# MATHEMATICAL TRIPOS: PART IA (Lent 2002)

**Probability — Example Sheet 2 (out of 4)** **GRG**

**Exercises**

**1.** A coin with probability $p$ of heads is tossed $n$ times. Let $E$ be the event 'a head is obtained on the first toss' and $F_k$ the event 'exactly $k$ heads are obtained'. For which pairs of integers $(n, k)$ are $E$ and $F_k$ independent?

**2.** The events $A$ and $B$ are independent. Show that the events $A^{\mathrm{c}}$ and $B$ are independent, and that the events $A^{\mathrm{c}}$ and $B^{\mathrm{c}}$ are independent.

**3.** Independent trials are performed, each with probability $p$ of success. Let $\pi_n$ be the probability that $n$ trials result in an even number of successes. Show that $\pi_n = \frac{1}{2}[1 + (1 - 2p)^n]$.

**4.** Two darts players $A$ and $B$ throw alternately at a board and the first to score a bull wins the contest. The outcomes of different throws are independent and on each of their throws $A$ has probability $p_A$ and $B$ has probability $p_B$ of scoring a bull. If $A$ has first throw, calculate the probability of $A$ winning the contest.

**5.** Suppose that $X$ and $Y$ are independent Poisson random variables with parameters $\lambda$ and $\mu$ respectively. Find the distribution of $X + Y$. Prove that the conditional distribution of $X$, given that $X + Y = n$, is binomial with parameters $n$ and $\lambda/(\lambda + \mu)$.

**6.** The number of misprints on a page has a Poisson distribution with parameter $\lambda$, and the numbers on different pages are independent. What is the probability that the second misprint will occur on page $r$?

**7.** $X_1, \dots, X_n$ are independent, identically distributed random variables with mean $\mu$ and variance $\sigma^2$. Find the mean of

$$S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2, \quad \text{where } \bar{X} = \frac{1}{n} \sum_{1}^{n} X_i.$$

**8.** In a sequence of $n$ independent trials the probability of a success at the $i$th trial is $p_i$. Show that mean and variance of the total number of successes are $n\bar{p}$ and $n\bar{p}(1 - \bar{p}) - \sum_i (p_i - \bar{p})^2$ where $\bar{p} = \sum_i p_i / n$. Notice that for a given mean, the variance is greatest when all $p_i$ are equal.

**9.** Let $(X, Y) = (\cos\theta, \sin\theta)$ where $\theta = \frac{K\pi}{4}$ and $K$ is a random variable such that $P\{K = r\} = 1/8$, $r = 0, 1, \dots, 7$. Show that $\operatorname{cov}(X, Y) = 0$, but that $X$ and $Y$ are not independent.

**10.** Let $a_1, a_2, \dots, a_n$ be a ranking of the yearly rainfalls in Cambridge over the next $n$ years: assume $a_1, a_2, \dots, a_n$ is a random permutation of $1, 2, \dots, n$. Say that $k$ is a record year if $a_i > a_k$ for all $i < k$ (thus the first year is always a record year). Let $Y_i = 1$ if $i$ is a record year and 0 otherwise. Find the distribution of $Y_i$ and show that $Y_1, Y_2, \dots, Y_n$ are independent. Calculate the mean and variance of the number of record years in the next $n$ years.

**11.** Hugo's bowl of spaghetti contains $n$ strands. He selects two ends at random and joins them together. He does this until no ends are left. What is the expected number of spaghetti hoops in the bowl?

**12.** Sarah collects figures from cornflakes packets. Each packet contains one figure, and $n$ distinct figures make a complete set. Show that the expected number of packets Sarah needs to buy to collect a complete set is $n \sum_{i=1}^{n} i^{-1}$.

**13.** $(X_k)$ is a sequence of independent identically distributed positive random variables where $E(X_k) = a$ and $E(X_k^{-1}) = b$ exist. Let $S_n = \sum_{k=1}^{n} X_k$. Show that $E(S_m/S_n) = m/n$ if $m \le n$, and $E(S_m/S_n) = 1 + (m - n)aE(S_n^{-1})$ if $m \ge n$.

**Problems**

**14.** You are on a game show and given the choice of three doors. Behind one is a car; behind the others are goats. You pick door 1, and the host opens door 3, which has a goat. He then asks if you want to pick door 2. Should you switch? [Consider two cases. First, suppose the host does not know where the car is. Secondly, suppose the host does know where the car is, and makes sure the door he opens shows a goat.]

**15.** Two cards are taken at random from an ordinary pack of 52 cards. Find the probabilities that:
(i) both cards are aces (event $A$)
(ii) the pair of cards includes an ace (event $B$)
(iii) the pair of cards includes the ace of hearts (event $C$).
Show that $P(A \mid B) \neq P(A \mid C)$.

**16.** Let $X$ be an integer-valued random variable with distribution

$$P(X = n) = n^{-s}/\zeta(s)$$

where $s > 1$, and $\zeta(s) = \sum_{n \geq 1} n^{-s}$, the Riemann zeta function. Let $p_1 < p_2 < p_3 < \cdots$ be the primes and let $A_k$ be the event $\{X \text{ is divisible by } p_k\}$. Find $P(A_k)$ and show that the events $A_1, A_2, \ldots$ are independent.

**17.** John chooses a sequence, such as HHH, and then Mary chooses a sequence, perhaps THH. A fair coin is tossed until one or other sequence occurs, when the coin is awarded to the person whose sequence has been observed. Advise Mary on which sequence she should choose for each (or at least some) of John's eight possible choices.

**18.** You are playing a match against an opponent in which at each point either you or your opponent serves. If you serve you win the point with probability $p_1$, but if your opponent serves you win the point with probability $p_2$. There are two possible conventions for serving:
(i) serves alternate;
(ii) the player serving continues to serve until she loses a point.
You serve first and the first player to reach $n$ points wins the match. Show that your probability of winning the match does not depend on the serving convention adopted.
[*Hint:* Under either convention you serve at most $n$ times and your opponent at most $n - 1$ times.]

**19.** Recall Sarah, collecting figures from cornflake packets (question **12**). How much easier is it to collect half the set than the complete set? Explore, using computer simulation.

**20.** Show that if a binary tree has $n$ leaves whose depths are $d_1, d_2, \ldots, d_n$ then $\sum_{i=1}^{n} 2^{-d_i} \leq 1$. Hence show that $d_1 + d_2 + \cdots + d_n \geq n \log_2 n$ [*Hint:* convexity].

Consider any algorithm for sorting $n$ keys, initially in a random order, by making pairwise comparisons. Obtain a lower bound on the expected number of comparisons.

**21.** What do you think of the following 'proof' by Lewis Carroll that an urn cannot contain two balls of the same colour? Suppose that the urn contains two balls, each of which is either black or white, thus, in the obvious notation $P(BB) = P(BW) = P(WB) = P(WW) = \frac{1}{4}$. We add a black ball, so that $P(BBB) = P(BBW) = P(BWB) = P(BWW) = \frac{1}{4}$. Next we pick a ball at random; the chance that the ball is black is (using conditional probabilities) $1 \cdot \frac{1}{4} + \frac{2}{3} \cdot \frac{1}{4} + \frac{2}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{1}{4} = \frac{2}{3}$. However, if there is probability $\frac{2}{3}$ that a ball, chosen randomly from three, is black, then there must be two black and one white, which is to say that originally there was one black and one white ball in the urn.