Probability in the Engineering and Informational Sciences, **20**, 2006, 95–102. Printed in the U.S.A.

CONDITIONAL POISSON DISTRIBUTIONS

M. Zhou, D. Yang, Y. Wang, and S. Nadarajah

Department of Statistics
University of Nebraska
Lincoln, NE 68583
E-mail: snadaraj@unlserve.unl.edu

Compared to the known univariate distributions for continuous data, there are relatively few available for discrete data. In this article, we derive a collection of 16 flexible discrete distributions by means of conditional Poisson processes. The calculations involve the use of several special functions and their properties.

1. INTRODUCTION

Compared to the great multitude of continuous univariate distributions, there are relatively few choices available with respect to univariate discrete distributions. This is evident from the length of the compendiums of distributions available in the literature; see Johnson, Kotz, and Balakrishnan [2,3] for continuous distributions and Johnson, Kotz, and Kemp [4] for discrete distributions.

In this article, we present a collection of *new* discrete distributions. These are generated by means of conditional Poisson processes (Ross [6]); suppose $\{N(t), t \geq 0\}$, where $N(t)$ denotes the number of events during a time period of length t, is a Poisson process with rate parameter Λ. If $g(\lambda)$ denotes the probability density function (p.d.f.) of Λ, then the unconditional distribution of $N(t)$ can be written as

$$\Pr\{N(t) = n\} = \int_0^\infty \frac{\exp(-\lambda t)(\lambda t)^n}{n!} g(\lambda)\, d\lambda. \tag{1}$$

Now a discrete distribution for $N(t)$ can be generated by substituting a valid form for $g(\lambda)$. In this article, we generate a collection of discrete distributions for $N(t)$

 0269-9648/06 $16.00

by taking $g(\lambda)$ to belong to 16 flexible families. The calculations use several special functions, including the integral cosine defined by

$$\mathrm{ci}(x) = -\int_x^\infty \frac{\cos t}{t}\,dt,$$

the integral sine defined by

$$\mathrm{si}(x) = -\int_x^\infty \frac{\sin t}{t}\,dt,$$

the incomplete gamma function defined by

$$\Gamma(a, x) = \int_x^\infty t^{a-1}\exp(-t)\,dt,$$

the error function defined by

$$\mathrm{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)\,dt,$$

the modified Bessel function of the third kind defined by

$$K_\nu(x) = \frac{x^\nu \Gamma(\frac{1}{2})}{2^\nu \Gamma(\nu + \frac{1}{2})}\int_1^\infty \exp(-xt)(t^2 - 1)^{\nu - 1/2}\,dt,$$

the parabolic cylinder function defined by

$$D_p(x) = \frac{\exp(-x^2/4)}{\Gamma(-p)}\int_0^\infty \exp\left\{-\left(tx + \frac{t^2}{2}\right)\right\} t^{-(p+1)}\,dt,$$

the ${}_1F_1$ hypergeometric function (also known as the confluent hypergeometric function) defined by

$${}_1F_1(a;b;x) = \sum_{k=0}^\infty \frac{(a)_k}{(b)_k}\frac{x^k}{k!},$$

the ${}_1F_2$ hypergeometric function defined by

$${}_1F_2(a;b,c;x) = \sum_{k=0}^\infty \frac{(a)_k}{(b)_k(c)_k}\frac{x^k}{k!},$$

and the Kummer function defined by

$$\Psi(a,b;x) = \frac{\Gamma(1-b)}{\Gamma(1+a-b)}\,{}_1F_1(a;b;x) + \frac{\Gamma(b-1)}{\Gamma(a)}\,x^{1-b}\,{}_1F_1(1+a-b;2-b;x),$$

where $(f)_k = f(f+1)\cdots(f+k-1)$ denotes the ascending factorial. The properties of these special functions can be found in Prudnikov, Brychkov, and Marichev [5] and Gradshteyn and Ryzhik [1].

