Supplementary Materials for
"Splitting models for multivariate count data"

Jean Peyhardi[a], Pierre Fernique[b], Jean-Baptiste Durand[c]

[a]*IMAG, Université de Montpellier, place Eugène Bataillon, 34090 Montpellier, France*
[b]*Research Centre of Limagrain Europe, Biostatistics Department, 63720 Chappes, France*
[c]*Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP*, LJK, 38000 Grenoble, France*
**Institute of Engineering Univ. Grenoble Alpes*

1. Univariate distributions

1.1. Continuous univariate distributions

Let us recall the definition of the (generalized) beta distribution with positive real parameters c, α and b, denoted by $\beta_c\,(a,b)$. Its probability density function described by Whitby [16] is given, for $x \in (0,c)$, by

$$f\,(x) = \frac{1}{B\,(a,b)} \frac{x^{a-1}\,(c-x)^{b-1}}{c^{a+b-1}}.$$

Note that $Z = dX$ with $d \in (0,\infty)$ and $X \sim \beta_c\,(a,b)$ implies that $Z \sim \beta_{cd}\,(a,b)$. The parameter c of the beta distribution can thus be interpreted as a rescaling parameter of the standard beta distribution. By convention the standard beta distribution (i.e., defined with $c = 1$) will be denoted by $\beta\,(a,b)$.

Let us introduce the definition of the (generalized) beta product distribution with parameters $(a_1, b_1, a_2, b_2) \in (0,\infty)^4$ and $c \in (0,\infty)$, denoted by $\beta^2_c(a_1, b_1, a_2, b_2)$, as the product of the two independent beta distributions $\beta(a_1, b_1)$ and $\beta(a_2, b_2)$ normalized on $(0,c)$; see [3] for details. It is named the standard beta product distribution when $c = 1$ and denoted by $\beta^2(a_1, b_1, a_2, b_2)$. More generally the product of m beta distributions could be defined.

1.2. Discrete univariate distributions

1.2.1. Power series distributions:

Let $(b_y)_{y\in\mathbb{N}}$ be a non-negative real sequence such that the series $\sum_{y\geq 0} b_y \theta^y$ converges toward $g(\theta)$ for all $\theta \in D = (0,R)$, where R is the radius of convergence. The discrete random variable Z is said to follow a power series distribution if for all $y \in \mathbb{N}$

$$P(Y = y) = \frac{b_y \theta^y}{g(\theta)},$$

and is denoted by $Y \sim PSD\{g(\theta)\}$. Several usual discrete distributions fall into the family of power series distributions:

1. The Poisson distribution $\mathcal{P}(\lambda)$ with $b_y = 1/y!$, $\theta = \lambda$, $g(\theta) = e^\theta$ and $D = (0,\infty)$.
2. The binomial distribution $\mathcal{B}_n(p)$ with $b_y = \binom{n}{y}\mathbf{1}_{y\leq n}$, $\theta = p/(1-p)$, $g(\theta) = (1+\theta)^n$ and $D = (0,\infty)$.
3. The negative binomial distribution $\mathcal{NB}(r,p)$ with $b_y = \binom{r+y-1}{y}$, $\theta = p$, $g(\theta) = (1-\theta)^{-r}$ and $D = (0,1)$.
4. The geometric distribution $\mathcal{G}(p)$ with $b_y = \mathbf{1}_{y\geq 1}$, $\theta = 1-p$, $g(\theta) = \theta/(1-\theta)$ and $D = (0,1)$.
5. The logarithmic series distribution $\mathcal{L}(p)$ with $b_y = \mathbf{1}_{y\geq 1}1/y$, $\theta = p$, $g(\theta) = -\ln(1-\theta)$ and $D = (0,1)$.

Let us define the zero modified logarithmic series distribution, denoted by $\mathcal{L}(p,\omega)$ with $p \in (0,1)$ and $\omega \in [0,\infty)$ and probability mass function (pmf) given by

$$p(y) = \begin{cases} \frac{\omega}{\omega - \ln(1-p)} & \text{if } y = 0 \\ \frac{p^y/y}{\omega - \ln(1-p)} & \text{otherwise} \end{cases}$$

It belongs to the family of power series distributions with

$$b_y = \begin{cases} \omega & \text{if } y = 0 \\ 1/y & \text{otherwise} \end{cases}$$

where the series $\sum_{y\geq 0} b_y p^y$ converges on $D = (0,1)$ towards $g(p) = \omega - \ln(1-p)$. Let us remark that $\omega = 0$ lead to the usual logarithmic series distribution $\mathcal{L}(p) = \mathcal{L}(p,0)$ with support $\mathbb{N}^*$.

