January 13, 1994

Derivation of a scaled binomial as an instance of a general discrete exponential distribution

Joseph Hilbe, *Arizona State University*

Derivation of a scaled binomial as an instance of a general discrete exponential distribution

Joseph Hilbe

The family of discrete distributions having the form

$$f(y) = N\, p^{by-a}\,(1-p)^{dy-c} \qquad (1)$$

where N is a normalization dependent upon a,b,c, and d.

(1) is a type of general exponential distribution containing several notable cases:

	a	b	c	d	->	f(y)
Bernoulli	0	1	-1	-1		$p^y(1-p)^{1-y}$
Binomial	0	1	-n	-1		$p^y(1-p)^{n-y}$
Geometric	-1	0	0	1		$p(1-p)^y$
Neg. Binomial	-1/k	0	0	1		$p^{1/k}(1-p)^y$

The negative binomial k is taken as the inverse of r (Hilbe 1994)

As a generalized exponential (1) becomes

$$\begin{aligned} f(y) &= \exp\{ (by-a)\ln p + (dy-c)\ln(1-p) + \ln(N) \} \\ &= \exp\{ y(b\ln p + d\ln(1-p)) - (a\ln p + c\ln(1-p)) + \ln(N) \} \\ &= \exp\{ y\theta - g(\theta) + \ln(N) \} \end{aligned}$$

where $\theta = b\ln p + d\ln(1-p)$, and
$g(\theta) = a\ln p + c\ln(1-p)$

$\mu = g'(\theta) = dp/d(\theta) * d(g)/dp$ using the chain rule

Hence

1) $$\begin{aligned} d(\theta)/dp &= b/p + d/(1-p) \\ &= \{b(1-p)+dp\}/\{p(1-p)\} \\ &= (b-bp+dp)/\{p(1-p)\} \\ &= \{b-(b+d)p\}/\{p(1-p)\} \end{aligned}$$

therefore $dp/d(\theta) = \{p(1-p)\}/\{b-(b+d)p\}$

2) $$\begin{aligned} d(g)/dp &= a/p - c/(1-p) \\ &= \{a(1-p)-cp\}/\{p(1-p)\} \\ &= (a-ap-cp)/\{p(1-p)\} \\ &= \{a-(c+a)p\}/\{p(1-p)\} \end{aligned}$$

therefore

$$\begin{aligned} \mu = g'(\theta) &= \{p(1-p)\}/\{b-(b+d)p\} * \{a-(c+a)p\}/\{p(1-p)\} \\ &= \{a-(c+a)p\}/\{b-(b+d)p\} \end{aligned}$$

$V = g''(\theta) = dp/d(\theta) * d(g')/dp = \{p(1-p)\}/\{b-(b+d)p\} * d(g')/dp$

Solving $d(g')/dp = (ad-bc)/\{b-(b+d)p\}^2$

therefore

$V = \{p(1-p)\}/\{b-(b+d)p\} * (ad-bc)/\{b-(b+d)p\}^2$

$= \{p(1-p)(ad-bc)\}/\{b-p(b+d)\}^3$

Summarizing, the generalized discrete exponential formulae for μ and V are:

$$\mu = \frac{a-(c+a)p}{b-(b+d)p} \qquad V = \frac{p(1-p)(ad-bc)}{\{b-p(b+d)\}^3}$$

with θ, the canonical link, as $b\ln p + d\ln(1-p)$

Testing these formulae for known members of the family:

i. Bernoulli a=0; b=1; c=-1; d=-1

$\mu = p \qquad V = p(1-p)$

ii. Binomial a=0; b=1; c=-n; d=-1

$\mu = np/\{1-(0)p\} = np$

$V = \{p(1-p)n\}/1 = np(1-p)$

iii. Geometric a=-1; b=0; c=0; d=1

$\mu = \{-(1-p)\}/-p = (1-p)/p$

$V = \{p(1-p)(-1)\}/(-p)^3 = (1-p)/p^2$

iv. Negative binomial a=-1/k; b=0; c=0; d=1

$\mu = \{(-1/k)-(-p/k)\}/-p = (-1/k+p/k)/-p = (1-p)/kp$

$V = \{p(1-p)(-1/k)\}/(-p)^3 = (1-p)/kp^2$

A scaled binomial distribution that is a member of the exponential family can be derived as follows:

let $\mu = np$ and $V = np(1-p)(1+k)$

Such properties should allow modeling of overdispersed binomial models; eg overdispersed logistic regression.

