Imperial College London

BSc, MSc and MSci EXAMINATIONS (MATHEMATICS)

May 2017

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science.

Probability and Statistics

Date: Wednesday, 24 May 2017

Time: 10:00 – 12:00

Time Allowed: 2 Hours

This paper has *4 Questions*.

Candidates should use ONE main answer book.

Supplementary books may only be used after the relevant main book(s) are full.

Statistical tables are provided.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Credit will be given for all questions attempted, but extra credit will be given for complete or nearly complete answers to each question as per the table below.

Raw Mark	Up to 12	13	14	15	16	17	18	19	20
Extra Credit	0	$1/2$	1	$1\,1/2$	2	$2\,1/2$	3	$3\,1/2$	4

- Each question carries equal weight.
- Calculators may not be used.

1. (a) State the three axioms of probability for events defined on a sample space Ω.

(b) Prove from the axioms that the probability of exactly one event occurring out of events E and F is given by $\mathsf{P}(E) + \mathsf{P}(F) - 2\mathsf{P}(E \cap F)$.

(c) If E and F are independent events, prove that E^C and F^C are also independent.

(d) Given that $\mathsf{P}(E) = 0.3$, $\mathsf{P}(E \cup F) = 0.7$ and $\mathsf{P}(F) = p$.

(i) Determine the range of values for p which satisfy the axioms of probability.

(ii) Determine the value of p for which E and F are independent.

(e) Let I_E be the indicator random variable for event E, i.e.

$$I_E = \begin{cases} 1, & \text{if } E \text{ occurs;} \\ 0, & \text{if } E \text{ does not occur.} \end{cases}$$

(i) Determine $f_{I_E}(\cdot)$ the probability mass function (pmf) of I_E.

(ii) Prove that

$$\mathsf{E}_{f_{I_E}}(I_E) = \mathsf{P}(E) \quad \text{and} \quad \mathsf{var}_{f_{I_E}}(I_E) = \mathsf{P}(E)(1 - \mathsf{P}(E)).$$

(iii) Show that $I_F = 1 - (1 - I_{E_1})(1 - I_{E_2})$ is the indicator random variable for the event $F = E_1 \cup E_2$ and $I_G = I_{E_1} I_{E_2}$ is the indicator random variable for the event $G = E_1 \cap E_2$. Use these results to prove that $\mathsf{P}(E_1 \cup E_2) = \mathsf{P}(E_1) + \mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2)$.

2. (a) Prove that $\mathsf{E}_{f_X}(X) = G_X^{(1)}(1)$ where $G_X(t)$ is the probability generating function (pgf) of the discrete random variable X.

(b) A discrete random variable Z has pgf given by:

$$G_Z(t) = \frac{t(1+t)^2}{5-t}.$$

(i) Determine $\mathsf{E}_{f_Z}(Z)$.

(ii) Show that Z is the sum of independent random variables X and Y with $X \sim Geometric(\alpha)$ and $Y \sim Binomial(n, \theta)$, where you should determine the values of n, θ and α.
Hint: to determine pgfs from the formula sheet, recall that $G_X(t) = M_X(\log(t))$.

(c) In a particular traffic study, N different types of vehicle are monitored, where N is a discrete random variable. Given N, the numbers of each type of vehicle, $X_1, X_2, \ldots, X_N$ are independently, identically distributed discrete random variables and their common distribution does not depend on N. Let S_N be the total number of vehicles observed (i.e. $S_N = \sum_{i=1}^{N} X_i$).

(i) Prove that

$$\mathsf{E}_{f_{S_N}}(t^{S_N}) = \sum_n \mathsf{E}_{f_{S_N|N}}(t^{S_N} \mid N = n)\mathsf{P}(N = n)$$

(ii) Hence show that

$$G_{S_N}(t) = G_N(G_{X_1}(t)).$$

(iii) Prove that $\mathsf{E}_{f_{S_N}}(S_N) = \mathsf{E}_{f_N}(N)\mathsf{E}_{f_{X_1}}(X_1)$.

(iv) If N follows a Geometric distribution with parameter θ, state the form of $G_N(t)$, the pgf of N. Hence determine $\mathsf{E}_{f_N}(N)$.

(v) If each X_i follows a Poisson distribution with parameter λ state the form of $G_{X_i}(t)$, the pgf of X_i. Hence determine $\mathsf{E}_{f_{X_i}}(X_i)$.

(vi) Determine $\mathsf{E}_{f_{S_N}}(S_N)$ if N and the X_i follow the distributions given in parts (iv) and (v).

3. The continuous random variable X has cumulative distribution function given by

$$F_X(x) = \begin{cases} 0, & x \leq 0; \\ 1-(1+x^c)^{-1}, & x > 0, \end{cases}$$

where $c \in \{2,3,4,5,\ldots\}$ i.e. $c \in \mathbb{N}$ and $c \geq 2$.

(a) Find $f_X(x)$, the probability density function of X.

(b) Prove that the median of X does not depend on c.

(c) Determine the smallest value of c for which $\mathsf{P}(X \leq 0.5) \leq 0.1$.

