MATHEMATICAL TRIPOS: PART IA **Lent 2015**

PROBABILITY **RRW**

Example Sheet 1 (of 4)

Each sheet contains about 12 fairly straightforward 'Exercises', a few more challenging and/or lengthy 'Problems', and also a couple 'Puzzles'. Please work through the Exercises and Problems. The Puzzles are for enthusiasts, and might be fun to talk about in supervision if you have done everything else.

Exercises

1. Four mice are chosen (without replacement) from a litter, two of which are white. The probability that both white mice are chosen is twice the probability that neither is chosen. How many mice are there in the litter?

2. A table-tennis championship for 2^n equally good players is organized as a knock-out tournament with n rounds, the last round being the final. Two players are chosen at random. Calculate probabilities they meet

(i) in the first round,
(ii) in the final,
(iii) in any round.

[Hint: Can the same sample space be used for all three calculations?]

3. A full deck of 52 cards is divided into half at random. Use Stirling's formula to estimate the probability that each half contains the same number of red and black cards.

4. You throw $6n$ dice at random. Show that the probability that each number appears exactly n times is

$$\frac{(6n)!}{(n!)^6}\left(\frac{1}{6}\right)^{6n}.$$

Use Stirling's formula to show that this is approximately $cn^{-5/2}$ for some constant c to be found.

5. The first axiom of probability was stated as "I. $0 \le P(A) \le 1$ for all $A \subseteq \Omega$". Show that equivalent axioms are ones in which this is changed to the weaker requirement: "I. $P(A) \ge 0$ for all $A \subseteq \Omega$".

6. (i) If A, B, C are three events, show that

$$P(A^c \cap (B \cup C)) = P(B) + P(C) - P(B \cap C) - P(C \cap A) - P(A \cap B) + P(A \cap B \cap C).$$

(ii) How many of the numbers $1, \dots, 500$ are not divisible by 7 but are divisible by 3 or 5?

7. A coin is tossed repeatedly. The outcomes of the tosses are independent. Let A_k be the event that the kth toss is a head, $k = 1, 2, \dots$.

(i) Explain in words what happens when the event $B = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c$ occurs.
(ii) Express, similarly as in (i), the event C, that 'an infinite number of heads occur'.
(iii) What do you think are the probabilities of B and C when $P(A_k) = p$ for all k and $0 < p < 1$? Can you prove this?

(iv) Use the properties of P as set out in Lecture 4 to give a rigorous calculation of $P(B)$ when the problem is changed so that tosses are independent, but are made with coins of different biases, such that $P(A_k) = p_k$ and $\sum_k p_k < \infty$.

8. Let $A_1, \dots, A_m \subseteq \{1, \dots, n\}$ be finite sets with $A_i \not\subset A_j$ and let $a_i = |A_i|$. Let $\sigma_1, \dots, \sigma_n$ be a randomly chosen permutation of $1, 2, \dots, n$ and let E_i be the event $\{\sigma_1, \dots, \sigma_{a_i}\} = A_i$. Are E_1 and E_2 disjoint? Are E_1 and E_2 independent?

Prove that

$$\sum_{i=1}^{m} \binom{n}{a_i}^{-1} \leq 1.$$

9. A committee of size r is chosen at random from a set of n people. Calculate the probability that m given people will all be on the committee (a) directly, and (b) using the inclusion-exclusion formula. Deduce that

$$\binom{n-m}{r-m} = \sum_{j=0}^{m} (-1)^j \binom{m}{j} \binom{n-j}{r}.$$

10. Let $A_1, \dots, A_n$ be events. Prove the following improvement of Boole's inequality.

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} P(A_i \cap A_{i+1}).$$

[Hint. Induction.] Deduce that

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) - \frac{2}{n} \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2}).$$

11. Examination candidates are graded into four classes known conventionally as I, II–1, II–2 and III, with probabilities 1/8, 2/8, 3/8 and 2/8 respectively. A candidate who misreads the rubric, — a common event with probability 2/3 —, generally does worse, his or her probabilities being 1/10, 2/10, 4/10 and 3/10. What is the probability:

(i) that a candidate who reads the rubric correctly is placed in the class II–1?
(ii) that a candidate who is placed in the class II–1 has read the rubric correctly?

12. Parliament contains a proportion p of Labour members, who are incapable of changing their minds about anything, and a proportion $1 - p$ of Conservative members who change their minds completely at random (with probability r) between successive votes on the same issue. A randomly chosen member is noticed to have voted twice in succession in the same way. What is the probability that this member will vote in the same way next time?

