The K-Truncated Poisson Distribution

Charles J. Geyer

June 18, 2006

1 Introduction

This document works through the details of the k-truncated Poisson distribution, a special case of which is the zero-truncated Poisson distribution. The k-truncated Poisson distribution is the distribution of a Poisson random variable Y conditional on the event $Y > k$. It has one parameter, which we may take to be $\mu = E(Y)$. Since μ is not the mean (or anything else simple) of the distribution of Y conditioned on the even $Y > k$, we do not call μ the mean, rather we call it the *original parameter*.

If f_μ is the probability mass function (PMF) of Y, then the PMF g_μ of the k-truncated Poisson distribution is defined by

$$g_\mu(x) = \frac{f_\mu(x)}{1 - \sum_{j=0}^{k} f_\mu(j)}, \qquad x = k+1, k+2, \ldots \tag{1}$$

Plugging in the formula for the Poisson PMF, we get

$$\begin{aligned} g_\mu(x) &= \frac{\frac{\mu^x}{x!} e^{-\mu}}{1 - \sum_{j=0}^{k} \frac{\mu^j}{j!} e^{-\mu}} \\ &= \frac{\mu^x}{x!(e^\mu - \sum_{j=0}^{k} \frac{\mu^j}{j!})} \end{aligned} \tag{2}$$

2 Exponential Family Properties

Clearly (1) is the PMF of a one parameter exponential family having canonical statistic x and canonical parameter $\theta = \log(\mu)$. Of course, the original parameter is $\mu = \exp(\theta)$.

The cumulant function for the family is then

$$\psi(\theta) = \log\left(e^{e^\theta} - \sum_{j=0}^{k} \frac{e^{j\theta}}{j!}\right) \tag{3}$$

and has derivatives

$$\begin{aligned}
\tau(\theta) = \psi'(\theta) &= \frac{e^\theta \cdot e^{e^\theta} - \sum_{j=1}^{k} \frac{e^{j\theta}}{(j-1)!}}{e^{e^\theta} - \sum_{j=0}^{k} \frac{e^{j\theta}}{j!}} \\
&= \frac{\mu - e^{-\mu} \sum_{j=1}^{k} \frac{\mu^j}{(j-1)!}}{1 - e^{-\mu} \sum_{j=0}^{k} \frac{\mu^j}{j!}}
\end{aligned} \tag{4}$$

and

$$\begin{aligned}
\psi''(\theta) &= \frac{(e^\theta + e^{2\theta}) \cdot e^{e^\theta} - \sum_{j=1}^{k} \frac{j e^{j\theta}}{(j-1)!}}{e^{e^\theta} - \sum_{j=0}^{k} \frac{e^{j\theta}}{j!}} - \tau(\theta)^2 \\
&= \frac{(\mu + \mu^2) - e^{-\mu} \sum_{j=1}^{k} \frac{j\mu^j}{(j-1)!}}{1 - e^{-\mu} - \sum_{j=0}^{k} \frac{\mu^j}{j!}} - \tau(\theta)^2
\end{aligned} \tag{5}$$

By exponential family theory we know $\psi'(\theta) = E_\theta(X)$ and $\psi''(\theta) = \text{var}_\theta(X)$, where X is the canonical statistic. Thus from our definition of $\tau(\theta)$ in (4) it follows that $\tau'(\theta) = \psi''(\theta) > 0$ for all θ. Hence the map τ is one-to-one and defines an invertible change of parameter. Since $\tau(\theta) = E_\theta(X)$, it is called the *mean value parameter*. It is the mean of the distribution under discussion: k-truncated Poisson.

