Math 5010 § 1. **Solutions to Seventh Homework**
Treibergs March 6, 2009

151[25] *Suppose that the probability of an insect laying n eggs is given by the Poisson distribution with mean $\mu > 0$, that is, by the probability distribution given over all the nonnegative integers defined by $e^{-\mu}\mu^n/n!$, $n \in D = \{0, 1, 2, 3, \ldots\}$. Suppose further, that the probability of an egg developing is p. Assuming mutual independence of the eggs, show that the probability distribution $f_Y(y)$ for the probability that there are y survivors is also of Poisson type, and find the mean.*

We are given two random variables, X, the number of eggs laid and Y, the number of eggs that survive. Both X and Y take values in D. The probability that $n \in D$ eggs are laid is

$$f_X(n) = \mathbf{P}(\{X = n\}) = \frac{e^{-\mu}\mu^n}{n!}$$

Given that n eggs are laid, then since each of the eggs survive independently, the number of these that survive is a binomial variable, so that the conditional probability that y survive is given by

$$f_Y(y|\{X = n\}) = \mathbf{P}(\{Y = y\}|\{X = n\}) = \begin{cases} \binom{n}{y} p^y\, q^{n-y}, & \text{if } y \in \{0, 1, 2, \ldots, n\}; \\ \\ 0, & \text{otherwise.} \end{cases}$$

Note that the number that survive can't be more than the number of eggs laid. Now condition on the number of eggs laid. That is, the sets $\{X = n\}$ for $n \in D$ partition Ω, that is they are mutually disjoint and exhaustive. We use the fact that the conditional mass function is zero if $y > n$ because more can't survive than are laid, $\mathbf{P}(\{Y = y\}|\{X = n\}) = 0$ if $y > n$. By the partitioning formula, if $y \in D$,

$$\begin{aligned} f_Y(y) = \mathbf{P}(\{Y = y\}) &= \mathbf{P}\left(\bigcup_{n=0}^{\infty}(\{Y = y\} \cap \{X = n\})\right) \\ &= \sum_{n=0}^{\infty} \mathbf{P}\left(\{Y = y\} \cap \{X = n\}\right) \\ &= \sum_{n=0}^{\infty} \mathbf{P}\left(\{Y = y\}|\{X = n\}\right)\, \mathbf{P}(\{X = n\}) \\ &= \sum_{n=y}^{\infty} \binom{n}{y} p^y\, q^{n-y}\, \frac{e^{-\mu}\mu^n}{n!} \\ &= \frac{e^{-\mu}\, p^y\, \mu^y}{y!} \sum_{n=y}^{\infty} \frac{q^{n-y}\mu^{n-y}}{(n-y)!} \\ &= \frac{e^{-\mu}\, p^y\, \mu^y}{y!} e^{q\mu} = \frac{e^{-(1-q)\mu}\, p^y\, \mu^y}{y!} = \frac{e^{-p\mu}\, (p\mu)^y}{y!}. \end{aligned}$$

Thus we see that the distribution is also Poisson, but this time the parameter is $p\mu$ instead of μ. Since the mean of a Poisson distribution is the parameter we get

$$\mathbf{E}(Y) = p\mu.$$

This should not come as a surprise. It says that if an insect lays on average $\mathbf{E}(X) = \mu$ eggs in a given period and p is the survival rate, then there should be on average $p\mu$ eggs surviving in the same period.

[A.] *Suppose that* $X \sim \text{Geom}(p)$ *is a geometric random variable with parameter* p*. Find*

(a) $\mathbf{P}(X \text{ is odd})$*;*

(b) $\mathbf{P}(X \text{ is even})$*;*

(c) $\mathbf{P}(X > k)$*;*

(d) Let k *be an integer such that* $1 \le k \le n$*. Find* $\mathbf{P}(X = k | X \le k)$*;*

(e) $\mathbf{P}(2 \le X \le 9 | X \ge 4)$*;*

(f) Let $k \in \mathbb{N}$*. Let* $g(x) = \min(x, k)$ *and* $Y = g(X)$*. Find the pmf* $f_Y(y)$ *and the expectation* $\mathbf{E}(Y)$*;*

(g) $\mathbf{E}(1/X)$*.*

The standard picture of a geometric variable is a sequence of independent coin flips where the probability of head is p and X is the number of flips to get the first head. It takes values in the natural numbers $D = \mathbb{N} = \{1, 2, 3, \ldots\}$, and its pmf for $x \in D$ is

$$f_X(x) = \mathbf{P}(X = x) = p\, q^{x-1}.$$

(a.) The event that X is odd is given by

$$\{X \text{ is odd}\} = \{X = 1\} \cup \{X = 3\} \cup \{X = 5\} \cup \cdots = \bigcup_{k=0}^{\infty} \{X = 2k + 1\}$$

