M2S1 PROBABILITY AND STATISTICS

SAMPLE EXAM QUESTIONS

1. (a) Let X, Y and Z be three continuous random variables taking values on set $\mathcal{X} \subseteq \mathcal{R}^3$ with joint density function $f_{X,Y,Z}$. Define the *conditional* density function of Z given $X = x$ and $Y = y$, $f_{Z|X,Y}(\,.\,|x, y)$, explaining carefully the terms or expressions that you use.

(i) Show that the marginal density of Z can be expressed as the expectation of the function $f_{Z|X,Y}(\,.\,|x, y)$ with respect to another joint density function.

(ii) Using the law of iterated expectation, give an expression for the the expectation of Z, $\mathrm{E}_{f_Z}[Z]$, in terms of an expectation with respect to the marginal density of X, an expectation with respect to the conditional density of Y given $X = x$, and an expectation with respect to $f_{Z|X,Y}(\,.\,|x, y)$.

(b) The joint density function of random variables X, Y and Z is given by

$$f_{X,Y,Z}(x, y, z) = \begin{cases} \dfrac{4}{y}\,\mathrm{e}^{-2x} & 0 < z < y < x < \infty \\ 0 & \text{otherwise} \end{cases}.$$

Calculate

(i) the conditional density of Z, given $X = x$ and $Y = y$,

(ii) the conditional density of Y, given $X = x$, and

(iii) the marginal density of X.

Hence, or otherwise, calculate the expectation of Z.

2. (a) Let X and Y be two random variables with joint density function specified as follows; suppose that the conditional density of Y, given $X = x > 0$ is given by

$$f_{Y|X}(y|x) = \frac{1}{2\sqrt{x}} \qquad -\sqrt{x} < y < \sqrt{x},$$

and the marginal density of X is given by

$$f_X(x) = \sqrt{\frac{x}{2\pi}}\, \mathrm{e}^{-x/2} \qquad x > 0$$

Find

(i) the joint density of X and Y,

(ii) the marginal density of Y,

(iii) the expectation of Y,

(iv) the variance of Y.

(b) Let U, V, and W be three random variables with joint density function $f_{U,V,W}$ defined by

$$f_{U,V,W}(u,v,w) = k \qquad 0 < u^2 < v < -2\lambda \log w$$

for some fixed $\lambda > 0$, and constant k. Find the marginal density of U, and deduce the value of k.

3. (a) Let X be a continuous random variable taking values on interval $\mathcal{X} \subseteq \mathcal{R}$ with pdf f_X and cdf F_X. Let $Y = H(X)$, where H is a monotonically increasing function from $\mathcal{X}$ to interval $\mathcal{Y}$. Show that the pdf of random variable Y is given by

$$f_Y(y) = f_X(x)\,\frac{d}{dy}\left\{H^{-1}(y)\right\} \qquad y \in \mathcal{Y}$$

where x is the unique solution of $y = H(x)$.

Show that, if function H is given by

$$H(x) = F_X(x) \qquad x \in \mathcal{X}$$

then Y has a *Uniform* distribution, $Y \sim Uniform(0,1)$.

Find the pdf of random variable Z defined by $Z = 1 - \mathrm{e}^{-Y}$.

(b) Let X and Y be independent and identically distributed $Gamma(1/2, 1/2)$ random variables, and let random variables U and V be defined by

$$\begin{aligned} U &= \sqrt{X+Y} \\ V &= \frac{X}{Y} \end{aligned}$$

Find the marginal pdfs of U and V, and show that U and V are independent.

