IEOR 6711: Stochastic Models I SOLUTIONS to First Midterm Exam, 2013 There are five questions, each with multiple parts.

Justify your answers; show your work.

2. Random Hats (15 points)

At a party n people each come wearing a hat. When they leave, a random hat is assigned to each person, with each hat being equally likely.

(a) What is the expected number of people who leave with the same hat they had when they arrived?

(b) What is the variance of the number of people who leave with the same hat they had when they arrived?

This is the matching problem; see Example 1.3(A) and 1.5(F). Let $X_i = 1$ if the i^{th} person gets his own hat. Then, for any i and j with $i \neq j$,

$$P(X_i = 1) = \frac{1}{n}$$
 and $P(X_i = 1, X_j = 1) = P(X_j = 1 | X_i = 1)P(X_i = 1) = \frac{1}{n-1} \frac{1}{n} = \frac{1}{n(n-1)}$,

so that

$$Var(X_i) = \frac{n-1}{n^2}$$
 and $Cov(X_i, X_j) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}$

Let $X \equiv X_1 + \cdots + X_n$. Then

$$E[X] = nE[X_1] = 1$$
 and $Var(X) = nVar(X_1) + 2\left(\frac{n!}{2!(n-2)!}\right)Cov(X_i, X_j) = 1.$

See pages 10-11.

(c) Suppose that this hat-matching experiment is repeated in independent experiments with larger and larger groups, with n people in the n^{th} experiment for all $n \ge 2$. Let N be the number of times among all these experiments that two designated people (assumed to be present in all experiments) both get their own hat back. What is $P(N < \infty)$ and why?

We have observed that in the experiment with n people,

$$P(X_i = 1, X_j = 1) = \frac{1}{n(n-1)}$$

Since

$$\sum_{n=2}^{\infty} \frac{1}{n(n-1)} < \infty,$$

we conclude that $P(N < \infty) = 1$ by the Borel-Cantelli lemma on p. 4.

However, in this case we can actually compute the mean. Note that

$$E[N] = \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \sum_{n=2}^{\infty} \left(\frac{1}{n-1} - \frac{1}{n}\right) = 1 < \infty.$$

Since $E[N] < \infty$, necessarily $P(N < \infty) = 1$.

3. Variance Formulas (20 points)

At a party n people each come wearing a hat. When they leave, a random hat is assigned to each person, with each hat being equally likely.

(a) Let X and Y be two real-valued random variables. Assume that $E[X^2] < 1$. Prove or disprove:

$$Var(X) = E[Var(X|Y)] + Var(E[X|Y]),$$

where Var(X|Y) is defined by

$$Var(X|Y) \equiv E[(X - E[X|Y])^2|Y].$$

This is the **conditional variance formula**; see Exercise 1.22 (answer in the back on p. 477. Care is needed here. It is good to write out E[Var(X|Y)] and Var(E[X|Y]) and add them. When we do this, we see that each contains the term $E[(E[X|Y])^2]$. But in the sum these terms cancel.

(b) Let $\{N(t) : t \ge 0\}$ be a Poisson process with rate λ . Let $\{X_n : n \ge 1\}$ be a sequence of independent random variables, each distributed as X, where $E[X^2] < \infty$. Prove or disprove:

$$Var\left(\sum_{i=1}^{N(t)} X_i\right) = \lambda t Var(X).$$

This is **false**. The correct statement should be

$$Var\left(\sum_{i=1}^{N(t)} X_i\right) = \lambda t E[X^2].$$

We can apply the conditional variance formula in part (a), conditioning on N(t), to get

$$Var\left(\sum_{i=1}^{N(t)} X_i\right) = E[N(t)]Var(X) + Var(N(t)E[X]) = \lambda t(Var(X) + E[X]^2) = \lambda tE[X^2].$$

3. The New Six (6) Subway Line. (20 points)

A new subway line has been added to the West Side for the convenience of Columbia students. It has six stations. There are stations at 86^{th} street (station 1), 96^{th} street (station

2), 106^{th} street (station 3), 116^{th} street (station 4), 126^{th} street (station 5) and 136^{th} street (station 6).

We consider only the northbound subway. A northbound subway arrives at station 1 every 10 minutes. The travel time between successive stations is constant, equal to 2 minutes. Suppose that the subway stations and the subway trains have unlimited capacity and that the time to load and unload passengers can be ignored. Suppose that the subway runs continuous, day and night.

