1. (a) Three red balls and three white balls are placed in a bag. The six balls are drawn out one by one, at random and without replacement. The random variable X is the number of white balls drawn before the first red ball is drawn.

   (i) Find the probabilities $P(X = x)$ for $x = 0, 1, 2, 3$. [5 marks]

   (ii) Find $E[X]$ and $\text{Var}[X]$. [3 marks]

(b) Suppose now that the balls are drawn out of the bag **with** replacement.

   (i) Let Y denote the number of balls drawn up to and including the first red ball drawn. Give an expression for the probability $P(Y = y)$ for $y = 1, 2, \ldots$, justifying your answer. Name the distribution of Y, and write down the values of $E[Y]$ and $\text{Var}[Y]$. [6 marks]

   (ii) Let V denote the number of balls drawn up to and including the **second** red ball drawn. Give an expression for the probability $P(V = v)$ for $v = 2, 3, \ldots$, justifying your answer. Name the distribution of V, and write down the values of $E[V]$ and $\text{Var}[V]$. [6 marks]

2. (a) Suppose the random variable T is exponentially distributed with parameter λ, so that T has probability density function

$$f_T(t) \quad = \quad \begin{cases} 0 & t < 0, \\ \lambda e^{-\lambda t} & t \geq 0. \end{cases}$$

Derive expressions for the (cumulative) distribution function $F_T(t)$ of T and the expectation $E[T]$.

Derive also expressions for the median and the upper and lower quartiles of the distribution of T. [6 marks]

(b) Suppose that T is exponentially distributed with parameter $\lambda = 4$, and that Y is defined by

$$Y \quad = \quad 1 - e^{-2T}$$

(i) Determine the range of Y. [2 marks]

(ii) Find the (cumulative) distribution function of Y. [6 marks]

(iii) Show that the probability density function of Y is given (within the range of non-zero density) by

$$f_Y(y) \quad = \quad 2(1 - y).$$

[2 marks]

(iv) Find $E[Y]$ and $\mathrm{Var}[Y]$, and the median of the distribution of Y. [4 marks]

3. (a) Give formulae defining the *covariance* and the *correlation* of two random variables X and Y. Explain how correlation values may be interpreted. [4 marks]

(b) Suppose X and Y are discrete random variables with the joint probability mass function given in the following table.

	Y=0	Y=1	Y=2
X=2	0.1	0.1	0.3
X=4	0.2	0.2	0.1

(i) Find the marginal probability mass functions of X and Y, and hence find $E[X]$, $E[Y]$, $\text{Var}[X]$, $\text{Var}[Y]$. [6 marks]

(ii) Find the covariance $\text{Cov}[X, Y]$ and the correlation $\text{Corr}[X, Y]$. [4 marks]

(iii) Comment on your computed correlation value. [2 marks]

(iv) Find $\text{Var}[X + Y]$ and $\text{Var}[X - Y]$. [4 marks]

4. Suppose that random variables X, Y are jointly continuous with joint density function

$$f_{X,Y}(x,y) \quad = \quad k\left(x(1-x)+y(1-y)\right), \qquad 0 \le x \le 1,\ 0 \le y \le 1,$$

for some constant k.

(a) Find the value of k, and the marginal densities $f_X(x)$ and $f_Y(y)$. [8 marks]

(b) Let random variables U, V be defined by

$$U = X - Y, \qquad V = X + Y.$$

Fnd the joint density $f_{U,V}(u,v)$. Sketch the region where the random vector (U, V) has positive density. [12 marks]

5. (a) For a discrete random variable U, the Probability Generating Function $G_U(s)$ of U is defined by $G_U(s) = E\left[s^U\right]$.