4. (a) Let X be a continuous random variable with pdf defined by

$$f_X(x) = \begin{cases} \dfrac{\beta}{\theta^\beta} x^{\beta-1} \mathrm{e}^{-(x/\theta)^\beta} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

for parameters $\theta, \beta > 0$. Using standard single variable transformation techniques, find the pdf and cdf of random variables

(i) $Y = \log X$,

(ii) $Z = (\log X)^2$

(b) Let X and Y be independent and identically distributed random variables, each with pdf of the form

$$\frac{1}{\theta} \mathrm{e}^{-t/\theta} \qquad t > 0$$

for parameter $\theta > 0$, and let random variables U and V be defined by

$$\begin{aligned} U &= X + Y \\ V &= \frac{X}{X+Y} \end{aligned}$$

Find the marginal pdfs of U and V, and show that U and V are independent.

Find the marginal pdf of random variable W defined by

$$W = \frac{Y}{X}$$

and show that U and W are independent.

5.(a) Let X_1 and X_2 be continuous random variables taking values on $\mathcal{X}_1$ and $\mathcal{X}_2$ respectively with joint density function f_{X_1,X_2}. Let random variable Y is defined by $Y = X_1 + X_2$. Prove that the marginal density of Y is given by

$$f_Y(y) = \int_{x_1 \in \mathcal{X}_1} f_{X_1,X_2}(x_1, y - x_1) dx_1$$

Using this result, show that, if X_1 and X_2 are independent and identically distributed $Exp(\lambda)$ random variables, then $Y = X_1 + X_2$ has a *gamma* distribution, $Y \sim Ga(2, \lambda)$.

(b) Prove, using the result in (a), that if random variables $X_1, X_2, ..., X_n$ are independent and identically distributed $Exp(\lambda)$ random variables, then random variable Y defined by

$$Y = X_1 + X_2 + ... + X_n$$

has a gamma distribution, that is, $Y \sim Ga(n, \lambda)$.

6.(a) Let $X_1, ..., X_k$ be independent and identically distributed random variables, each taking values on $\mathcal{X} \subseteq \mathcal{R}$ with pdf f_X and cdf F_X. Define the k *order statistics* derived from $X_1, ..., X_k$, denoted $Y_1, ..., Y_k$.

Write down the form of the joint pdf of $Y_1, ..., Y_k$, and the form of the marginal pdf of the jth order statistic, Y_j.

Suppose that $X_1, ..., X_k$ are iid $Uniform(0, 1)$ random variables. Derive the marginal pdf of the jth order statistic, Y_j.

(b) Suppose that $X_1, ..., X_k$ are iid $Exp(\lambda)$ random variables. Show that, for $j = 1, ..., k$, the jth order statistic has pdf

$$f_{Y_j}(y_j) = \frac{\lambda k!}{(k-j)!} \sum_{i=0}^{j-1} \frac{(-1)^i}{i!(j-i-1)!} \, \mathrm{e}^{-\lambda(k-j+i+1)y_j} \qquad y_j > 0$$

7. (a) Let X be a discrete random variable taking values on the set $\{1, 2, ...\}$ with mass function f_X. Define the *moment generating function* (mgf) of X, $M_X(t)$, for real-valued t

Show that, if $M_X^{(r)}(t)$ is the rth derivative of $M_X(t)$ with respect to t, then

$$M_X^{(r)}(0) = E_{f_X}[X^r]$$

If $X \sim Geometric(\theta)$, find M_X, and hence calculate the expectation and variance of X.

(b) Let X_1 and X_2 be independent $Geometric(\theta)$ random variables. Find the mgf of random variable $Y = X_1 + X_2$, stating any properties of mgfs that you use.