For $1 \leq i \leq 5$, customers arrive at station *i* to use the northbound subway according to a Poisson process with rate λ_i per minute. Suppose that each customer entering station *i* gets off at station *j* with probability $P_{i,j}$, independently of all other customers, where $P_{i,j} > 0$ if and only if j > i and

$$\sum_{j=i+1}^{6} P_{i,j} = 1 \text{ for all } i, \quad 1 \le i \le 5.$$

We remark that this problem is a minor modification and elaboration of assigned homework exercise 2.14.

(a) Give an expression for the expected number of customers to get on the subway (necessarily going north) at each visit to station i.

Let $N_i(t)$ be the number of customers that arrive at station *i* in the interval [0, t], where the initial time 0 is chosen to be an instant at which the subway arrives at station *i*. Since the subway arrives at each station every 10 minutes, we want

$$E[N_i(10)] = 10\lambda_i.$$

(b) Give an expression for the probability generating function of the number of customers to get on the subway at each visit to station i.

Let N_i be the number that get on the subway at each visit to station *i*. It is Poisson with mean $m_i = 10\lambda_i$. Hence,

$$P(N_i = k) = \frac{m_i^k e^{-m_i}}{k!} \quad \text{where} \quad m_i = 10\lambda_i$$

and the probability generating function is

$$\hat{P}_{N_i}(z) = E[z^{N_i}] = \sum_{k=0}^{\infty} z^k P(N_i = k) = e^{10\lambda_i(z-1)}.$$

(c) Give an expression for the expected value of the sum of the waiting times of all customers to get on the subway at each visit to station i.

This is a minor variant of Example 2.3(A) on p. 68. Let S be the sum of all the waiting times. We exploit the **conditional-uniform property**, discussed in §2.3. We condition on N_i and then uncondition. We use the fact that, conditional on $N_i = k$, the k arrival times are distributed as the order statistics of k i.i.d. random variables, each uniformly distributed over the interval [0, 10]. Hence,

$$E[S|N_i = k] = 5k$$

so that

$$E[S] = \sum_{k=0}^{\infty} E[S|N_i = k] P(N_i = k) = \sum_{k=0}^{\infty} 5kP(N_i = k) = 5E[N_i] = 50\lambda_i$$

(d) Suppose that 8 customers get on the subway at station 1 at one specified time. What is the probability that exactly 3 of these customers had to wait more than 4 minutes before getting on the subway?

Let A be the event in question. We again exploit the conditional uniform property. Conditional on 8 arrivals, each can be regarded as arriving according to i.i.d. uniform variables in [0, 10]. The probability that each has to wait more than 4 minutes is p = 0.6. The probability that 3 have to wait more than 4 minutes is then **binomial**:

$$P(A) = \left(\frac{8!}{3!5!}\right) p^3 (1-p)^5 = \left(\frac{8!}{3!5!}\right) (0.6)^3 (0.4)^5.$$

(no need to go further in the calculation)

(e) Give an expression for the probability that the number of customers getting off the northbound subway at a visit to station 4 is exactly j.

Let D_i be the number of departures (getting off the subway) at each visit to station *i*, By the **independent thinning and the independent superposition of Poisson processes**, D_4 has a **Poisson distribution** with mean

$$\delta_4 \equiv E[D_4] = 10\lambda_1 P_{1,4} + 10\lambda_2 P_{2,4} + 10\lambda_3 P_{3,4},$$

Hence,

$$P(D_4 = j) = \frac{\delta_4^j e^{-\delta_4}}{j!}$$

for δ_4 defined above.

To elaborate, the numbers $N_{i,j}$ to get on at station *i* with destination *j* are independent Poisson variables by independent thinning. Then the sum of independent Poisson variables is Poisson.

(f) Give an expression for the probability that, simultaneously, the number of customers getting off the northbound subway at a visit to station 4 is j and the number getting off at the next stop, at station 5, is k.