When the support is a subset of $\mathbb{N}$, the b_y values can be weighted by an indicator function as for binomial, geometric and logarithmic distributions. The b_y must be independent of θ but they may depend on other parameters as for binomial and negative binomial distributions.

1.2.2. Beta compound distributions:

Usual characteristics of the standard beta binomial [14], standard beta negative binomial - also described by Xekalaki [17] as the univariate generalized Waring distribution (UGWD) - and the beta Poisson distributions [5] are first recalled in Table 2. Then we introduce these beta compound distributions in a general way, i.e. using the generalized beta distribution, described at the begining of this Section. For the Poisson case we obtain the same distribution since $\mathcal{P}(\lambda p) \underset{p}{\wedge} \beta(a,b) = \mathcal{P}(\theta) \underset{\theta}{\wedge} \beta_\lambda(a,b)$. The two other case lead us to new distributions (see Table 3 for detailed characteristics). Let us remark that if $\pi = 1$ then, the generalized beta binomial (resp. generalized beta negative binomial) turns out to be the standard beta binomial (resp. standard negative binomial distribution). In opposition, if $\pi < 1$, the non-standard beta binomial distribution (respectively non-standard beta negative binomial distribution) is obtained.

Generalized beta compound distributions. Let $n \in \mathbb{N}$, $a \in (0,\infty)$, $b \in (0,\infty)$ and $\pi \in (0,1)$ and consider the compound distribution $\mathcal{B}_n(p) \underset{p}{\wedge} \beta_\pi(a,b)$ denoted by $\beta_\pi\mathcal{B}_n(a,b)$. Considering π as a rescaling parameter, we have $\beta_\pi\mathcal{B}_n(a,b) = \mathcal{B}_n(\pi p) \underset{p}{\wedge} \beta(a,b)$. According to Property 4 we have $\mathcal{B}_n(\pi p) = \mathcal{B}_N(\pi) \underset{N}{\wedge} \mathcal{B}_n(p)$. Finally, using the Fubini theorem we obtain

$$\begin{aligned} \beta_\pi\mathcal{B}_n(a,b) &= \left\{\mathcal{B}_N(\pi) \underset{N}{\wedge} \mathcal{B}_n(p)\right\} \underset{p}{\wedge} \beta(a,b), \\ &= \mathcal{B}_N(\pi) \underset{N}{\wedge} \left\{\mathcal{B}_n(p) \underset{p}{\wedge} \beta(a,b)\right\}, \\ \beta_\pi\mathcal{B}_n(a,b) &= \mathcal{B}_N(\pi) \underset{N}{\wedge} \beta\mathcal{B}_n(a,b). \end{aligned}$$

This is a binomial damage distribution whose the latent variable N follows a standard beta binomial distribution. The equation (10) of the paper can thus be used to compute the probability mass function. The y^{th} derivative of the probability generating function (pgf) of the standard beta binomial distribution is thus needed

$$G_N^{(y)}(s) = \frac{(b)_n}{(a+b)_n} \frac{(-n)_y(a)_y}{(-b-n+1)_y} {}_2F_1\{(-n+y, a+y); -b-n+1+y; s\},$$

obtained by induction on $y \in \mathbb{N}$. The moments are obtained with the total law of expectation and variance given the latent variable N of the binomial damage distribution. In the same way, we obtain the pgf as $G_Y(s) = G_N(1-\pi+\pi s)$. A similar proof holds for the generalized beta negative binomial case.

Generalized beta product compound distributions. It is also possible to define the (generalized) beta product distribution, as the product of two independent beta distributions [3], and then define the (generalized) beta product compound distributions.

