Generic Component Structure

Consider a combinatorial object with a component structure $(C_1, \ldots, C_n)$ whose joint distribution is of the form

$$P(C_1 = c_1, \ldots, C_n = c_n) = \begin{cases} \frac{f(c_1)\cdots f(c_n)}{g(t,n)} & \substack{c_1+\ldots+c_n=t \\ c_j \in \{\tau,\tau+1,\ldots\}\ \forall j} \\ 0 & \text{else} \end{cases}$$

where

$$g(t,n) \;=\; \sum_{\substack{(c_1,\ldots,c_n)\ni \\ c_1+\ldots+c_n=t \\ c_j\in\{\tau,\tau+1,\ldots\}\ \forall j}} \!\!\cdots\!\! \sum f(c_1)\cdots f(c_n).$$

for $\tau \in \{0, 1, \ldots\}$ and for some function f such that $f(x) \geq 0$ for $x \geq \tau$. Then for $t \geq n\tau$

$$\mathrm{E}(\Psi(C_1, \ldots, C_n)) \;=\; \frac{1}{g(t,n)\, t!} \left(\frac{d^t}{d\theta^t} \left(\left(\prod_{j=1}^{n} \eta(\theta, \delta_j) \right) \mathrm{E}(\Psi(Z_1, \ldots, Z_n)) \right) \Bigg|_{\theta=0} \right)$$

where $Z_1, \ldots, Z_n$ are independent random variables such that for each $j \in \{1, \ldots, t\}$

$$P(Z_j = z) = \frac{f(z, \delta_j)}{\eta(\theta, \delta_j)}\, \theta^z \qquad z \in \{\tau, \tau+1, \ldots\}$$

for real $\theta \geq 0$ and $\eta(\theta, \delta_j) = \sum_{z=\tau}^{\infty} f(z, \delta_j)\, \theta^z$.

Consider a fixed set of nonnegative integers $(q_1, \ldots, q_n)$. If in the above model $C_i \geq q_i$, we will say type i has reached its quota (by time t). Let $W_{r:Q}$ represent the *waiting time* (i.e. the smallest value of t) until exactly r different types have reached their quota.

Let

(1) $\mathcal{A}_j$ be the event that $Z_j < q_j$.

(2) $\mathbb{A}_{Q:r}$ be the event that at least $n-r+1$ of the (independent) events $\mathcal{A}_1,\ldots,\mathcal{A}_n$ occur

It follows that

$$W_{r:Q} > t \Leftrightarrow (C_1,\ldots,C_n) \in \mathbb{A}_{Q:r}$$

$$\mathrm{E}\left(W_{r:Q}^{[k]}\right) = k\sum_{t=0}^{\infty} P(W_{r:Q} > t)\frac{(t+k-1)!}{t!}$$

$$= k\sum_{t=0}^{\infty} P((C_1,\ldots,C_n) \in \mathbb{A}_{Q:r})\frac{(t+k-1)!}{t!}$$

$$= k\sum_{t=0}^{\infty}\frac{1}{g(t,n)\,t!}\left(\frac{d^t}{d\theta^t}\left(\left(\prod_{j=1}^{n}\eta(\theta,\delta_j)\right)P((Z_1,\ldots,Z_n)\in\mathbb{A}_{Q:r})\right)\Bigg|_{\theta=0}\right)\frac{(t+k-1)!}{t!}$$

$$= k\frac{d^t}{d\theta^t}\left(\sum_{t=0}^{\infty}\frac{1}{g(t,n)\,t!}\left(\prod_{j=1}^{n}\eta(\theta,\delta_j)\right)P((Z_1,\ldots,Z_n)\in\mathbb{A}_{Q:r})\frac{(t+k-1)!}{t!}\right)\Bigg|_{\theta=0}$$

$$\mathrm{E}\left(W_{r:Q}^{[k]}\right) = k\sum_{t=0}^{\infty} P((C_1,\ldots,C_n) \in \mathbb{A}_{Q:r})\left(\frac{d^r}{d\theta^r}\left(h(\theta,k)\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t\right)\Bigg|_{\theta=0}\right)$$

$$= k\frac{d^r}{d\theta^r}\left(h(\theta,k)\sum_{t=0}^{\infty}P((C_1,\ldots,C_n)\in\mathbb{A}_{Q:r})\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t\right)\Bigg|_{\theta=0}$$

$$= k\frac{d^r}{d\theta^r}(h(\theta,k)P((Z_1,\ldots,Z_n)\in\mathbb{A}_{Q:r}))|_{\theta=0}$$

$$\mathrm{E}\left(W_{r:Q}^{[k]}\right) = k\sum_{t=0}^{\infty}P(W_{r:Q}>t)\frac{(t+k-1)!}{t!}$$

$$= k\sum_{t=0}^{\infty}P((C_1,\ldots,C_n)\in\mathbb{A}_{Q:r})\frac{(t+k-1)!}{t!}$$

Now suppose $h(\theta,k,n)$ and Δ are chosen such that

$$\int_{\theta\in\Delta}h(\theta,k,n)\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t d\theta = \frac{(t+k-1)!}{t!}.$$

