Imperial College London

M2S1

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May-June 2018

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

Probability and Statistics II

Date: Wednesday, 23 May 2018

Time: 10:00 AM - 12:00 PM

Time Allowed: 2 hours

This paper has 4 questions.

Candidates should start their solutions to each question in a new main answer book.

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

All required additional material will be 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.
- Each question carries equal weight.
- Calculators may not be used.

1. The bivariate normal vector $\mathbf{Z} = (Z_1, Z_2)$ has density function

$$f_{Z_1 Z_2}(z_1, z_2) = \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(z_1^2 - 2\rho z_1 z_2 + z_2^2\right)\right), \quad z_1, z_2 \in \mathbb{R},$$

where $\quad -1 < \rho < 1.$

(a) Obtain the marginal density function of Z_1.

(b) Obtain the conditional density of $Z_2 | Z_1 = z_1$.

(c) Find the covariance between Z_1 and Z_2.

Let $Y_1 = Z_1$ and $Y_2 = \frac{Z_2 - \rho Z_1}{\sqrt{1-\rho^2}}$.

(d) Determine the joint distribution of Y_1 and Y_2, and explain whether or not these variables are independent.

(e) Stating clearly any general results used, give the distribution of

(i) $\bar{Z} = \frac{1}{2}(Z_1 + Z_2)$.

(ii) $U = Y_1^2 + Y_2^2$.

(iii) $V = (Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2$, where $\bar{Y} = \frac{1}{2}(Y_1 + Y_2)$.

(f) Show that

$$\Pr(Z_1 > 0, Z_2 > 0) = \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\rho.$$

2. (a) What is meant by saying a collection $\mathcal{F}$ of subsets of a set Ω is a field? What is meant by saying that a field $\mathcal{G}$ is a sigma algebra?

(b) For each of the following statements, state whether it is true or false, giving justifications in each case.

(i) If $\mathcal{F}$ is a field and $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

(ii) If $A \subset \Omega$ and $B \subset \Omega$ with $A \cap B \in \mathcal{F}$, where $\mathcal{F}$ is a field, then $A \in \mathcal{F}$ and $B \in \mathcal{F}$.

(c) Let $\Omega = \{1, 2, \ldots\}$ and define $\mathcal{F}$ to be the set of all finite or cofinite subsets of Ω.

(i) Give an example of a subset of Ω that is not contained in $\mathcal{F}$.

(ii) Show that $\mathcal{F}$ is a field.

(iii) State with brief justification whether or not $\mathcal{F}$ is a sigma algebra.

(d) Suppose that $B_1 \supset B_2 \supset \ldots$ is a sequence of sets in a sigma algebra $\mathcal{G}$, and $\Pr$ is a probability function on $\mathcal{G}$. By considering a suitable disjoint union, show that

$$\Pr(\cap_{n=1}^{\infty} B_n) = \lim_{n\to\infty} \Pr(B_n).$$

(e) Show that if the sequence C_n of sets in a sigma algebra $\mathcal{G}$ is such that $\Pr(C_n) = 1$ for all $n = 1, 2, \ldots$, then $\Pr(\cap_{n=1}^{\infty} C_n) = 1$.

3. The distribution of the random variable X is specified hierarchically in terms of the random variable $Z \sim \textsc{Bernoulli}(p)$ as follows

$$\Pr(X = k|Z = 1) = \begin{cases} 1 & k = 0 \\ 0 & k \geq 1, \end{cases}$$

and

$$\Pr(X = k|Z = 0) = \frac{e^{-\lambda}\lambda^k}{k!} \quad k \geq 0,$$

where $0 < p < 1$ and $\lambda > 0$.

(a) Write down the unconditional probability mass function of X.

(b) Derive $\Pr(Z = 0|X = 0)$.

(c) State the law of iterated expectation and use it to compute the mean of X.