(d) Prove that

$$\mathsf{E}_{f_X}(X) = \int_0^\infty (1 - F_X(x))\ \mathsf{d}x.$$

(e) By considering the substitution $y = (1+x^c)^{-1}$ or otherwise, prove that

$$\mathsf{E}_{f_X}(X) = \Gamma\left(1+\frac{1}{c}\right)\Gamma\left(1-\frac{1}{c}\right).$$

Hint: the normalizing constant of the Beta distribution is given by:

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\ \mathsf{d}x = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

4. (a) For continuous random variables Y and Z, prove that

$$\mathsf{E}_{f_Y}(Y) = \mathsf{E}_{f_Z}(\mathsf{E}_{f_{Y|Z}}(Y|Z)).$$

(b) The continuous random variables X and Y have joint pdf given by

$$f_{X,Y}(x,y) = \frac{\sqrt{x}\mathsf{e}^{-y\sqrt{x}}}{(1+x)^2}, \quad x > 0; y > 0.$$

(i) Determine $f_X(x)$, the marginal pdf of X.

(ii) Determine $f_{Y|X}(y|X=x)$, the conditional pdf of Y given $X=x$. Name this conditional distribution and state its parameter.

(iii) Determine the conditional expectation of Y given $X=x$.

(iv) Find the pdf of $Z=(1+X)^{-1}$ (Note: consider the range of Z carefully).

(v) By considering $F_{Y|Z}(y \mid Z=z) = \mathsf{P}(Y \leq y \mid (1+X)^{-1} = z)$ show that

$$f_{Y|Z}(y|Z=z) = \sqrt{\frac{1-z}{z}}\mathsf{e}^{-y\sqrt{\frac{1-z}{z}}}.$$

(vi) Use the result in part (a), or otherwise, to find $\mathsf{E}_{f_Y}(Y)$.
You may find the hint given in question 3 part (e) useful.

DISCRETE DISTRIBUTIONS

	RANGE	PARAMETERS	MASS FUNCTION f_X	CDF F_X	$\mathsf{E}_{f_X}[X]$	$\mathsf{Var}_{f_X}[X]$	MGF M_X
$Bernoulli(\theta)$	$\{0,1\}$	$\theta \in (0,1)$	$\theta^x(1-\theta)^{1-x}$		θ	$\theta(1-\theta)$	$1-\theta+\theta e^t$
$Binomial(n,\theta)$	$\{0,1,...,n\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n}{x}\theta^x(1-\theta)^{n-x}$		$n\theta$	$n\theta(1-\theta)$	$(1-\theta+\theta e^t)^n$
$Poisson(\lambda)$	$\{0,1,2,...\}$	$\lambda \in \mathbb{R}^+$	$\dfrac{e^{-\lambda}\lambda^x}{x!}$		λ	λ	$\exp\{\lambda(e^t-1)\}$
$Geometric(\theta)$	$\{1,2,...\}$	$\theta \in (0,1)$	$(1-\theta)^{x-1}\theta$	$1-(1-\theta)^x$	$\dfrac{1}{\theta}$	$\dfrac{(1-\theta)}{\theta^2}$	$\dfrac{\theta e^t}{1-e^t(1-\theta)}$
$NegBinomial(n,\theta)$	$\{n,n+1,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{x-1}{n-1}\theta^n(1-\theta)^{x-n}$		$\dfrac{n}{\theta}$	$\dfrac{n(1-\theta)}{\theta^2}$	$\left(\dfrac{\theta e^t}{1-e^t(1-\theta)}\right)^n$
or	$\{0,1,2,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n+x-1}{x}\theta^n(1-\theta)^x$		$\dfrac{n(1-\theta)}{\theta}$	$\dfrac{n(1-\theta)}{\theta^2}$	$\left(\dfrac{\theta}{1-e^t(1-\theta)}\right)^n$

For **CONTINUOUS** distributions (see over), define the **GAMMA FUNCTION**

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx \qquad \Gamma(\alpha+1) = \alpha\Gamma(\alpha) \qquad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

and the LOCATION/SCALE transformation $Y = \mu + \sigma X$ gives

$$f_Y(y) = f_X\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} \qquad F_Y(y) = F_X\left(\frac{y-\mu}{\sigma}\right) \qquad M_Y(t) = e^{\mu t}M_X(\sigma t) \qquad \mathsf{E}_{f_Y}[Y] = \mu + \sigma\mathsf{E}_{f_X}[X] \qquad \mathsf{Var}_{f_Y}[Y] = \sigma^2\mathsf{Var}_{f_X}[X]$$