2.1 Check

If these are correct, then (4) should be $E(X)$ and the fraction in (5) should be $E(X^2)$ when X has the k-truncated Poisson distribution.

$$\begin{aligned}
E(X) &= \sum_{x=k+1}^{\infty} x g_\mu(x) \\
&= \sum_{x=k+1}^{\infty} \frac{x f_\mu(x)}{1 - \sum_{j=0}^{k} f_\mu(j)} \\
&= \frac{\mu - \sum_{j=0}^{k} j f_\mu(j)}{1 - \sum_{j=0}^{k} f_\mu(j)}
\end{aligned}$$

agrees with (4).

$$
\begin{aligned}
E(X^2) &= \sum_{x=k+1}^{\infty} \frac{x^2 f_\mu(x)}{1 - \sum_{j=0}^{k} f_\mu(j)} \\
&= \frac{\mu + \mu^2 - \sum_{j=0}^{k} j^2 f_\mu(j)}{1 - \sum_{j=0}^{k} f_\mu(j)}
\end{aligned}
$$

agrees with (5).

2.2 Computing

As always, we wish to compute things, in this case the cumulant function and its first two derivatives, without overflow or cancellation error. Problems arise when μ is nearly zero or when μ is very large.

2.2.1 Cumulant Function

From (3) we get, using $\mu = \exp(\theta)$,

$$
\begin{aligned}
\psi(\theta) &= \mu + \log\left(1 - e^{-\mu} \sum_{j=0}^{k} \frac{\mu^j}{j!}\right) \\
&= \mu + \log \Pr_\mu\{Y > k\}
\end{aligned}
\tag{6}
$$

where $Y \sim \text{Poi}(\mu)$. This looks fairly stable whether μ is large or small. We will leave the calculation of the log Poisson probability to R.

2.2.2 First Derivative of Cumulant Function

From (4) we get, using $\mu = \exp(\theta)$,

$$
\begin{aligned}
\tau(\theta) &= \frac{\mu - e^{-\mu} \sum_{j=1}^{k} \frac{\mu^j}{(j-1)!}}{\Pr_\mu\{Y > k\}} \\
&= \frac{\mu \left[1 - e^{-\mu} \sum_{j=0}^{k-1} \frac{\mu^j}{j!}\right]}{\Pr_\mu\{Y > k\}} \\
&= \frac{\mu \Pr_\mu\{Y \geq k\}}{\Pr_\mu\{Y > k\}} \\
&= \mu \left(1 + \frac{\Pr_\mu\{Y = k\}}{\Pr_\mu\{Y > k\}}\right)
\end{aligned}
\tag{7a}
$$

While this looks good for large μ it is not at all clear that it behaves well when μ is small. As $\mu \to 0$ (and $\theta \to -\infty$) we have $\tau(\theta) \to k + 1$. Let us see if we can get a computationally stable way to compute that without using L'Hospital's rule.

$$
\begin{aligned}
\tau(\theta) &= \mu + \frac{\mu \Pr_\mu\{Y = k\}}{\Pr_\mu\{Y > k\}} \\
&= \mu + \frac{\mu^{k+1} e^{-\mu} / k!}{\mu^{k+1} e^{-\mu} / (k+1)! + \Pr_\mu\{Y > k + 1\}} \\
&= \mu + \frac{k + 1}{1 + \dfrac{\Pr_\mu\{Y > k + 1\}}{\Pr_\mu\{Y = k + 1\}}}
\end{aligned}
\tag{7b}
$$

When μ is nearly zero, then the fraction in the denominator is also nearly zero and we get nearly $k + 1$ with no chance of overflow. Oops! It can produce `NaN` (IEEE not a number) when the fraction in the denominator is 0/0, actually underflow over underflow. If we special case this, then everything works.

Actually, our second formula (7b), seems too work just as well as (7a), even when μ is very large. In calculating $\tau(\theta)$ from zero to 1000 in steps of 0.1 both formulas give the same answers to within machine precision (relative error about 10^{-16}) whenever they do not give `Inf`, which they do for precisely the same arguments $\theta \le 709.7$.

In hindsight, this is no surprise. When $\mu \simeq \infty$, then $\Pr_\mu\{Y > k+1\} \approx 1$ and $\Pr_\mu\{Y = k+1\} \approx 0$ and the quotient in the denominator of (7b) either is very large or overflows giving `Inf` when IEEE arithmetic is in use (what happens on ancient computers without it is problematic), and the whole fraction is nearly zero. Hence, when $\mu \simeq \infty$, (7b) adds something very small or zero to μ.

