SOLUTIONS to the First Midterm Exam, October 7, 2012
IEOR 6711: Stochastic Models I,

1. Poisson Process and Transforms (30 points)

[**grading scheme:** On all questions, partial credit will be given. Up to 4 points off for errors on parts (a)-(g); up to 8 points off for errors on part (h). But minimum possible score is 0. Parts (a)-(g) are judged relatively easy, but a full answer to part (h) is harder.]

Let $\{N(t) : t \geq 0\}$ be a **Poisson process** with rate (intensity) λ.

(a) Give expressions for: (i) the **probability mass function (pmf)** of $N(t)$, $p_{N(t)}(k) \equiv P(N(t) = k)$; (ii) the **probability generating function (pgf)** of $N(t)$, $\hat{p}_N(t)(z)$; the **moment generating function (mgf)**, $\psi_{N(t)}(u)$ and (iv) the **characteristic function (cf)** of $N(t)$, $\phi_{N(t)}(u)$.

$$
\begin{aligned}
p_{N(t)}(k) &\equiv \frac{e^{-\lambda t}(\lambda t)^k}{k!} \\
\hat{p}_N(t)(z) &\equiv E[z^{N(t)}] = \sum_{k=1}^{\infty} z^k p_{N(t)}(k) = e^{\lambda t(z-1)} \\
\psi_{N(t)}(u) &\equiv E[e^{uN(t)}] = \sum_{k=1}^{\infty} e^{uk} p_{N(t)}(k) = \hat{p}_N(t)(e^u) = e^{\lambda t(e^u-1)} \\
\phi_{N(t)}(u) &\equiv E[e^{iuN(t)}] = \sum_{k=1}^{\infty} e^{iuk} p_{N(t)}(k) = \hat{p}_N(t)(e^{iu}) = e^{\lambda t(e^{iu}-1)}
\end{aligned}
$$

(b) Show how the mgf $\psi_{N(t)}(u)$ can be used to derive the mean and variance of $N(t)$.

Use the fact that $\psi_N(t)^{(k)}(u) = E[N(t)^k e^{uN(t)}]$, where is the k^{th} derivative of the mgf $\psi_N(t)(u)$ with respect to u, so that $\psi_N(t)^{(k)}(0) = E[N(t)^k]$. We directly get $\psi_N(t)^{(1)}(u) = \lambda t e^u \psi_N(t)(u)$, so that $E[N(t)] = \psi_N(t)^{(1)}(0) = \lambda t$ and

$$\psi_N(t)^{(2)}(u) = (\lambda t e^u)^2 \psi_N(t)(u) + \lambda t e^u \psi_N(t)(u),$$

so that $E[N(t)^2] = (\lambda t)^2 + \lambda t$. Hence,

$$Var(N(t)) = E[N(t)^2] - E[N(t)] = (\lambda t)^2 + \lambda t - (\lambda t)^2 = \lambda t.$$

(c) Use the mgf $\psi_{N(t)}(z)$ to prove or disprove the claim: If $[a, b]$ and $[c, d]$ are two disjoint subintervals of the positive halfline $[0, \infty)$, then the sum $(N(d) - N(c)) + (N(b) - N(a))$ has a Poisson distribution.

We draw on §2.1 (definition of a Poisson process) and §1.4 (transforms). We exploit the fact (axiom) that a Poisson process has independent increments, and that each increment has a mean proportional to its length. We also exploit the fact that the transform of a sum of independent random variables is the product of the transforms. Thus,

$$\begin{aligned}\psi_{(N(d)-N(c))+(N(b)-N(a))}(u) &= \psi_{(N(d)-N(c))}(u)\psi_{(N(b)-N(a))}(u)\\ &= \psi_{N(d-c)}(u)\psi_{N(b-a))}(u) = e^{\lambda(d-c)(e^u-1)}e^{\lambda(b-a)(e^u-1)}\\ &= e^{\lambda[(d-c)+(b-a)](e^u-1)},\end{aligned}$$

which implies that $(N(d)-N(c))+(N(b)-N(a))$ has a Poisson distribution with mean equal to the sum of the two means, $\lambda[(d-c)+(b-a)]$. In the last step we exploit the property $e^a \times e^b = e^{a+b}$. Note that deriving higher moments from the pgf is somewhat more complicated.