The details of the derivation for (1) are not given in this article and can be obtained from the authors. The structural properties of $N(t)$ are also not given since they can be obtained directly from those of Λ. For example, the mean and the variance of $N(t)$ are

$$E(N(t)) = tE(\Lambda)$$

and

$$\mathrm{Var}(N(t)) = tE(\Lambda) + t^2\mathrm{Var}(\Lambda),$$

respectively. Thus, these follow directly by knowing $E(\Lambda)$ and $E(\Lambda^2)$; see Johnson et al. [2,3].

2. DISCRETE MODELS

In this section, we provide a collection of formulas for $\Pr\{N(t) = n\}$ by taking g to belong to 16 flexible families.

Beta distribution: If g takes the form

$$g(\lambda) = \frac{\lambda^{p-1}(1-\lambda)^{q-1}}{B(p,q)}$$

for $0 < \lambda < 1$, then

$$\Pr\{N(t) = n\} = \frac{t^n B(p+n,q)}{n!B(p,q)}\, {}_1F_1(p+n;p+q+n;-t).$$

If g takes the form of the generalized beta distribution given by

$$g(\lambda) = \frac{(\lambda-a)^{p-1}(b-\lambda)^{q-1}}{B(p,q)(b-a)^{p+q-1}}$$

for $a < \lambda < b$, then

$$\Pr\{N(t) = n\} = \frac{\exp(-at)t^n}{n!B(p,q)} \sum_{k=0}^{n} \binom{n}{k} a^{n-k}(b-a)^k B(k+p,q)$$
$$\times\, {}_1F_1(k+p;k+p+q;-t(b-a)).$$

Uniform distribution: If g takes the form

$$g(\lambda) = \frac{1}{b-a}$$

https://doi.org/10.1017/S0269964806060050 Published online by Cambridge University Press

for $a < \lambda < b$, then

$$\Pr\{N(t) = n\} = \frac{\Gamma(n+1, ta) - \Gamma(n+1, tb)}{tn!(b-a)}.$$

Inverted beta distribution: If g takes the form

$$g(\lambda) = \frac{\lambda^{\alpha-1}(1+\lambda)^{-\alpha-\beta}}{B(\alpha, \beta)}$$

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{t^n \Gamma(n+\alpha)\Psi(n+\alpha, n-\beta+1; t)}{n! B(\alpha, \beta)}.$$

Exponential distribution: If g takes the form

$$g(\lambda) = \beta \exp(-\beta\lambda)$$

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{\beta t^n}{(t+\beta)^{n+1}}.$$

Gamma distribution: If g takes the form

$$g(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} \exp(-\beta\lambda)}{\Gamma(\alpha)}$$

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{\beta^\alpha t^n \Gamma(n+\alpha)}{\Gamma(\alpha) n! (t+\beta)^{n+\alpha}}.$$

Rayleigh distribution: If g takes the form

$$g(\lambda) = 2\beta^2 \lambda \exp\{-(\beta\lambda)^2\}$$

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{\beta\sqrt{\pi} t^n (-1)^{n+1}}{n!} \frac{\partial^{n+1}}{\partial q^{n+1}} \left[\exp\left(\frac{q^2}{4\beta^2}\right) \operatorname{erfc}\left(\frac{q}{2\beta}\right) \right] \Bigg|_{q=t}.$$

Stacy distribution ($c = 2$): If g takes the form

$$g(\lambda) = \frac{2\beta^{2\alpha} \lambda^{2\alpha-1} \exp\{-(\lambda\beta)^2\}}{\Gamma(\alpha)}$$

https://doi.org/10.1017/S0269964806060050 Published online by Cambridge University Press

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{2^{1-\alpha-n/2} t^n \Gamma(n+2\alpha)}{n! \beta^n \Gamma(\alpha)} \exp\left(\frac{t^2}{8\beta^2}\right) D_{-n-2\alpha}\left(\frac{t}{\sqrt{2}\beta}\right).$$