Since these are disjoint events, we may add using the geometric sum $\sum_{k=0}^{\infty} r^k = (1 - r)^{-1}$ with $r = q^2$,

$$\mathbf{P}(X \text{ is odd}) = \sum_{k=0}^{\infty} \mathbf{P}(X = 2k + 1) = \sum_{k=0}^{\infty} p\, q^{2k} = \frac{p}{1 - q^2} = \frac{1}{1 + q}.$$

(b.) The complementary event is

$$\mathbf{P}(X \text{ is even}) = \mathbf{P}(\{X \text{ is odd}\}^c) = 1 - \mathbf{P}(X \text{ is odd}) = 1 - \frac{1}{1 + q} = \frac{q}{1 + q}.$$

(c.) The event that X is greater than k is

$$\{X > k\} = \bigcup_{x \,\in\, D \text{ and } x \,>\, k} \{X = x\}$$

Since these are disjoint events, we may add using the geometric sum. If $k + 1 \in D$,

$$\begin{aligned}\mathbf{P}(X > k) &= \sum_{x=k+1}^{\infty} p\, q^{x-1} = p\left(q^k + q^{k+1} + q^{k+2} + \cdots\right) \\ &= p\, q^k \left(1 + q + q^2 + \cdots\right) = \frac{p\, q^k}{1 - q} = q^k.\end{aligned}$$

Note that this implies for $k \in D$, $\mathbf{P}(X \ge k) = \mathbf{P}(X > k - 1) = q^{k-1}$. Also the cumulative distribution function for $x \in D$,

$$F_X(x) = \mathbf{P}(X \le x) = \mathbf{P}(\{X > x\}^c) = 1 - \mathbf{P}(X > x) = 1 - q^x. \tag{1}$$

(d.) If k is an integer so that $1 \le k \le n$, then the event

$$\{X = k\} \subset \{X \le n\}.$$

Compute conditional probabilities as usual using (1)

$$\mathbf{P}(X = k | X \le n) = \frac{\mathbf{P}(\{X = k\} \cap \{X \le n\})}{\mathbf{P}(X \le n)} = \frac{\mathbf{P}(X = k)}{\mathbf{P}(X \le n)} = \frac{p\, q^{k-1}}{1 - q^n}.$$

(e.) Using the fact that the latter events are disjoint

$$\{X > 3\} = \{4 \le X \le 9\} \cup \{X > 9\}$$

we get the probability using (c.),

$$\mathbf{P}(4 \le X \le 9) = \mathbf{P}(X > 3) - \mathbf{P}(X > 9) = q^3 - q^9.$$

Equivalently, $\mathbf{P}(4 \le X \le 9) = \mathbf{P}(X \le 9) - \mathbf{P}(X \le 3) = F_X(9) - F_X(3)$. We compute conditional probabilities as usual:

$$\mathbf{P}(2 \le X \le 9 | X \ge 4) = \frac{\mathbf{P}(\{2 \le X \le 9\} \cap \{X \ge 4\})}{\mathbf{P}(X \ge 4)} = \frac{\mathbf{P}(4 \le X \le 9)}{\mathbf{P}(X > 3)} = \frac{q^3 - q^9}{q^3} = 1 - q^6.$$

(f.) For the natural number k, the function

$$g(x) = \min(x, k) = \begin{cases} x, & \text{if } x < k; \\ k, & \text{if } x \ge k. \end{cases}$$

g maps D to $D' = \{1, 2, 3, \ldots, k\}$. Thus if $y \in D'$ and $y < k$ then there is exactly one $x \in D$ such that $y = g(x)$, namely, $x = y$. If $y = k$ then the set of x's that map to k is $\{k, k+1, k+2, \ldots\} = \{X \ge k\}$. This set is also called the preimage $g^{-1}(\{k\})$. Using the formula for the pmf of the new random variable $Y = g(X)$ we have

$$f_Y(y) = \sum_{x \,\in\, D \text{ such that } g(x) \,=\, y} f_X(x).$$

For this $g(X)$, using (c.), it becomes for $y \in D'$,

$$f_Y(y) = \begin{cases} f_X(y) = p\, q^{y-1}, & \text{if } y < k; \\ \sum_{x=k}^{\infty} f_X(x) = \mathbf{P}(X \ge k) = q^{k-1}, & \text{if } x = k. \end{cases}$$

Of course, $f_Y(y) = 0$ if $y \notin D'$.

To find the expectation we may use the definition or Theorem 4.3.4, which give the same expression. Thus if $k > 1$,

$$\begin{aligned} \mathbf{E}(Y) &= \sum_{y \in D'} y\, f_Y(y) = \left(\sum_{y=1}^{k-1} y\, p\, q^{y-1} \right) + kq^{k-1} = p \frac{d}{dq} \left(\sum_{y=0}^{k-1} q^y \right) + kq^{k-1} \\ &= p \frac{d}{dq} \left(\frac{1 - q^k}{1 - q} \right) + kq^{k-1} = p \left(\frac{-kq^{k-1}}{1 - q} + \frac{1 - q^k}{(1 - q)^2} \right) + kq^{k-1} = \frac{1 - q^k}{p}. \end{aligned}$$

If $k = 1$ there is one term and $E(Y) = 1$ so the formula works for $k \ge 1$ as well. Since $g(x) \le x$, it is no surprise that this is close to but less than $\mathbf{E}(X) = 1/p$.