Problems

Some of these are more challenging. I hope you will learn and have fun by attempting them.

13. Mary tosses two coins and John tosses one coin. What is the probability that Mary gets more heads than John? Answer the same question if Mary tosses three coins and John tosses two. Make a conjecture for the probability when Mary tosses $n + 1$ and John tosses n. Can you prove your conjecture?

14. [*This simply stated problem requires thought in modelling the sample space.*] Suppose that n balls are tossed independently and at random into n boxes. What is the probability that exactly one box is empty? Check your formula gives the correct answer for $n = 2$ and $n = 3$ (for which the probabilities are 1/2 and 2/3, respectively).

[Hint. In a problem like this it is up to you to decide whether you will imagine that the balls and/or boxes are (or are not) distinguishable. You may need to experiment until you find the model that best enables you to calculate the desired probability.]

15. The Polya urn model for contagion is as follows. We start with an urn which contains one white ball and one black ball. At each second we choose a ball at random from the urn and replace it together with one more ball of the same colour. Calculate the probability that when n balls are in the urn, i of them are white. (You might like to carry out a computer simulation — do you think the proportion of white balls might tend to a limit?)

16. Suppose a die is rolled n times. Show that the probability of a roll of i appearing k_i times, $i = 1, \dots, 6$, is

$$\phi(k_1, \dots, k_6) = \frac{n!}{k_1! \cdots k_6!} \frac{1}{6^n}.$$

The expected value of the total of n rolls is $3.5n$. Suppose $\sum_i ik_i = \rho n$, $1 \le \rho \le 6$, and $k_i = np_i$. Use Stirling's formula to show that subject to these constraints, ϕ is maximized by choosing the p_i as nonnegative numbers that solve the optimization problem

$$\underset{p_1,\dots,p_6}{\text{maximize}} - \textstyle\sum_i p_i \log p_i, \quad \text{subject to } \sum_i p_i = 1, \text{ and } \sum_i ip_i = \rho.$$

For $\rho = 4$ the maximizer is $p^* \approx (0.103, 0.123, 0.146, 0.174, 0.207, 0.247)$, so the most likely way of obtaining a total of $4n$ is when about 24.7% of the dice rolls are a 6. What do you guess would be the solution to the optimization problem if $\rho = 3.5$? If $\rho = 3$?

Puzzles

These are for enthusiasts, or to discuss in supervision when you have done everything else.

17. I am playing poker with three friends and from a well-shuffled deck we have each been dealt five cards. I have a hand consisting of the four kings and the two of hearts. Being a poker wizard I know exactly the probability that I have a winning hand.

But then I suddenly discover that earlier in the day the family's children were playing with the cards and fed six of them to their pet goat (but I don't know which). How does this information change the probability with which I believe my hand will beat my opponents? Why?

18. Suppose that there are 42 bags, labeled 0 though 41. Bag i contains i red balls and $41 - i$ blue balls. Suppose that you pick a bag at random, then pull out three balls without replacement. What is the probability that all 3 balls are the same colour?

MATHEMATICAL TRIPOS: PART IA **Lent 2015**

PROBABILITY **RRW**

Example Sheet 2 (of 4)

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^C and B are independent, and that the events A^C and B^C are independent.

3. Independent trials are performed, each with probability p of success. Let P_n be the probability that n trials result in an even number of successes. Show that

$$P_n = \tfrac{1}{2}[1 + (1 - 2p)^n].$$

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 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 $\lambda/(\lambda + \mu)$.

6. (i) 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?

(ii) A proofreader studies a single page looking for misprints. She catches each misprint (independently of others) with probability $1/2$. Let X be the number of misprints she catches. Find $P(X = k)$. Given that she has found $X = 10$ misprints, what is the distribution of Y, the number of misprints she has not caught? How useful is X in predicting Y?

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

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

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.

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

10. Let $a_1, a_2, \ldots, a_n$ be a ranking of the yearly rainfalls in Cambridge over the next n years: assume $a_1, a_2, \ldots, a_n$ is a random permutation of $1, 2, \ldots, n$. 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 $Y_1, Y_2, \ldots, Y_n$ are independent. Calculate the mean and variance of the number of record years in the next n years.