CONTINUOUS DISTRIBUTIONS

	$\mathbb{X}$	PARAMS.	PDF f_X	CDF F_X	$\mathsf{E}_{f_X}[X]$	$\mathsf{Var}_{f_X}[X]$	MGF M_X
$Uniform(\alpha,\beta)$ (stand. model $\alpha=0,\beta=1$)	(α,β)	$\alpha<\beta\in\mathbb{R}$	$\dfrac{1}{\beta-\alpha}$	$\dfrac{x-\alpha}{\beta-\alpha}$	$\dfrac{(\alpha+\beta)}{2}$	$\dfrac{(\beta-\alpha)^2}{12}$	$\dfrac{e^{\beta t}-e^{\alpha t}}{t\,(\beta-\alpha)}$
$Exponential(\lambda)$ (stand. model $\lambda=1$)	$\mathbb{R}^+$	$\lambda\in\mathbb{R}^+$	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	$\left(\dfrac{\lambda}{\lambda-t}\right)$
$Gamma(\alpha,\beta)$ (stand. model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$		$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$	$\left(\dfrac{\beta}{\beta-t}\right)^\alpha$
$Weibull(\alpha,\beta)$ (stand. model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\alpha\beta x^{\alpha-1}e^{-\beta x^\alpha}$	$1-e^{-\beta x^\alpha}$	$\dfrac{\Gamma(1+1/\alpha)}{\beta^{1/\alpha}}$	$\dfrac{\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma\left(1+\frac{1}{\alpha}\right)^2}{\beta^{2/\alpha}}$	
$Normal(\mu,\sigma^2)$ (stand. model $\mu=0,\sigma=1$)	$\mathbb{R}$	$\mu\in\mathbb{R}$ $\sigma\in\mathbb{R}^+$	$\dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}$		μ	σ^2	$e^{\{\mu t+\sigma^2t^2/2\}}$
$Student(\nu)$	$\mathbb{R}$	$\nu\in\mathbb{R}^+$	$\dfrac{(\pi\nu)^{-\frac{1}{2}}\Gamma\left(\dfrac{\nu+1}{2}\right)}{\Gamma\left(\dfrac{\nu}{2}\right)\left\{1+\dfrac{x^2}{\nu}\right\}^{(\nu+1)/2}}$		0 (if $\nu>1$)	$\dfrac{\nu}{\nu-2}$ (if $\nu>2$)	
$Pareto(\theta,\alpha)$	$\mathbb{R}^+$	$\theta,\alpha\in\mathbb{R}^+$	$\dfrac{\alpha\theta^\alpha}{(\theta+x)^{\alpha+1}}$	$1-\left(\dfrac{\theta}{\theta+x}\right)^\alpha$	$\dfrac{\theta}{\alpha-1}$ (if $\alpha>1$)	$\dfrac{\alpha\theta^2}{(\alpha-1)(\alpha-2)}$ (if $\alpha>2$)	
$Beta(\alpha,\beta)$	$(0,1)$	$\alpha,\beta\in\mathbb{R}^+$	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$		$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	

Imperial College London

SOLUTIONS

BSc, MSc and MSci EXAMINATIONS (MATHEMATICS)

May/June 2017

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science.

Probability and Statistics

Date: Wednesday, 24 May 2017

Time: 10:00 – 12:00

Time Allowed: 2 Hours

This paper has *4 Questions*.

Candidates should use ONE main answer book.

Supplementary books may only be used after the relevant main book(s) are full.

Statistical tables are provided.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Credit will be given for all questions attempted, but extra credit will be given for complete or nearly complete answers to each question as per the table below.

Raw Mark	Up to 12	13	14	15	16	17	18	19	20
Extra Credit	0	$1/2$	1	$1\,1/2$	2	$2\,1/2$	3	$3\,1/2$	4

- Each question carries equal weight.
- Calculators may not be used.

seen $\Downarrow$

1. (a) **Axioms of Probability**

Given a $\sigma-$field, $\mathcal{F}$ (a set of subsets of the sample space Ω.) For events $E, E_1, E_2, \ldots \in \mathcal{F}$, then the probability function, $\mathsf{P}(\cdot)$, must satisfy:

(I) $\mathsf{P}(E) \geq 0$.
(II) $\mathsf{P}(\Omega) = 1$.
(III) If $E_1, E_2, \ldots$, are pairwise disjoint then $\mathsf{P}(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mathsf{P}(E_i)$ (Countable additivity).

(Do not need to specify $\sigma-$field, could instead say: for events $E, E_1, \ldots \subseteq \Omega$. Lose 1 mark if finite rather than countable additivity specified, but they do need to specify the meaning of finite/countable additivity). 3

(b) The probability of both exactly one of E and F occurring is (using Axiom III),

$$
\begin{aligned}
\mathsf{P}((E \cap F^C) \cup (E^C \cap F)) &= \mathsf{P}(E \cap F^C) + \mathsf{P}(E^C \cap F) \\
\mathsf{P}(E) &= P((E \cap F) \cup (E \cap F^C)) = \mathsf{P}(E \cap F) + \mathsf{P}(E \cap F^C) \\
\Rightarrow \mathsf{P}(E \cap F^C) &= \mathsf{P}(E) - \mathsf{P}(E \cap F) \\
\mathsf{P}(F) &= P((F \cap E) \cup (F \cap E^C)) = \mathsf{P}(F \cap E) + \mathsf{P}(F \cap E^C) \\
\Rightarrow \mathsf{P}(F \cap E^C) &= \mathsf{P}(F) - \mathsf{P}(E \cap F) \\
\Rightarrow \mathsf{P}((E \cap F^C) \cup (E^C \cap F)) &= \mathsf{P}(E) - \mathsf{P}(E \cap F) + \mathsf{P}(F) - \mathsf{P}(E \cap F) \\
&= P(E) + \mathsf{P}(F) - 2\mathsf{P}(E \cap F).
\end{aligned}
$$

as required. 3

(c) If E and F are independent, then $\mathsf{P}(E \cap F) = \mathsf{P}(E)\mathsf{P}(F)$.

