UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Mathematical Methods for Computer Science

Computer Science Tripos, Part IB
Michaelmas Term 2014/15

R.J. Gibbens

Problem sheets
for
Probability methods

William Gates Building
15 JJ Thomson Avenue
Cambridge
CB3 0FD

http://www.cl.cam.ac.uk/teaching/1415/MathMforCS/

Problem sheet #1

1. Suppose that $X_1, X_2, \ldots, X_n$ is a sequence of independent and identically distributed random variables with $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$. Define the *sample mean*, $\overline{X}_n$, by

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \, .$$

Show that $\mathbb{E}(\overline{X}_n) = \mu$ and $\mathrm{Var}(\overline{X}_n) = \sigma^2/n$. Define the *sample variance*, $\overline{S}_n$, by

$$\overline{S}_n = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 \, .$$

Show that $\mathbb{E}(\overline{S}_n) = \sigma^2$. [*You don't need to use PGFs for this question but it makes for good revision of first year probability.*]

2. Let X be a random variable having a geometric distribution with parameter p and let $q = 1 - p$. Show that for $|z| < 1/q$, X has a probability generating function given by $G_X(z) = pz/(1 - qz)$. Using the probability generating function $G_X(z)$ calculate the mean and variance of X.

3. Suppose that X and Y are independent Poisson random variables with parameters λ_1 and λ_2, respectively.
   (a) Show that $X + Y \sim \mathrm{Pois}(\lambda_1 + \lambda_2)$.
   (b) Find the probability distribution of X conditional on the event that $X + Y = n$ where $n = 0, 1, 2, \ldots$ is a fixed non-negative integer.

4. Suppose that X is a continuous random variable with the $U(-1, 1)$ distribution. Find the exact value of $\mathbb{P}(|X| \geq a)$ for each $a > 0$ and compare it to the upper bounds obtained from the Markov and Chebychev inequalities.

5. Let X be the random variable giving the number of heads obtained in a sequence of n independent fair coin flips. Compare the upper bounds on $\mathbb{P}(X \geq 3n/4)$ obtained from the Markov and Chebychev inequalities.

6. Suppose that α, μ and δ are real constants such that $0 \leq \alpha \leq 1$ and $\delta > 0$. Consider the random variable X with probability mass function

$$\mathbb{P}(X = x) = \begin{cases} \alpha & \text{if } x = \mu + \delta \\ 1 - 2\alpha & \text{if } x = \mu \\ \alpha & \text{if } x = \mu - \delta \, . \end{cases}$$

Show that $\mathbb{E}(X) = \mu$ and find an expression for $\mathrm{Var}(X)$ in terms of α and δ. Hence show that

$$\mathbb{P}(|X - \mu| \geq \delta) = \frac{\mathrm{Var}(X)}{\delta^2} \, .$$

Note that this result shows that the upper bound in Chebychev's inequality can not be improved in general. Can you construct an example to show the tightness of Markov's inequality?

7. Let A_i $(i = 1, 2, \ldots, n)$ be a collection of random events and set $N = \sum_{i=1}^{n} \mathbb{I}(A_i)$. By considering Markov's inequality applied to $\mathbb{P}(N \geq 1)$ show Boole's inequality, namely,

$$\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)\,.$$

Problem sheet #2

1. Let $h : \mathbb{R} \to [0, \infty)$ be a non-negative function. Show that

$$\mathbb{P}(h(X) \geq a) \leq \frac{\mathbb{E}(h(X))}{a} \qquad \text{for all} \qquad a > 0\,.$$

By making suitable choices of $h(x)$, show that we may obtain the Markov and Chebychev inequalities as special cases.

2. Show the following properties of the moment generating function.

   (a) If X has mgf $M_X(t)$ then $Y = aX + b$ has mgf $M_Y(t) = e^{bt} M_X(at)$.

   (b) If X and Y are independent then $X + Y$ has mgf $M_{X+Y}(t) = M_X(t) M_Y(t)$.

   (c) $\mathbb{E}(X^n) = M_X^{(n)}(0)$ where $M_X^{(n)}$ is the n^{th} derivative of M_X.

   (d) If X is a discrete random variable taking values $0, 1, 2, \ldots$ with probability generating function $G_X(z) = \mathbb{E}(z^X)$ then $M_X(t) = G_X(e^t)$.

3. Let X be a random variable with moment generating function $M_X(t)$ which you may assume exists for any real value of t. Show that for all $a > 0$

$$\mathbb{P}(X \leq a) \leq e^{-ta} M_X(t) \qquad \text{for all} \qquad t < 0\,.$$

4. Show that, if $X_n \xrightarrow{D} X$, where X is a degenerate random variable (that is, $\mathbb{P}(X = \mu) = 1$ for some constant μ) then $X_n \xrightarrow{P} X$.

5. Suppose that you estimate your monthly phone bill by rounding all amounts to the nearest pound. If all rounding errors are independent and identically distributed as $U(-0.5, 0.5)$, estimate the probability that the total error exceeds one pound when your bill has 12 items. How might this procedure suggest an approximate method for constructing Normal random variables?

6. Consider a sequence of independent identically distributed random variables $Y_1, Y_2, \ldots$ with $\mathbb{P}(Y_i = 1) = p$ and $\mathbb{P}(Y_i = -1) = 1 - p$ with $p \in [0, 1]$. Define the *simple random walk* X_n by

$$X_n = X_0 + Y_1 + Y_2 + \cdots + Y_n$$

where $X_0 \in \mathbb{R}$.