(g.) To find the expectation of $Z = h(X)$ where $h(x) = 1/x$, we use Theorem 4.3.4.

$$\mathbf{E}(Z) = \sum_{x \in D} h(x)\, f_X(x) = \sum_{x=1}^{\infty} \frac{p\, q^{x-1}}{x} = \frac{p}{q} \sum_{x=1}^{\infty} \frac{q^x}{x} = -\frac{p\, \log(1 - q)}{q} = -\frac{p\, \log p}{q}.$$

See problem 151[18] and page 23.

[B.] *Suppose that an unfair coin is tossed repeatedly. Suppose that the tosses are independent and the probability of each head is p. Let X denote the number of tosses it takes to get three heads. Derive the formulas for* $\mathbf{E}(X)$ *and* $\mathbf{Var}(X)$.

X is distributed according to the negative binomial distribution with parameters $k = 3$ and p. The variable takes values in $D = \{k, k+1, k+2, \ldots\}$ since one must toss the coin k times at least in order to have k heads. In order that the k-th head occur at the x-th toss, there must be $k-1$ heads in the first $x-1$ tosses and the x-th toss has to be a head. Thus the negative binomial pmf is for $x \in D$,

$$f_X(x) = \binom{x-1}{k-1} p^k\, q^{x-k}.$$

The expectation is given by the sum

$$\mathbf{E}(X) = \sum_{x \in D} x\, f_X(x) = \sum_{x=k}^{\infty} x \binom{x-1}{k-1} p^k\, q^{x-k}.$$

Using the formula for binomial coefficients,

$$x\binom{x-1}{k-1} = \frac{x\,(x-1)!}{(k-1)!\,(x-k)!} = \frac{k\,x!}{k\,(k-1)!\,(x-k)!} = \frac{k\,x!}{k!\,(x-k)!} = k\binom{x}{x-k}.$$

Inserting and changing dummy index by $j = x - k$,

$$\mathbf{E}(X) = \sum_{x=k}^{\infty} k \binom{x}{x-k} p^k\, q^{x-k} = k\,p^k \sum_{j=0}^{\infty} \binom{k+j}{j} q^j.$$

From page 22, the negative binomial series is

$$\sum_{j=0}^{\infty} \binom{m+j-1}{j} z^j = (1-z)^{-m}$$

which makes sense even if $m > 0$ is a real number. In our series, the role of m is played by $m = k+1$. Thus

$$\mathbf{E}(X) = k\,p^k\,(1-q)^{-(k+1)} = \frac{k}{p}.$$

To compute the variance, we use the computation formula

$$\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2.$$

Using the equation $x^2 = (x+1)x - x$, the expectation of the square is

$$\begin{aligned}
\mathbf{E}(X^2) &= \sum_{x \in D} x^2\, f_X(x) = \sum_{x=k}^{\infty} x^2 \binom{x-1}{k-1} p^k\, q^{x-k} \\
&= \sum_{x=k}^{\infty} (x+1)x \binom{x-1}{k-1} p^k\, q^{x-k} - \sum_{x=k}^{\infty} x \binom{x-1}{k-1} p^k\, q^{x-k}.
\end{aligned}$$

The second sum is $-\,\mathbf{E}(X)$. We have another binomial coefficient identity

$$\begin{aligned}
(x+1)x\binom{x-1}{k-1} &= \frac{(x+1)\,x\,(x-1)!}{(k-1)!\,(x-k)!} = \frac{(k+1)\,k\,(x+1)!}{(k+1)\,k\,(k-1)!\,(x-k)!} \\
&= \frac{(k+1)\,k\,(x+1)!}{(k+1)!\,(x-k)!} = (k+1)k\binom{x+1}{x-k}.
\end{aligned}$$

Inserting this, changing dummy index to $j = x - k$ and using $m = k + 2$ in the negative binomial series, the first sum in $\mathbf{E}(X^2)$ yields

$$\sum_{x=k}^{\infty}(x+1)x\binom{x-1}{k-1}p^k\,q^{x-k} = \sum_{x=k}^{\infty}(k+1)k\binom{x+1}{x-k}p^k\,q^{x-k}$$
$$= (k+1)kp^k\sum_{j=0}^{\infty}\binom{k+j+1}{j}q^j = (k+1)kp^k\,(1-q)^{-(k+2)} = \frac{(k+1)k}{p^2}.$$

Assembling,

$$\mathbf{Var}(X) = \frac{(k+1)k}{p^2} - \frac{k}{p} - \frac{k^2}{p^2} = \frac{(k+1)k - pk - k^2}{p^2} = \frac{kq}{p^2}.$$