(d) What is the probability $P(N(t)\,\text{is even}) \equiv P(N(t) \in \{2k : k \geq 0\})$? Prove that this probability is always greater than 1/2 and converges to 1/2 as $t \to \infty$.

This is part (d) of problem 1.11 in the textbook, assigned in our first homework. We use the pgf. If necessary, we directly derive the following:

$$P(N(t)\,\text{is even}) = \frac{\hat{p}_N(t)(1) + \hat{p}_N(t)(-1)}{2} = \frac{(1+e^{-\lambda t})}{2},$$

from which it is immediate that it is always greater than 1/2 and converges to 1/2 as $t \to \infty$.

Let $\{X_k : k \geq 1\}$ be a sequence of i.i.d. continuous real-valued random variables with **probability density function (pdf)** $f(x)$, mean m and variance σ^2. Let

$$Y(t) \equiv \sum_{k=1}^{N(t)} X_k, \quad t \geq 0.$$

(e) Give an expression for the cf of $Y(t)$.

First note that $\{Y(t) : t \geq 0\}$ is a compound Poisson process, as discussed in §2.5 of the textbook and the lecture notes of September 27. Hence, paralleling p. 82, we have

$$\begin{aligned}\phi_{Y(t)}(u) &\equiv E[e^{iuY(t)}] = E[e^{iu\sum_{k=1}^{N(t)} X_k}]\\ &= E\left(\sum_{n=0}^{\infty} e^{iu\sum_{k=1}^{n} X_k} P(N(t)=n)\right)\\ &= \sum_{k=0}^{\infty} \phi_X(u)^n P(N(t)=n)\\ &= \sum_{k=0}^{\infty} \phi_X(u)^n \frac{e^{-\lambda t}(\lambda t)^n}{n!}\\ &= \exp\{\lambda t(\phi_X(u)-1)\} = \hat{p}_{N(t)}(\phi_X(u)).\end{aligned}$$

(f) Derive the mean and variance of $Y(t)$.

Again, see the lecture notes of September 27. As noted on p. 83, the mean and variance can easily be found by differentiating the transform, yielding

$$E[Y(t)] = \lambda t E[X] \quad \text{and} \quad Var(Y(t)) = \lambda t E[X^2].$$

Note that the variance involves $E[X^2]$, not $Var(X)$.

(g) Does there exist a finite random variable L with a non-degenerate probability distribution (such that $P(L = c) \neq 1$ for any c) and constants a and b such that

$$\frac{Y(t) - at}{\sqrt{bt}} \Rightarrow L \quad \text{as} \quad t \to \infty, \tag{1}$$

where $\Rightarrow$ denotes convergence in distribution? If so, what are a, b and L?

Yes, we can apply the CLT. We can let $a = E[Y(1)] = \lambda E[X]$, $b = Var(Y(1)) = \lambda E[X^2]$ and $L = N(0, 1)$, a standard normal random variable.

(h) Give a detailed proof to support your answer in part (g).

We can apply the proof using cf's and a Taylor series expansion, just as for the ordinary CLT on p. 11 of the lecture notes for September 11. See the lecture notes of September 27.

---

2. Conditional Remaining Lifetimes (30 points)

Let a **random lifetime** be represented by a nonnegative continuous random variable X with probability density function (pdf) $f(x)$, cumulative distribution function (cdf) $F(x) \equiv \int_0^x f(s)\, ds$ and complementary cdf (ccdf) $F^c(x) \equiv 1 - F(x) \equiv P(X > x)$ satisfying $F^c(x) > 0$ for all $x \geq 0$. For $t \geq 0$, the associated **conditional remaining lifetimes** are the random variables $X(t)$ with ccdf (a function of x for $x \geq 0$ which depends on t)

$$F^c(x; t) \equiv P(X(t) > x) \equiv P(X > t + x | X > t), \quad t \geq 0, x \geq 0,$$

and associated pdf $f(x; t)$ (with $F(x; t) \equiv \int_0^x f(s; t)\, ds$). Let $r(t) \equiv f(0; t)$.