(d) State the law of total variance and use it to show that the variance of X is

$$\lambda(1 - p)(1 + \lambda p).$$

(e) Use the method of moments to find estimators of the two parameters p and λ, based on a random sample $x_1, \ldots, x_n$.

(f) Can the method of moments estimator of p ever be outside the parameter space?

4. (a) Let $U \sim \textsc{Uniform}[0, 1]$ and let the strictly increasing function F_X be the cumulative distribution function of the continuous random variable X. Show that $F_X^{-1}(U)$ has cumulative distribution function F_X.

(b) Explain in detail how a random sample $U_1, \ldots U_n \sim \textsc{Uniform}[0, 1]$ can be used to generate a random sample of size n from the random variable Y, whose cumulative distribution function is

$$F_Y(y) = e^{-e^{-y}}, \quad y \in \mathbb{R}.$$

(c) Define what it means for a sequence of random variables to converge in distribution.

(d) Find the cumulative distribution function of the random variable X, whose density function is given by

$$f_X(x) = \frac{e^x}{(1 + e^x)^2}, \quad x \in \mathbb{R}.$$

(e) If $X_1, X_2, \ldots, X_n$ is a random sample from the distribution defined in part (d), find the cumulative distribution function of $Z_n = \max\{X_1, \ldots X_n\}$ and show that $F_{Z_n}(z) \to 0$ as $n \to \infty$ for each $z \in \mathbb{R}$.

(f) Find a constant (i.e. non-random) sequence a_n such that $Y_n = Z_n - a_n$ converges in distribution to the variable Y defined in part (b).

DISCRETE DISTRIBUTIONS

	range $\mathbb{X}$	*parameters*	*pmf* f_X	*cdf* F_X	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	*mgf* M_X
$Bernoulli(\theta)$	$\{0,1\}$	$\theta \in (0,1)$	$\theta^x(1-\theta)^{1-x}$		θ	$\theta(1-\theta)$	$1-\theta+\theta e^t$
$Binomial(n,\theta)$	$\{0,1,\ldots,n\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n}{x}\theta^x(1-\theta)^{n-x}$		$n\theta$	$n\theta(1-\theta)$	$(1-\theta+\theta e^t)^n$
$Poisson(\lambda)$	$\{0,1,2,\ldots\}$	$\lambda \in \mathbb{R}^+$	$\dfrac{e^{-\lambda}\lambda^x}{x!}$		λ	λ	$\exp\{\lambda(e^t-1)\}$
$Geometric(\theta)$	$\{1,2,\ldots\}$	$\theta \in (0,1)$	$(1-\theta)^{x-1}\theta$	$1-(1-\theta)^x$	$\dfrac{1}{\theta}$	$\dfrac{(1-\theta)}{\theta^2}$	$\dfrac{\theta e^t}{1-e^t(1-\theta)}$
$NegBinomial(n,\theta)$	$\{n,n+1,\ldots\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{x-1}{n-1}\theta^n(1-\theta)^{x-n}$		$\dfrac{n}{\theta}$	$\dfrac{n(1-\theta)}{\theta^2}$	$\left(\dfrac{\theta e^t}{1-e^t(1-\theta)}\right)^n$
or	$\{0,1,2,\ldots\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n+x-1}{x}\theta^n(1-\theta)^x$		$\dfrac{n(1-\theta)}{\theta}$	$\dfrac{n(1-\theta)}{\theta^2}$	$\left(\dfrac{\theta}{1-e^t(1-\theta)}\right)^n$

The PDF of the *multivariate normal distribution* is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{K/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}.$$

for $\boldsymbol{x} \in \mathbb{R}^K$ with Σ a $(K \times K)$ variance-covariance matrix and $\boldsymbol{\mu}$ a $(K \times 1)$ mean vector.

