MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Fall 2005 — 6.436J/15.085J

Final exam (3 hours; 100 pts) — 12/20/05

All parts can be done independently

Problem 1: (? points)
Let A, B, $A_1, A_2, \ldots$ be events. Suppose that for each k, we have $A_k \subseteq A_{k+1}$, and that B is independent of A_k. Let $A = \cup_{k=0}^{\infty} A_k$. Show that B is independent of A.

Solution:
$\left(\bigcup_{k\geq 1} A_k\right) \cap B = \bigcup_{k\geq 1} (A_k \cap B)$. Furthermore, $(A_k \cap B)$ is an increasing sequence of events. Using continuity of probability in the first and last equality, we have

$$\mathbf{P}(A \cap B) = \lim_{k\to+\infty} \mathbf{P}(A_k \cap B) = \lim_{k\to+\infty} \mathbf{P}(A_k)\mathbf{P}(B) = \mathbf{P}(A)\mathbf{P}(B).$$

Problem 2: (?? points)
Let $X_1, X_2, \ldots$ be continuous random variables with probability density functions (PDFs) $f_{X_1}, f_{X_2}, \ldots$.

(a) Suppose that $\lim_{k\to\infty} f_{X_k}(x) = g(x)$, for all $x \in \Re$. Invoke a certain result on integration to show that $\int_{-\infty}^{\infty} g(x)\, dx \leq 1$, and give an example to show that g need not be a PDF.

(b) Suppose that:
(i) $\lim_{k\to\infty} f_{X_k}(x) = f_X(x)$, for all $x \in \Re$, where f_X is the PDF of some random variable X, and
(ii) we have $f_{X_k}(x) \leq h(x)$, for all $x \in \Re$, where h is a function that satisfies $\int_{-\infty}^{\infty} h(x)\, dx < \infty$.

Show that X_k converges to X, in distribution.

Solution:

(a) $\lim_{k\to\infty} f_{X_k}(x) = g(x)$, for all $x \in \Re$. Hence by Fatou lemma,

$$\int g(x)dx \leq \liminf_k \int f_{X_k}(x)dx = 1.$$

Let X_k be the density of a normal random variable with mean k and variance 1. Then $f_{X_k}(x) \to 0$ for all x. But $g = 0$ is not a pdf.

(b) Let $F_k = \int_{-\infty}^{t} f_{X_k}(x)dx$ be the distribution function of X_k and F be the distribution function of X. Since $f_{X_k}(x) \to f_X(x)$ for all x and the density f_{X_k} are dominated by an integrable function h, the dominated convergence theorem yields for all t

$$\lim_k F_k(t) = \int_{-\infty}^{t} \lim f_{X_k}(x)dx = F(t).$$

Problem 3: (? points)
The following fact is known, and can be used in this problem: if a sequence of normal random variables X_k converges in distribution to a random variable X, then X is normal.

Suppose that for every k, the pair (X_k, Y) has a bivariate normal distribution. Furthermore, suppose that the sequence X_k converges to X, almost surely. Show that (X, Y) has a bivariate normal distribution. *Hint:* Use the "right" definition of the bivariate normal.

Solution:

For any $a, b \in \Re$, $aX_k + bY \xrightarrow{a.s} aX + bY \;\Rightarrow\; aX_k + bY \xrightarrow{D} aX + bY$. Since $aX_k + bY$ is a sequence of normal random variables, $aX + bY$ is normal by the given fact. Hence, (X, Y) has a bivariate normal distribution.

Problem 4: (? points)
Let X_1, X_2, X_3 be independent exponential random variables with mean 1. Let

$$\alpha = \mathbf{P}(X_1 > X_2 + X_3).$$

(a) Find α, without calculating any integrals.

(b) Find the probability that the largest of the three random variables X_1, X_2, X_3 is larger than the sum of the other two. [You can express your answer in terms of the constant α from part (a).]

Solution:

(a) Consider the merging of two independent Poisson processes, both with rate 1. α is the probability that the first two arrivals of the merged Poisson process are from the, say second process. Since each arrival of the merged process belongs independently to the second processes with probability $1/2$, $\alpha = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$.

(b) Let $\sigma \in \{1, 2, 3\}$ be the (random) index of largest random variable out of X_1, X_2, X_3.

$$\begin{aligned}
&\mathbf{P}(\max(X_1, X_2, X_3) > X_1 + X_2 + X_3 - \max(X_1, X_2, X_3)) \\
=&\mathbf{P}(X_1 > X_2 + X_3, \sigma = 1) + \mathbf{P}(X_2 > X_1 + X_3, \sigma = 2) + \mathbf{P}(X_3 > X_1 + X_2, \sigma = 3) \\
=&3\alpha = 3/4.
\end{aligned}$$

Problem 5: (? points)
Fast and slow customers arrive at a 24 hour store according to independent Poisson processes, each with rate 1 per minute. Fast customers stay in the bookstore for 1 minute, slow customers stay in the store for 2 minutes.

(a) What is the PMF of the total number of customer arrivals during a one minute interval?

(b) Find the variance of the number of customers in the store at 3 p.m.

(c) At 3 p.m., there is only one customer present in the store.

   (i) What is the probability, β, that the customer is a fast one?

   (ii) What is the PDF that this customer will depart before a new customer arrives? [You may express your answer in terms of the constant β from part (i). Also, you may leave your answer as a formula involving integrals – you do not have to evaluate the integrals.]

Let N_t be the number of fast customer arrivals during $[0, t]$.

