IEOR 6711: Stochastic Models I

Fall 2012, Professor Whitt

Topic for Discussion, Thursday, September 27

Compound Poisson Process

1. Withdrawals from an ATM

Exercise 5.88 in Green Ross. Customers arrive at an automatic teller machine (ATM) in accordance with a Poisson process with rate 12 per hour. The amount of money withdrawn on each transaction is a random variable with mean \$30 and standard deviation \$50. (A negative withdrawal means that money was deposited.) Suppose that the machine is in use 15 hours per day. Approximate the probability that the total daily withdraw is less than \$6000.

ANSWER: Let $X(t)$ be the total amount withdrawn in the interval $[0,t]$, where time t is measured in hours. Assuming that the successive amounts withdrawn are independent and identically distributed random variables, the stochastic process $\{X(t) : t \geq 0\}$ is a compound Poisson process. Let $X(15)$ denote the daily withdrawal. Its mean and variance can be calculated as follows using the equations on top of p. 83.

$$E[X(15)] = 12 \times 15 \times 30 = 5400 \quad \text{and} \quad Var[X(15)] = 12 \times 15 \times [30 \times 30 + 50 \times 50] = 612,000$$

Now using the CLT for the compound Poisson process, with $\sqrt{612,000} \approx 782$ and $600/782 \approx 0.767$, we obtain the approximating probability

$$P(X(15) \leq 6000) = P\left(\frac{X(15) - 5400}{\sqrt{612000}} \leq \frac{600}{\sqrt{612000}}\right) = P(N(0,1) \leq 0.767) = 0.78,$$

where $N(0,1)$ is a standard normal random variable. We have used a table of the standard normal distribution for the actual numerical value.

2. Definition of a compound Poisson process

Let $N \equiv \{N(t) : t \geq 0\}$ be a Poisson process with rate λ, so that $E[N(t)] = \lambda t$ for $t \geq 0$. Let $X_1, X_2, \cdots$ be IID random variables independent of N. Let $D(t)$ be the random sum

$$D(t) \equiv \sum_{i=1}^{N(t)} X_i, \quad t \geq 0 \ . \tag{1}$$

Then $D \equiv \{D(t) : t \geq 0\}$ is a *compound Poisson process.* See Chapter 2 of Ross.

3. Lévy process

Theorem 0.1 *The stochastic process D is a Lévy process, i.e., it has stationary independent increments.*

Proof. An increment of D is $D(t_2) - D(t_1)$ for $t_2 > t_1 > 0$. The independent-increments property is: For all k and for all time points $0 < t_1 < t_2 < \cdots < t_k$, the k increments $D(t_{i+1}) - D(t_i)$ are k mutually independent random variables. The stationary-increment property is: The distribution of $D(t_2 + h) - D(t_1 + h)$ for $t_2 > t_1 > 0$ and $h > 0$ is independent of h, and similarly for the joint distribution of k increments. We prove the special case of the independent-increments property for $k = 2$; the general case k is proved in the same way. It suffices to show that

$$P(D(t_1) \le a_1, D(t_2) - D(t_1) \le a_2) = P(D(t_1) \le a_1) P(D(t_2) - D(t_1) \le a_2) \tag{2}$$

for all $a_1 > 0$ and $a_2 > 0$ (for all $0 < t_1 < t_2$). To do so, just apply the definition of $D(t)$ and condition on the values of the counting process for the times t_1 and t_2. In particular,

$$P(D(t_1) \le a_1, D(t_2) - D(t_1) \le a_2) = P\left(\sum_{i=1}^{N(t_1)} X_i \le a_1, \sum_{i=1}^{N(t_2)} X_i - \sum_{i=1}^{N(t_1)} X_i \le a_2\right)$$