If 2α is an integer, then the above reduces to the simpler form

$$\Pr\{N(t) = n\} = \frac{\sqrt{\pi}(-1)^{n+2\alpha-1} \beta^{2\alpha-1}}{n! \Gamma(\alpha)} \frac{\partial^{n+2\alpha+1}}{\partial q^{n+2\alpha+1}} \left[\exp\left(\frac{q^2}{4\beta^2}\right) \operatorname{erfc}\left(\frac{q}{2\beta}\right)\right]\Bigg|_{q=t}.$$

Pareto distribution of the first kind: If g takes the form

$$g(\lambda) = a k^a \lambda^{-a-1}$$

for $\lambda > k$, then

$$\Pr\{N(t) = n\} = \frac{a(kt)^a \Gamma(n-a, kt)}{n!}.$$

Pareto distribution of the second kind: If g takes the form

$$g(\lambda) = a\lambda^a (\lambda + c)^{-a-1}$$

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{a\Gamma(n+a+1)(tc)^n}{n!} \Psi(n+a+1, n+1; tc).$$

Inverse Gaussian distribution: If g takes the form

$$g(\lambda) = \sqrt{\frac{\phi}{2\pi}} \exp(\phi) \lambda^{-3/2} \exp\left\{-\frac{\phi}{2}\left(\lambda + \frac{1}{\lambda}\right)\right\}$$

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{\sqrt{2\phi} \exp(\phi) t^n}{\sqrt{\pi} n!} \left(\frac{\phi}{2t+\phi}\right)^{(2n-1)/4} K_{n-1/2}(\sqrt{\phi(2t+\phi)}).$$

Half Normal distribution: If g takes the form

$$g(\lambda) = \frac{\sqrt{2}}{\sqrt{\pi}\sigma} \exp\left\{-\frac{\lambda^2}{2\sigma^2}\right\}$$

for $\lambda > 0$, then

$$\Pr\{N(t) = n\} = \frac{(-t)^n}{n!} \frac{\partial^n}{\partial q^n} \left[\exp\left(\frac{q^2\sigma^2}{2}\right) \operatorname{erfc}\left(\frac{q\sigma}{\sqrt{2}}\right)\right]\Bigg|_{q=t}.$$

https://doi.org/10.1017/S0269964806060050 Published online by Cambridge University Press

Half logistic distribution: If g takes the form

$$g(\lambda) = \frac{2\beta \exp(-\beta\lambda)}{[1+\exp(-\beta\lambda)]^2}$$

for $\lambda > 0$, then

$$\Pr\{N(t)=n\} = 2\beta t^n \sum_{k=0}^{\infty} \binom{-2}{k} \{t+(k+1)\beta\}^{-n-1}.$$

Half Cauchy distribution: If g takes the form

$$g(\lambda) = \frac{2}{\pi\alpha}\left\{1+\frac{\lambda^2}{\alpha^2}\right\}^{-1}$$

for $\lambda > 0$, then

$$\Pr\{N(t)=n\} = \frac{2\alpha t^n}{\pi n!}\left[\alpha^{n-1}\left\{\sin\left(t\alpha - \frac{\pi n}{2}\right)\mathrm{ci}(t\alpha) - \cos\left(t\alpha - \frac{\pi n}{2}\right)\mathrm{si}(t\alpha)\right\}\right.$$
$$\left. + t^{1-n}\sum_{k=1}^{[n/2]}(n-2k)!(-t^2\alpha^2)^{k-1}\right].$$