The location/scale transformation $Y = \mu + \sigma X$ gives

$$f_Y(y) = \frac{1}{\sigma} f_X\left(\frac{y-\mu}{\sigma}\right) \qquad F_Y(y) = F_X\left(\frac{y-\mu}{\sigma}\right)$$

$$M_Y(t) = e^{\mu t} M_X(\sigma t) \qquad \mathrm{E}[Y] = \mu + \sigma \mathrm{E}[X] \qquad \mathrm{Var}[Y] = \sigma^2 \mathrm{Var}[X]$$

The *gamma function* is given by $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$.

CONTINUOUS DISTRIBUTIONS

		parameters	*pdf*	*cdf*	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	*mgf*
$Uniform(\alpha,\beta)$ (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	$e^{(\mu t-\sigma^2t^2/2)}$
$Student(\nu)$	$\mathbb{R}$	$\nu\in\mathbb{R}^+$	$\dfrac{(\pi\nu)^{-\frac{1}{2}}\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left\{1+\frac{x^2}{\nu}\right\}^{(\nu+1)/2}}$		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 $\alpha>1$)	$\dfrac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$ (if $\alpha>2$)	
$Beta(\alpha,\beta)$	$(0,1)$	$\alpha,\beta\in\mathbb{R}^+$	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$		$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	

1. (a) *[Seen]* Note first that

$$z_1^2 - 2\rho z_1 z_2 + z_2^2 = (z_2 - \rho z_1)^2 + (1-\rho^2)z_1^2.$$

For any $z_1 \in \mathbb{R}$,

$$\begin{aligned} f_{Z_1}(z_1) &= \int_{-\infty}^{\infty} f_{Z_1,Z_2}(z_1,z_2)\,dz_2 = \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)} - \frac{z_1^2}{2}\right) dz_2 \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)}\right) dz_2 = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2}, \end{aligned}$$

where the final equality follows because the integrand is the density function of a $N(\rho z_1, 1-\rho^2)$variable.

(b) *[Seen]* Using the factorization given in part (a) and the definition of the conditional probability density,

$$\begin{aligned} f_{Z_2|Z_1}(z_2|z_1) = \frac{f_{Z_1,Z_2}(z_1,z_2)}{f_{Z_1}(z_1)} &= \frac{\frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)} - \frac{z_1^2}{2}\right)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2}} \\ &= \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)}\right), \qquad z_2 \in \mathbb{R}. \end{aligned}$$

Hence $Z_2|Z_1 = z_1 \sim N(\rho z_1, 1-\rho^2)$.

(c) *[Seen]* Using the factorization in part (a),

$$\begin{aligned} \mathrm{E}(Z_1Z_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{z_1 z_2}{2\pi\sqrt{(1-\rho^2)}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)} - \frac{z_1^2}{2}\right) dz_2 dz_1 \\ &= \int_{-\infty}^{\infty} \frac{z_1}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2}\right) \int_{-\infty}^{\infty} \frac{z_2}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)}\right) dz_2 dz_1 \end{aligned}$$

The inner integral can be seen to be the expectation of a $N(\rho z_1, 1-\rho^2)$ variable, so this simplifies to

$$\mathrm{E}(Z_1Z_2) = \int_{-\infty}^{\infty} \frac{\rho z_1^2}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2}\right) dz_1 = \rho \mathrm{E}(Z_1^2) = \rho.$$

Equivalently, using iterated expectations and part (b),

$$\mathrm{E}(Z_1Z_2) = \mathrm{E}(\mathrm{E}(Z_1Z_2|Z_1)) = \mathrm{E}(\rho Z_1^2) = \rho.$$

Now $\mathrm{Cov}(Z_1, Z_2) = \mathrm{E}(Z_1Z_2) - \mathrm{E}(Z_1)\mathrm{E}(Z_2) = \rho - 0 = \rho$.

(d) *[Seen Method]* As (Y_1, Y_2) is a linear transformation of a multivariate normal vector, it also has the multivariate normal distribution.

To determine expectations, note that $\mathrm{E}(Y_1) = \mathrm{E}(Z_1) = 0$ and $\mathrm{E}(Y_2) = \frac{\mathrm{E}(Z_2)-\rho\mathrm{E}(Z_1)}{\sqrt{1-\rho^2}} = 0$.