- The standard beta product binomial distribution is defined as $\mathcal{B}_n(p) \underset{p}{\wedge} \beta^2(a_1, b_1, a_2, b_2)$ and denoted by $\beta^2\mathcal{B}_n(a_1, b_1, a_2, b_2)$.
- The standard beta product negative binomial distribution is defined as $\mathcal{NB}(r, p) \underset{p}{\wedge} \beta^2(a_1, b_1, a_2, b_2)$ and denoted by $\beta^2\mathcal{NB}(r, a_1, b_1, a_2, b_2)$.
- The generalized beta product binomial distribution is defined as $\mathcal{B}_n(p) \underset{p}{\wedge} \beta^2_\pi(a_1, b_1, a_2, b_2)$ and denoted by $\beta^2_\pi\mathcal{B}_n(a_1, b_1, a_2, b_2)$
- The generalized beta product negative binomial distribution is defined as $\mathcal{NB}(r, p) \underset{p}{\wedge} \beta^2_\pi(a_1, b_1, a_2, b_2)$ and denoted by $\beta^2_\pi\mathcal{NB}(r, a_1, b_1, a_2, b_2)$
- The beta product Poisson distribution is defined as $\mathcal{P}(\theta) \underset{\theta}{\wedge} \beta^2_\lambda(a_1, b_1, a_2, b_2)$ and denoted by $\beta^2_\lambda\mathcal{P}(a_1, b_1, a_2, b_2)$.

Table 1: References of parameter inference procedures for seven usual univariate discrete distributions.

Distribution	Notation	Parameter Inference
Binomial	$\mathcal{B}_n(p)$	See [1]
Negative binomial	$\mathcal{NB}(r, p)$	See [2]
Poisson	$\mathcal{P}(\lambda)$	See [8]
Logarithmic series	$\mathcal{L}(p)$	See [8]
Beta binomial	$\beta\mathcal{B}_n(a, b)$	See [9, 14] for n known
Beta negative binomial	$\beta\mathcal{NB}(r, a, b)$	See [7]
Beta Poisson	$\beta_\lambda\mathcal{P}(a, b)$	See [5, 15]

Table 2: Usual characteristics of the standard beta compound binomial, the standard beta compound negative binomial and the beta Poisson distributions.

Name	Standard beta binomial	Standard beta negative binomial	Beta Poisson
Definition	$\mathcal{B}_n(p) \underset{p}{\wedge} \beta(a,b)$	$\mathcal{NB}(r,p) \underset{p}{\wedge} \beta(a,b)$	$\mathcal{P}(\lambda p) \underset{p}{\wedge} \beta(a,b)$
Notation	$\beta\mathcal{B}_n(a,b)$	$\beta\mathcal{NB}(r,a,b)$	$\beta\mathcal{P}_\lambda(a,b)$
Re-parametrization		UGWD(r,b,a)	
Supp(Y)	$\{0,1,\ldots,n\}$	$\mathbb{N}$	$\mathbb{N}$
$P(Y=y)$	$\frac{(b)_n}{(a+b)_n}\frac{(-n)_y(a)_y}{(-b-n+1)_y}\frac{1}{y!}$	$\frac{(a)_b}{(a+r)_b}\frac{(r)_y(b)_y}{(r+a+b)_y}\frac{1}{y!}$	$\frac{(a)_y}{(a+b)_y}\frac{\lambda^y}{y!}{}_1F_1(a+y;a+b+y;-\lambda)$
E(Y)	$n\frac{a}{a+b}$	$r\frac{b}{a-1}$ (define if $a>1$)	$\lambda\frac{a}{a+b}$
V(Y)	$n\frac{ab(a+b+n)}{(a+b)^2(a+b+1)}$	$r\frac{b(a+r-1)(a+b-1)}{(a-1)^2(a-2)}$ [1]	$\lambda\frac{a}{a+b}\left\{1+\lambda\frac{b}{(a+b)(a+b+1)}\right\}$
$G_Y(s)$	$\frac{(b)_n}{(a+b)_n}\,{}_2F_1\{(-n,a);-b-n+1;s\}$	$\frac{(a)_r}{(a+b)_r}\,{}_2F_1\{(r,b);r+a+b;s\}$	${}_1F_1\{a;a+b;\lambda(s-1)\}$

Table 3: Usual characteristics of the generalized beta binomial and the generalized beta negative binomial distributions.