[**grading scheme:** Up to 5 points off for errors on all parts, but the minimum possible score is 0.]

(a) Give an explicit expression for $r(t)$ in terms of the pdf f of X.

The function $r(t)$ is the **failure (or hazard) rate function**, which is defined on p.38 of the textbook. The pdf is derived there.

$$f(x;t) = \frac{f(t+x)}{F^c(t)} \quad \text{and} \quad r(t) \equiv f(0;t) = \frac{f(t)}{F^c(t)} = \frac{f(t)}{\int_t^\infty f(s)\,ds} \tag{2}$$

(b) Suppose that $r(t) = 7$, $t \geq 0$. Does that imply that the pdf f of X is well defined and that we know it? If so, what is it?

As shown on pp. 38-39 of the textbook, the failure rate function uniquely characterizes the cdf F. To see that, we can integrate r over $[0, t]$ to get

$$\int_0^t r(s)\,ds = -\log\{F^c(t)\} + c,$$

yielding the explicit expression

$$F^c(t) = \exp\{-\int_0^t r(s)\,ds\}. \quad t \geq 0. \tag{3}$$

If we substitute $r(t) = 7$ into (3), then we get

$$F^c(t) = e^{-7t}, \quad t \geq 0,$$

which is the ccdf of the exponential pdf with rate $\lambda = 7$, i.e.,

$$f(t) = 7e^{-7t}, \quad t \geq 0.$$

A lifetime distribution is exponential if and only if its failure rate function is constant.

(c) Suppose that $r(t) = t$, $t \geq 0$. Does that imply that the pdf f of X is well defined and that we know it? If so, what is it?

We reason as in part (b), substituting $r(t) = t$ into (3) to get

$$F^c(t) = e^{-t^2/2}, \quad t \geq 0,$$

which, we see by differentiating with respect to t, is the ccdf of the pdf

$$f(t) = te^{-t^2/2}, \quad t \geq 0.$$

Now let X_i, $i = 1, 2$, be two independent random lifetimes defined as above, having pdf's $f_i(x)$, cdf's $F_i(x)$, ccdf's $F_i^c(x)$. Let $X_i(t)$ be the associated conditional remaining lifetimes with cdf's $F_i^c(x;t)$ and pdf's $f_i(x;t)$. Let $r_i(t) \equiv f_i(0;t)$

(d) Prove or disprove:

$$P(X_1 < X_2 | \min\{X_1, X_2\} = t) = \frac{r_1(t)}{r_1(t) + r_2(t)}.$$

This formula is valid. This is Problem 1.34 in the textbook. To prove it, we first observe that

$$P(X_1 < X_2 | \min\{X_1, X_2\} = t) = \frac{f_1(t)F_2^c(t)}{f_1(t)F_2^c(t) + f_2(t)F_1^c(t)}. \tag{4}$$

To approach (4) more formally, we might exploit basic properties of conditional pdf's (which can be separately justified). For that purpose, let $M \equiv \min\{X_1, X_2\}$. You can start with the joint distribution of (X_1, X_2, M) and the associated conditional density

$$f_{X_1,X_2|M}(x, y|t) = \frac{f_{X_1,X_2,M}(x, y, t)}{f_M(t)},$$

where we must have $t = \min\{x, y\}$. We then can derive

$$f_M(t) = f_1(t)F_2^c(t) + f_2(t)F_1^c(t), \quad t \geq 0,$$

and then integrate to obtain (4). Given (4), we divide both the numerator and denominator of (4) by $F_1^c(t)F_2^c(t)$ and apply the representation for r in (2).

Now suppose that $r_1(t) \leq r_2(t)$ for all $t \geq 0$. Prove or disprove each of the following statements:

(e) $F_1(t) \leq F_2(t) \quad \text{for all} \quad t \geq 0,$

This results is valid too. The statement is tantamount to the conclusion that hazard rate (or failure rate) ordering (expressed by the condition) implies stochastic ordering (expressed by the conclusion). See §9.3 and §9.1 of the textbook, respecitvely. The conclusion here is equivalent to the stochastic ordering $X_1 \geq_{st} X_2$. (The random variables are ordered stochastically in the opposite order of the cdf ordering as functions. See the extra lecture notes of September 27.)