To determine variances, $\mathrm{Var}(Y_1) = \mathrm{Var}(Z_1) = 1$ and

$$\mathrm{Var}(Y_2) = \frac{\mathrm{Var}(Z_2) + \rho^2\mathrm{Var}(Z_1) - 2\rho\mathrm{Cov}(Z_1, Z_2)}{1-\rho^2} = \frac{1+\rho^2-2\rho^2}{1-\rho^2} = 1.$$

For the covariance,

$$\mathrm{Cov}(Y_1, Y_2) = \mathrm{Cov}\left(Z_1, \frac{Z_2 - \rho Z_1}{\sqrt{1-\rho^2}}\right) = \frac{1}{\sqrt{1-\rho^2}}\left(\mathrm{Cov}(Z_1, Z_2) - \rho\mathrm{Cov}(Z_1, Z_1)\right) = 0.$$

As multivariate normal random variables with zero correlation, Y_1 and Y_2 are independent. (Joint density factorizes).

(e) (i) *[Seen Method]* As a linear combination of normal variables, $\bar{Z}$ is normally distributed. By linearity of expectation,

$$\frac{1}{2}\left(\mathrm{E}(Z_1) + \mathrm{E}(Z_2)\right) = 0.$$

By the addition formula for variances,

$$\mathrm{Var}(\bar{Z}) = \frac{1}{4}\left(\mathrm{Var}(Z_1) + 2\mathrm{Cov}(Z_1, Z_2) + \mathrm{Var}(Z_2)\right) = \frac{1}{2}(1+\rho).$$

(ii) *[Seen Similar]* The variables $Y_1 = Z_1$ and $Y_2 = \frac{Z_2 - \rho Z_1}{\sqrt{1-\rho^2}}$ are independent, standard normal variables, so each $Y_i^2 \sim \chi^2(1)$, so that $Y_1^2 + Y_2^2 \sim \chi^2(2)$.

(iii) *[Seen Similar]* For a random sample $X_1 \ldots X_n \sim \mathrm{Norm}(\mu, \sigma^2)$, defining $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$,

$$\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1).$$

Here, $n = 2$ and $\sigma^2 = 1$, so $V \sim \chi^2(1)$.

(f) *[Unseen]*

$$\begin{aligned}
\Pr(Z_1 > 0, Z_2 > 0) &= \Pr\left(\rho Y_1 + \sqrt{1-\rho^2}Y_2 > 0, Y_1 > 0\right) \\
&= \int_0^\infty \int_{\frac{-\rho y_1}{\sqrt{1-\rho^2}}}^\infty \frac{1}{2\pi} e^{-\frac{1}{2}(y_1^2 + y_2^2)}\, dy_2 dy_1.
\end{aligned}$$

The region of the (y_1, y_2) plane to be integrated over can be written in polar coordinates,

$$\begin{aligned}
\{Z_1 > 0, Z_2 > 0\} &= \{\rho Y_1 + \sqrt{1-\rho^2}Y_2 > 0, Y_1 > 0\} = \left\{\frac{Y_2}{Y_1} > \frac{-\rho}{\sqrt{1-\rho^2}} > 0, Y_1 > 0\right\} \\
&= \left\{\frac{-\rho}{\sqrt{1-\rho^2}} < \tan\Theta < \infty, R > 0\right\} = \left\{\tan^{-1}\frac{-\rho}{\sqrt{1-\rho^2}} < \Theta < \frac{\pi}{2}, R > 0\right\} \\
&= \left\{-\sin^{-1}\rho < \Theta < \frac{\pi}{2}, R > 0\right\},
\end{aligned}$$

since if $\rho = \sin\alpha$, then $\tan\alpha = \frac{\rho}{\sqrt{1-\rho^2}}$.