Name	Generalized beta binomial	Generalized beta negative binomial
Definition	$\mathcal{B}_n(p) \underset{p}{\wedge} \beta_\pi(a,b)$	$\mathcal{NB}(r,p) \underset{p}{\wedge} \beta_\pi(a,b)$
Notation	$\beta_\pi\mathcal{B}_n(a,b)$	$\beta_\pi\mathcal{NB}(r,a,b)$
Supp(Y)	$\{0,1,\ldots,n\}$	$\mathbb{N}$
$P(Y=y)$	$\frac{(b)_n}{(a+b)_n}\frac{(-n)_y(a)_y}{(-b-n+1)_y}\frac{\pi^y}{y!}{}_2F_1\{(-n+y,a+y);-b-n+1+y;1-\pi\}$	$\frac{(a)_r}{(a+b)_r}\frac{(r)_y(b)_y}{(r+a+b)_y}\frac{\pi^y}{y!}{}_2F_1\{(r+y,b+y);r+a+b+y;1-\pi\}$
E(Y)	$n\pi\frac{a}{a+b}$	$r\pi\frac{b}{a-1}$ [2]
V(Y)	$n\pi\frac{a}{a+b}\left\{\pi\frac{b(a+b+n)}{(a+b)(a+b+1)}+1-\pi\right\}$	$r\pi\frac{b}{a-1}\left\{\pi\frac{(a+r-1)(a+b-1)}{(a-2)(a-1)}+1-\pi\right\}$ [3]
$G_Y(s)$	$\frac{(b)_n}{(a+b)_n}\,{}_2F_1\{(-n,a);-b-n+1;1+\pi(s-1)\}$	$\frac{(a)_r}{(a+b)_r}\,{}_2F_1\{(r,b);r+a+b;1+\pi(s-1)\}$

2. Characteristics of specific convolution splitting distributions

Table 4: Characteristics of multinomial splitting (a) beta binomial, (b) beta negative binomial and (c) beta Poisson distribution.

(a)

Distribution	$\mathcal{M}_{\Delta_n}(\boldsymbol{\pi}) \underset{n}{\wedge} \beta\mathcal{B}_m(a,b)$
Supp($\boldsymbol{Y}$)	$\blacktriangle_m$
$P(\boldsymbol{Y}=\boldsymbol{y})$	$\binom{m}{\boldsymbol{y}} \frac{B(\lvert\boldsymbol{y}\rvert + a, m - \lvert\boldsymbol{y}\rvert + b)}{B(\alpha, b)} \prod_{j=1}^{J} \pi_j^{y_j}$
E($\boldsymbol{Y}$)	$m\frac{a}{a+b} \cdot \boldsymbol{\pi}$
Cov($\boldsymbol{Y}$)	$m\frac{a}{a+b} \cdot \left\{ \mathrm{diag}(\boldsymbol{\pi}) + \frac{b(n-1) - a(a+b+1)}{(a+b)(a+b+1)} \cdot \boldsymbol{\pi}\boldsymbol{\pi}^t \right\}$
$G_{\boldsymbol{Y}}(\boldsymbol{s})$	${}_2F_1\{(-m, a); \alpha + b; 1 - \boldsymbol{\pi}^t \boldsymbol{s}\}$
Marginals	$Y_j \sim \beta_{\pi_j}\mathcal{B}_m(a,b)$

(b)

Distribution	$\mathcal{M}_{\Delta_n}(\boldsymbol{\pi}) \underset{n}{\wedge} \beta\mathcal{NB}(r,a,b)$
Supp($\boldsymbol{Y}$)	$\mathbb{N}^J$
$P(\boldsymbol{Y}=\boldsymbol{y})$	$\binom{\lvert\boldsymbol{y}\rvert + r - 1}{\boldsymbol{y}, r-1} \frac{B(r+a, \lvert\boldsymbol{y}\rvert + b)}{B(a,b)} \prod_{j=1}^{J} \pi_j^{y_j}$
E($\boldsymbol{Y}$)	$r\frac{b}{a-1} \cdot \boldsymbol{\pi}$ [4]
Cov($\boldsymbol{Y}$)	$r\frac{b}{a-1} \cdot \left\{ \mathrm{diag}(\boldsymbol{\pi}) + \frac{a(b+r+1) + r(b-1) - b - 1}{(a-1)(a-2)} \cdot \boldsymbol{\pi}\boldsymbol{\pi}^t \right\}$ [5]
$G_{\boldsymbol{Y}}(\boldsymbol{s})$	$\frac{a_{(b)}}{(r+a)_{(b)}} {}_2F_1 \left\{ (r,b); r+a+b; \boldsymbol{\pi}^t \boldsymbol{s} \right\}$
Marginals	$Y_j \sim \beta_{\pi_j}\mathcal{NB}(r,a,b)$

(c)