$$\begin{aligned}
&= P\left(\sum_{i=1}^{N(t_1)} X_i \le a_1, \sum_{i=N(t_1)+1}^{N(t_2)} X_i \le a_2\right) \quad \text{apply (1)} \\
&= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} P\left(\sum_{i=1}^{m_1} X_i \le a_1, \sum_{i=m_1+1}^{m_1+m_2} X_i \le a_2 | N(t_1) = m_1, N(t_2) - N(t_1) = m_2\right) \\
&\qquad \times P(N(t_1) = m_1, N(t_2) - N(t_1) = m_2) \quad \text{(condition)} \\
&= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} P\left(\sum_{i=1}^{m_1} X_i \le a_1, \sum_{i=m_1+1}^{m_1+m_2} X_i \le a_2\right) P\left(N(t_1) = m_1, N(t_2) - N(t_1) = m_2\right) \\
&= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} P\left(\sum_{i=1}^{m_1} X_i \le a_1\right) P\left(\sum_{i=m_1+1}^{m_1+m_2} X_i \le a_2\right) P\left(N(t_1) = m_1\right) P\left(N(t_2) - N(t_1) = m_2\right) \\
&= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} P\left(\sum_{i=1}^{m_1} X_i \le a_1\right) P\left(\sum_{i=1}^{m_2} X_i \le a_2\right) P\left(N(t_1) = m_1\right) P\left(N(t_2 - t_1) = m_2\right) \\
&= \sum_{m_1=0}^{\infty} P\left(\sum_{i=1}^{m_1} X_i \le a_1\right) P\left(N_1(t) = m_1\right) \sum_{m_2=0}^{\infty} P\left(\sum_{i=1}^{m_2} X_i \le a_2\right) P\left(N(t_2 - t_1) = m_2\right) \\
&= P\left(\sum_{i=1}^{N(t_1)} X_i \le a_1\right) P\left(\sum_{i=1}^{N(t_2-t_1)} X_i \le a_2\right) \\
&= P\left(D(t_1) \le a_1\right) P\left(D(t_2 - t_1) \le a_2\right) = P\left(D(t_1) \le a_1\right) P\left(D(t_2) - D(t_1) \le a_2\right) ,
\end{aligned} \tag{3}$$

using independence in line 5 and stationarity in line 6. ∎

We now state and prove the central limit theorem (CLT) for compound Poisson processes. We use the fact that

$$E[D(t)] = \lambda t m_1 \quad \text{and} \quad Var(D(t)) = \lambda t m_2 , \tag{4}$$

where $m_i \equiv E[X_1^i]$. Note in particular that m_2 in (4) is the second moment, not the variance of X_1.

4. Variance formula

We pause to prove the variance formula in (4) above. To do so, we use the conditional variance formula; see Exercise 1.22 of Ross. We use the fact that $E[E[X|Y]] = E[X]$.

Theorem 0.2 *Suppose that X is a random variable with finite second moment $E[X^2]$. (Aside: As a consequence $E[|X|] < \infty$, and for any random variable Y, $E[X^2|Y]$ and $E[X|Y]$ are well* *defined with* $E[E[X^2|Y]] = E[X^2] < \infty$ *and* $E[E[|X||Y]] = E[|X|] < \infty$.*) Then*

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

where $Var(X|Y)$ *is defined to mean*

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

Proof. Find expressions for the two terms on the right:

$$E[Var(X|Y)] = E[E[X^2|Y]] - E[(E[X|Y])^2] = E[X^2] - E[(E[X|Y])^2]$$

and

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

Combining these last two expressions gives the result. ■

Now we consider the implications for a random sum of i.i.d. random variables.

Theorem 0.3 *Suppose that* N *is a nonnegative integer-valued random variable with finite second moment* $E[N^2]$. *Let* $\{X_n : n \geq 1\}$ *be a sequence of i.i.d random variables with mean* $\mu = E[X_1]$ *and finite variance* $\sigma^2 = Var(X_1)$. *Suppose that* N *is independent of the sequence* $\{X_n : n \geq 1\}$. *Then*

$$Var\left(\sum_{n=1}^{N} X_n\right) = \sigma^2 E[N] + \mu^2 Var(N) \ .$$

Proof. We apply the conditional variance formula. Let S denote the random sum in question. Then

$$Var(S|N = n) = n\sigma^2 \quad \text{and} \quad E[S|N = n] = n\mu \ .$$

Therefore,

$$Var(S|N) = N\sigma^2 \quad \text{and} \quad E[S|N] = N\mu \ .$$

Then the conditional variance formula gives

$$Var(S) = E[Var(S|N)] + Var(E[S|N]) = E[N]\sigma^2 + Var(N)\mu^2 \ . \quad \blacksquare$$