(f) $E[X_1^3] \geq E[X_2^3]$,

This results is valid too. The statement follows from part (e) by the representation

$$E[X^3] = \int_0^\infty 3x^2 F^c(x)\, dx$$

given in 1 (d) (iii) on page 3 of homework assignment 1. It also follows from a basic property of the stochastic ordering: $X_1 \geq_{st} X_2$ if and only if

$$E[h(X_1)] \geq E[h(X_2)] \tag{5}$$

for all nondecreasing integrable real-valued functions h. See Proposition 9.1.2 of the textbook and the extra lecture notes of September 27.

Both approaches are easily understood by recognizing that, given $X_1 \geq_{st} X_2$, it is possible to construct alternative random variables Y_i, $i = 1, 2$, with Y_i distributed the same as X_i and having the property $P(Y_1 \geq Y_2) = 1$. The construction can be in terms of a random variable U uniformly distributed on $[0, 1]$. We let $Y_i \equiv F_i^{-1}(U)$. (This construction is discussed in homework 1.)

(g) $P(X_1 \geq X_2) = 1)$.

This statement is invalid. In fact, we know nothing about the joint distribution of the vector (X_1, X_2). The statement *is* valid for the specially constructed random variables (Y_1, Y_2) in part (f), but not for the original random variables. For a concrete example, suppose that X_1 and X_2 are i.i.d. exponential random variables with mean 1. let these random variables initially be defined on the space $[0, 1]$ using $X_i \equiv F_i^{-1}(U)$, so that the conclusion would be true; indeed we would have $P(X_1 = X_2) = 1)$. But then change the way X_2 is defined. In particular, let X_2 be defined over $[0, 1/2]$ as it was initially over $[1/2, 1]$ and let X_2 be defined over $[1/2, 1]$ as it was initially over $[0, 1/2]$. The distribution is unchanged, but the w.p.1 order is lost. Indeed, after this modification we have $P(X_1 < X_2) = P(X_1 > X_2) = 1/2)$. And yet, the two random variables after this modification do still have the same distribution.

---

3. Independent Random Variables. (20 points)

Let $\{X_n : n \geq 1\}$ and $\{Y_n : n \geq 1\}$ be independent sequences of independent random variables with X_n distributed the same as Y_n for all $n \geq 1$ and

$$P(X_n = n) = 1 - P(X_n = 0) = \frac{1}{n} \quad \text{for all} \quad n \geq 1.$$

Let

$$Z_n \equiv X_n Y_n \quad \text{and} \quad D_n \equiv X_n - Y_n, \quad n \geq 1.$$

[**grading scheme:** Part (a) 2 points; parts (b)-(d) 6 points each.]

(a) What are the mean and variance of Z_n?

By independence, $P(X_n Y_n = n^2) = P(X_n = n, Y_n = n) = (1/n)^2 = 1/n^2$. Otherwise $Z_n \equiv X_n Y_n = 0$. Hence we have

$$P(Z_n = n^2) = 1 - P(Z_n = 0) = \frac{1}{n^2},$$

so that

$$E[Z_n] = 1, \quad E[Z_n^2] = n^2 \quad \text{and} \quad Var(Z_n) = E[Z_n^2] - (E[Z_n])^2 = n^2 - 1.$$

(b) What is the probability that the sum $Z_1 + Z_2 + \cdots + Z_n$ converges to a finite limit as $n \to \infty$?

By the Borel-Cantelli Lemma in Proposition 1.1.2,

$$P(Z_n \neq 0 \quad \text{infinitely often}) = 0,$$

because

$$\sum_{n=1}^{\infty} P(Z_n \neq 0) = \sum_{n=1}^{\infty} (1/n^2) < \infty.$$

Since there almost surely only finitely many non-zero terms in the series, the series necessarily converges to a finite limit with probability 1.