Distribution	$\mathcal{M}_{\Delta_n}(\boldsymbol{\pi}) \underset{n}{\wedge} \beta_\lambda\mathcal{P}(a,b)$
Supp($\boldsymbol{Y}$)	$\mathbb{N}^J$
$P(\boldsymbol{Y}=\boldsymbol{y})$	$\frac{(a)_{\lvert\boldsymbol{y}\rvert}\lambda^{\lvert\boldsymbol{y}\rvert}}{(a+b)_{\lvert\boldsymbol{y}\rvert}} {}_1F_1(a + \lvert\boldsymbol{y}\rvert, a+b+\lvert\boldsymbol{y}\rvert; -\lambda) \prod_{j=1}^{J} \frac{\pi_j^{y_j}}{y_j!}$
E($\boldsymbol{Y}$)	$\lambda\frac{a}{a+b} \cdot \boldsymbol{\pi}$
Cov($\boldsymbol{Y}$)	$\lambda\frac{a}{a+b} \cdot \left\{ \mathrm{diag}(\boldsymbol{\pi}) + \lambda\frac{b}{(a+b)(a+b+1)} \cdot \boldsymbol{\pi}\boldsymbol{\pi}^t \right\}$
$G_{\boldsymbol{Y}}(\boldsymbol{s})$	${}_1F_1 \left\{ a; a+b; \lambda(\boldsymbol{\pi}^t \boldsymbol{s} - 1) \right\}$
Marginals	$Y_j \sim \beta_{\pi_j\lambda}\mathcal{P}(a,b)$

Table 5: Usual characteristics of Dirichlet multinomial splitting beta Poisson distribution respectively without constraint and with $a = |\alpha|$.

Distribution	$\mathcal{DM}_{\Delta_N}(\boldsymbol{\alpha}) \underset{N}{\wedge} \beta_\lambda\mathcal{P}(a,b)$	
Constraint	no constraint	$a = \lvert\boldsymbol{\alpha}\rvert$
Supp($\boldsymbol{Y}$)	$\mathbb{N}^J$	$\mathbb{N}^J$
$P(\boldsymbol{Y}=\boldsymbol{y})$	$\frac{(a)_{\lvert\boldsymbol{y}\rvert}\lambda^{\lvert\boldsymbol{y}\rvert}}{(a+b)_{\lvert\boldsymbol{y}\rvert}(\lvert\boldsymbol{\alpha}\rvert)_{\lvert\boldsymbol{y}\rvert}}\prod_{j=1}^{J}\frac{(\alpha_j)_{y_j}}{y_j!}\,{}_1F_1(a+\lvert\boldsymbol{y}\rvert;a+b+\lvert\boldsymbol{y}\rvert;-\lambda)$	$\frac{\lambda^{\lvert\boldsymbol{y}\rvert}}{(a+b)_{\lvert\boldsymbol{y}\rvert}}\prod_{j=1}^{J}\frac{(\alpha_j)_{y_j}}{y_j!}\,{}_1F_1(a+\lvert\boldsymbol{y}\rvert;a+b+\lvert\boldsymbol{y}\rvert;-\lambda)$
E($\boldsymbol{Y}$)	$\frac{\lambda a}{\lvert\boldsymbol{\alpha}\rvert(a+b)}\cdot\boldsymbol{\alpha}$	$\frac{\lambda}{\lvert\boldsymbol{\alpha}\rvert+b}\cdot\boldsymbol{\alpha}$
Cov($\boldsymbol{Y}$)	$\frac{\lambda a}{\lvert\boldsymbol{\alpha}\rvert(\lvert\boldsymbol{\alpha}\rvert+1)(a+b)}\cdot\Big[\Big\{\frac{\lambda b}{(a+b)(a+b+1)}+\frac{\lambda a}{a+b}+\lvert\boldsymbol{\alpha}\rvert+1\Big\}\cdot\mathrm{diag}(\boldsymbol{\alpha})+\Big\{\frac{\lambda b}{(a+b)(a+b+1)}-\frac{\lambda a}{\lvert\boldsymbol{\alpha}\rvert(a+b)}\Big\}\cdot\boldsymbol{\alpha}\boldsymbol{\alpha}^t\Big]$	$\frac{\lambda}{(\lvert\boldsymbol{\alpha}\rvert+1)(a+b)}\cdot\Big[\Big\{\frac{\lambda b}{(\lvert\boldsymbol{\alpha}\rvert+b)(\lvert\boldsymbol{\alpha}\rvert+b+1)}+\frac{\lambda\lvert\boldsymbol{\alpha}\rvert}{\lvert\boldsymbol{\alpha}\rvert+b}+\lvert\boldsymbol{\alpha}\rvert+1\Big\}\cdot\mathrm{diag}(\boldsymbol{\alpha})+\Big\{\frac{\lambda b}{(\lvert\boldsymbol{\alpha}\rvert+b)(\lvert\boldsymbol{\alpha}\rvert+b+1)}-\frac{\lambda}{a+b}\Big\}\cdot\boldsymbol{\alpha}\boldsymbol{\alpha}^t\Big]$
$G_{\boldsymbol{Y}}(\boldsymbol{s})$	$\sum_{\boldsymbol{y}\in\mathbb{N}^J}\sum_{k\in\mathbb{N}}\frac{(a)_{\lvert\boldsymbol{y}\rvert+k}\prod_{j=1}^{J}(\alpha_j)_{y_j}}{(a+b)_{\lvert\boldsymbol{y}\rvert+k}(\lvert\boldsymbol{\alpha}\rvert)_{\lvert\boldsymbol{y}\rvert}}\frac{(-\lambda)^k}{k!}\prod_{j\in\mathcal{J}}\frac{(\lambda s_j)^{y_j}}{y_j!}$	${}_0^{J+1}F_1\{(\boldsymbol{\alpha},a+\lvert\boldsymbol{y}\rvert);a+b;(\lambda\cdot\boldsymbol{s},-\lambda)\}$
Marginals	$Y_j\sim\beta_\lambda^2\mathcal{P}\big(\alpha_j,\lvert\boldsymbol{\alpha}_{-j}\rvert,a,b\big)$	$Y_j\sim\beta_\lambda\mathcal{P}\big(\alpha_j,\lvert\boldsymbol{\alpha}_{-j}\rvert+b\big)$