11. Liam's bowl of spaghetti contains n strands. He selects two ends at random and joins them together. He repeats 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 of n distinct figures. Each type of figure is equally likely. Show that the expected number of packets Sarah needs to buy to collect a complete set of n is

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

[*After doing this, you might like to visit the Wikipedia article about the 'Coupon collector's problem'.*]

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 $m \leq n$, and $E(S_m/S_n) = 1 + (m-n)aE(S_n^{-1})$ if $m \geq n$. [*This was a Cambridge entrance exam question c. 1970.*]

Problems

These next questions are more challenging. I hope you will learn and have fun by attempting them.

14. You wish to use a fair coin to simulate occurrence or not of an event A that happens with probability 1/3. One method is to start by tossing the coin twice. If you see HH say that A occurred, if you see HT or TH say that A has not occurred, and if you see TT then repeat the process. Show that this enables you to simulate the event using an expected number of tosses equal to 8/3.

Can you do better? (i.e. simulate something that happens with probability 1/3 using a fair coin and with a smaller expected number of tosses) Hint. The binary expansion of 1/3 is 0.0101010101 . . . .]

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

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

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

$$\prod_{k=1}^{\infty} (1 - p_k^{-s}) = 1/\zeta(s).$$

16. 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. Recall Pascal and Fermat's 'problem of the points', treated in lectures.]

Puzzles

This section is for enthusiasts — or for discussion in supervision when you have done everything else. The following puzzles have been communicated to me by Peter Winkler.

17. Let $k < n$, k even, n odd. Joey is to play n chess games against his parents, alternating between his father and mother. To receive his allowance he must win k games in a row. Prove that given the choice, he should start against the stronger parent.

[Hint: start by solving the cases $k = 2, n = 3$, and $k = n - 1$.]

18. Let $X_1, \dots, X_6$ be i.i.d. $B(1, p)$ with $p = 0.4$. Let $S_n = X_1 + \cdots + X_n$. Argue that

$$P(S_4 \geq 3) = P(S_6 \geq 4)$$

without explicitly computing either the left or right hand sides.

[Hint. Compare $P(S_4 \geq 3 \mid S_5 = i)$ and $P(S_6 \geq 4 \mid S_5 = i)$ for each of $i = 0, \dots, 5$.]

MATHEMATICAL TRIPOS: PART IA — Lent 2015

PROBABILITY — RRW

Example Sheet 3 (of 4)

Exercises

1. Let $x_1, x_2, \dots, x_n$ be positive real numbers. Then the geometric mean lies between the harmonic mean and the arithmetic mean:

$$\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}\right)^{-1} \le \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \le \frac{1}{n}\sum_{i=1}^{n} x_i.$$

The second inequality is the AM–GM inequality: establish the first inequality.

2. Let X be a positive random variable taking only finitely many values. Show that

$$E\left(\frac{1}{X}\right) \ge \frac{1}{EX}$$

and that the inequality is strict unless $P\{X = EX\} = 1$.

3. Let X be a random variable for which $EX = \mu$ and $E(X-\mu)^4 = \beta_4$. Prove that for $t > 0$,

$$P\{|X-\mu| \ge t\} \le \frac{\beta_4}{t^4}\,.$$

4. Consider a random sample taken from a distribution. Use Chebyshev's inequality to determine a sample size that will be sufficient, whatever the distribution, for the probability to be at least 0.99 that the sample mean will be within two standard deviations of the mean of the distribution.

5. In a sequence of Bernoulli trials, X is the number of trials up to and including the a^{th} success. Show that

$$P\{X = r\} = \binom{r-1}{a-1} p^a q^{r-a}, \quad r = a, a+1, \dots\,.$$

Verify that the probability generating function for this distribution is $p^a t^a (1-qt)^{-a}$. Show that $EX = a/p$ and $\mathrm{Var}(X) = aq/p^2$. Show how X can be represented as the sum of a independent random variables, all with the same distribution. Use this representation to derive again the mean and variance of X.

6. For a random variable X with mean μ and variance σ^2 define the function

$$V(x) = E(X-x)^2\,.$$

Express the random variable $V(X)$ in terms of μ, σ^2 and X, and hence show that

$$\sigma^2 = \tfrac{1}{2}E(V(X))\,.$$

7. Suppose X is a real-valued random variable and $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are two nondecreasing functions. Prove the 'Chebyshev order inequality':

$$E[f(X)]E[g(X)] \leq E[f(X)g(X)].$$

[Hint. Consider $[f(X_1) - f(X_2)][g(X_1) - g(X_2)]$ where X_1 and X_2 are i.i.d.]