$$
\begin{aligned}
\mathsf{P}(E^C \cap F^C) &= \mathsf{P}((E \cup F)^C) = 1 - \mathsf{P}(E \cup F) = 1 - (\mathsf{P}(E) + F(F) - \mathsf{P}(E \cap F) \\
&= 1 - \mathsf{P}(E) - \mathsf{P}(F) - \mathsf{P}(E)\mathsf{P}(F) = (1 - \mathsf{P}(E))(1 - \mathsf{P}(F)) = \mathsf{P}(E^C)\mathsf{P}(F^C)
\end{aligned}
$$

Hence, E^C and F^C are also independent. 2

(d) $\mathsf{P}(E) = 0.3, \mathsf{P}(E \cup F) = 0.7, \mathsf{P}(F) = p$. Hence sim seen $\Downarrow$

$$
\mathsf{P}(E \cap F) = \mathsf{P}(E) + \mathsf{P}(F) - \mathsf{P}(E \cup F) = 0.3 + p - 0.7 = p - 0.4.
$$

(i) We must have $0 \leq \mathsf{P}(E \cap F) \leq \mathsf{P}(F)$ and $1 \geq P(E \cup F) \geq \mathsf{P}(F)$. 1
So, $0 \leq p - 0.4 \leq p$ and $1 \geq 0.7 \geq p \Rightarrow 0.4 \leq p \leq 0.7$.

(ii) For E and F to be independent, we must have $\mathsf{P}(E \cap F) = \mathsf{P}(E)\mathsf{P}(F)$. 1
Hence $p - 0.4 = 0.3p \Rightarrow 0.7p = 0.4 \Rightarrow p = 4/7$.

(e) (i) $f_{I_E}(x) = \mathsf{P}(I_E = x)$, so

$$f_{I_E}(x) = \begin{cases} \mathsf{P}(E), & x = 1; \\ \mathsf{P}(E^C), & x = 0. \end{cases}$$

$\boxed{1}$

(ii)

$$\mathsf{E}_{f_{I_E}}(I_E) = \sum_{x=0}^{1} x f_{I_E}(x) = 0 \times \mathsf{P}(E^C) + 1 \times \mathsf{P}(E) = \mathsf{P}(E),$$

as required. $\boxed{1}$

$$\begin{aligned} \mathsf{var}_{f_{I_E}}(I_E) &= \mathsf{E}_{f_{I_E}}(I_E^2) - E^2_{f_{I_E}}(I_E) \\ \mathsf{E}_{f_{I_E}}(I_E^2) &= \sum_{x=0}^{1} x^2 f_{I_E}(x) = 0 \times \mathsf{P}(E^C) + 1 \times \mathsf{P}(E) = \mathsf{P}(E) \\ \Rightarrow \mathsf{var}_{f_{I_E}}(I_E) &= \mathsf{P}(E) - \mathsf{P}(E)^2 = \mathsf{P}(E)(1 - \mathsf{P}(E)), \end{aligned}$$

as required. $\boxed{2}$

(iii) $I_F = 1 - (1 - I_{E_1})(1 - I_{E_2})$, $I_G = I_{E_1} I_{E_2}$. $\boxed{\text{unseen} \Downarrow}$

Values for I_F and I_G:

I_F	E_2	E_2^C
E_1	1	1
E_1^C	1	0

I_G	E_2	E_2^C
E_1	1	0
E_1^C	0	0

We know that $E_1 \cup E_2 = (E_1 \cap E_2) \cup (E_1 \cap E_2^C) \cup (E_1^C \cap E_2^C)$.
Hence, we see from the table for I_F that if $(E_1 \cup E_2)$ occurs then $I_F = 1$ and if $(E_1 \cup E_2)^C = E_1^C \cap E_2^C$ occurs, then $I_F = 0$.
So, we have

$$I_F = \begin{cases} 1, & \text{if } E_1 \cup E_2 \text{ occurs;} \\ 0, & \text{if } E_1 \cup E_2 \text{ does not occur.} \end{cases}$$

Hence, I_F is the indicator random variable for the event $E_1 \cup E_2$. $\boxed{2}$

Also, if $E_1 \cap E_2$ occurs, then $I_G = 1$ and if $(E_1 \cap E_2)^C$ occurs, then $I_G = 0$.
So, we have

$$I_G = \begin{cases} 1, & \text{if } E_1 \cap E_2 \text{ occurs;} \\ 0, & \text{if } E_1 \cap E_2 \text{ does not occur.} \end{cases}$$

Hence, I_G is the indicator random variable for the event $E_1 \cap E_2$. $\boxed{2}$