Again, by the independent thinning and the independent superposition of Poisson processes, D_4 and D_5 are **independent Poisson random variables**. (The variables $N_{i,j}$ and $N_{l,k}$ are independent for all i and l if $j \neq k$.) Hence,

$$P(D_4 = j, D_5 = k) = P(D_4 = j)P(D_5 = k)$$
$$= \left(\frac{\delta_4^j e^{-\delta_4}}{j!}\right) \left(\frac{\delta_5^k e^{-\delta_5}}{k!}\right)$$

where δ_4 is given in the previous part and

$$\delta_5 \equiv E[D_5] = 10\lambda_1 P_{1,5} + 10\lambda_2 P_{2,5} + 10\lambda_3 P_{3,5} + 10\lambda_4 P_{4,5}.$$

(g) Suppose that $\lambda_i = 12 - 2i$ for all $i, 1 \le i \le 5$, $P_{i,j} = 1/(6-i)$ for all i and j with j > i $(1 \le i \le 5 \text{ and } 2 \le j \le 6)$. Determine a convenient accurate approximation for the probability that the number of customers getting off the northbound subway at one specified visit to station 6 is greater than 130? Is that probability more than 1/20? Why is your approximation justified?

We are interested in D_6 . Under these extra assumptions $E[D_6] = 100$. Since the variance equals the mean, we have $Var(D_6) = 100$ and $\sqrt{Var(D_6)} = 10$. Hence, we use a **normal approximation**:

$$P(D_6 > 130) = P\left(\frac{D_6 - E[D_6]}{\sqrt{Var(D_6)}} > \frac{130 - E[D_6]}{\sqrt{Var(D_6)}}\right)$$

$$\approx P\left(N(0, 1) > \frac{130 - E[D_6]}{\sqrt{Var(D_6)}}\right)$$

$$= P\left(N(0, 1) > \frac{130 - 100]}{\sqrt{100}}\right)$$

$$= P\left(N(0, 1) > \frac{30}{10}\right) \approx P(N(0, 1) > 3.0) \approx 0.0013$$

For the final numerical details, we only need to know that P(N(0,1) > 3.0) < 0.05. Everybody should know that without calculating. The approximation is justified by the **central limit theorem**. It applies because a Poisson random variable with mean nm can be written as the sum of n i.i.d. Poisson random variables, each of mean m.

4. modes of convergence (15 points)

Consider a sequence of real-valued random variables $\{X_n : n \ge 1\}$.

(a) Define convergence of X_n to a random variable X in (i) mean square, (ii) in probability and (iii) with probability 1.

The modes of convergence are discussed in the lecture notes of Thursday, September 5.

- (i) $E[|X_n X|^2] \to 0$ as $n \to \infty$
- (ii) For all $\epsilon > 0$, $P(|X_n X| > \epsilon) \to 0$ as $n \to \infty$

(iii) $P(\lim_{n\to\infty} X_n = X) \equiv P(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1,$

where Ω is the underlying sample space and ω is an element in the space Ω . In the above definitions we assume that the notion of convergence for a sequence of real numbers is understood.

(b) Prove or disprove: Convergence of X_n to a random variable X in mean square implies convergence of X_n to a random variable X in probability.

This conclusion is valid. It can be proved by Markov's inequality, Lemma 1.7.1 on p. 39. For any ϵ and n,

$$P(|X_n - X| > \epsilon) \le \frac{E[|X_n - X|^2]}{\epsilon^2} \to 0 \text{ as } n \to \infty.$$

by the mean-square convergence.

(c) Prove or disprove: Convergence of X_n to a random variable X in mean square implies convergence of X_n to a random variable X with probability 1.

This conclusion is **not valid**. See Counterexample 2 on page 4 of the notes for September 5. Such examples are harder to construct than it might seem.

5. peaks (15 points)

Let $\{X_n : n \ge 1\}$ be a sequence of i.i.d. random variables with a continuous cdf F. We say that a peak occurs at time n if $X_{n-1} < X_n > X_{n+1}$. Let N_n be the number of peaks among the first n variables. Prove or disprove:

 $\frac{N_n}{n} \to \frac{1}{3} \quad \text{as} \quad n \to \infty \quad \text{with probability 1.}$

This is homework exercise 1.37, assigned in the first homework. See the written solutions. We break up the random variables into three groups and apply the SLLN to each group and then combine the results to get the overall result. In particular, we let $Y_n = 1$ if X_n is a peak and $Y_n = 0$ otherwise. Since X_n has a continuous cdf, $P(Y_n = 1) = 1/3$. Then we look at the three sequences $\{Y_{3n-2} : n \ge 1\}$, $\{Y_{3n-1} : n \ge 1\}$ and $\{Y_{3n} : n \ge 1\}$. Since the random variables in each sequence are i.i.d. and bounded, the SLLN applies to each sequence. We then combine the limits to get the result. We use: If $a_n \to a$ and $b_n \to b$ as $n \to \infty$, then $a_n + b_n \to a + b$ as $n \to \infty$ for sequences of real numbers.