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

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?

1A. *Pairs (n, k) such that $np = k$. I expect them to use the definition of independence, $P(E \cap F_k) = P(E)P(F_k)$. Interesting to note that $P(E \mid F_k) = k/n$, whatever p.*

2. The events A and B are independent. Show that the events A^c and B are independent, and that the events A^c and B^c are independent.

2A. *As an extension they might show that if the events $A_1, A_2, \dots, A_n$ are independent then the random variables $I[A_1], I[A_2], \dots, I[A_n]$ are independent.*

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

3A. *Induction, recursion, or $\pi_n = \frac{1}{2}[(1-p+p)^n + (1-p-p)^n]$ and binomial theorem.*

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.

4A. *$p_A/(p_A + p_B - p_A p_B)$. Extension — deduce probability B wins, and check sum of two probabilities is one.*

5. Suppose that X and Y are independent Poisson random variables with parameters λ and μ 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 $\lambda/(\lambda + \mu)$.

5A. *$X + Y \sim \text{Po}(\lambda + \mu)$, by summation (generating functions later). Then conditional probabilities evaluated by noting $P\{X = r, X + Y = n\} = P\{X = r, Y = n - r\} = P\{X = r\}P\{Y = n - r\}$. Hence, after reduction,*

$$P\{X = r \mid X + Y = n\} = \binom{n}{r} \left(\frac{\lambda}{\lambda+\mu}\right)^r \left(\frac{\mu}{\lambda+\mu}\right)^{n-r}.$$

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

6A. *Translation from natural language, plus disjointness and independence.*

$$\begin{aligned} P\{\text{2nd misprint on page } r\} =& P\{\text{1 misprint on pages 1 to } r-1, \text{ and } \geq 1 \text{ misprint on page } r\} \\ &+ P\{\text{0 misprints on pages 1 to } r-1, \text{ and } \geq 2 \text{ misprints on page } r\} \\ =& \lambda(r-1)e^{-\lambda(r-1)}(1 - e^{-\lambda}) + e^{-\lambda(r-1)}(1 - e^{-\lambda} - \lambda e^{-\lambda}). \end{aligned}$$

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

$$S^2 = \sum_{i=1}^n (X_i - \bar{X})^2, \quad \text{where } \bar{X} = \frac{1}{n}\sum_1^n X_i.$$

7A. *Brute force, using properties of expectation and variance. First*

$$E\bar{X} = \mu, \quad \text{var}\bar{X} = \frac{1}{n^2}\sum_{i=1}^n \text{var}X_i = \frac{\sigma^2}{n}.$$

Then, for example,

$$\begin{aligned}
ES^2 &= E\left[\sum_{i=1}^{n} X_i^2 - 2\bar{X}\sum_{i=1}^{n} X_i + n\bar{X}^2\right] \\
&= E\left[\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right] = E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right] \\
&= \sum_{i=1}^{n} EX_i^2 - nE\bar{X}^2 = n(\mu^2+\sigma^2) - n((E\bar{X})^2 + \mathrm{var}\bar{X}) \\
&= n(\mu^2+\sigma^2) - n\left(\mu^2 + \frac{\sigma^2}{n}\right) = (n-1)\sigma^2.
\end{aligned}$$

Common mistakes involve assuming X_i and $\bar{X}$ independent. Interesting to replace X_i by $(X_i-\mu)/\sigma$: this gives lack of dependence on μ, and form of dependence on σ^2. Note that random variable $S^2/(n-1)$ has mean σ^2: it is an 'unbiased estimate' of σ^2.

8. In a sequence of n independent trials the probability of a success at the ith 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.

8A. *Mainly algebra: By independence,* $\mathrm{var}N = \Sigma \mathrm{var}\, X_i = \Sigma p_i(1-p_i) = \dots$.

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 $\mathrm{cov}(X,Y) = 0$, but that X and Y are not independent.

9A. *Explicit calculations — diagram helpful — note failure of converse to result in lectures.*

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.

10A. *Several ways to show independence. For example, ranking of first i yearly rainfalls is a random permutation of $1, 2, \dots, i$. Thus $P\{Y_i = 1\} = \frac{1}{i}$, $P\{Y_i = 0\} = \frac{i-1}{i}$. Let X_i be rank of year i amongst first i years. Then $1 \le X_i \le i$. Further, $(X_1, X_2, \dots, X_n)$ determines and is determined by the permutation $(a_1, a_2, \dots, a_n)$, with each of the $n!$ possibilities equally likely. Thus*

$$P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} = \frac{1}{n!} = \prod_{i=1}^{n} P\{X_i = x_i\}$$

and so $X_1, X_2, \dots, X_n$ are independent. Now $Y_i = I[X_i = 1]$, and so $Y_1, Y_2, \dots, Y_n$ are independent.

$$E\sum_1^n Y_i = \sum_1^n EY_i = \sum_1^n \frac{1}{i}. \qquad \mathrm{var}\left(\sum_1^n Y_i\right) = \sum_1^n \mathrm{var}\, Y_i = \sum_1^n \frac{1}{i}\left(1 - \frac{1}{i}\right).$$

Note that 8, 10, 11 are exercises with indicator functions.