Half *t* distribution: If g takes the form

$$g(\lambda) = \frac{2}{\sqrt{\nu}B(\frac{1}{2},\nu/2)}\left(1+\frac{\lambda^2}{\nu}\right)^{-(1+\nu)/2}$$

for $\lambda > 0$, then

$$\Pr\{N(t)=n\}$$
$$= \frac{2\nu^{\nu/2}t^n}{n!B(\frac{1}{2},\nu/2)}\left[t^{\nu-n}\Gamma(n-\nu)\,{}_1F_2\left(\frac{1+\nu}{2};1+\frac{\nu-n}{2},\frac{1+\nu-n}{2};-\frac{t^2\nu}{4}\right)\right.$$
$$+\frac{\nu^{(n-\nu)/2}}{2}B\left(\frac{\nu-n}{2},\frac{n+1}{2}\right)$$
$$\times\,{}_1F_2\left(\frac{1+n}{2};\frac{1}{2},1+\frac{n-\nu}{2};-\frac{t^2\nu}{4}\right)$$
$$-\frac{t\nu^{(1+n-\nu)/2}}{2}B\left(\frac{\nu-n-1}{2},\frac{n+2}{2}\right)$$
$$\left.\times\,{}_1F_2\left(\frac{2+n}{2};\frac{3}{2},\frac{3+n-\nu}{2};-\frac{t^2\nu}{4}\right)\right].$$

If $(1+\nu)/2$ is an integer, then the above reduces to the simpler form

$$\Pr\{N(t)=n\} = \frac{2^{(3-\nu)/2}\nu^{\nu/2}(-1)^{n+(\nu-1)/2}t^n}{((\nu-1)/2)!\,B(\frac{1}{2},\nu/2)n!}\,\frac{\partial^n}{\partial p^n}\left(\frac{\partial}{z\partial z}\right)^{(\nu-1)/2} g(p,z)\bigg|_{p=t,\,z=\sqrt{\nu}},$$

where

$$g(p,z) = \frac{\sin(pz)\mathrm{ci}(pz) - \cos(pz)\mathrm{si}(pz)}{z}.$$

Fréchet distribution: If g takes the form

$$g(\lambda) = \frac{\gamma}{\lambda^2}\exp\left(-\frac{\gamma}{\lambda}\right)$$

for $\lambda > 0$, then

$$\Pr\{N(t)=n\} = \frac{2(\gamma t)^{(n+1)/2}}{n!}K_{n-1}(2\sqrt{\gamma t}).$$

Pearson type VI distribution: If g takes the form

$$g(\lambda) = \frac{\Gamma(p)(b-a)^{p-q-1}(\lambda-b)^q}{\Gamma(p-q-1)\Gamma(q+1)(\lambda-a)^p}$$

for $\lambda \geq b > a > 0$, then

$$\Pr\{N(t)=n\} = \frac{\Gamma(p)t^n\exp(-bt)}{\Gamma(p-q-1)\Gamma(q+1)n!}\sum_{k=0}^{n}\binom{n}{k}b^{n-k}(b-a)^k\Gamma(q+k+1) \times \Psi(q+k+1, q-p+k+2; t(b-a)).$$

3. CONCLUSIONS

We have generated a collection of 16 flexible discrete distributions. The definition of the conditional Poisson process is used as the mathematical tool. We hope that this work will help to address the inadequacy of the number of distributions available to model discrete data.

References

1. Gradshteyn, I.S. & Ryzhik, I.M. (2000). *Table of integrals, series, and products*, 6th ed. San Diego: Academic Press.
2. Johnson, N.L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions*, Vol. 1, 2nd ed. New York: Wiley.

3. Johnson, N.L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions*, Vol. 2, 2nd ed. New York: Wiley.
4. Johnson, N.L., Kotz, S., & Kemp, A.W. (1992). *Univariate discrete distributions*, 2nd ed. New York: Wiley.
5. Prudnikov, A.P., Brychkov, Y.A., & Marichev, O.I. (1986). *Integrals and series*, Vols. 1, 2, and 3. Amsterdam: Gordon and Breach Science.
6. Ross, S.M. (2000). *Introduction to probability models*, 7th ed. Burlington, MA: Harcourt/Academic Press.