(c) What is the probability that the sum $D_1 + D_2 + \cdots + D_n$ converges to a finite limit as $n \to \infty$?

Unlike part (b), the probability that the sum $D_1 + D_2 + \cdots + D_n$ converges to a finite limit as $n \to \infty$ is 0.

To see that, first note that $P(D_n \neq 0) = 2(1/n)((1-(1/n)) = 2(n-1)/n^2 > 1/n$ for $n \geq 2$. Since $\{D_n : n \geq 1\}$ is a sequence of independent random variables, we can apply the converse of the Borel-Cantelli Lemma in Proposition 1.1.3 to deduce that

$$P(D_n \neq 0 \quad \text{infinitely often}) = 1,$$

because

$$\sum_{n=1}^{\infty} P(D_n \neq 0) = \sum_{n=1}^{\infty} 2(n-1)/n^2 = \infty.$$

Hence $D_n = n$ or $D_n = -n$ infinitely often almost surely. That almost surely violates the criterion for convergence. In particular, for such a limit to exist, there must exist for each ϵ an $n_0 \equiv n_0(\epsilon)$ such that

$$\sum_{j=n_0}^{\infty} D_j < \epsilon.$$

But that clearly is violated.

(d) Are there deterministic constants a_n and b_n with $a_n \to \infty$, $b_n \to \infty$ and $a_n/b_n \to \infty$ such that

$$\frac{(Z_1 + \cdots + Z_n) - a_n}{b_n} \Rightarrow L \quad \text{as} \quad n \to \infty,$$

where L is a nondegenerate random variable (with a probability distribution not concentrating on a single value) and $\Rightarrow$ denotes convergence in distribution? If so, what are the se constants a_n and b_n and what is L?

No, there do not exist such constants and nondegenerate random variable L. Since the random variables Z_n are not identically distributed and since the variance of Z_n is order n^2, clearly the standard CLT does not apply. Since $Z_1 + \cdots Z_n \to U$ as $n \to \infty$ w.p.1 for some (unknown) nondegenerate random variable U by part (b) and $b_n \to \infty$, it is evident that

$$\frac{(Z_1 + \cdots + Z_n)}{b_n} \to 0 \quad \text{as} \quad n \to \infty \quad w.p.1$$

So we can disregard that term. On the other hand, $a_n/b_n \to \infty$, by assumption.

---

4. Cars on a Highway Segment During Rush Hour (20 points)

Two car-detection devices have been placed on a segment of a one-way highway without exits or entrances in between. Let $N(t)$ be the number of cars that pass the first detection point during $[0, t]$. Suppose we consider a rush hour period in the morning. Thus, we let $\{N(t) : t \geq 0\}$ be a nonhomogeneous Poisson process with rate function $\lambda(t) = 12t$ over an initial time interval $[0, 6]$. However, the first detection device does not work perfectly. Indeed, each car is detected by the initial detection device only with probability $2/3$, independently of the history up to the time it passes the detection device. However, The second detection device works perfectly.

As a simplifying assumption, assume that the cars do not interact, so that the length of time that the cars remain in the highway segment can be regarded as i.i.d. random variables, independent of the times that they pass the detection device. These times are regarded as random variables, because the cars travel at different speeds. Suppose that the length of time each car remains in the highway segment is uniformly distributed over the interval $[1, 3]$.

[**grading scheme:** Up to 4 points off for errors on all parts, but the minimum possible score is 0.]

(a) What is the expected number of cars that pass the first detection point on the highway during the interval $[0, 5]$ and are detected by the detection device?

Let $N_d(t)$ be the number of cars that pass the initial detection device and are detected in the interval $[0, t]$. By independent thinning, $\{N_d(t) : t \geq 0\}$ is also a nonhomogeneous Poisson process with rate function $\lambda_d(t) = (2/3)\lambda(t) = 8t$ over the initial time interval $[0, 6]$.

$$m(5) \equiv E[N_d(5)] = \int_0^5 8t\, dt = 100$$

(b) What is the probability that precisely 50 cars pass the first detection point on the highway during the interval $[0, 5]$ and are detected by the detection device?