Corollary 0.1 *If* N *is Poisson with mean* λ, *then*

$$Var(S) = \lambda\sigma^2 + \lambda\mu^2 = \lambda E[X_1^2] \ . \quad \blacksquare$$

For the compound Poisson process, $D(t)$ *has mean* λt, *so that*

$$Var(D(t)) = \lambda t\sigma^2 + \lambda t\mu^2 = \lambda t E[X_1^2] \ . \quad \blacksquare$$

5. Central Limit Theorem (CLT)

Theorem 0.4 *If* $m_2 < \infty$, *then*

$$\frac{D(t) - \lambda t m_1}{\sqrt{\lambda t m_2}} \Rightarrow N(0,1) \quad as \quad t \to \infty \ . \tag{5}$$

Proof. The simple high-level proof is to apply the standard CLT after observing that D is a process with stationary and independent increments. For any t, we can thus represent $D(t)$ is the sum of n IID increments, each distributed as $D(t/n)$. By the condition $m_2 < \infty$, and the variance formula in (4), these summands have finite variance and second moment. There is a bit of technical difficulty in writing a full proof, because we want to let $t \to \infty$ in the desired statement. If we divide the interval $[0, t]$ into n subintervals and let $n \to \infty$, leaving t fixed, then we obviously do not get the desired convergence; then the sum of the n increments is $D(t)$ for all n; we do not get larger t.

What we want to do now is give a direct proof using transforms. The cleaner mathematical approach is to use characteristic functions, which involves complex variables, but for simplicity we will use *moment generating functions* (mgf's). We would use characteristic functions, because a mgf is not always well defined (finite). However, we assume that the mgf of X is well defined (finite) for all positive θ); i.e.,

$$\psi_X(\theta) \equiv E[e^{\theta X}] = \int_{-\infty}^{\infty} e^{\theta x} f_X(x)\, dx < \infty \tag{6}$$

for all positive θ. A sufficient condition is for X to be a bounded random variable, but that is stronger than needed. The condition would be satisfied if X were normally distributed, for example. However, it would suffice for the mgf to be defined for all θ with $|\theta| < \epsilon$ for some $\epsilon > 0$. The argument is essentially as follows for characteristic functions.

We have acted as if X has a probability density function (pdf) $f_X(x)$; that is not strictly necessary. The general idea is that the convergence in distribution in (5) is equivalent to the convergence of mgf's for all θ. (That means the usual pointwise convergence of a sequence of functions.) We are using a *continuity theorem for mgf's*: There is convergence in distribution of random variables if and only if there is convergence of the mgf's for all θ. What we want to show then is that

$$\psi_{[D(t)-\lambda t m_1]/\sqrt{\lambda t m_2}}(\theta) \to \psi_{N(0,1)}(\theta) = e^{\theta^2/2} \quad \text{as} \quad t \to \infty \tag{7}$$

for all θ.

Just as in the proof of the ordinary CLT (in previous lecture notes), we use a Taylor expansion of the component mgf $\psi_X(\theta)$ about $\theta = 0$. We can write

$$\psi_X(\theta) = 1 + \theta m_1 + \frac{\theta^2 m_2}{2} + o(\theta^2) \quad \text{as} \quad \theta \downarrow 0 \ . \tag{8}$$

(Recall that $o(\theta)$ is a quantity $f(\theta)$ such that $f(\theta)/\theta \to 0$ as $\theta \to 0$. Thus $o(\theta^2)$ is a quantity $f(\theta)$ such that $f(\theta)/\theta^2 \to 0$ as $\theta \to 0$. The last term in (8) is asymptotically negligible compared to the previous terms in the limit.)

We also exploit the mgf of $D(t)$, which is discussed at length in Ross, but perhaps not enough. We can use the transform of $D(t)$ to compute the exact distribution numerically by applying numerical transform inversion. If $X \geq 0$, then $D(t) \geq 0$ and we can work with Laplace transforms. Then you can apply the inversion algorithm you are writing.

