

**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  $\hat{p}$  of  $p$ ?

Now suppose that a Bayesian statistician has the  $\text{Beta}(M, N)$  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  $\text{Beta}(M, N)$  distribution has density  $\frac{\Gamma(M + N)}{\Gamma(M)\Gamma(N)}x^{M-1}(1 - x)^{N-1}$  for  $0 < x < 1$  and mean  $\frac{M}{M + N}$ .]

**Paper 2, Section I****8H Statistics**

Let  $X_1, \dots, X_n$  be random variables with joint density function  $f(x_1, \dots, x_n; \theta)$ , where  $\theta$  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  $X_1, \dots, X_n$  be independent random variables with probability mass function  $f(x; \theta)$ , where  $\theta$  is an unknown parameter.

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

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

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

(iii) Find a one-dimensional sufficient statistic  $T$  for  $\theta$ .

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

(v) Find another unbiased estimator  $\hat{\theta}$  which is a function of the sufficient statistic  $T$  and that has smaller variance than  $\tilde{\theta}$ . You may use the following fact without proof:  $X_1 + \dots + X_n$  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,  $\beta$  is an unknown  $p \times 1$  vector of parameters, and  $\epsilon$  is an  $n \times 1$  vector of independent  $N(0, \sigma^2)$  random variables with unknown variance  $\sigma^2$ . Assume the  $p \times p$  matrix  $X^T X$  is invertible.

- (i) Derive the least squares estimator  $\hat{\beta}$  of  $\beta$ .
- (ii) Derive the distribution of  $\hat{\beta}$ . Is  $\hat{\beta}$  an unbiased estimator of  $\beta$ ?
- (iii) Show that  $\frac{1}{\sigma^2} \|Y - X\hat{\beta}\|^2$  has the  $\chi^2$  distribution with  $k$  degrees of freedom, where  $k$  is to be determined.
- (iv) Let  $\tilde{\beta}$  be an unbiased estimator of  $\beta$  of the form  $\tilde{\beta} = CY$  for some  $p \times n$  matrix  $C$ . By considering the matrix  $\mathbb{E}[(\hat{\beta} - \tilde{\beta})(\hat{\beta} - \tilde{\beta})^T]$  or otherwise, show that  $\hat{\beta}$  and  $\hat{\beta} - \tilde{\beta}$  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  $X_1, \dots, X_n$  with the  $N(\mu_X, \sigma_X^2)$  distribution and  $Y_1, \dots, Y_n$  with the  $N(\mu_Y, \sigma_Y^2)$  distribution, where the means  $\mu_X, \mu_Y$  and variances  $\sigma_X^2, \sigma_Y^2$  are unknown. Derive the generalised likelihood ratio test of size  $\alpha$  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  $T = \frac{S_{XX}}{S_{XX} + S_{YY}}$  and the quantiles of a beta distribution, where

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

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

**Paper 1, Section I**
**7E Statistics**

Suppose  $X_1, \dots, X_n$  are independent  $N(0, \sigma^2)$  random variables, where  $\sigma^2$  is an unknown parameter. Explain carefully how to construct the uniformly most powerful test of size  $\alpha$  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:

	$\chi_1^2$	$\chi_2^2$	$\chi_3^2$	$\chi_4^2$	$\chi_5^2$	$\chi_6^2$
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  $x_1, \dots, x_n$  are known, the independent random variables  $\epsilon_1, \dots, \epsilon_n$  have the  $N(0, \sigma^2)$  distribution, and the parameters  $\beta$  and  $\sigma^2$  are unknown. Find the maximum likelihood estimator for  $\beta$ .

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

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

$$\mathbb{E}_{\beta, \sigma^2}[(\hat{\beta} - \beta)^4] \leq \mathbb{E}_{\beta, \sigma^2}[(\tilde{\beta} - \beta)^4]$$

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

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

**Paper 3, Section II**
**20E Statistics**

Let  $X_1, \dots, X_n$  be independent  $\text{Exp}(\theta)$  random variables with unknown parameter  $\theta$ . Find the maximum likelihood estimator  $\hat{\theta}$  of  $\theta$ , 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  $\theta$  of the form  $[0, C\hat{\theta}]$  for a constant  $C > 0$  depending on  $\alpha$ .

Now, taking a Bayesian point of view, suppose your prior distribution for the parameter  $\theta$  is  $\Gamma(k, \lambda)$ . Show that your Bayesian point estimator  $\hat{\theta}_B$  of  $\theta$  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  $\alpha$  such that the posterior probability that  $\theta \leq C_B \hat{\theta}_B$  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  $X_1, \dots, X_n$  of independent random variables with common density function  $f(x; \theta)$  depending on a real parameter  $\theta$ . What does it mean to say  $T$  is a sufficient statistic for  $\theta$ ? Prove that if the joint density of  $X_1, \dots, X_n$  satisfies the factorisation criterion for a statistic  $T$ , then  $T$  is sufficient for  $\theta$ .