8. Let N be a non-negative integer-valued random variable with mean μ_1 and variance σ_1^2, and let $X_1, X_2, \ldots$ be identically distributed random variables, each with mean μ_2 and variance σ_2^2; furthermore, assume that $N, X_1, X_2, \ldots$ are independent. Calculate the mean and variance of the random variable $S_N = X_1 + \cdots + X_N$.

9. At time 0, a blood culture starts with one red cell. At the end of one minute, the red cell dies and is replaced by one of the following combinations with probabilities as indicated:

$$2 \text{ red cells } \tfrac{1}{4}; \quad 1 \text{ red, } 1 \text{ white } \tfrac{2}{3}; \quad 2 \text{ white } \tfrac{1}{12}.$$

Each red cell lives for one minute and gives birth to offspring in the same way as the parent cell. Each white cell lives for one minute and dies without reproducing. Assume the individual cells behave independently.

(a) At time $n + \frac{1}{2}$ minutes after the culture began, what is the probability that no white cells have yet appeared?
(b) What is the probability that the entire culture dies out eventually?

10. (a) A mature individual produces offspring according to the probability-generating function $F(s)$. Suppose we start with a population of k immature individuals, each of which grows to maturity with probability p and then reproduces, independently of the other individuals. Find the probability generating function of the number of (immature) individuals at the next generation.

(b) Find the probability generating function of the number of mature individuals at the next generation, given that there are k mature individuals in the parent generation.

Show that the distributions in (a) and (b) have the same mean, but not necessarily the same variance.

11. A slot machine operates so that at the first turn the probability for the player to win is $\frac{1}{2}$. Thereafter the probability for the player to win is $\frac{1}{2}$ if he lost at the last turn, but is $p(< \frac{1}{2})$ if he won at the last turn. If u_n is the probability that the player wins at the n^{th} turn, show that, provided $n > 1$,

$$u_n + (\tfrac{1}{2} - p)u_{n-1} = \tfrac{1}{2}\,.$$

Observe that this equation also holds for $n = 1$, if u_0 is suitably defined. Solve the equation, showing that

$$u_n = \frac{1 + (-1)^{n-1}(\frac{1}{2} - p)^n}{3 - 2p}\,.$$

12. A fair coin is tossed n times. Let U_n be the probability that the sequence of tosses never has 'head' followed by 'head'. Show that

$$U_n = \tfrac{1}{2}U_{n-1} + \tfrac{1}{4}U_{n-2}\,.$$

Find U_n, using the condition $U_0 = U_1 = 1$. Check that the value for U_2 is correct.

Problems

Some of these are more challenging. I hope you will learn and have fun by attempting them.

13. Let $b_1, b_2, \ldots, b_n$ be a rearrangement of the positive real numbers $a_1, a_2, \ldots, a_n$. Prove that

$$\sum_{i=1}^{n} \frac{a_i}{b_i} \geq n \, .$$

14. Let $F(s) = 1 - p(1-s)^\beta$, where p and β are constants and $0 < p < 1$, $0 < \beta < 1$. Prove that $F(s)$ is a probability generating function and that its iterates are

$$F_n(s) = 1 - p^{1+\beta+\cdots+\beta^{n-1}}(1-s)^{\beta^n} \quad \text{for } n = 1, 2, \ldots \, .$$

Find the mean m of the associated distribution and the extinction probability, $q = \lim_{n\to\infty} F_n(0)$, for a branching process with offspring distribution determined by F.

15. Let $(X_n)_{n\geq 0}$ be a branching process such that $X_0 = 1$, $EX_1 \equiv \mu$. If $Y_n = X_0 + X_1 + \cdots + X_n$, and for $0 \leq s \leq 1$

$$\Psi_n(s) \equiv E s^{Y_n} \, ,$$

prove that

$$\Psi_{n+1}(s) = s\phi(\Psi_n(s)) \, ,$$

where $\phi(s) \equiv E s^{X_1}$. Deduce that, if $Y = \sum_{n\geq 0} X_n$, then $\Psi(s) \equiv E s^Y$ satisfies

$$\Psi(s) = s\phi(\Psi(s)), \qquad 0 \leq s \leq 1 \, ,$$

where $s^\infty \equiv 0$. If $\mu < 1$, prove that $EY = (1-\mu)^{-1}$.

16. A particle moves at each step two units in the positive direction, with probability p, or one unit in the negative direction, with probability $q = 1 - p$. If the starting position is $z > 0$, find the probability a_z that the particle will ever reach the origin. Deduce that if a fair coin is tossed repeatedly the probability that the number of heads ever exceeds twice the number of tails is $(\sqrt{5} - 1)/2$.

