Non-negative Integer Valued Random Variables

H. Krieger, Mathematics 156, Harvey Mudd College

Fall, 2008

Let N be a random variable with values from the set $\{+\infty, 0, 1, 2, \ldots\}$. Then N is referred to as a non-negative integer valued random variable. Letting $p_k = P(N = k)$ for $k = 0, 1, 2, \ldots$, we see that

$$\sum_{k=0}^{\infty} p_k = P(N < +\infty) \text{ and } P(N = +\infty) = 1 - \sum_{k=0}^{\infty} p_k.$$

If $P(N = +\infty) = 0$, then N is said to be finite with probability one or simply finite valued.

Convolutions: If N and M are *independent* non-negative integer valued random variables, with $p_k = P(N = k)$ for $k = 0, 1, 2, \ldots$ and $q_j = P(M = j)$ for $j = 0, 1, 2, \ldots$, then $N + M$ is another non-negative integer valued random variable with distribution given by

$$r_i = P(N + M = i) = \sum_{k+j=i} p_k q_j = \sum_{k=0}^{i} p_k q_{i-k}$$

for $i = 0, 1, 2, \ldots$. In this case we write $\{r_k\} = \{p_k\} * \{q_k\}$ and say that $\{r_k\}$ is the *convolution* of $\{p_k\}$ and $\{q_k\}$.

Expectations: If $P(N = +\infty) > 0$, then $E(N) = +\infty$. Otherwise,

$$E(N) = \sum_{k=0}^{\infty} k p_k = \sum_{k=1}^{\infty} k P(N = k).$$

Note that $0 \le E(N) \le +\infty$, with $E(N) = 0$ if and only if $P(N > 0) = 0$. Moreover, we have the alternate formula

$$E(N) = \sum_{k=1}^{\infty} P(N \ge k) = \sum_{k=0}^{\infty} P(N > k),$$

which also holds even if $P(N = +\infty) > 0$.

Stopping Times: Let $\{X_n : n = 1, 2, 3, \ldots\}$ be a sequence of random variables defined on a space Ω and let τ be a non-negative integer valued random variable also defined on Ω. Then τ is called a stopping time with respect to $\{X_n\}$ iff $\forall n \geq 0$ the event $\{\tau = n\}$ is independent of $\{X_k : k > n\}$. In other words, the event $\{\tau = n\}$ can depend at most on the random variables $\{X_1, X_2, \dots, X_n\}$. Note in particular that the event $\{\tau \geq n\} = \{\tau < n\}^c$ is independent of X_n.

Clearly, τ is a stopping time if τ is independent of all of the $\{X_n\}$. Another common way to construct a stopping time is to let A be an event in the range of the $\{X_n\}$ and let τ_A be the *hitting time* of A, namely:

$$\tau_A(\omega) = \inf\{n : X_n(\omega) \in A\}.$$

One of the most important results involving stopping times is the following:

Wald's Lemma: Suppose the $\{X_n\}$ are identically distributed with common finite mean $E(X)$. Let τ be a stopping time with respect to $\{X_n\}$ with $E(\tau) < \infty$. Define the *random sum* $\sum_{k=1}^{\tau} X_k$ to be 0 when $\tau = 0$ and to be $\sum_{k=1}^{n} X_k$ when $\tau = n \geq 1$. Then,

$$E\left(\sum_{k=1}^{\tau} X_k\right) = E(X)E(\tau).$$

Proof: We give the argument only for the case in which the $\{X_n\}$ are non-negative (in which case the finiteness of the expectations is irrelevant), so that interchanges of infinite sums and expectations can always be done. Therefore,

$$\begin{aligned} E\left(\sum_{k=1}^{\tau} X_k\right) &= E\left(\sum_{k=1}^{\infty} X_k 1_{\{k \leq \tau\}}\right) \\ &= \sum_{k=1}^{\infty} E\left(X_k 1_{\{k \leq \tau\}}\right) \\ &= \sum_{k=1}^{\infty} E(X_k) E(1_{\{k \leq \tau\}}) \end{aligned}$$

since $\{\tau \geq k\}$ is independent of X_k. But the $\{X_n\}$ are identically distributed with common mean $E(X)$ and hence

$$E\left(\sum_{k=1}^{\tau} X_k\right) = E(X) \sum_{k=1}^{\infty} P(\tau \geq k) = E(X)E(\tau).$$