So we have $\mathsf{P}(E_1 \cup E_2) = \mathsf{E}_{f_{I_F}}(I_F)$ and $\mathsf{P}(E_1 \cap E_2) = \mathsf{E}_{f_{I_G}}(I_G) = \mathsf{E}_{f_{I_{E_1},I_{E_2}}}(I_{E_1} I_{E_2})$.
Now,

$$
\begin{aligned}
I_F &= 1 - (1 - I_{E_1})(1 - I_{E_2}) \\
\Rightarrow \mathsf{E}_{f_{I_F}}(I_F) &= 1 - \mathsf{E}_{f_{I_{E_1},I_{E_2}}}((1 - I_{E_1})(1 - I_{E_2})) \text{ (linearity of expectation)} \\
&= 1 - \mathsf{E}_{f_{I_{E_1},I_{E_2}}}(1 - I_{E_1} - I_{E_2} + I_{E_1} I_{E_2}) \\
&= 1 - 1 + E_{f_{I_{E_1}}}(I_{E_1}) + E_{f_{I_{E_2}}}(I_{E_2}) - \mathsf{E}_{f_{I_{E_1},I_{E_2}}}(I_{E_1} I_{E_2}) \text{ (linearity of expectation)} \\
&= \mathsf{P}(E_1) + \mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2).
\end{aligned}
$$

$\boxed{2}$

Question 1: commentary:
Parts (a), (b) and (c) are all seen/bookwork (8 marks).
Parts (d) is straightforward and should be accessible to anyone that has engaged with the material (2 marks).
Part (e)(i) and (ii) are relatively straightforward, but contain unfamiliar notation, they have met Bernoulli random variables but these have not been described in the context of indicator random variables (4 marks).
Part (e)(iii) given its unseen nature, the students will find this more challenging and requires a good understanding of the concept of random variables and their properties (6 marks).

seen ⇓

2. (a)

$$G_X(t) = \mathsf{E}_{f_X}(t^X) = \sum_x t^x f_X(x)$$
$$G_X^{(1)}(t) = \sum_x x t^{x-1} f_X(x)$$
$$G_X^{(1)}(1) = \sum_x x f_X(x) = \mathsf{E}_{f_X}(X),$$

as required. 2

(b) (i) meth seen ⇓

$$G_Z(t) = \frac{t(1+t)^2}{5-t}$$
$$G_Z^{(1)}(t) = t(1+t)^2(5-t)^{-2} + (5-t)^{-1}(2t(1+t) + (1+t)^2)$$
$$\Rightarrow \mathsf{E}_{f_Z}(Z) = G_Z^{(1)}(1) = 1 \cdot 2^2 \cdot 4^{-2} + 4^{-1}(2 \cdot 1 \cdot 2 + 2^2) = \frac{9}{4}.$$

2

unseen ⇓

(ii) If Z is the sum of two independent random variables X and Y where $Y \sim Binomial(n, \theta)$ and $X \sim Geometric(\alpha)$.

$$G_Z(t) = G_X(t)G_Y(t) = \frac{t(1+t)^2}{5-t} = \frac{4t(0.5+0.5t)^2}{5-t} = \frac{4t}{5(1-t/5)}(0.5-0.5t)^2$$
$$= \frac{(4/5)t}{1-(1-4/5)t}(0.5-0.5t)^2 = G_X(t)G_Y(t).$$

Where $X \sim Geometric(0.8)$ and $Y \sim Binomial(2, 0.5)$, $\alpha = 0.8$, $n = 2$ and $\theta = 0.5$. 3

(c) (i) Using the theorem of total probability, we have,

$$\mathsf{E}_{f_{S_N}}(t^{S_N}) = \sum_s t^s \mathsf{P}(S_N = s) = \sum_s t^s \sum_n \mathsf{P}(S_N = s \mid N = n)\mathsf{P}(N = n)$$
$$= \sum_n \left\{ \sum_s t^s \mathsf{P}(S_N = s \mid N = n) \right\} \mathsf{P}(N = n)$$
$$= \sum_n \mathsf{E}_{f_{S_N|N}}(t^{S_N} \mid N = n)\mathsf{P}(N = n).$$

2

(ii)

$$G_{S_N}(t) = \mathsf{E}_{f_{S_N}}(t^{S_N}) = \sum_n \mathsf{E}_{f_{S_N|N}}(t^{S_N} \mid N = n)\mathsf{P}(N = n)$$
$$= \sum_n \mathsf{E}_{f_{S_N|N}}(t^{S_n})\mathsf{P}(N = n) = \sum_n \mathsf{E}_{S_n}(t^{X_1+X_2+\ldots X_n})\mathsf{P}(N = n)$$
$$= \sum_n \mathsf{E}_{f_{X_1}}(t^{X_1})\mathsf{E}_{f_{X_2}}(t^{X_2})\ldots\mathsf{E}_{f_{X_n}}(t^{X_n})\mathsf{P}(N = n) \text{ as } X_i \text{ indep.}$$
$$= \sum_n G_{X_1}(t)^n \mathsf{P}(N = n) = G_N(G_{X_1}(t)) \text{ as } X_i \text{ are iid.}$$