17. Let (X_k) be a sequence of independent, identically distributed random variables, each with mean μ and variance σ^2. Show that

$$\sum_{k=1}^{n} (X_k - \bar{X})^2 = \sum_{k=1}^{n} (X_k - \mu)^2 - n(\bar{X} - \mu)^2 \, ,$$

where $\bar{X} = \frac{1}{n} \sum_1^n X_k$. Prove that, if $E(X_1 - \mu)^4 < \infty$, then for every $\epsilon > 0$

$$P\left\{ \left| \frac{1}{n} \sum_{k=1}^{n} (X_k - \bar{X})^2 - \sigma^2 \right| > \epsilon \right\} \to 0$$

as $n \to \infty$.

Puzzles

These are for enthusiasts, or to discuss in supervision when you have done everything else.

18. (a) Show that it is impossible to load a die so that the sum of two rolls of this die will take all values $\{2, 3, \dots, 12\}$ with equal probability.

(b) Could you load the die so that the totals $\{2, 3, 4, 5, 6, 7\}$ are obtained with equal probabilities?

(c) Can you construct a distribution whose support is the nonnegative integers and is such that if X_1 and X_2 are independent r.v.s with this distribution then $X_1 + X_2$ has a geometric distribution with parameter p, i.e. $P(X_1 + X_2 = i) = p(1-p)^i$. $i = 0, 1, \dots$?

[Hint. $(1-x)^{-1/2} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 + \frac{63}{256}x^5 + \frac{231}{1024}x^6 + \cdots$.]

19. £X is placed in an envelope, according to the probability distribution $P(X = 2^n) = (1/3)(2/3)^n$, $n = 0, 1, 2, \dots$. In a second identical envelope is placed £Y, where $Y = 2X$. You select an envelope at random and open it. Let E_i be the event you find it contains £2^i, where $i > 0$. You now know that either $(X, Y) = (2^i, 2^{i+1})$ or $(X, Y) = (2^{i-1}, 2^i)$.

Show that $P((X, Y) = (2^i, 2^{i+1}) \mid E_i) = 2/5$.

Show that, conditional on E_i, the expected amount of money in the unopened envelope is greater than the amount of money in the opened envelope. Is this surprising?

Having made these calculations, you may like to look at
`http://en.wikipedia.org/wiki/Two_envelopes_problem` and
`http://en.wikipedia.org/wiki/St._Petersburg_paradox`.

MATHEMATICAL TRIPOS: PART IA **Lent 2015**

PROBABILITY **RRW**

Example Sheet 4 (of 4)

Exercises

1. Alice and Bob agree to meet in the Copper Kettle after their Saturday lectures. They arrive at times that are independent and uniformly distributed between 12.00 and 1.00 pm. Each is prepared to wait 10 minutes before leaving. Find the probability they meet.

2. A stick is broken in two places, independently uniformly distributed along its length. What is the probability that the three pieces will make a triangle?

3. The radius of a circle is exponentially distributed with parameter λ. Determine the probability density function of the area of the circle.

4. The random variables X and Y are independent and exponentially distributed with parameters λ and μ respectively. Find the distribution of $\min\{X, Y\}$, and the probability that X exceeds Y.

5. How large a random sample should be taken from a normal distribution in order for the probability to be at least 0.99 that the sample mean will be within one standard deviation of the mean of the distribution? Hint. $\Phi(2.58) = 0.995$.

6. The random variable X has a *log-normal distribution* if $Y = \log X$ is normally distributed. If $Y \sim N(\mu, \sigma^2)$, calculate the mean and variance of X. (The log-normal distribution is sometimes used to represent the size of small particles after a crushing process, or as a model for future commodity prices. Why?)

7. X and Y are independent random variables, each distributed normally, as $N(0, 1)$. Show that, for any fixed θ, the random variables

$$U = X\cos\theta + Y\sin\theta \qquad V = -X\sin\theta + Y\cos\theta$$

are independent and find their distributions.

8. The random variables X and Y are independent and exponentially distributed, each with parameter λ. Show that the random variables $X+Y$ and $X/(X+Y)$ are independent and find their distributions.

9. A shot is fired at a circular target. The vertical and horizontal coordinates of the point of impact (taking the centre of the target as origin) are independent random variables, each distributed normally $N(0, 1)$.