Generating Functions: For $0 \leq s \leq 1$, let

$$\Phi(s) = p_0 + s p_1 + s^2 p_2 + \cdots = \sum_{k=0}^{\infty} s^k p_k.$$

Then $\Phi(s)$ is called the generating function of the sequence $\{p_k\}$ or, less precisely, of the random variable N. Note that, for $0 < s < 1$, $\Phi(s) = E(s^N)$ since in this case $s^{+\infty} = 0$. Furthermore, note that $\Phi(0) = p_0$ and $\Phi(1) = \sum_{k=0}^{\infty} p_k = P(N < +\infty)$. Since each $p_k \geq 0$ and $0 \leq \sum_{k=0}^{\infty} p_k \leq 1$, this power series converges uniformly and is infinitely differentiable for $0 \leq s < 1$. Generating functions have the following properties:

1. $\Phi(s)$ uniquely determines $\{p_k\}$, since $\Phi^{(k)}(0)/k! = p_k$ for $k = 0, 1, 2, \ldots$.

2. For $0 \leq s < 1$, $\Phi'(s) = \sum_{k=1}^{\infty} k s^{k-1} p_k$, so that as $s \uparrow 1$, $\Phi'(s) \uparrow \sum_{k=1}^{\infty} k p_k$, which is $E(N)$ in the case that N is finite valued. But in that case, we also see that

$$\begin{aligned} \frac{1-\Phi(s)}{1-s} &= \frac{\Phi(s)-\Phi(1)}{s-1} \\ &= \sum_{k=1}^{\infty}(1+s+\cdots+s^{k-1})p_k \\ &= \sum_{j=0}^{\infty} s^j \sum_{k=j+1}^{\infty} p_k \\ &= \sum_{j=0}^{\infty} s^j P(N > j). \end{aligned}$$

   Thus $\frac{1-\Phi(s)}{1-s}$ is the generating function for the sequence $\{P(N > k)\}$ and we have

$$\frac{\Phi(s)-\Phi(1)}{s-1} \uparrow E(N) \text{ as } s \uparrow 1.$$

   In other words, Φ is differentiable from below at 1 with

$$\Phi'(1) = \lim_{s \uparrow 1} \Phi'(s) = E(N).$$

3. Similarly, for $n > 1$, we get

$$\Phi^{(n)}(1) = \lim_{s \uparrow 1} \Phi^{(n)}(s) = E(N(N-1)\cdots(N-n+1)),$$

   so, for example:

$$\Phi''(1) = E(N(N-1)) = E(N^2) - E(N).$$

4. If N and M are *independent* non-negative integer valued random variables, with generating functions Φ_N and Φ_M respectively, then the generating function Φ_{N+M} of $N+M$ is given by $\Phi_{N+M}(s) = \Phi_N(s)\Phi_M(s)$. To see this note that

$$\Phi_{N+M}(s) = E(s^{N+M}) = E(s^N s^M) = E(s^N)E(s^M) = \Phi_N(s)\Phi_M(s).$$

   In other words, the generating function of the convolution of two sequences is the product of the generating functions of the individual sequences, a familiar result for many transforms.

5. Suppose $\{X_n\}$ are independent, identically distributed, finite non-negative integer valued random variables which have a common generating function $\Phi_{X_1}(s) = E(s^{X_1})$. Let N be non-negative integer valued, independent of the $\{X_n\}$, with generating function $\Phi_N(s) = E(s^N)$. Then the random sum $S = \sum_{k=1}^{N} X_k$ has generating function Φ_S given by

$$\Phi_S(s) = \Phi_N\left(\Phi_{X_1}(s)\right).$$

This result follows by conditioning on the value of N and using its independence from the $\{X_n\}$ as follows:

$$\begin{aligned}
\Phi_S(s) &= E(s^{\sum_{k=1}^{N} X_k}) = P(N=0) + \sum_{n=1}^{\infty} E(s^{\sum_{k=1}^{N} X_k} | N=n) P(N=n) \\
&= P(N=0) + \sum_{n=1}^{\infty} E(s^{\sum_{k=1}^{n} X_k} | N=n) P(N=n) \\
&= P(N=0) + \sum_{n=1}^{\infty} E(s^{\sum_{k=1}^{n} X_k}) P(N=n) \\
&= P(N=0) + \sum_{n=1}^{\infty} \left[\Phi_{X_1}(s)\right]^n P(N=n) = \Phi_N\left(\Phi_{X_1}(s)\right).
\end{aligned}$$

One important use of this result is to give an alternate proof of Wald's Lemma in this case. Using the chain rule we see that

$$E(S) = \Phi_S'(1) = \Phi_N'\left(\Phi_{X_1}(1)\right)\Phi_{X_1}'(1) = \Phi_N'(1)\Phi_{X_1}'(1) = E(N)E(X_1).$$