2.2.3 Second Derivative of Cumulant Function

We start our computation of $\psi''(\theta)$ by noting that $\psi''(\theta) = \tau'(\theta)$, and, because $d\mu/d\theta = \mu$,

$$
\tau'(\theta) = \mu \frac{\tau(\mu)}{d\mu}.
$$

Thus we differentiate our "good" expression (7b) for τ expressed in terms of μ. It will simplify notation if we also define

$$
\beta = \frac{\Pr_\mu\{Y > k + 1\}}{\Pr_\mu\{Y = k + 1\}} = \frac{e^\mu \Pr_\mu\{Y > k + 1\}}{e^\mu \Pr_\mu\{Y = k + 1\}}
$$

and note that (7b) says

$$\tau = \mu + \frac{k+1}{1+\beta}.$$

so

$$\frac{d\tau}{d\mu} = 1 - \frac{k+1}{(1+\beta)^2} \cdot \frac{d\beta}{d\mu}.$$

To calculate $d\beta/d\mu$ we first figure out

$$\frac{d}{d\mu} e^{\mu} \Pr_{\mu}\{Y > k+1\} = \frac{d}{d\mu} \sum_{y=k+2}^{\infty} \frac{\mu^y}{y!} = \sum_{y=k+2}^{\infty} \frac{\mu^{y-1}}{(y-1)!} = e^{\mu} \Pr_{\mu}\{Y > k\}$$

and

$$\frac{d}{d\mu} e^{\mu} \Pr_{\mu}\{Y = k+1\} = \frac{d}{d\mu} \frac{\mu^{k+1}}{(k+1)!} = \frac{\mu^k}{k!} = e^{\mu} \Pr_{\mu}\{Y = k\}$$

Then

$$\begin{aligned}
\frac{d\beta}{d\mu} &= \frac{d}{d\mu} \frac{e^{\mu} \Pr_{\mu}\{Y > k+1\}}{e^{\mu} \Pr_{\mu}\{Y = k+1\}} \\
&= \frac{e^{\mu} \Pr_{\mu}\{Y > k\}}{e^{\mu} \Pr_{\mu}\{Y = k+1\}} - \frac{e^{\mu} \Pr_{\mu}\{Y > k+1\}}{\left(e^{\mu} \Pr_{\mu}\{Y = k+1\}\right)^2} \cdot e^{\mu} \Pr_{\mu}\{Y = k\} \\
&= \frac{\Pr_{\mu}\{Y > k\}}{\Pr_{\mu}\{Y = k+1\}} - \frac{\Pr_{\mu}\{Y > k+1\}}{\Pr_{\mu}\{Y = k+1\}} \cdot \frac{\Pr_{\mu}\{Y = k\}}{\Pr_{\mu}\{Y = k+1\}} \\
&= \frac{\Pr_{\mu}\{Y > k+1\}}{\Pr_{\mu}\{Y = k+1\}} + 1 - \frac{\Pr_{\mu}\{Y > k+1\}}{\Pr_{\mu}\{Y = k+1\}} \cdot \frac{\Pr_{\mu}\{Y = k\}}{\Pr_{\mu}\{Y = k+1\}} \\
&= \beta + 1 - \beta \cdot \frac{\Pr_{\mu}\{Y = k\}}{\Pr_{\mu}\{Y = k+1\}} \\
&= \beta + 1 - \beta \cdot \frac{k+1}{\mu}
\end{aligned}$$

So finally

$$\psi''(\theta) = \mu \left(1 - \frac{k+1}{1+\beta} \left(1 - \frac{k+1}{\mu} \cdot \frac{\beta}{1+\beta} \right) \right) \tag{8}$$

3 Random Variate Generation

To simulate a k-truncated Poisson distribution, the simplest method is to simulate ordinary Poisson random variates (using the `rpois` function in R) and reject all of the simulations less than or equal to k. This works well unless $\mu = \exp(\theta)$, the mean of the untruncated Poisson distribution is nearly zero, in which case the acceptance rate is also nearly zero. In that case, another simple rejection sampling scheme, simulates $Y \sim \text{Poi}(\mu)$ and uses $X = Y + m$ as the rejection sampling proposal, where m is a nonnegative integer (the case $m = 0$ is the case already discussed).