Performing the integration,

$$\begin{aligned}
\Pr\left(\rho Y_1 + \sqrt{1-\rho^2} Y_2 > 0, Y_1 > 0\right) &= \int_0^{\infty} \int_{\frac{-\rho y_1}{\sqrt{1-\rho^2}}}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}(y_1^2+y_2^2)}\, dy_2 dy_1 \\
&= \int_{-\sin^{-1}\rho}^{\frac{\pi}{2}} \int_0^{\infty} \frac{r}{2\pi} e^{-\frac{1}{2}r^2}\, dr d\theta \\
&= \int_{-\sin^{-1}\rho}^{\frac{\pi}{2}} \int_0^{\infty} \frac{1}{2\pi} \left[-e^{-\frac{1}{2}r^2}\right]_0^{\infty} dr d\theta \\
&= \frac{1}{2\pi}\left(\frac{\pi}{2} + \sin^{-1}\rho\right) = \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\rho.
\end{aligned}$$

2. (a) *[Seen]* $\mathcal{F}$ is a field if

   * $\emptyset \in \mathcal{F}$
   * $A^c \in \mathcal{F}$ whenever $A \in \mathcal{F}$
   * $A \cup B \in \mathcal{F}$ whenever $A \in \mathcal{F}$ and $B \in \mathcal{F}$.

   The field $\mathcal{G}$ is a sigma algebra if whenever $A_1, A_2, \ldots$ is a sequence of sets in $\mathcal{G}$, $\cup_{i=1}^{\infty} A_i \in \mathcal{G}$, i.e. $\mathcal{G}$ is closed under countable unions.

(b) (i) *[Seen]* The statement is true. Let $A, B \in \mathcal{F}$. Then by the second field axiom, $A^c, B^c \in \mathcal{F}$. Now by the third field axiom, $A^c \cup B^c \in \mathcal{F}$. Again using the second field axiom, $(A^c \cup B^c)^c \in \mathcal{F}$. But by de Morgan's laws, this is the same as $A \cap B$, so $A \cap B \in \mathcal{F}$.

(ii) *[Unseen]* The statement is false. Consider $\Omega = \{0, 1\}$, with the collection $\mathcal{F} = \{\emptyset, \Omega\}$, which is a field. Let $A = \{0\}$ and $B = \{1\}$, then $A \cap B = \emptyset \in \mathcal{F}$, but $A \notin \mathcal{F}$ and $B \notin \mathcal{F}$.

(c) (i) *[Seen]* Any set which is neither finite nor cofinite, e.g. $\{2, 4, 6, \ldots\}$.

(ii) *[Seen]*

   - Clearly $\emptyset$ is finite so $\emptyset \in \mathcal{F}$.
   - If $A \in \mathcal{F}$, then A is either finite or cofinite; so A^c is cofinite or finite, respectively, so $A^c \in \mathcal{F}$.
   - If $A, B \in \mathcal{F}$, consider distinct cases. If both A and B are finite, then $A \cup B$ is finite so $A \cup B \in \mathcal{F}$. If A (say) is cofinite, then $(A \cup B)^c = A^c \cap B^c \subset A^c$, which is finite, so $A \cup B \in \mathcal{F}$.

(iii) *[Seen Method]* $\mathcal{F}$ is not a sigma algebra, since the singleton sets $A_k = \{2k\} \in \mathcal{F}$ for each $k = 1, 2, \ldots$, but $\cup_{k=1}^{\infty} A_k = \{2, 4, \ldots\}$ is neither finite nor cofinite.