Using the mgf, find the expectation and variance of Y, and hence verify the additive properties of these quantities, that is, that

$$\begin{aligned}
\mathrm{E}_{f_Y}[Y] &= \mathrm{E}_{f_{X_1}}[X_1] + \mathrm{E}_{f_{X_2}}[X_2] \\
\mathrm{Var}_{f_Y}[Y] &= \mathrm{Var}_{f_{X_1}}[X_1] + \mathrm{Var}_{f_{X_2}}[X_2]
\end{aligned}$$

8. (a) Let X be a continuous random variable taking values on $\mathcal{R}$ with density function f_X, and let $M_X(t)$ be the moment generating function of X, for $t \in \mathcal{R}$. Prove that, if random variable Y is defined by $Y = aX + b$, for real constants a and b, then the moment generating function of Y is given by

$$M_Y(t) = \mathrm{e}^{bt} M_X(at) \qquad t \in \mathcal{R}$$

Hence show that, if $E_{f_X}[X] = \mu$, then

$$\mathrm{E}_{f_Y}[Y] = a\mu + b$$

(b) Let X be a continuous random variable taking values on $\mathcal{R}$ with density function given by

$$f_X(x) = \begin{cases} \dfrac{1}{2\theta}\,\mathrm{e}^{(x-\mu)/\theta} & x < \mu \\ \\ \dfrac{1}{2\theta}\,\mathrm{e}^{-(x-\mu)/\theta} & x \geq \mu \end{cases}$$

for parameters $\mu \in \mathcal{R}$ and $\theta > 0$. Calculate the mgf of X, and the expectation of X. Hence find constants a and b such that random variable $Y = aX + b$ has expectation zero.

9. (a) Let $X_1, ..., X_n$ be a sequence of random variables, and let the cdf of X_i be denoted F_{X_i} for $i = 1, ..., n$, so that

$$F_{X_i}(x) = \mathrm{P}[X_i \leq x] \qquad x \in \mathcal{R}.$$

Explain what is meant by the statement

"*the sequence* $X_1, ..., X_n$ *converges in distribution to random variable* X *with limiting distribution* F_X"

If $X_1, ..., X_n$ is a continuous sequence of random variables, and F_{X_i} is given by

$$F_{X_i}(x) = 1 - \left(1 + \frac{x}{n}\right)^{-n} \qquad x > 0$$

for $i = 1, ..., n$, find the limiting distribution, F_X. What is this distribution ?

[Hint : $\mathrm{e}^x = 1 + x + x^2/2 + x^3/6 + ...$]

(b) Let $X_n \sim Bin(n, \theta)$, so that the mgf of X_n is given by

$$M_{X_n}(t) = (1 - \theta + \theta \mathrm{e}^t)^n \qquad t \in \mathcal{R}.$$

Calculate the mgf of random variable Z_n, defined by

$$Z_n = \frac{X_n - \mu_n}{\sigma_n}$$

where $\mu_n = n\theta$, $\sigma_n = \sqrt{n\theta(1-\theta)}$. Hence show that, as $n \longrightarrow \infty$,

$$M_{Z_n}(t) \longrightarrow \mathrm{e}^{t^2/2}.$$

Show that this result is merely a special case of the Central Limit Theorem.

10. (a) Let $X_1, ..., X_n$ be a random sample from a density function f_X depending on a single parameter θ in parameter space Θ. Let $x_1, ..., x_n$ be a set of observed values of the random variables. Define the *likelihood function*, and the *maximum likelihood estimate* of θ based on $x_1, ..., x_n$. Explain how the maximum likelihood estimate is calculated.

Calculate the maximum likelihood estimate of θ based on $x_1, ..., x_n$ if the distribution of X is specified by

(i) $f_X(x|\theta) = \theta x^{\theta-1}$, $0 < x < 1$ for $\theta > 0$.

(ii) $f_X(x|\theta) = \theta^2 x e^{-\theta x}$, $0 < x$ for $\theta > 0$.

(b) Let $X_1, ..., X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution. Find the sampling distributions of the statistics

$$\begin{aligned} T_1 &= \sum_{i=1}^{n} X_i \\ T_2 &= \sum_{i=1}^{n} (X_i - \bar{X})^2. \end{aligned}$$

Find the maximum likelihood estimates of μ and σ^2 based on a observed sample $x_1, ..., x_n$. Hence decide whether the corresponding estimators of these parameters are unbiased estimators.