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 $\alpha = 1/100$.