2

(iii) Differentiating both sides of the result in part (ii):

$$\begin{aligned}
G^{(1)}_{S_N}(t) &= G^{(1)}_{X_1}(t)G^{(1)}_N(G_{X_1}(t)) \\
\mathsf{E}_{f_{S_N}}(S_N) = G^{(1)}_{S_N}(1) &= G^{(1)}_{X_1}(1)G^{(1)}_N(G_{X_1}(1)) \\
&= \mathsf{E}_{f_{X_1}}(X_1)G^{(1)}_N(1) = \mathsf{E}_{f_{X_1}}(X_1)\mathsf{E}_{f_N}(N).
\end{aligned}$$

2

seen ⇓

(iv) Full marks for stating the form of the pgf, derivation given here for info.

$$\begin{aligned}
G_N(t) &= \sum_{n=1}^{\infty} t^n f_N(n) = \sum_{n=1}^{\infty} t^n \theta(1-\theta)^{n-1} \\
&= \theta t \sum_{n=1}^{\infty} (t(1-\theta))^{n-1} = \frac{t\theta}{1-(1-\theta)t} \\
G^{(1)}_N(t) &= t\theta(1-\theta)(1-(1-\theta)t)^{-2} + (1-(1-\theta)t)^{-1}\theta \\
\mathsf{E}_{f_N}(N) = G^{(1)}_N(t) &= \theta(1-\theta)(1-(1-\theta))^{-2} + (1-(1-\theta))^{-1}\theta = \frac{\theta(1-\theta)}{\theta^2} + 1 = \frac{1}{\theta}.
\end{aligned}$$

3

(v) Full marks for stating the form of the pgf, derivation given here for info.

$$\begin{aligned}
G_{X_i}(t) &= \sum_{x=0}^{\infty} t^x f_{X_i}(x) = \sum_{x=0}^{\infty} t^x \frac{\mathsf{e}^{-\lambda}\lambda^x}{x!} \\
&= \mathsf{e}^{-\lambda} \sum_{x=0}^{\infty} \frac{(t\lambda)^x}{x!} = \mathsf{e}^{-\lambda}\mathsf{e}^{t\lambda} = \mathsf{e}^{\lambda(t-1)}. \\
G^{(1)}_{X_i}(t) &= \lambda \mathsf{e}^{\lambda(t-1)} \\
\mathsf{E}_{f_{X_i}}(X_i) &= G^{(1)}_{X_i}(1) = \lambda.
\end{aligned}$$

3

unseen ⇓

(vi)

$$\mathsf{E}_{f_{S_N}}(S_N) = \mathsf{E}_{X_i}(X_i)\mathsf{E}_{f_N}(N) = \frac{\lambda}{\theta}.$$

1

Question 2: commentary:
Part (a) seen/bookwork (2 marks).
Part (b)(i) straightforward and should be accessible to anyone that has engaged with the material (2 marks).
Part (b)(ii) slightly less straightforward, they have not directly seen the factorization in this direction, I would still expect most good students to be able to find the parameters (3 marks).
Parts (c)(i) and (ii) and (iii) are more challenging as they have not seen this result in the case of random N. (6 marks)
Parts (c)(iv) and (v) seen bookwork (6 marks).
Part (vi) straightforward, but the unseen context may prove challenging (1 mark).

meth seen ⇓

3. (a)

$$f_X(x) = \frac{\mathsf{d}}{\mathsf{d}x} F_X(x) = \frac{\mathsf{d}}{\mathsf{d}x}(1 - (1 + x^c)^{-1})$$
$$= \frac{cx^{c-1}}{(1+x^c)^2}, \quad x > 0.$$

(need to specify the range of x for full marks) $\boxed{3}$

(b) Median is the value of x such that $F_X(x) = 0.5$,

$$F_X(x) = 0.5 \Rightarrow 1 - (1 + x^c)^{-1} = 0.5 \Rightarrow x^c = 1 \Rightarrow x = 1.$$ $\boxed{3}$

The median is 1 and does not depend on c.

(c)

$$\mathsf{P}(X \leq 0.5) \leq 0.1 \Rightarrow F_X(0.5) \leq 0.1 \Rightarrow 1 - \left(1 + \frac{1}{2^c}\right)^{-1} \leq 0.1$$
$$\Rightarrow 0.9 \leq \left(1 + \frac{1}{2^c}\right)^{-1} \Rightarrow \left(1 + \frac{1}{2^c}\right) \leq \frac{10}{9}$$
$$\Rightarrow \frac{1}{2^c} \leq \frac{1}{9} \Rightarrow 2^c \geq 9 \Rightarrow c \geq 4.$$ $\boxed{3}$

