UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Friday 8 June 2001 1.30 to 3.30

PAPER 26

LARGE DEVIATIONS AND QUEUEING THEORY

*Attempt **THREE** questions. The questions carry equal weight.*

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 (a) State Cramér's Theorem.

Let W be an exponential random variable with mean λ^{-1}, and let X_n be the average of n independent copies of W. Prove that X_n satisfies a large deviations principle with good rate function

$$I(x) = \begin{cases} \lambda x - 1 - \log(\lambda x) & \text{for } x > 0 \\ \infty & \text{for } x \leqslant 0. \end{cases}$$

(You may use the following fact: $\mathbb{E}e^{\theta W}$ is equal to $\lambda/(\lambda - \theta)$ for $\theta < \lambda$, and equal to ∞ otherwise.)

(b) Let $Y_n(1), \ldots, Y_n(k)$ be independent copies of X_n defined above. Let M_n be the minimum of $Y_n(1), \ldots, Y_n(k)$. Prove that M_n satisfies a large deviations principle with good rate function

$$J(m) = \begin{cases} kI(m) & \text{for } m \geqslant \lambda^{-1} \\ I(m) & \text{for } m < \lambda^{-1}. \end{cases}$$

(c) State Varadhan's Integral Lemma.

With M_n as defined above, let $Z_n = \min(b, \max(a, M_n))$ for some $0 < a < \lambda^{-1} < b$. Prove that, for sufficiently large k,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}(Z_n)^n = (k+1)\log\Big(\frac{k+1}{k}\Big) - \log\lambda - 1.$$

2 (a) What does it mean to say that a sequence of random variables X^L satisfies a large deviations principle with rate function I?

(b) State and prove the contraction principle.

(c) What does it mean to say that the sequence of random variables X^L is exponentially tight?

(d) Recall that the sequence of random variables X^L is said to satisfy a weak large deviations principle if the large deviations upper bound is required to hold only for compact sets. Suppose that the sequence X^L is exponentially tight, and satisfies a weak large deviations principle with rate function I. Show that it satisfies a large deviations principle with good rate function I.

(You may use the following result without proof: Let $\mathcal{X}$ be a topological space, and let $f : \mathcal{X} \to \mathbb{R}$ have compact level sets. Then f attains its infimum in any closed set.)

3 (a) Let X be a Poisson random variable with mean λ. Let $X^{\oplus L}$ be the sum of L independent copies of X. Prove that for all $0 < \beta < 1$, $X^{\oplus L}$ satisfies the following moderate deviations principle: for all open sets $B \subset \mathbb{R}$,

$$\lim_{L\to\infty} \frac{1}{L^\beta} \log \mathbb{P}\big(L^{(1-\beta)/2}(L^{-1}X^{\oplus L} - \lambda) \in B\big) = -\inf_{x\in B} \tfrac{1}{2}x^2/\lambda.$$

(You may use the following fact: $\mathbb{E}s^X = e^{\lambda(s-1)}$.)

(b) Consider a bufferless queue, whose input at each timestep has distribution $X^{\oplus L}$, and whose service rate is $L\lambda + L^{(1+\beta)/2}C$ for some $C > 0$. We say that overflow occurs (at any given timestep) if $W^L > 0$, where W^L is the amount of work lost (at that timestep),

$$W^L = \big(X^{\oplus L} - (L\lambda + L^{(1+\beta)/2}C)\big) \vee 0.$$

Prove that

$$\lim_{L\to\infty} \frac{1}{L^\beta} \log \mathbb{P}(\text{overflow}) = -\tfrac{1}{2}C^2/\lambda.$$

(c) Let Y^L be the amount of work that is served by the queue (in a given timestep):

$$Y^L = X^{\oplus L} - W^L.$$

Using the notion of exponential equivalence, or otherwise, prove the following:

If Y^L is fed into another bufferless queue, which has service rate $L\lambda + L^{(1+\alpha)/2}B$ for some $0 < \alpha < \beta$ and $B > 0$, then for this queue

$$\lim_{L\to\infty} \frac{1}{L^\alpha} \log \mathbb{P}(\text{overflow}) = -\tfrac{1}{2}B^2/\lambda.$$

(d) Let Y^L be as in part (c). Suppose that Y^L is instead fed into a different bufferless queue, one which has service rate $L\lambda + L^{(1+\gamma)/2}D$, for some $\beta < \gamma < 1$ and $D > 0$. Prove that for this queue

$$\lim_{L\to\infty} \frac{1}{L^\gamma} \log \mathbb{P}(\text{overflow}) = -\infty.$$

(e) Comment briefly on the implication of (c) and (d) for the relative burstiness of the input $X^{\oplus L}$ and the output Y^L.

4 Write an essay on the scaling properties of queues. In your answer, you should give a heuristic derivation of at least three different scaling results, for queues in which one or more of the following grow large (in an appropriate sense): the buffer size, the service rate, and the input process. Discuss how you would choose which of these results to use, in order to describe a given queueing system.