(i) Show that the distance of the point of impact from the centre has p.d.f. $re^{-r^2/2}$ for $r \geq 0$.
(ii) Show that the mean of this distance is $\sqrt{\pi/2}$, the median is $\sqrt{\log 4}$, and the mode is 1.

10. A radioactive source emits particles in a random direction (with all directions being equally likely). It is held at a distance d from a vertical infinite plane photographic plate.

(i) Show that, given the particle hits the plate, the horizontal coordinate of its point of impact (with the point nearest the source as origin) has p.d.f. $d/\pi(d^2 + x^2)$. (This distribution is known as the Cauchy distribution).

(ii) Can you compute the mean of this distribution?

11. A random sample is taken in order to find the proportion of Labour voters in a population. Find a sample size such that the probability of a sampling error less than 0.04 will be 0.99 or greater.

12. The random variables $Y_1, Y_2, \ldots, Y_n$ are independent, with $\mathbb{E}Y_i = \mu_i$, $\mathrm{Var}(Y_i) = \sigma^2$, $1 \le i \le n$. For constants $a_i, b_i, 1 \le i \le n$, show that

$$\mathrm{cov}\Big(\sum_i a_i Y_i, \sum_i b_i Y_i\Big) = \sigma^2 \sum_i a_i b_i.$$

Prove that if $Y_1, Y_2, \ldots, Y_n$ are independent normal random variables, then $\sum_i a_i Y_i$ and $\sum_i b_i Y_i$ are independent if and only if $\sum_i a_i b_i = 0$.

Problems

Some of these are more challenging. I hope you will learn and have fun by attempting them.

13. Show that the mean and variance of the number shown upon rolling a fair die are 7/2 and 35/12 respectively. Use the central limit theorem to estimate the probability q that the total of 10 rolls of a die is at least 45. (answer: $1 - \Phi(2\sqrt{6/7}) = 0.0320$).

Find q exactly. Hint: use *Mathematica* to compute the p.g.f. and find your answers thereby. Try

```
p[z_] = (1/6) (z + z^2 + z^3 + z^4 + z^5 + z^6)
mean = p'[z] /. z -> 1
var = p''[z] + p'[z] - p'[z]^2 /. z -> 1
q = Sum[SeriesCoefficient[p[z]^10, {z, 0, i}], {i, 45, 60}]
```

14. You wish to determine π by repeatedly dropping a straight pin of length ℓ ($< L$) onto a floor marked with parallel lines spaced L apart. Estimate how closely you could determine the value of π by devoting 50 years to full-time pin dropping. What pin length, ℓ, would you prefer?

15. A random sample of size $2n + 1$ is taken from the uniform distribution on $[0, 1]$. Find the distribution of the sample median.

16. Suppose that n items are being tested simultaneously and that the items have independent lifetimes, each exponentially distributed with parameter λ. Determine the mean and variance of the length of time until r items have failed.

17. (i) X and Y are independent random variables, with continuous symmetric distributions, with p.d.f.s f and g respectively. Show that the p.d.f. of $Z = X/Y$ is

$$h(a) = 2 \int_0^\infty x f(ax) g(x) dx\,.$$

(ii) X and Y are independent random variables distributed $N(0, \sigma^2)$ and $N(0, \tau^2)$. Show that X/Y has p.d.f. $f(x) = d/\pi(d^2 + x^2)$, where $d = \sigma/\tau$.

18. Derive the distribution of the sum of n independent random variables each having the Poisson distribution with parameter 1. Use the central limit theorem to prove that

$$e^{-n}\left(1+\frac{n}{1!}+\frac{n^2}{2!}+\cdots+\frac{n^n}{n!}\right)\to 1/2 \quad \text{as } n\to\infty\,.$$

19. If X, Y and Z are independent random variables each uniformly distributed on $(0,1)$, show that $(XY)^Z$ is also uniformly distributed on $(0,1)$.

[Hint. What is the distribution of $-\log X$? How might Exercise 8 help?]

Puzzle

This is for enthusiasts — or for discussion in supervision when you have done everything else.

20. Imagine there are a 100 people in line to board a plane that seats 100. The first person in line realizes he lost his boarding pass so when he boards he decides to take a random seat instead. Every person that boards the plane after him will either take their “proper” seat, or if that seat is taken, a random seat instead.

Assuming the order of boarding is random, what is the probability that the last person that boards will end up in his/her proper seat?

(You might begin by finding the probability when 100 is replaced by 2 or 3. This may help you to guess the answer.)