There are many amusing extensions. For example, find the probability the second record year occurs at year i, i.e. $p_i = P\{Y_1 = 1, Y_2 = \ldots = Y_{i-1} = 0, Y_i = 1\}$. Hence, letting $n \to \infty$, find the expected number of years until the second record year occurs, i.e. $\sum_2^\infty (i-1)p_i$. If you were an Olympic athlete, would you find this result demoralizing?

To answer this, recall that $P\{Y_i = 1\} = 1/i$ and that Y_is independent. Thus $p_i = 1 \times \frac{1}{2} \times \frac{2}{3} \times \cdots \times \frac{i-2}{i-1} \times \frac{1}{i} = \frac{1}{(i-1)i}$, and so the requested expectation is infinite.

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?

11A. *Let $X_i = 1$ or 0 according as ith join makes a hoop or not.*

$$E\sum_1^n X_i = \sum_1^n P\{X_i = 1\} = \sum_1^n \frac{1}{2(n-i)+1} = \sum_1^n \frac{1}{2i-1}$$

($\sim \frac{1}{2}\log n$: quite few hoops, it is interesting to note.)

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 $n\sum_{i=1}^{n} i^{-1}$.

12A. *$N = 1 + \sum_{i=1}^{n-1}(Y_i + 1)$, where Y_i geometric with mean $i/(n-i)$. Thus $EN = n\sum_1^n \frac{1}{i} \sim n\log n$. The able might be amused to note that the Y_is are independent, hence $\operatorname{var} N = n^2\sum_1^n \frac{1}{i^2} - n\sum_1^n \frac{1}{i} \sim cn^2$, and so (standard deviation)/mean approaches zero. A similar conclusion can be drawn in questions 10 and 11: question 19 encourages computer experimentation.*

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 $m \ge n$.

13A. For $m \le n$, *$E(S_m/S_n) = E\sum_{k=1}^{m}(X_k/S_n) = \sum_{k=1}^{m} E(X_k/S_n) = mE(X_1/S_n)$. Put $m = n$, to obtain $1 = E(S_n/S_n) = nE(X_1/S_n)$. Hence $E(S_m/S_n) = m/n$. If $m \ge n$,*

$$E(S_m/S_n) = E(1 + \sum_{k=n+1}^{m}(X_k/S_n)) = 1 + \sum_{k=n+1}^{m} E(X_k)E(S_n^{-1}),$$

from independence of S_n and X_k, $k > m$. They may reasonably assume all expectations exist: implied by assumption that $E(X_k), E(X_k^{-1})$ exist.

Problems

Note: These problems should not be tackled at the expense of examples on later sheets.

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.

14A. *First case, indifferent: conditional probability of winning is 1/2, whether you switch or not. Second case, switch: increases probability of winning from 1/3 to 2/3. For the student who finds this easy, consider a further extension. Suppose the host knows where the car is, but may or may not open a door. If he is attempting to minimize the cost of the show, what should be his strategy, and what should be yours?*

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 $P(A \mid B) \neq P(A \mid C)$.

15A. *The answers to this question seem paradoxical if we think of the following game. I pick two cards at random and look at them. First I tell you event B occurs: i.e. the cards include an ace. Then I tell you event C occurs: i.e. the cards include the ace of hearts. Does my second statement really give you any additional information about whether both are aces? Suppose you thought that in making the statement about C occurring I was simply naming at random some ace that was included in the pair of cards.*

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

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

where $s > 1$, and $\zeta(s) = \sum_{n\ge1} n^{-s}$, 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.

16A. $P(A_k) = \sum_j P(X = jp_k) = p_k^{-s}$. Similarly, $P(A_{i_1} \cap \cdots \cap A_{i_n}) = (p_{i_1} \cdots p_{i_n})^{-s}$ if $i_1, i_2, \ldots, i_n$ are distinct, and independence follows.

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.

17A. *What should John do? What should John do if the game involves sequences $n(>3)$ long? A first-year undergraduate answered this question in 1992; for his solution see J.A. Csirik 'Optimal strategy for the first player in the Penney ante game',* Combinatorics, Probability and Computing, **2** (1993).

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.]

18A. *Let $\Omega = \{0,1\}^{2n-1} = \{(a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_{n-1})\}$, where $a_i = 1$ if you win your ith serve, and $b_j = 1$ if you win your opponent's jth serve.*

$$P\{(a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_{n-1})\} = p_1^{\sum_1^n a_i}(1-p_1)^{n-\sum_1^n a_i}(1-p_2)^{\sum_1^{n-1} b_j} p_2^{n-1-\sum_1^{n-1} b_j}$$

under either convention (as in the problem of the points, some serves may follow end of match!) The event {you win} is determined by the outcome $\omega \in \Omega$, as it is simply the event $\{\sum_1^n a_i + \sum_1^{n-1} b_j \geq n\}$, under either convention. Same event, same probability distribution over Ω, and hence result. This is an example of a useful technique known as coupling: using a single sample space and probability distribution over it to model various apparently distinct situations. The random variable [loser's final score], for example, is a function defined on Ω that does *depend on the convention.*

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.

19A. *Recalling* **12**, *the expected number of packets needed to collect half the set is $\sum_0^{n/2-1} \frac{n}{n-i} \sim n/\log 2$, and the complete set $\sim n \log n$. From computer experimentation they may notice small variability about these means: this can be explained by a variance calculation and Chebyshev's inequality.*

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.