(d) *[Seen Similar]*

$$\Pr\left(\cap_{i=1}^{\infty} B_i\right) = 1 - \Pr\left(\left(\cap_{i=1}^{\infty} B_i\right)^c\right) = 1 - \Pr\left(\left(\cup_{i=1}^{\infty} B_i^c\right)\right) \text{ (de Morgan)}.$$

The complementary sets form an increasing sequence $B_1^c \subset B_2^c \subset \ldots$. We can now write $\cup_{i=1}^{\infty} B_i^c$ as a disjoint union $B_1^c \cup (B_2^c \backslash B_1^c) \cup \ldots$, so that by countable additivity of Pr,

$$\Pr(\cup_{i=1}^{\infty} B_i^c) = \Pr(B_1^c) + \sum_{i=2}^{\infty} \Pr(B_i^c) - \Pr(B_{i-1}^c) = \Pr(B_1^c) + \lim_{n\to\infty} \sum_{i=2}^{n} \Pr(B_i^c) - \Pr(B_{i-1}^c).$$

Considering the cancellations in successive terms, this gives

$$\Pr(\cup_{i=1}^{\infty} B_i^c) = \lim_{n\to\infty} \Pr(B_n^c),$$

so that on taking complements,

$$\Pr\left(\cap_{i=1}^{\infty} B_i\right) = 1 - \Pr\left(\left(\cap_{i=1}^{\infty} B_i\right)^c\right) = 1 - \lim_{n\to\infty} \Pr(B_n^c) = \lim_{n\to\infty} \Pr(B_n).$$

(e) *[Unseen]* For any events A and B, we have that

$$\Pr(A \cap B) = \Pr(A) + \Pr(B) - \Pr(A \cup B) \geq \Pr(A) + \Pr(B) - 1.$$

Applying this result to C_1 and C_2 gives

$$\Pr(C_1 \cap C_2) = \Pr(C_1) + \Pr(C_2) - \Pr(C_1 \cup C_2) \geq \Pr(C_1) + \Pr(C_2) - 1 = 1,$$

from which it follows that we must in fact have $\Pr(C_1 \cap C_2) = 1$.

Arguing inductively, we can apply the first result with $A = \cap_{i=1}^{n-1} C_i$ and $B = C_n$ to conclude that

$$\Pr(\cap_{i=1}^{n} C_i) \geq 1 \qquad \text{for all } n,$$

so then we must in fact have

$$\Pr(\cap_{i=1}^{n} C_i) = 1 \qquad \text{for all } n.$$

By the property established in the previous part, applied to the decreasing sequence $B_n = \cap_{i=1}^{n} C_i$, this gives

$$\Pr\left(\cap_{i=1}^{\infty} C_i\right) = \lim_{n\to\infty} \Pr\left(\cap_{i=1}^{n} C_i\right) \geq 1.$$

Hence $\Pr\left(\cap_{i=1}^{\infty} C_i\right) = 1$.

3. (a) *[Seen Method]*

$$\Pr(X = k) = \Pr(X = k|Z = 0)\Pr(Z = 0) + \Pr(X = k|Z = 1)\Pr(Z = 1).$$

For $X = 0$, both terms are non-zero and we have

$$\Pr(X = 0) = e^{-\lambda}(1-p) + 1 \times p.$$

For $X > 0$, only the first term is non-zero and we have

$$\Pr(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}(1-p).$$

(b) *[Seen Method]*

$$\Pr(Z = 0|X = 0) = \frac{\Pr(X = 0|Z = 0)\Pr(Z = 0)}{\Pr(X = 0)} = \frac{e^{-\lambda}(1-p)}{p + e^{-\lambda}(1-p)}.$$

(c) *[Seen]* The law of iterated expectation states that for random variables X and Z,

$$\mathrm{E}(X) = \mathrm{E}(\mathrm{E}(X|Z)),$$

where the inner expectation is with respect to the conditional distribution of $X|Z$ and the outer expectation is with respect to the distribution of Z.
In this case, $X|Z = (1-Z)Y$, where $Y \sim \textsc{Poisson}(\lambda)$, so

$$\mathrm{E}(X|Z) = (1-Z)\lambda.$$

Since now $E(Z) = p$, the law of iterated expectations gives

$$\mathrm{E}(X) = \lambda\mathrm{E}(1-Z) = \lambda(1-p).$$

(d) *[Seen]* The law of total variance states that for random variables X and Z,

$$\mathrm{Var}(X) = \mathrm{E}(\mathrm{Var}(X|Z)) + \mathrm{Var}(\mathrm{E}(X|Z)).$$