(a) Find $\mathbb{E}(X_n)$ and $\mathrm{Var}(X_n)$ when $X_0 = 0$.

(b) Find $\mathbb{P}(X_n = n + k)$ when $X_0 = k$.

7. (a) Consider the Gambler's ruin problem studied in lectures and construct both $\mathbb{P}(\text{A is ruined})$ and $\mathbb{P}(\text{B is ruined})$. What is $\mathbb{P}(\text{A is ruined}) + \mathbb{P}(\text{B is ruined})$?

   (b) Check that the solution given in lectures for the expected duration of the Gambler's ruin problem solves the stated difference equation.

8. Using the model for Bitcoin given in the lectures consider the attacker's success probability

$$s(q, n) = \sum_{k=0}^{\infty} \mathbb{P}(X = k) \begin{cases} \left(\frac{q}{p}\right)^{n-k} & k \leq n \\ 1 & k > n \end{cases}$$

where X is the random number of blocks constructed by the attacker in the time taken for the honest block chain to grow by n blocks. In the case that $X \sim \mathrm{Pois}(\lambda)$ show that $s(q, n)$ can be written in the computationally more convenient form as a finite sum

$$s(q, n) = 1 - \sum_{k=0}^{n} \frac{\lambda^k e^{-\lambda}}{k!} \left(1 - \left(\frac{q}{p}\right)^{n-k}\right).$$

You may like to use the `C` program given in the Bitcoin white paper (referenced in the lectures) or a program of your own to investigate the function $s(q, n)$. Consider plotting graphs of $s(q, n)$ for $q = 0.1, 0.2, 0.3$ and for $n = 0, 1, 2, \ldots, 50$. How would you explain the choice of $n = 6$ as implemented in the Bitcoin mechanism?

Problem sheet #3

1. Suppose that (X_n) is a Markov chain with n-step transition matrix, $P^{(n)}$, and let $\lambda_i^{(n)} = \mathbb{P}(X_n = i)$ be the elements of a row vector $\lambda^{(n)}$ $(n = 0, 1, 2, \ldots)$. Show that

   (a) $P^{(m+n)} = P^{(m)} P^{(n)}$ for $m, n = 0, 1, 2, \ldots$

   (b) $\lambda^{(n)} = \lambda^{(0)} P^{(n)}$ for $n = 0, 1, 2, \ldots$.

2. Suppose that (X_n) is a Markov chain with transition matrix P. Define the relations *"state j is accessible from state i"* and *"states i and j communicate"*. Show that the second relation is an equivalence relation and define the communicating classes as the equivalence classes under this relation. What is meant by the terms *closed class*, *absorbing class* and *irreducible*?

3. Show that

$$P_{ij}(z) = \delta_{ij} + F_{ij}(z) P_{jj}(z)$$

where

$$P_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}^{(n)} z^n, \qquad F_{ij}(z) = \sum_{n=0}^{\infty} f_{ij}^{(n)} z^n$$

and $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. [You should assume that $p_{ij}^{(n)}$ and $f_{ij}^{(n)}$ are as defined in lectures with $p_{ij}^{(0)} = \delta_{ij}$ and $f_{ij}^{(0)} = 0$ for all states i, j.]

4. Suppose that (X_n) is a finite state Markov chain and that for some state i and for all states j
$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$$
for some collection of numbers (π_j). Show that $\pi = (\pi_j)$ is a stationary distribution.

5. Consider the Markov chain with transition matrix
$$P = \begin{pmatrix} 0.128 & 0.872 \\ 0.663 & 0.337 \end{pmatrix}$$
for Markov's example of a chain on the two states $\{\text{vowel}, \text{consonant}\}$ for consecutive letters in a passage of text. Find the stationary distribution for this Markov chain. What are the mean recurrence times for the two states?

6. Define what is meant by saying that (X_n) is a reversible Markov chain and write down the local balance conditions. Show that if a vector π is a distribution over the states of the Markov chain that satisfies the local balance conditions then it is a stationary distribution.

7. Consider the Erhenfest model for m balls moving between two containers with transition matrix
$$p_{i,i+1} = 1 - \frac{i}{m}, \qquad p_{i,i-1} = \frac{i}{m}$$
where i $(0 \le i \le m)$ is the number of balls in a given container. Show that the Markov chain is irreducible and periodic with period 2. Derive the stationary distribution.

8. Consider a random walk, (X_n), on the states $i = 0, 1, 2, \ldots$ with transition matrix
$$\begin{aligned} p_{i,i-1} &= p \qquad i = 1, 2, \ldots \\ p_{i,i+1} &= 1 - p \qquad i = 0, 1, \ldots \\ p_{0,0} &= p \end{aligned}$$
where $0 < p < 1$. Show that the Markov chain is irreducible and aperiodic. Find a condition on p to make the Markov chain positive recurrent and find the stationary distribution in this case.

9. Describe PageRank as a Markov chain model for the motion between nodes in a graph. Explain the main mathematical results that underpin PageRank's connection to a notion of web page *"importance"*.

Revision questions

1. 2006 Paper 3 Question 10
2. 2007 Paper 4 Question 5
3. 2008 Paper 4 Question 4
4. 2009 Paper 2 Question 8 (Probability IA)
5. 2010 Paper 6 Question 8
6. 2011 Paper 2 Question 7 (Probability IA)
7. 2011 Paper 6 Question 7
8. 2012 Paper 2 Question 7 (Probability IA)
9. 2012 Paper 6 Question 8
10. 2013 Paper 2 Question 7 (Probability IA)
11. 2013 Paper 6 Question 8
12. 2014 Paper 6 Question 8