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  $\theta$ , or otherwise, find an unbiased estimator of  $\theta$  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  $\theta$  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,  $\beta$  is an unknown  $p$ -vector of parameters,  $p < n$ , and  $\varepsilon$  is an  $n$ -vector of independent  $N(0, \sigma^2)$  random variables. Find the maximum-likelihood estimator  $\hat{\beta}$  of  $\beta$ , 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  $\chi_3^2$  distribution is 7.8147, and the 95% point of a  $\chi_4^2$  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  $X_1, \dots, X_n$  are independent with common  $\text{Exp}(\lambda)$  distribution, and  $0 < \lambda_0 < \lambda_1$ , find the form of the most powerful size- $\alpha$  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  $\lambda$ . Deduce that the test you have constructed is a size- $\alpha$  test of  $H_0 : \lambda \leq \lambda_0$  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  $\mu$  and covariance matrix  $\Sigma$ ?

Suppose that  $X \sim N_d(0, \sigma^2 I_d)$ , and that for each  $j = 1, \dots, J$ ,  $A_j$  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  $Y \equiv (Y_1^T, \dots, Y_J^T)^T$  has a multivariate normal distribution.

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

If  $Z_1, \dots, Z_n$  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  $X_1, \dots, X_n$  are independent random variables uniformly distributed over  $[a, b]$ , where the parameters  $a < b$  are not known, and  $n \geq 2$ . Find a sufficient statistic for the parameter  $\theta \equiv (a, b)$  based on the sample  $X_1, \dots, X_n$ . Based on your sufficient statistic, derive an unbiased estimator of  $\theta$ .

**1/I/7H Statistics**

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

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

where  $\alpha_0, \beta_0, \mu_0$  and  $K_0$  are constants which he chooses. Show that after observing  $X_1, \dots, X_n$  his posterior density  $\pi_n(\mu, \tau)$  is again of the form

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

where you should find explicitly the form of  $\alpha_n, \beta_n, \mu_n$  and  $K_n$ .

**1/II/18H Statistics**

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

- (i) Find (with careful justification) the form of the size- $\alpha$  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- $\alpha$  likelihood-ratio test of the hypothesis

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

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

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

and prove that the maximum-likelihood estimators of  $\alpha$  and  $\beta$  based on a sample of size  $n$  drawn from the distribution are

$$\hat{\alpha} = \frac{\bar{X}}{1 + \bar{X} + \bar{Y}}, \quad \hat{\beta} = \frac{\bar{Y}}{1 + \bar{X} + \bar{Y}},$$

where  $\bar{X}$  (respectively,  $\bar{Y}$ ) 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  $n \rightarrow \infty$ ,  $\hat{\alpha} \rightarrow \alpha$ , 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  $\theta$  unknown, what is a 95% confidence set for  $\theta$ ?

In the case where the  $X_i$  are independent  $N(\mu, \sigma^2)$  random variables with  $\sigma^2$  known,  $\mu$  unknown, find (in terms of  $\sigma^2$ ) how large the size  $n$  of the sample must be in order for there to exist a 95% confidence interval for  $\mu$  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  $Y_i$ ,  $i = 1, \dots, n$ , depend on known explanatory variables  $x_i$ ,  $i = 1, \dots, n$ , and independent  $N(0, \sigma^2)$  random variables  $\varepsilon_i$ ,  $i = 1, \dots, n$ .

Derive the maximum-likelihood estimators of  $\alpha$ ,  $\beta$  and  $\sigma^2$ .

Stating clearly any results you require about the distribution of the maximum-likelihood estimators of  $\alpha$ ,  $\beta$  and  $\sigma^2$ , 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  $\theta$  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:

$\theta$ (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

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

is proposed. Using the theory of part (i) above, find expressions for the maximum-likelihood estimators of  $\alpha$  and  $\beta$ .

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  $X_1, \dots, X_n$  be independent, identically distributed random variables from the  $N(\mu, \sigma^2)$  distribution where  $\mu$  and  $\sigma^2$  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  $X_1, \dots, X_n$  be independent, identically distributed random variables with

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

where  $\theta$  is an unknown parameter,  $0 < \theta < 1$ , and  $n \geq 2$ . It is desired to estimate the quantity  $\phi = \theta(1 - \theta) = n\text{Var}((X_1 + \dots + X_n)/n)$ .

- (i) Find the maximum-likelihood estimate,  $\hat{\phi}$ , of  $\phi$ .
- (ii) Show that  $\hat{\phi}_1 = X_1(1 - X_2)$  is an unbiased estimate of  $\phi$  and hence, or otherwise, obtain an unbiased estimate of  $\phi$  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  $\theta$ ,  $\pi(\theta)$ , is taken to be the uniform distribution on  $(0, 1)$ . Compute the Bayes point estimate of  $\phi$  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

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

where  $\Gamma(\theta) = \int_0^\infty y^{\theta-1} e^{-y} dy$  and  $\theta \geq 1$  is unknown. Find the most powerful test of size  $\alpha$ ,  $0 < \alpha < 1$ , of the hypothesis  $H_0 : \theta = 1$  against the alternative  $H_1 : \theta = 2$ . Express the power of the test as a function of  $\alpha$ .

