EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

GRADUATE DIPLOMA, 2009

(Modular format)

MODULE 1 : Probability Distributions

Time Allowed: Three Hours

Candidates should answer **FIVE** *questions.*

All questions carry equal marks.
The number of marks allotted for each part-question is shown in brackets.

Graph paper and Official tables are provided.

Candidates may use calculators in accordance with the regulations published in the Society's "Guide to Examinations" (document Ex1).

The notation log *denotes logarithm to base* **e**.
Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

Note also that $\binom{n}{r}$ *is the same as* ${}^{n}C_{r}$.

This examination paper consists of 5 printed pages, **each printed on one side only**.
This front cover is page 1.
Question 1 starts on page 2.

There are 8 questions altogether in the paper.

©RSS 2009

1. (i) The events $\{E_i : i = 1, 2, ..., n\}$ partition a sample space, and A is an event with non-zero probability $P(A)$. Write down Bayes' Theorem for the conditional probability $P(E_1 \mid A)$ in terms of conditional probabilities $P(A \mid E_i)$ and unconditional probabilities $P(E_i)$, $i = 1, 2, \ldots, n$.

(4)

(ii) There is a dispute as to whether customer Brian gave barmaid Anna a £10 note or a £20 note to pay for his drink. Both are honest, but may make mistakes. Consider the following three pieces of information.

(A) The till contains ten £10 notes and twenty £20 notes, all mixed up at random.

(B) Anna correctly identifies notes 90% of the time, and states that Brian used a £10 note.

(C) Brian, who correctly identifies £20 notes 80% of the time, and correctly identifies £10 notes 70% of the time, claims to have paid using a £20 note.

Explaining your reasoning, estimate the probability that Brian used a £20 note, if you are given each of the following.

(a) Information (A) alone. (2)

(b) Both (A) and (B). (6)

(c) All of (A), (B) and (C). (8)

2. (i) Suppose A and B are independent events, and $\bar{A}, \bar{B}$ denote the complementary events to A, B. Show that $\bar{A}$ and B are independent; deduce that $\bar{A}$ and $\bar{B}$ are independent.

(4)

(ii) The point (X, Y) has the uniform distribution over the unit square $\{ 0 \le x \le 1,\ 0 \le y \le 1 \}$, and u is a given value with $0 < u < 0.5$. Events A and B are defined as $A : \{X \ge 2Y\}$ and $B : \{Y \le X \le Y + u\}$.

(a) Sketch three separate diagrams of the unit square, illustrating the events A, B and $A \cap B$ respectively. (4)

(b) Find the probabilities of these three events, and show that A and B are independent if, and only if, $u = 2/5$. (10)

(c) Deduce the value of $P(\bar{A} \cap \bar{B})$ when $u = 2/5$. (2)

Turn over

3. Suppose X and Y are independent random variables having Poisson distributions with respective means λ (> 0) and μ (> 0).

(i) Show that $X + Y$ also follows a Poisson distribution.

(5)

(ii) Find $P(X = k \mid X + Y = n)$ when k and n are integers with $0 \le k \le n$. For given fixed $n > 0$, name the distribution you have obtained.

(7)

(iii) Telephone calls arriving at a computer helpline are classed as urgent or standard; urgent calls average 8 per hour, standard calls average 24 per hour. Ten calls arrive within 30 minutes; find (to two significant figures) the probability that at most two of them are urgent, stating any assumptions you make.

(8)

4. (i) Consider a family of non-negative random variables with the property that the variance of any member of the family is proportional to the square of its expected value. Let X be a member of this family, with $E(X) = \mu$. Use a Taylor series expansion to show that, if $Y = \log(X)$, then the variance of Y is approximately constant no matter which X is chosen from the original family of random variables.

(8)

(ii) Show that, when X has the Gamma distribution $\Gamma(\alpha, \nu)$, i.e. its probability density function is $f(x) = \frac{\nu^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\nu x}$ for $x > 0$, then $E(X^k) = \frac{\Gamma(k+\alpha)}{\nu^k \, \Gamma(\alpha)}$ when $k > 0$. Hence verify that, for fixed $\alpha > 0$, this family of random variables has the property described in (i) as ν varies.

(12)

[You may use without proof the results $\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$ and $\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$.]

5. The continuous random variables X and Y have joint probability density function $f(x, y) = kxy$ if $0 < x < y < 1$, with $f(x, y) = 0$ elsewhere, where k is a constant.

(i) Evaluate k, and find the marginal probability densities of X and Y. Say, with a reason, whether or not X and Y are independent.

(10)

(ii) Show that, for all non-negative integers r and s, $E(X^r Y^s) = \dfrac{8}{(r+2)(r+s+4)}$.

Hence find the correlation between X and Y.

(10)

6. (i) State the *Central Limit Theorem* for sums of independent identically distributed random variables.

(4)

(ii) Explain carefully what is meant by the phrase *continuity correction* when using a continuous distribution to approximate the probability $P(X \le k)$ when X takes integer values only, including the value k.

(4)

(iii) Give a detailed argument involving the Central Limit Theorem that justifies the use of a Normal distribution to give an approximation to a probability of the form $P(a \le Y \le b)$ in the following cases.

(a) Y has the binomial distribution with parameters n and p ($0 < p < 1$), with n large.

(6)

(b) Y has a Poisson distribution with parameter λ, when λ is large.

(6)

[You may use without proof standard properties of binomial or Poisson random variables, and need not mention any continuity correction in your answers to (iii).]

Turn over

7. (i) Suppose X and Y are independent random variables, each following the chi-squared distribution with four degrees of freedom; this distribution has probability density function (pdf) $we^{-w/2}/4$ on $w > 0$.

Define new random variables $U = X/Y$ and $V = Y$. Obtain the joint pdf of U and V, and hence show that U has pdf $h(u) = 6u/(1+u)^4$ for $u > 0$.

(10)

[You may use without proof the result $\int_0^\infty t^k e^{-t} dt = k!$ when k is a positive integer.]

(ii) Name the distribution followed by U. Using the density function, show that $E(U) = 2$, and hence verify that $E(X/Y) > E(X)/E(Y)$. Explain briefly why this is no surprise.

(10)

8. It is desired to use the rejection technique to simulate values from the beta distribution with density $f(x) = 6x(1-x)$ over $0 < x < 1$.

(i) Sketch the graph of the function $f(x)$ and, on the same diagram, draw the trapezium whose four vertices are A = (0, 0), B = (0.25, 1.5), C = (0.75, 1.5) and D = (1, 0). Prove that the given density function is entirely contained within this trapezium.

(6)

(ii) Given a method that generates points uniformly distributed inside the trapezium $ABCD$, show that, in the long run, 1/9 of these points will be rejected when the rejection technique is used.

(4)

(iii) Give the outcomes when the rejection technique is applied to the four points (0.452, 0.697), (0.120, 0.066), (0.914, 0.871) and (0.657, 1.345). In each case, state the reason for your answer.

(4)

(iv) Describe an efficient method (i.e. one that wastes no input) for using a stream $\{y_i : i = 1, 2, 3, \ldots\}$ of independent random numbers, each uniformly distributed over the interval (0, 1), to generate points uniformly distributed inside $ABCD$.

(6)