The ratio of target density to proposal density is

$$\begin{aligned}
\frac{g_\mu(x)}{f_\mu(y)} &= \frac{g_\mu(x)}{f_\mu(x-m)} \\
&= \frac{f_\mu(x) I(x > k)}{f_\mu(x-m)\left(1 - \sum_{j=0}^{k} f_\mu(j)\right)} \\
&= \frac{(x-m)! \cdot \mu^m I(x > k)}{x! \left(1 - \sum_{j=0}^{k} f_\mu(j)\right)}
\end{aligned} \tag{9}$$

where $I(x > k)$ is one when $x > k$ and zero otherwise. This achieves its upper bound (considered as a function of x) when $x = \max(m, k + 1)$. To avoid the "max" let us impose the condition that $m \leq k + 1$, so the max is achieved when $x = k + 1$ and is

$$\frac{(k+1-m)! \cdot \mu^m}{(k+1)! \left(1 - \sum_{j=0}^{k} f_\mu(j)\right)} \tag{10}$$

Thus we accept proposals with probability (9) divided by (10), which is

$$\frac{(x-m)!(k+1)!}{x!(k+1-m)!} \cdot I(x > k) \tag{11}$$

As noted at the beginning of the discussion, when $m = 0$ we accept a proposal x with probability $I(x > k)$. When $m = 1$ we accept a proposal x with probability

$$\frac{k+1}{x} \cdot I(x > k)$$

and so forth.

To understand the performance of this algorithm, hence to understand how to chose m, we need to calculate the acceptance rate

$$\begin{aligned}
a(k,m) &= E\left\{\frac{(X-m)!(k+1)!}{X!(k+1-m)!}\cdot I(X>k)\right\} \\
&= \frac{(k+1)!}{(k+1-m)!}\cdot E\left\{\frac{Y!}{(Y+m)!}\cdot I(Y>k-m)\right\} \\
&= \frac{(k+1)!}{(k+1-m)!}\sum_{y=k+1-m}^{\infty}\frac{y!}{(y+m)!}\cdot\frac{\mu^y}{y!}e^{-\mu} \\
&= \frac{(k+1)!}{(k+1-m)!}\cdot\frac{1}{\mu^m}\sum_{y=k+1-m}^{\infty}\frac{\mu^{y+m}}{(y+m)!}e^{-\mu} \\
&= \frac{(k+1)!}{(k+1-m)!}\cdot\frac{1}{\mu^m}\sum_{w=k+1}^{\infty}\frac{\mu^w}{w!}e^{-\mu} \\
&= \frac{(k+1)!}{(k+1-m)!}\cdot\frac{1}{\mu^m}\cdot\Pr(Y>k)
\end{aligned}$$

Everything is fixed in our formula for acceptance rate except m which we many choose to be any integer $0 \leq m \leq k+1$. Consider

$$\frac{a(k,m+1)}{a(k,m)} = \frac{(k+1-m)}{\mu}$$

this is greater than one (so it pays to increase m) when

$$k+1-m < \mu$$

which suggests we make

$$m = \lceil k+1-\mu \rceil$$

so long as this denotes a nonnegative integer (otherwise we set $m = 0$).

The performance of this algorithm seems to be fine for small k. However the worst case acceptance rate, which occurs for μ between $k/4$ and $k/2$, does seem to go to zero as k goes to infinity. For a zero-truncated Poisson distribution the worst case acceptance rate is 63.2%. For a two-truncated Poisson distribution the worst case acceptance rate is 48.2%. For a twenty-truncated Poisson distribution the worst case acceptance rate is 21.6%. For a one-hundred-truncated Poisson distribution the worst case acceptance rate is 10.2%.