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  $\theta$  of being defective independently of all other bulbs.

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

Distribution	$\chi_1^2$	$\chi_2^2$	$\chi_3^2$	$\chi_4^2$	$t_1$	$t_2$	$t_3$	$t_4$
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

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

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

- (i) Find the least-squares estimates  $\hat{\alpha}$  and  $\hat{\beta}$  of  $\alpha$  and  $\beta$ , respectively, and explain why in this case they are the same as the maximum-likelihood estimates.
- (ii) Determine the maximum-likelihood estimate  $\hat{\sigma}^2$  of  $\sigma^2$  and find a multiple of it which is an unbiased estimate of  $\sigma^2$ .
- (iii) Determine the joint distribution of  $\hat{\alpha}$ ,  $\hat{\beta}$  and  $\hat{\sigma}^2$ .
- (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  $X_1, \dots, X_n$  is taken from a normal distribution having unknown mean  $\theta$  and variance 1. Find the maximum likelihood estimate  $\hat{\theta}_M$  for  $\theta$  based on  $X_1, \dots, X_n$ .

Suppose that we now take a Bayesian point of view and regard  $\theta$  itself as a normal random variable of known mean  $\mu$  and variance  $\tau^{-1}$ . Find the Bayes' estimate  $\hat{\theta}_B$  for  $\theta$  based on  $X_1, \dots, X_n$ , 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  $\theta$ . Explain what is meant by a sufficient statistic  $T(X)$  for  $\theta$ .

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

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

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

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

2/II/19C     **Statistics**

Suppose that  $X_1, \dots, X_n$  are independent normal random variables of unknown mean  $\theta$  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  $X_1, \dots, X_m$ , the second resulting in observations  $Y_1, \dots, Y_n$ . We assume that all observations are independent and normally distributed, with unknown means  $\mu_X$  in the first series and  $\mu_Y$  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.

<i>Distribution</i>	$t_1$	$t_2$	$t_3$	$\chi_1^2$	$\chi_2^2$	$\chi_3^2$	$F_{1,2}$	$F_{2,3}$
<i>95% percentile</i>	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

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

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  $\alpha$ , where  $0 < \alpha < 1$ , and express its power in terms of  $\Phi$  and  $\alpha$ .

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

**2/II/19D Statistics**

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

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

where  $\gamma$  and  $\delta$  are given positive constants.

Calculate the estimate explicitly in the case when  $f(x | \theta)$  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  $X_1, \dots, X_n$  be a random sample from a normal distribution with mean  $\mu$  and variance  $\sigma^2$ , where  $\mu$  and  $\sigma^2$  are unknown. Derive the form of the size- $\alpha$  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  $\alpha$ .

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

**4/II/19D Statistics**

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

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

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

- (i) Determine the maximum-likelihood estimates  $(\hat{\beta}, \hat{\sigma}^2)$  of  $(\beta, \sigma^2)$ .
- (ii) Find the distribution of  $\hat{\beta}$ .
- (iii) By showing that  $Y_i - \hat{\beta}x_i$  and  $\hat{\beta}$  are independent, or otherwise, determine the joint distribution of  $\hat{\beta}$  and  $\hat{\sigma}^2$ .
- (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  $X_1, X_2, \dots, X_n$  are independent, identically-distributed random variables with distribution

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

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  $\chi^2$  distribution.]

	$\chi_1^2$	$\chi_2^2$	$\chi_3^2$	$\chi_4^2$	$\chi_5^2$	$\chi_6^2$
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

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

for an unknown real parameter  $\theta$ . Find the best test of size  $\alpha$ ,  $0 < \alpha < 1$ , 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  $\theta_0$  and  $\theta_1$  will the power of the best test be at least 0.95?

**4/I/9H Statistics**

Suppose that  $Y_1, \dots, Y_n$  are independent random variables, with  $Y_i$  having the normal distribution with mean  $\beta x_i$  and variance  $\sigma^2$ ; here  $\beta, \sigma^2$  are unknown and  $x_1, \dots, x_n$  are known constants.

Derive the least-squares estimate of  $\beta$ .

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  $\theta$  after observing  $X$ , a single random variable with probability density function  $f(x | \theta)$ ; the parameter  $\theta$  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  $f(x | \theta) = \theta e^{-\theta x}$ ,  $x > 0$  and  $\pi(\theta) = \mu e^{-\mu\theta}$ ,  $\theta > 0$ , where  $\mu > 0$  is known. Determine the posterior distribution of  $\theta$  given  $X$ .

Determine the optimal Bayesian point estimate of  $\theta$  in the cases when

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

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