(d) Does $(N_{2t} - N_t)/t$ converge in probability, as $t \to \infty$? With probability 1? If yes, to what? Outline a rigorous justification for your answers. You can start with t integer-valued and then argue for $t \in \Re$.

(e) Find (approximately) a time k such that

$$\mathbf{P}(N_k \geq 100) \approx 0.758.$$

Note that if Z is a standard normal random variable, then $\mathbf{P}(Z \leq 0.7) = 0.758$. [You do not need to be rigorous in deriving your answer. You may leave your answer in the form of an equation for k, which you do not need to solve numerically.]

Solution:

(a) The merged arrival process is Poisson with arrival rate $2/min$. Hence, the desired pmf is Poisson with parameter 2.

(b) Let N, N_f, N_s be respectively the total number of people of any type, of fast type and of slow type in the store at 3 p.m. N_f and N_s are independent. N_s is the random number of arrivals of slow customers in $[2:58, 3:00]$ and has pmf Poisson with parameter 2. Similarly, N_f has pmf Poisson with parameter 1.

$$var(N) = var(N_s + N_f) = var(N_s) + var(N_f) = 2 + 1 = 3.$$

(c) At 3 p.m., there is only one customer present in the store. Denote by S this event.

(i) Let T be the random type of the last customer to arrive before 3pm. By Bayes rule,

$$\beta = \mathbf{P}(S, T = \text{fast})/[\mathbf{P}(S, T = \text{fast}) + \mathbf{P}(S, T = \text{slow})]$$

The event $(S, T = \text{fast})$ occurs iff the last arrival is a fast customer and exactly one of those occurred in [2:59,3:00] and no arrival from slow customers occurred in [2:58,3:00]. Using the (unconditional) independence of the arrival processes,
$\mathbf{P}(S, T = \text{fast}) = \frac{1}{2}\exp(-2)\exp(-1)$. Similarly, $\mathbf{P}(S, T = \text{slow}) = \frac{1}{2}2\exp(-2)\exp(-1)$. Hence, $\beta = 1/3$.

(ii) Note Q the desired probability. The residual time Z of shopping for a customer observed at 3pm is uniformly distributed over $[0, 2]$ if the customer type is slow and over $[0, 1]$ if the customer type is fast. Given that the residual life is z, the probability that no customer arrives before the current customer leaves is $\exp(-2z)$. Hence,

$$Q = \beta \int_0^1 \exp(-2z)dz + (1 - \beta) \int_0^2 \exp(-2z)\frac{1}{2}dz.$$

Let N_t be the number of fast customer arrivals during $[0, t]$.

(d) $(N_{2t} - N_t)/t$ converge to 1 as $t \to \infty$ with probability one (a fortiori in probability). Let X_i be the number of arrivals of fast customers in the time interval $[i, i+1]$. X_i are iid with Poisson distribution with parameter 1 (mean and variance is 1). For integer-valued t, $(N_{2t} - N_t)/t = \frac{1}{t}\sum_{i=t}^{2t-1} X_i$, which converges to one with probability one by the strong law of large number. For real-valued t, let n_t be the largest integer such that $n_t \leq t$.

$$0 \leq (n_t/t)(N_{t+n_t} - N_t)/n_t \leq (N_{2t} - N_t)/t \leq (n_t+1/t)(N_{t+n_t+1} - N_t)/(n_t+1).$$

By the previous argument, we have that the lower bound and the upper bound converge to one with probability one.

(e) Let k be an integer. $N_k = \sum_{i=0}^{k-1} X_i$ and X_i are iid with mean and variance 1.

$$\mathbf{P}(N_k \geq 100) = \mathbf{P}\left(\sum_{i=0}^{k-1} X_i \geq 100\right) = \mathbf{P}\left(\sqrt{k}\frac{\sum_{i=0}^{k-1} X_i/k - 1}{\sqrt{1}} \geq \sqrt{k}\frac{100/k - 1}{\sqrt{1}}\right)$$

For large k, the central limit theorem states that $\mathbf{P}\left(\sqrt{k}\frac{\sum_{i=0}^{k-1} X_i/k-1}{\sqrt{1}} \geq q\right) \simeq \mathbf{P}(Z \geq q)$.

Hence, $\mathbf{P}(N_k \geq 100) \approx 0.758$ iff $\mathbf{P}(Z \geq \sqrt{k}\frac{100/k-1}{\sqrt{1}}) \approx 0.758 \approx \mathbf{P}(Z \geq -0.7)$.

k such that $100 - k + 0.7\sqrt{k} = 0$ is the answer.

Problem 6: (? points)
A 6-state Markov chain has a transition probability matrix of the following form, where a $*$ indicates a positive entry.

$$\begin{pmatrix} 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ * & * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 & 0 \end{pmatrix}$$

Let $r_{ij}(n) = \mathbf{P}(X_n = j \mid X_0 = i)$.

(a) Does $r_{ij}(n)$ converge as the integer variable n tends to infinity? If yes, does the limit depend on i?

(b) Does $r_{ij}(3n)$ converge as the integer variable n tends to infinity? If yes, does the limit depend on i?

Solution:

(a) The Markov chain has no transient state and a single recurrent class with period 3. Hence, $r_{ij}(n)$ does not converge for any i, j as $n \to +\infty$.

(b) $r_{ij}(3n)$ is the probability of a transition from state i to state j for the Markov chain with transition probability matrix P^3. This Markov chain has three aperiodic recurrent classes and no transient state. Hence, $r_{ij}(n)$ has a limit l_{ij} as n goes to infinity but the limit does depend on i. For example, $l_{ij} = 0$ iff i and j are not in the same recurrent class.