By part (a),

$$P(N_d(5) = 50) = \frac{e^{-m(5)} m(5)^{50}}{50!} = \frac{e^{-100}(100)^{50}}{50!}$$

(c) Give a convenient approximate expression for the probability that more than 120 cars pass the first detection point on the highway during the interval $[0, 5]$ and are detected by the detection device.

Since $m(5)$ is large, we can use a normal approximation. Let $N(0, 1)$ be a standard normal random variable.

$$\begin{aligned} P(P(N_d(5) > 120) &= P\left(\frac{N_d(5) - E[N_d(5)]}{\sqrt{Var(N_d(5))}} > \frac{120 - E[N_d(5)]}{\sqrt{Var(N_d(5))}}\right) \\ &\approx P\left(N(0, 1) > \frac{120 - E[N_d(5)]}{\sqrt{Var(N_d(5))}}\right) \end{aligned}$$

$$= P\left(N(0,1) > \frac{120 - 100]}{\sqrt{100}}\right) = P(N(0,1) > 2.0) \approx 0.023$$

(You are not required to know or calculate the final normal probability, but actually you should have some idea, because it is such a common case.)

(d) Let $C(t)$ be the number of cars that are detected by the detection device up to time t and remain in the highway segment at time t. Give an expression for the probability distribution of $C(4)$?

We use the $M_t/GI/\infty$ queueing model. Let S be the random time each car remains in the highway segment. From the "Physics paper" or from the textbook, we know that $C(t)$ has a Poisson distribution for each t with mean

$$m(t) = E[C(t)] = \int_0^t \lambda(s) G^c(t-s)\, ds,$$

where $\lambda(s) = 8s$, $G(x) \equiv P(S \le x) = (x-1)/2$ $(1 \le x \le 3)$, $G(x) = 0$ $(x < 1)$, $G(x) = 1$ $(x > 3)$, and $G^c(x) \equiv P(S > x) = 1 - G(x)$. Hence,

$$\begin{aligned} m(4t) &\equiv E[C(4)] = \int_0^4 8sG^c(4-s)\, ds \\ &= \int_1^3 8s\frac{(s-1)}{2}\, ds + \int_3^4 8s\, ds = \int_1^3 (4s^2 - 4s)\, ds + 28 \\ &= [(4/3)(3^3) - 2(3)^2 - (4/3)(1^3) + 2(1)^2] + 28 = [36 - 18 - (4/3) + 2] + 28 = 46\frac{2}{3} \end{aligned}$$

(e) What is the covariance between $C(3)$ and $C(6)$?

The covariance is 0 because the random variables are independent. See Theorem 2 of the Physics paper, but notice that the service times are uniform on $[1,3]$ here. Since the service times are at most 3, the $C(3)$ cars will all be gone by time 6. We have the covariance expression $Cov(A+B, B+C) = Cov(B,B) = Var(B) = E[B]$, where A, B and C are the numbers of events in the three regions in the plane in Figure 3 of the Physics paper, but necessarily $B = 0$.

(f) Let $D(t)$ be the number of cars that are detected by the detection device and have departed by time t. Give an expression for the joint distribution of $C(4)$ and $D(4)$.

By Theorem 1 of the "Physics paper," the random variables $C(4)$ and $D(4)$ are independent Poisson random variables. Clearly, $C(t) + D(t) = N_d(t)$, where

$$E[N_d(4)] = \int_0^4 8t\, dt = 64,$$

reasoning as in part (a). Hence, $E[D(4)] = 64 - E[C(4)]$, but by part (d) $E[C(4) = \frac{54}{3}$. so that

$$E[D(4)] = 64 - \frac{54}{3} = \frac{192 - 54}{3} = \frac{138}{3}.$$

Thus, finally we can write

$$P(C(4) = j, D(4) = k) = P(C(4) = j)P(D(4) = k) = \left(\frac{e^{-(54/3)}(54/3)^j}{j!}\right)\left(\frac{e^{-(138/3)}(138/3)^k}{k!}\right)$$