic.*]

4/II/19D **Statistics**

Let $Y_1, \ldots, Y_n$ be observations satisfying

$$Y_i = \beta x_i + \epsilon_i, \quad 1 \leqslant i \leqslant n,$$

where $\epsilon_1, \ldots, \epsilon_n$ are independent random variables each with the $N(0, \sigma^2)$ distribution. Here $x_1, \ldots, x_n$ are known but β and σ^2 are unknown.

(i) Determine the maximum-likelihood estimates $(\widehat{\beta}, \widehat{\sigma}^2)$ of (β, σ^2).

(ii) Find the distribution of $\widehat{\beta}$.

(iii) By showing that $Y_i - \widehat{\beta} x_i$ and $\widehat{\beta}$ are independent, or otherwise, determine the joint distribution of $\widehat{\beta}$ and $\widehat{\sigma}^2$.

(iv) Explain carefully how you would test the hypothesis $H_0 : \beta = \beta_0$ against $H_1 : \beta \neq \beta_0$.

1/I/10H **Statistics**

Use the generalized likelihood-ratio test to derive Student's t-test for the equality of the means of two populations. You should explain carefully the assumptions underlying the test.

1/II/21H **Statistics**

State and prove the Rao–Blackwell Theorem.

Suppose that $X_1, X_2, \ldots, X_n$ are independent, identically-distributed random variables with distribution

$$P\left(X_1 = r\right) = p^{r-1}(1-p), \quad r = 1, 2, \ldots,$$

where p, $0 < p < 1$, is an unknown parameter. Determine a one-dimensional sufficient statistic, T, for p.

By first finding a simple unbiased estimate for p, or otherwise, determine an unbiased estimate for p which is a function of T.

2/I/10H **Statistics**

A study of 60 men and 90 women classified each individual according to eye colour to produce the figures below.

	Blue	Brown	Green
Men	20	20	20
Women	20	50	20

Explain how you would analyse these results. You should indicate carefully any underlying assumptions that you are making.

A further study took 150 individuals and classified them both by eye colour and by whether they were left or right handed to produce the following table.

	Blue	Brown	Green
Left Handed	20	20	20
Right Handed	20	50	20

How would your analysis change? You should again set out your underlying assumptions carefully.

[*You may wish to note the following percentiles of the* χ^2 *distribution.*

	χ_1^2	χ_2^2	χ_3^2	χ_4^2	χ_5^2	χ_6^2
95% percentile	3.84	5.99	7.81	9.49	11.07	12.59
99% percentile	6.64	9.21	11.34	13.28	15.09	16.81

]

2/II/21H **Statistics**

Defining carefully the terminology that you use, state and prove the Neyman–Pearson Lemma.

Let X be a single observation from the distribution with density function

$$f(x \mid \theta) = \tfrac{1}{2}e^{-|x-\theta|}, \quad -\infty < x < \infty,$$

for an unknown real parameter θ. Find the best test of size α, $0 < \alpha < 1$, of the hypothesis $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$, where $\theta_1 > \theta_0$.

When $\alpha = 0.05$, for which values of θ_0 and θ_1 will the power of the best test be at least 0.95?

4/I/9H **Statistics**

Suppose that $Y_1, \ldots, Y_n$ are independent random variables, with Y_i having the normal distribution with mean βx_i and variance σ^2; here β, σ^2 are unknown and $x_1, \ldots, x_n$ are known constants.

Derive the least-squares estimate of β.

Explain carefully how to test the hypothesis $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$.

4/II/19H **Statistics**

It is required to estimate the unknown parameter θ after observing X, a single random variable with probability density function $f(x \mid \theta)$; the parameter θ has the prior distribution with density $\pi(\theta)$ and the loss function is $L(\theta, a)$. Show that the optimal Bayesian point estimate minimizes the posterior expected loss.

Suppose now that $f(x \mid \theta) = \theta e^{-\theta x}$, $x > 0$ and $\pi(\theta) = \mu e^{-\mu\theta}$, $\theta > 0$, where $\mu > 0$ is known. Determine the posterior distribution of θ given X.

Determine the optimal Bayesian point estimate of θ in the cases when

(i) $L(\theta, a) = (\theta - a)^2$, and

(ii) $L(\theta, a) = |(\theta - a)/\theta|$.