Show that (i) $E[U] = G'_U(1)$ and (ii) $\text{Var}[U] = G''(1) + G'(1) - (G'_U(1))^2$, where $G'_U(s)$ and $G''(s)$ denote the first and second derivatives, respectively, of $G_U(s)$ with respect to s. [7 marks]

(b) Suppose V is a discrete random variable with probability mass function

$$\Pr(V = k) \quad = \quad \frac{1}{2^k \; k \; \ln(2)} \qquad k = 1, 2, \ldots.$$

Show that the Probability Generating Function $G_V(s)$ of V is given by

$$G_V(s) \quad = \quad 1 - \frac{\ln(2 - s)}{\ln(2)} \qquad \text{for } -2 \le s < 2.$$

[You may use without proof the result that for $-1 \le x < 1$,

$$\sum_{i=1}^{\infty} \frac{x^i}{i} = \ln\left(\frac{1}{1-x}\right).$$
]

[7 marks]

(c) Using the probability generating function $G_V(s)$ from part (b) above, find the expectation $E[V]$ and the variance $\text{Var}[V]$. [6 marks]

6. Suppose the random variables $X_1, X_2, \ldots, X_n$ are independent and identically distributed, each with mean $\mu = 9$ and variance $\sigma^2 = 9$. Denote by $\bar{X}$ the sample mean, so that

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}.$$

(a) Suggest an approximation to the distribution of $\bar{X}$ valid in the limit as $n \to \infty$, justifying your suggestion. [2 marks]

Hence

(b) Find (approximately) the smallest n for which $P\left(\bar{X} > 9.05\right) < 0.01$. [4 marks]

(c) Show that for $c > 0$,

$$P\left(\left|\bar{X} - 9\right| < c\right) \approx 2\Phi\left(\frac{c\sqrt{n}}{3}\right) - 1,$$

where $\Phi(x)$ is the standard normal distribution function. [4 marks]

Suppose now that for $i = 1, 2, \ldots, n$ each X_i follows the Poisson distribution with mean 9, and that the random variables $X_1, X_2, \ldots, X_n$ are mutually independent as before. With $n = 5$ and $c = 0.3$,

(d) compute the probability $P\left(\left|\bar{X} - 9\right| < 0.3\right)$ exactly; [5 marks]

(e) compute also an approximation to $P\left(\left|\bar{X} - 9\right| < 0.3\right)$ using the formula of part (c) above; [2 marks]

(f) comment on your results of parts (d) and (e). [3 marks]

7. (a) Let Z be a standard normal random variable, so that $Y = Z^2$ follows the chi-squared distribution with 1 degree of freedom. Show that the probability density function $f_Y(y)$ of Y is given by $f_Y(y) = \dfrac{1}{\sqrt{2\pi y}} \exp(-y/2)$ for $y \geq 0$.

[Recall that the standard normal random variable Z has probability density function $f_Z(z) = \dfrac{1}{\sqrt{2\pi}} \exp\left(-z^2/2\right)$ for $-\infty < z < \infty$.] [7 marks]

(b) Given that the chi-squared distribution with 1 degree of freedom is identical to the gamma distribution with parameters $\left(\frac{1}{2}, \frac{1}{2}\right)$, that is, $\chi_1^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$, write down an expression for the moment generating function $M_Y(t)$ of the random variable Y of part (a) above.

[You may use without proof the result that for $\lambda, \alpha > 0$ the gamma distributed random variable $U \sim \Gamma(\lambda, \alpha)$ has moment generating function given by

$$M_U(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}.$$

]

[2 marks]

(c) State the definition of the chi-squared distribution with n degrees of freedom for $n \geq 1$.

Hence, denoting by V a random variable following the chi-squared distribution with n degrees of freedom, derive an expression for the moment generating function of V.

Given that the chi-squared distribution with n degrees of freedom is identical to the gamma distribution with parameters λ_n, α_n for some $\lambda_n, \alpha_n > 0$, give expressions for λ_n and α_n in terms of n. [6 marks]

(d) Suppose that for $n = 1, 2, \ldots$, the random variable V_n follows the chi-squared distribution with n degrees of freedom, and define $W_n = V_n/n$. Denote by $M_n(t)$ the moment generating function of W_n. Show that the limiting function $M(t) = \lim_{n\to\infty} M_n(t)$ is given by $M(t) = e^t$. From this, what can you say about the limit of the distribution of the random variable W_n as $n \to \infty$?

[You may use without proof the result that for $-\infty < \mu < \infty$,

$$\lim_{n\to\infty} \left(1 + \frac{\mu}{n}\right)^n = e^{\mu}.$$

]

[5 marks]