(d)

$$\mathsf{E}_{f_X}(X) = \int_0^\infty x f_X(x)\,\mathsf{d}x = \int_0^\infty \frac{cxx^{c-1}}{(1+x^c)^2}\,\mathsf{d}x \qquad u = x \quad \frac{\mathrm{d}v}{\mathrm{d}x} = \frac{cx^{c-1}}{(1+x^c)^2}, \quad \frac{\mathrm{d}u}{\mathrm{d}x} = 1 \quad v = -\frac{1}{1+x^c}$$
$$= \left[\frac{-x}{1+x^c}\right]_0^\infty + \int_0^\infty \frac{1}{1+x^c}\,\mathsf{d}x$$
$$= \int_0^\infty \frac{1}{1+x^c}\,\mathsf{d}x = \int_0^\infty (1 - F_X(x))\,\mathsf{d}x.$$

seen ⇓

Alternatively, could prove in general:

$$\int_0^\infty (1 - F_X(x))\,\mathsf{d}x = \int_0^\infty u\frac{\mathrm{d}v}{\mathrm{d}x}\,\mathsf{d}x \qquad u = 1 - F_X(x) \quad \frac{\mathrm{d}v}{\mathrm{d}x} = 1, \quad \frac{\mathrm{d}u}{\mathrm{d}x} = -f_X(x) \quad v = x$$
$$= [x(1 - F_X(x))]_0^\infty + \int_0^\infty x f_X(x)\,\mathsf{d}x = \mathsf{E}_{f_X}(X).$$

Or, by swapping order of integration:

$$\int_0^\infty (1 - F_X(x))\,\mathsf{d}x = \int_0^\infty \mathsf{P}(X \geq x)\,\mathsf{d}x = \int_0^\infty \int_x^\infty f_X(y)\,\mathsf{d}y\,\mathsf{d}x$$
$$= \int_0^\infty \int_0^y f_X(y)\,\mathsf{d}x\,\mathsf{d}y = \int_0^\infty f_X(y) \int_0^y 1\,\mathsf{d}x\,\mathsf{d}y$$
$$= \int_0^\infty f_X(y)\,[x]_0^y\,\mathsf{d}y = \int_0^\infty y f_X(y)\,\mathsf{d}y = \mathsf{E}_{f_X}(X).$$ $\boxed{4}$

unseen $\Downarrow$

(e)

$$\mathsf{E}_{f_X}(X) = \int_0^\infty (1 - F_X(x))\,\mathsf{d}x = \int_0^\infty \frac{1}{1+x^c}\,\mathsf{d}x$$

Using the substitution $y = (1+x^c)^{-1}$,

$$\begin{aligned}
y &= (1+x^c)^{-1} \qquad x = \left(\frac{1-y}{y}\right)^{1/c}\\
\frac{\mathrm{d}y}{\mathrm{d}x} &= -cx^{c-1}(1+x^c)^{-2}\\
\mathsf{d}x &= \frac{-(1+x^c)^2}{cx^{c-1}}\,\mathsf{d}y = \frac{-y^{-2}}{c}\left(\frac{1-y}{y}\right)^{(1-c)/c}\,\mathsf{d}y\\
x &= 0, y = 1 \quad x = \infty, y = 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathsf{E}_{f_X}(X) &= \int_0^1 y\frac{y^{-2}}{c}\left(\frac{1-y}{y}\right)^{(1-c)/c}\,\mathsf{d}y\\
&= \frac{1}{c}\int_0^1 y^{-1/c}(1-y)^{1/c-1}\,\mathsf{d}y = \frac{1}{c}\Gamma\left(1-\frac{1}{c}\right)\Gamma\left(\frac{1}{c}\right)\\
&= \Gamma\left(1-\frac{1}{c}\right)\Gamma\left(1+\frac{1}{c}\right) \quad \text{as } \Gamma(\alpha+1) = \alpha\Gamma(\alpha).
\end{aligned}$$

$\boxed{7}$

Question 3: commentary:

Parts (a), (b) and (c) are straightforward and accessible to any student that has engaged with the material (9 marks).

Part (d) general proof is bookwork (4 marks).

Part (e) is more challenging, they have seen the Gamma function, but only related to the normalizing constant of the Gamma distribution, the change of limits is also tricky (7 marks).

seen $\Downarrow$

4. (a)

$$\mathsf{E}_{f_Z}(\mathsf{E}_{f_{Y|Z}}(Y|Z)) = \int_z \mathsf{E}_{f_{Y|Z}}(Y|Z) f_Z(z)\,\mathsf{d}z = \int_z \left\{ \int_y y f_{Y|Z}(y|Z=z)\,\mathsf{d}y \right\} f_Z(z)\,\mathsf{d}z$$
$$= \int_z \left\{ \int_y y \frac{f_{Y,Z}(y,z)}{f_Z(z)}\,\mathsf{d}y \right\} f_Z(z)\,\mathsf{d}z$$
$$= \int_y y \left\{ \int_x f_{Y,Z}(y,z)\,\mathsf{d}z \right\}\,\mathsf{d}y = \int_y y f_Y(y)\,\mathsf{d}y = \mathsf{E}_{f_Y}(Y).$$