Table 6: Usual characteristics of Dirichlet multinomial splitting binomial, negative binomial and Poisson distribution.

Distribution	$\mathcal{DM}_{\Delta_N}(\boldsymbol{\alpha}) \underset{N}{\wedge} \mathcal{L}(\psi)$		
$\mathcal{L}(\psi)$	$\mathcal{B}_n(p)$	$\mathcal{NB}(r,p)^6$	$\mathcal{P}(\lambda)$
Supp($\boldsymbol{Y}$)	$\blacktriangle_n$	$\mathbb{N}^J$	$\mathbb{N}^J$
$P(\boldsymbol{Y}=\boldsymbol{y})$	$\frac{\Gamma(n+1)p^{\lvert\boldsymbol{y}\rvert}(1-p)^{n-\lvert\boldsymbol{y}\rvert}}{\Gamma(n-\lvert\boldsymbol{y}\rvert+1)(\lvert\boldsymbol{\alpha}\rvert)_{\lvert\boldsymbol{y}\rvert}}\prod_{j=1}^{J}\frac{(\alpha_j)_{y_j}}{y_j!}$	$(1-p)^r\frac{(r)_{\lvert\boldsymbol{y}\rvert}p^{\lvert\boldsymbol{y}\rvert}}{(\lvert\boldsymbol{\alpha}\rvert)_{\lvert\boldsymbol{y}\rvert}}\prod_{j=1}^{J}\frac{(\alpha_j)_{y_j}}{y_j!}$	$e^{-\lambda}\frac{\lambda^{\lvert\boldsymbol{y}\rvert}}{(\lvert\boldsymbol{\alpha}\rvert)_{\lvert\boldsymbol{y}\rvert}}\prod_{j=1}^{J}\frac{(\alpha_j)_{y_j}}{y_j!}$
E($\boldsymbol{Y}$)	$\frac{np}{\lvert\boldsymbol{\alpha}\rvert}\cdot\boldsymbol{\alpha}$	$\frac{rp}{\lvert\boldsymbol{\alpha}\rvert(1-p)}\cdot\boldsymbol{\alpha}$	$\frac{\lambda}{\lvert\boldsymbol{\alpha}\rvert}\cdot\boldsymbol{\alpha}$
Cov($\boldsymbol{Y}$)	$\frac{np}{\lvert\boldsymbol{\alpha}\rvert(\lvert\boldsymbol{\alpha}\rvert+1)}\cdot\{\{(n-1)p+\lvert\boldsymbol{\alpha}\rvert+1\}\cdot\mathrm{diag}(\boldsymbol{\alpha})-\frac{p(n+\lvert\boldsymbol{\alpha}\rvert)}{\lvert\boldsymbol{\alpha}\rvert}\cdot\boldsymbol{\alpha}\boldsymbol{\alpha}^t\}$	$\frac{rp}{\lvert\boldsymbol{\alpha}\rvert(\lvert\boldsymbol{\alpha}\rvert+1)(1-p)}\cdot[\frac{(r-\lvert\boldsymbol{\alpha}\rvert)p+\lvert\boldsymbol{\alpha}\rvert+1}{1-p}\cdot\mathrm{diag}(\boldsymbol{\alpha})+\frac{(\lvert\boldsymbol{\alpha}\rvert-r)p}{\lvert\boldsymbol{\alpha}\rvert(1-p)}\cdot\boldsymbol{\alpha}\boldsymbol{\alpha}^t]$	$\frac{\lambda}{\lvert\boldsymbol{\alpha}\rvert(\lvert\boldsymbol{\alpha}\rvert+1)}\cdot\{(\lambda+\lvert\boldsymbol{\alpha}\rvert+1)\cdot\mathrm{diag}(\boldsymbol{\alpha})-\frac{\lambda}{\lvert\boldsymbol{\alpha}\rvert}\cdot\boldsymbol{\alpha}\boldsymbol{\alpha}^t\Big\}$
$G_{\boldsymbol{Y}}(\boldsymbol{s})$	$(1-p)^n\,{}_1^JF_1(-n;\boldsymbol{\alpha};\lvert\boldsymbol{\alpha}\rvert;-\frac{p}{1-p}\cdot\boldsymbol{s})$	$(1-p)^r\,{}_1^JF_1(r;\boldsymbol{\alpha};\lvert\boldsymbol{\alpha}\rvert;p\cdot\boldsymbol{s})$	$e^{-\lambda}\,{}_0^JF_1(\boldsymbol{\alpha};\lvert\boldsymbol{\alpha}\rvert;\lambda\cdot\boldsymbol{s})$
Marginals	$Y_j\sim\beta_p\mathcal{B}_n(\alpha_j,\lvert\boldsymbol{\alpha}_{-j}\rvert)$	$Y_j\sim\beta_p\mathcal{NB}(r,\alpha_j,\lvert\boldsymbol{\alpha}_{-j}\rvert)$	$Y_j\sim\beta_\lambda\mathcal{P}(\alpha_j,\lvert\boldsymbol{\alpha}_{-j}\rvert)$