In this case, for Y as in the previous part,

$$\mathrm{Var}(X|Z) = \mathrm{Var}((1-Z)Y|Z) = (1-Z)^2\mathrm{Var}(Y) = (1-Z)^2\lambda,$$

so

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{E}((1-Z)^2\lambda) + \mathrm{Var}((1-Z)\lambda) \\ &= (1-p)\lambda + \lambda^2 p(1-p) = \lambda(1-p)(1+\lambda p). \end{aligned}$$

(e) *[Seen Method]* In the method of moments, we assume $\bar{x} \approx \mathrm{E}(X)$ and $s^2 \approx \mathrm{Var}(X)$.

Need to solve the two equations

$$\bar{x} = \lambda(1-p) \quad \text{and } s^2 = \lambda(1-p)(1+\lambda p).$$

Substituting the first equation into the second gives

$$\frac{s^2}{\bar{x}} = 1 + \lambda p.$$

Since we also have

$$\bar{x} = \lambda - \lambda p,$$

adding gives

$$\frac{s^2}{\bar{x}} + \bar{x} = 1 + \lambda,$$

so that

$$\hat{\lambda} = \frac{s^2}{\bar{x}} + \bar{x} - 1.$$

Substituting then gives

$$\begin{aligned} \hat{p} &= \frac{\frac{s^2}{\bar{x}} - 1}{\hat{\lambda}} \\ &= \frac{\frac{s^2}{\bar{x}} - 1}{\frac{s^2}{\bar{x}} + \bar{x} - 1}. \end{aligned}$$

(f) *[Seen Similar]* For the sample $(2, 2)$, clearly $\bar{x} = 2$ and $s^2 = 0$.
Substituting for $\hat{p}$ then gives

$$\hat{p} = \frac{0 - 1}{0 + 2 - 1} = -1 < 0.$$

This is clearly outside the domain of p.

4. (a) *[Seen]* Let $Y = F_X^{-1}(U)$.
Then since F_X is an increasing, and therefore 1-1, function, whose range is $[0, 1]$,

$$\Pr(Y \le y) = \Pr(F_X^{-1}(U) \le y) = \Pr(U \le F_X(y)) = F_X(y).$$

Hence Y has the same cumulative distribution function as X.

(b) *[Seen Similar]* By the previous result, it suffices to find the inverse function F_Y^{-1}.
If $u = e^{-e^{-y}}$ then $\log u = -e^{-y}$, so $y = -\log(-\log u)$.
Hence, if we define $Y_i = -\log(-\log U_i)$, then $Y_1, Y_2, \ldots Y_n$ is a random sample from the distribution given.

(c) *[Seen]* A sequence $X_1, X_2, \ldots$ converges in distribution to a random variable X if

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

at all points of continuity of F_X.

(d) *[Seen Method]*

$$F_X(x) = \int_{-\infty}^{x} \frac{e^t}{(1+e^t)^2}\,dt = \int_{u=0}^{u=e^x} \frac{1}{(1+u)^2}\,du = \left[\frac{-1}{1+u}\right]_0^{u=e^x} = 1 - \frac{1}{1+e^x} = \frac{e^x}{1+e^x} = \frac{1}{1+e^{-x}}, \quad x \in \mathbf{R}.$$

(e) *[Seen Method]* Note that $Z_n \le z$ if and only if $X_i \le x$ for all $i = 1, 2, \ldots n$. Hence, since the X_i are independent,

$$Pr(Z_n \le z) = \Pr(X_1 \le z, \ldots, X_n \le z) = \prod_{i=1}^{n} \Pr(X_i \le z) = \frac{1}{(1+e^{-z})^n}.$$

For any $z \in \mathbf{R}$, $1 + e^{-z} > 1$, hence $\Pr(Z_n \le z) \to 0$ as $n \to \infty$.