3

(b) (i)

meth seen $\Downarrow$

$$f_X(x) = \int_0^\infty f_{X,Y}(x,y)\,\mathsf{d}y = \int_0^\infty \frac{\sqrt{x}\mathsf{e}^{-y\sqrt{x}}}{(1+x)^2}\,\mathsf{d}y$$
$$= \frac{1}{(1+x)^2}\int_0^\infty \sqrt{x}\mathsf{e}^{-y\sqrt{x}}\,\mathsf{d}y = \frac{1}{(1+x)^2}\left[-\mathsf{e}^{-y\sqrt{x}}\right]_0^\infty$$
$$= \frac{1}{(1+x)^2}, \quad x > 0.$$

3

(ii)

$$f_{Y|X}(y|X=x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\sqrt{x}\mathsf{e}^{-y\sqrt{x}}}{(1+x)^2} \times (1+x)^2 = \sqrt{x}\mathsf{e}^{-y\sqrt{x}}, \;\; y > 0, x > 0.$$

2

The conditional distribution of Y given $X = x$ is $Exponential(\sqrt{x})$, i.e an exponential distribution with parameter $\sqrt{x}$. 1

(iii)

$$\mathsf{E}_{f_{Y|X}}(Y|X=\lambda) = \int_y y f_{Y|X}(y|X=x)\,\mathsf{d}y$$
$$= \int_0^\infty y\sqrt{x}\mathsf{e}^{-\sqrt{x}y}\,\mathsf{d}y \qquad \begin{array}{ll} u = y & \frac{\mathrm{d}v}{\mathrm{d}y} = \sqrt{x}\mathsf{e}^{-y\sqrt{x}} \\ \frac{\mathrm{d}u}{\mathrm{d}y} = 1 & v = -\mathsf{e}^{-y\sqrt{x}} \end{array}$$
$$= \left[-y\mathsf{e}^{-y\sqrt{x}}\right]_0^\infty + \int_0^\infty \mathsf{e}^{-y\sqrt{x}}\,\mathsf{d}y = \left[\frac{-\mathsf{e}^{-y\sqrt{x}}}{\sqrt{x}}\right] = \frac{1}{\sqrt{x}}.$$

3

Or could state this as $Y|X=x$ is $Exponential(\sqrt{x})$

sim seen ⇓

(iv) $Z=(1+X)^{-1}$, so the range of Z is $(0,1)$. Consider the cdf of Z:

$$\begin{aligned}
\mathsf{P}(Z\le z) &= \mathsf{P}\left(\frac{1}{1+X}\le z\right)=\mathsf{P}\left(1\le z(1+X)\right)\\
&= \mathsf{P}\left(X\ge\frac{1-z}{z}\right)=1-F_X\left(\frac{1-z}{z}\right)\\
\Rightarrow f_Z(z) &= \frac{1}{z^2}f_X\left(\frac{1-z}{z}\right)\\
&= \frac{1}{z^2}\left(1+\frac{1-z}{z}\right)^{-2}=\frac{1}{z^2}(z^2)=1,\ 0<z<1.
\end{aligned}$$

So $Z\sim U(0,1)$. 2

(v) unseen ⇓

$$\begin{aligned}
F_{Y|Z}(y|Z=z) &= \mathsf{P}(Y\le y\mid(1+X)^{-1}=z)=\mathsf{P}\left(Y\le y\mid X=\frac{1-z}{z}\right)\\
&= F_{Y|X}\left(y|X=\frac{1-z}{z}\right)\\
\Rightarrow f_{Y|Z}(y|Z=z) &= f_{Y|X}\left(y|X=\frac{1-z}{z}\right)=\sqrt{\frac{1-z}{z}}\mathsf{e}^{-y\sqrt{\frac{1-z}{z}}},\ y>0;\ 0<z<1.
\end{aligned}$$

2

(vi) Note that

$$Y|Z=z\sim \mathit{Exponential}\left(\sqrt{\frac{1-z}{z}}\right)\Rightarrow \mathsf{E}_{f_{Y|Z}}(Y|Z=z)=\sqrt{\frac{z}{1-z}}.$$

$$\begin{aligned}
\mathsf{E}_{f_Y}(Y) &= \mathsf{E}_{f_Z}(\mathsf{E}_{f_{Y|Z}}(Y|Z))=\int_0^1\mathsf{E}_{f_{Y|Z}}(Y|Z)f_Z(z)\,\mathsf{d}z\\
&= \int_0^1\sqrt{\frac{z}{1-z}}\,\mathsf{d}z=\int_0^1 z^{1/2}(1-z)^{-1/2}\,\mathsf{d}z\\
&= \Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{1}{2}\right)=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)^2=\frac{\pi}{2}.
\end{aligned}$$

4

Question 4: commentary:
Part (a) seen/bookwork (3 marks).
Parts (b)(i) and (ii) straightforward for any student that has engaged with the course - similar to seen material (6 marks).
Part (b)(iii) slightly more challenging given the notation (3 marks).
Part (b)(iv) Relatively straightforward transformation, though students often forget to consider the range (2 marks).
Part (b)(v) This is unseen, the good students should be able to map from standard marginal transformations, but this will be challenging for weaker students (2 marks).
Part (b)(vi) Straightforward application of iterated expectation, but it does require a good understanding of the concepts to be able to develop the expectation (4 marks).