3. Inference of singular and univariate regressions

Table 7: References of inference procedures for (a) singular regressions and (b) univariate regressions.

(a)

Regression	Notation	Canonical link function	Inference
Multinomial	$\mathcal{M}_{\Delta_n}\{\boldsymbol{\pi}(\boldsymbol{x})\}$	$\pi_j = \dfrac{\exp(\boldsymbol{x}^t\boldsymbol{\beta}_j)}{1+\sum_{k=1}^{J-1}\exp(\boldsymbol{x}^t\boldsymbol{\beta}_k)},\ j=1,\ldots,J-1$	See [11]
Dirichlet multinomial	$\mathcal{DM}_{\Delta_n}\{\boldsymbol{\alpha}(\boldsymbol{x})\}$	$\alpha_j = \exp(\boldsymbol{x}^t\boldsymbol{\beta}_j),\ j=1,\ldots,J$	See [19]
Generalized Dirichlet multinomial	$\mathcal{GDM}_{\Delta_n}\{\boldsymbol{\alpha}(\boldsymbol{x}),\boldsymbol{\beta}(\boldsymbol{x})\}$	$a_j = \exp(\boldsymbol{x}^t\boldsymbol{\beta}_{1,j}),\ j=1,\ldots,J-1$ $b_j = \exp(\boldsymbol{x}^t\boldsymbol{\beta}_{2,j}),\ j=1,\ldots,J-1$	See [19]
Logistic normal multinomial	$\mathcal{LNM}_{\Delta_n}\{\boldsymbol{\mu}(\boldsymbol{x}),\Sigma\}$	$\mu_j = \boldsymbol{x}^t\boldsymbol{\beta}_j,\ j=1,\ldots,J-1$ $\pi_j = \dfrac{\exp(\mu_j)}{1+\sum_{k=1}^{J-1}\exp(\mu_k)},\ j=1,\ldots,J-1$	See [18]

(b)