(f) *[Seen Similar]* Define $a_n = \log n$. Then

$$\Pr(Y_n \le y) = \Pr(Z_n \le y + \log n) = \frac{1}{(1+e^{-y-\log n})^n} = \frac{1}{(1+\frac{e^{-y}}{n})^n} = \left(1 + \frac{e^{-y}}{n}\right)^{-n}.$$

Applying the limit definition $\lim_{n\to\infty}(1+\frac{x}{n})^{-n} = e^{-x}$ then gives

$$\lim_{n\to\infty} \Pr(Y_n \le y) = e^{-e^{-y}}, y \in \mathbf{R}.$$

Marks:

1. (a) 3 marks. 1 mark for factorizing quadratic, 1 mark for identifying kernel of normal density, 1 mark for specifying range for z_1.
   (b) 2 marks. 1 mark for definition of conditional density and 1 mark for numerical answer.
   (c) 2 marks. 1 mark for correct definition of covariance, 1 mark for correct answer.
   (d) 3 marks. 1 mark for correctly determining that the Y_i follow a multivariate normal distribution, 1 mark for correctly determining its covariance, 1 mark for justification of independence.
   (e) (i) 2 marks. 1 mark for correct distribution, 1 mark for justification.
   (ii) 2 marks. 1 mark for correct distribution, 1 mark for justification.
   (iii) 2 marks. 1 mark for correct distribution, 1 mark for justification.
   (f) 4 marks. 2 marks for correctly determining region of integration in terms of independent variables, 1 mark for reasonable use of polar coordinates, 1 mark for correct answer.

2. (a) 3 marks. 2 marks for all field properties correct, 1 mark for sigma algebra condition.
   (b) (i) 2 marks. 1 mark for correct use of field axioms, 1 mark for use of de Morgan's law.
   (ii) 2 marks. 1 mark for correct example, 1 mark for reasoning.
   (c) (i) 1 mark for any correct example.
   (ii) 2 marks. 1 mark for first two field properties, 1 mark for third property.
   (iii) 2 marks. 1 mark for correct example, 1 mark for reasoning.
   (d) 4 marks. 2 marks for identifying correct disjoint union, 1 mark for telescoping sum, 1 mark for final answer.
   (e) 4 marks. 3 marks for proving valid inequality for finite n, 1 mark for using continuity to pass to limit.

3. (a) 2 marks. 1 mark for $X = 0$, 1 mark for $X > 0$.

   (b) 3 marks. 1 mark for attemping to use Bayes' theorem, 1 mark for correct denominator, 1 mark for correct answer.

   (c) 3 marks. 1 mark for correctly stating the law of iterated expectation, 2 marks for correctly applying it.

   (d) 4 marks. 1 mark for correctly stating the law of total variance, 1 mark for computing each inner term correctly, 1 mark for correctly taking expectations over Z.

   (e) 5 marks. 1 mark for forming each moment equation correctly, 1 mark for reasonable attempt to solve equations, 1 mark for each correct estimator.

   (f) 3 marks. 1 mark for correct example, 2 marks for justification.

4. (a) 3 marks. 1 mark for asserting F is 1-1, 1 mark for correctly manupulating inequality, 1 mark for correct use of uniform cdf.

   (b) 3 marks. 1 mark for attempting to invert cdf, 1 mark for correct inverse, 1 mark for conclusion.

   (c) 2 marks. 1 mark for correct limit, 1 mark for "at all points of continuity".

   (d) 3 marks. 1 mark for attempting integration of pdf, 1 mark for correct substitution, 1 mark for final answer.

   (e) 5 marks. 2 marks for condition that $Z_n \leq z$, 1 mark for using independence to factorize joint cdf, 1 mark for final answer, 1 mark for behaviour as $n \to \infty$.

   (f) 4 marks. 1 mark for reasonable attempt to determine a_n, 1 mark for correctly identifying $\log n$, 1 mark for using limit definition of e, 1 mark for correct final form.