Setting a=0; b=1; c=-n; d=-1, we have, as in the standard binomial,

$$\mu = np = \frac{a-(c+a)p}{b-(b+d)p}$$

Since d = -b, c/b=-n, and c=-nb, we can formulate the desired distribution as:

$$\begin{aligned} f(y)_{sb} &= N\ p^{by-a}\ (1-p)^{dy-c} \\ &= N\ p^{by}\ (1-p)^{-by-c} \\ &= N\ p^{by}\ (1-p)^{-by+nb} \\ &= N\ p^{by}\ (1-p)^{bn-by} \\ &= N\ p^{by}\ (1-p)^{b(n-y)} \end{aligned}$$

Depending on the choice of b, there is a different variance:

$$V = \frac{p(1-p)(-bc)}{(b)^3} = \frac{p(1-p)(-b*-nb)}{b^3} = \frac{p(1-p)(nb^2)}{b^3} = \frac{np(1-p)}{b}$$

Since we wanted V = np(1-p)(1+k) then b=1/(1+k), giving

$$f(y) = N\ p^{y/(1+k)}\ (1-p)^{(n-y)/(1+k)}$$

The correct mormalization may be found by scaling the variable terms of the binomial normalization by 1/(1+k). We do this since we want the binomial if k=0. Given

$$\text{Binomial } N = \frac{\Gamma(n+1)}{\Gamma(n-y+1)\Gamma(y+1)}$$

the scaled binomial normalization will appear as

$$\text{Scaled Binomial} = \frac{1}{1+k}\ \frac{\Gamma\{n/(1+k) + 1\}}{\Gamma\{(n-y)/(1+k)+1\}\Gamma\{y/(1+k)+1\}}$$

The full scaled binomial probability density function is then

$$f(y)_{sb} \quad \frac{1}{1+k}\ \frac{\Gamma\{n/(1+k) + 1\}}{\Gamma\{(n-y)/(1+k)+1\}\Gamma\{y/(1+k)+1\}}\ p^{y/(1+k)}\ (1-p)^{(n-y)/(1+k)}$$

Recalling the general form canonical link

$$\theta = b\ln p + d\ln(1-p)$$

and given that, for the scaled binomial, $d=-b$

$$\begin{aligned}\theta &= b\ln p - b\ln(1-p)\\ &= b\ln(p/(1-p))\end{aligned}$$

which is simply the binomial link with b as a multiplicative factor.

It follows that the inverse link, in terms of p, can be derived as:

$$\begin{aligned}
\theta &= b\ln(p/(1-p))\\
\theta/b &= \ln(p/(1-p))\\
\theta(1+k) &= \ln(p/(1-p))\\
e^{\theta(1+k)} &= p/(1-p)\\
e^{-\theta(1+k)} &= (1-p)/p\\
e^{-\theta(1+k)} &= 1/p - 1\\
1+e^{-\theta(1+k)} &= 1/p\\
1/(1+e^{-\theta(1+k)}) &= p
\end{aligned}$$

Hence $$p = \frac{1}{1+e^{-\theta(1+k)}} \quad \text{or} \quad \frac{1}{1+e^{-\eta(1+k)}}$$

Multiplying p by n gives μ. Hence, the inverse link has n as a numerator for grouped data models.

In terms of the table on the first page, the exponential properties of p in the pdf may summarized as:

	a	b	c	d	f(y)
Scaled Binomial	0	$1/(1+k)$	$-n$	-1	$p^{y/(1+k)}(1-p)^{(n-y)/(1+k)}$

Logistic overdispersion caused by heterogeniety may be accommodated by using a scaled binomial regression model. k is adjusted such that the chi2 dispersion approximates 1.0. Unlike many quasilikelihood models, the scaled binomial, as a full member of the exponential family, allows meaningful interpretation and full use of related diagnostics.

Jan 13, 1994