Regression	Notation	Canonical link function	Parameter inference
Poisson	$\mathcal{P}\{\lambda(\boldsymbol{x})\}$	$\lambda = \exp(\boldsymbol{x}^t\boldsymbol{\beta})$	See [10]
Binomial	$\mathcal{B}_n\{p(\boldsymbol{x})\}$	$p = \dfrac{\exp(\boldsymbol{x}^t\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}^t\boldsymbol{\beta})}$	See [10] for n known
Negative binomial	$\mathcal{NB}\{r,p(\boldsymbol{x})\}$	$p = \exp(\boldsymbol{x}^t\boldsymbol{\beta})$	See [6]
Beta Poisson	$\beta\mathcal{P}\{a(\boldsymbol{x}),b(\boldsymbol{x})\}$	$\dfrac{a}{a+b} = \dfrac{\exp(\boldsymbol{x}^t\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}^t\boldsymbol{\beta})}$	See [15]
Beta binomial	$\beta\mathcal{B}_n\{a(\boldsymbol{x}),b(\boldsymbol{x})\}$	$\dfrac{a}{a+b} = \dfrac{\exp(\boldsymbol{x}^t\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}^t\boldsymbol{\beta})}$	See [4] and [9] for n known
Beta negative binomial	$\beta\mathcal{NB}\{r,a(\boldsymbol{x}),b(\boldsymbol{x})\}$	$\dfrac{a}{a+b} = \dfrac{\exp(\boldsymbol{x}^t\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}^t\boldsymbol{\beta})}$	See [12, 13]

References

[1] Blumenthal, S., Dahiya, R.C., 1981. Estimating the binomial parameter n. Journal of the American Statistical Association 76, 903–909.

[2] Dai, H., Bao, Y., Bao, M., 2012. Maximum likelihood estimate for the dispersion parameter of the negative binomial distribution. Statistics & Probability Letters 83, 21–27.

[3] Dunkl, C.F., 2013. Products of Beta distributed random variables. ArXiv e-prints `arXiv:1304.6671`.

[4] Forcina, A., Franconi, L., 1988. Regression analysis with the beta-binomial distribution. Rivista di Statistica Applicata 21.

[5] Gurland, J., 1958. A generalized class of contagious distributions. Biometrics 14, 229–249.

[6] Hilbe, J.M., 2011. Negative binomial regression. Cambridge University Press.

[7] Irwin, J.O., 1968. The generalized Waring distribution applied to accident theory. Journal of the Royal Statistical Society. Series A (General) , 205–225.

[8] Johnson, N., Kemp, A., Kotz, S., 1993. Univariate discrete distributions. Wiley-Interscience.

[9] Lesnoff, M., Lancelot, R., 2012. aod: Analysis of Overdispersed Data. URL: `http://cran.r-project.org/package=aod`. r package version 1.3.

[10] McCullagh, P., Nelder, J., 1989. Generalized linear models. Chapman & Hall, London.

[11] Peyhardi, J., Trottier, C., Guédon, Y., 2015. A new specification of generalized linear models for categorical responses. Biometrika 102, 889–906.

[12] Rodríguez-Avi, J., Conde-Sánchez, A., Sáez-Castillo, A., Olmo-Jiménez, M., Martínez-Rodríguez, A.M., 2009. A generalized Waring regression model for count data. Computational Statistics & Data Analysis 53, 3717–3725.

[13] Saez-Castillo, A.J., Vilchez-Lopez, S., Olmo-Jimenez, M.J., Rodriguez-Avi, J., Conde-Sanchez, A., Martinez-Rodriguez, A.M., 2017. GWRM: Generalized Waring Regression Model for Count Data. URL: `https://cran.r-project.org/package=GWRM`. gWRM R package version 2.1.0.3.

[14] Tripathi, R.C., Gupta, R.C., Gurland, J., 1994. Estimation of parameters in the beta binomial model. Annals of the Institute of Statistical Mathematics 46, 317–331.

[15] Vu, T.N., Wills, Q.F., Kalari, K.R., Niu, N., Wang, L., Rantalainen, M., Pawitan, Y., 2016. Beta-Poisson model for single-cell rna-seq data analyses. Bioinformatics 32, 2128–2135.

[16] Whitby, O., 1972. Estimation of parameters in the generalized beta distribution. Ph.D. thesis. Stanford University.

[17] Xekalaki, E., 1981. Chance mechanisms for the univariate generalized Waring distribution and related characterizations, in: Statistical distributions in scientific work. Springer, pp. 157–171.

[18] Xia, F., Chen, J., Fung, W.K., Li, H., 2013. A logistic normal multinomial regression model for microbiome compositional data analysis. Biometrics 69, 1053–1063.

[19] Zhang, Y., Zhou, H., Zhou, J., Sun, W., 2017. Regression models for multivariate count data. Journal of Computational and Graphical Statistics 26, 1–13.