There is a general form for the mgf of an integer random sum that is important to know about. Thus, suppose that N is an arbitrary integer-valued random variable and let $\{X_i : i \geq 1\}$ be a sequence of IID random variables independent of N. (This is our case with $N = N(t)$ for some t.) The important point is the the mgf of the random sum is the *composition* of the generating function of N and the mgf of X; i.e., the generating function is evaluated at the mgf of X. Let $\tilde{g}_N(z)$ be the *generating function* of the random variable N; i.e.,

$$\tilde{g}_N(z) \equiv \sum_{n=0}^{\infty} z^n P(N = n) \ . \tag{9}$$

Note that the generating function of the Poisson distribution is an exponential function: If N has a Poisson distribution with mean α, then

$$\tilde{g}_N(z) \equiv \sum_{n=0}^{\infty} z^n \frac{\alpha^n e^{-\alpha}}{n!} = e^{\alpha(z-1)} \tag{10}$$

So, in our case,

$$\tilde{g}_{N(t)}(z) = e^{\lambda t(z-1)} \ . \tag{11}$$

Now we want to compute the mgf of $D(t)$.

$$\begin{aligned}
\psi_{D(t)}(\theta) &\equiv E[e^{\theta D(t)}] \\
&= E[\exp \theta \sum_{i=1}^{N(t)} X_i] \\
&= \sum_{m=1}^{\infty} E[\exp \theta \sum_{i=1}^{m} X_i] P(N(t) = m) \\
&= \sum_{m=1}^{\infty} E[\exp \theta \sum_{i=1}^{m} X_i] \frac{(\lambda t)^m e^{-\lambda t}}{m!} \\
&= \sum_{m=1}^{\infty} E[e^{\theta X_1}]^m \frac{(\lambda t)^m e^{-\lambda t}}{m!} \\
&= \sum_{m=1}^{\infty} \psi_X(\theta)^m \frac{(\lambda t)^m e^{-\lambda t}}{m!} \\
&= e^{\lambda t(\psi_X(\theta)-1)} \ .
\end{aligned} \tag{12}$$

Now we combine all the results above to carry out the desired proof. Note that we have derived the mgf of $D(t)$, but we want to work with the mgf of the scaled random variable $[D(t) - \lambda t m_1]/\sqrt{\lambda t m_2}$. We thus need to calculate the mgf of the scaled random variable, but that is easy. It just requires being careful (accounting).

$$\begin{aligned}
\psi_{[D(t)-\lambda t m_1]/\sqrt{\lambda t m_2}}(\theta) &= E[e^{\frac{\theta D(t)}{\sqrt{\lambda t m_2}}}] e^{\frac{-\theta \lambda t m_1}{\sqrt{\lambda t m_2}}} \\
&= e^{\lambda t(\psi_X(\theta/\sqrt{\lambda t m_2})-1)} e^{\frac{-\theta \lambda t m_1}{\sqrt{\lambda t m_2}}} \\
&= e^{\lambda t(\psi_X(\theta/\sqrt{\lambda t m_2})-1-(\theta \lambda t m_1/\sqrt{\lambda t m_2}))} \ .
\end{aligned} \tag{13}$$

Note that we obtain a simple exponential function as a final expression. It suffices to do asymptotics for the exponent of this exponential function. We now insert the Taylor expansion in (8) for $\psi_X(\theta/\sqrt{\lambda t m_2})$ as $t \to \infty$. In other words, we replace the argument θ by $\theta/\sqrt{\lambda t m_2}$ and let $t \to \infty$. Note that, as $t \to \infty$, the argument of the mgf is going to 0, so the Taylor expansion applies. Starting to consider only the first term of the exponent only, we have

$$\begin{aligned}
\psi_X(\theta/\sqrt{\lambda t m_2}) &= 1 + (\theta/\sqrt{\lambda t m_2}) m_1 + (\theta/\sqrt{\lambda t m_2})^2 (m_2/2) + o((\theta/\sqrt{\lambda t m_2})^2) \\
&= 1 + (\theta/\sqrt{\lambda t m_2}) m_1 + (\theta^2/2\lambda t) + o((1/t)) \quad \text{as} \quad t \to \infty \ .
\end{aligned} \tag{14}$$

Now note that (when exponentials are used), the first two terms in (14) exactly cancel the last two terms on the right in the final exponent of (13), and we are left with

$$\psi_{[D(t)-\lambda t m_1]/\sqrt{\lambda t m_2}}(\theta) = e^{(\theta^2/2)+o(1)} \to e^{(\theta^2/2)} \quad \text{as} \quad t \to \infty \ . \quad \blacksquare \tag{15}$$