Then

$$\mathrm{E}\left(W_{r:Q}^{[k]}\right) = k\sum_{t=0}^{\infty}P(N_Q>n-r)\left(\int_{\theta\in\Delta}h(\theta,k,n)\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t d\theta\right)$$

$$= k\int_{\theta\in\Delta}h(\theta,k,n)\left(\sum_{t=0}^{\infty}P(N_Q>n-r)\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t\right)d\theta$$

$$= k\int_{\theta\in\Delta} h(\theta,k,n)\; P\Big(N_Q^{psd} > n-r\Big) d\theta$$

Or suppose $h(\theta,k,n)$ and r are chosen such that

$$\frac{d^r}{d\theta^r}\left(h(\theta,k)\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t\right)\Bigg|_{\theta=0} = \frac{(t+k-1)!}{t!}.$$

Then

$$\mathrm{E}\Big(W_{r:Q}^{[k]}\Big) = k\sum_{t=0}^{\infty} P((C_1,\ldots,C_n)\in\mathbb{A}_{Q:r})\left(\frac{d^r}{d\theta^r}\left(h(\theta,k)\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t\right)\Bigg|_{\theta=0}\right)$$

$$= k\frac{d^r}{d\theta^r}\left(h(\theta,k)\sum_{t=0}^{\infty} P((C_1,\ldots,C_n)\in\mathbb{A}_{Q:r})\frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t\right)\Bigg|_{\theta=0}$$

$$= k\frac{d^r}{d\theta^r}(h(\theta,k)P((Z_1,\ldots,Z_n)\in\mathbb{A}_{Q:r}))|_{\theta=0}$$

$$g(n,t) = \sum\cdots\sum_{\substack{(x_1,\ldots,x_n)\ni\\ x_1+\ldots+x_n=t\\ x_j\in\{\tau,\tau+1,\ldots\}\ \forall j}} a_{x_1}\cdots a_{x_n}$$

$$b_1(\theta)\cdots b_n(\theta) = \sum_{t=\tau n}^{\infty} g(n,t)\,\theta^t$$

$$P\left(\sum_{i=1}^{n} X_i = t\right) = \left(\sum_{\substack{(x_1,\ldots,x_n)\ni \\ x_1+\ldots+x_n=t \\ x_j\in\{\tau,\tau+1,\ldots\}\,\forall\, j}} \!\!\cdots\! \sum a_{x_1}\cdots a_{x_n} \right) \frac{1}{b_1(\theta)\cdots b_n(\theta)}\theta^t = \frac{g(n,t)}{b_1(\theta)\cdots b_n(\theta)}\theta^t$$

$$b_1(\theta)\cdots b_n(\theta) = \sum_{t=\tau n}^{\infty} \left(\sum_{\substack{(x_1,\ldots,x_n)\ni \\ x_1+\ldots+x_n=t \\ x_j\in\{\tau,\tau+1,\ldots\}\,\forall\, j}} \!\!\cdots\! \sum a_{x_1}\cdots a_{x_n} \right) \theta^t = \sum_{t=\tau n}^{\infty} g(n,t)\,\theta^t$$

If X follows the power series family of distributions, that is, if

$$P(X = x) = \frac{a_x\theta^x}{b(\theta)} \qquad x \in \mathbb{T} = \{\tau, \tau+1, \ldots\}$$

for some nonnegative integer τ, real $\theta \geq 0$, $a_x \geq 0$ for all $x \in \mathbb{T}$ and $b(\theta) = \sum_{x=\tau}^{\infty} a_x\theta^x$, then

$$P\left(\sum_{i=1}^{n} X_i = t\right) = \left(\sum_{\substack{(x_1,\ldots,x_n)\ni \\ x_1+\ldots+x_n=t \\ x_j\in\{\tau,\tau+1,\ldots\}\,\forall\, j}} \!\!\cdots\! \sum a_{x_1}\cdots a_{x_n} \right) \frac{1}{b_1(\theta)\cdots b_n(\theta)}\theta^t$$

$$g(n,t) = \sum_{\substack{(c_1,\ldots,c_n)\ni \\ c_1+\ldots+c_n=t \\ c_j\in\{\tau,\tau+1,\ldots\}\,\forall\, j}} \!\!\cdots\! \sum a(c_1)\cdots a(c_n).$$

Example

Polya

$$P(Y_i = y) = \binom{y + m_i - 1}{m_i - 1} p^{m_i} (1-p)^y \qquad y \in \{0, 1, \dots\} \text{ and } 0 \le p \le 1$$

Reparametrize with $\theta = 1 - p$

$$P(Y_i = x) = \binom{x + m_i - 1}{m_i - 1} (1-\theta)^{m_i} \theta^x \qquad x \in \{0, 1, \dots\} \text{ and } 0 \le \theta \le 1$$

$a_x = \binom{x+m_i-1}{m_i-1}$

$b_i(\theta) = (1-\theta)^{-m_i}$

$b_1(\theta)\cdots b_t(\theta) = (1-\theta)^{-m_1-\ldots-m_t} = (1-\theta)^{-M}$

$$g(n,t) \; = \; \sum_{\substack{(x_1,\ldots,x_n)\ni \\ x_1+\ldots+x_n=t \\ x_j\in\{\tau,\tau+1,\ldots\}\,\forall\, j}} \cdots \sum \; a(x_1)\cdots a(x_n)$$

$$= \; \sum_{\mathbb{S}_t^n} \binom{x_1+m_1-1}{m_1-1} \cdots \binom{x_t+m_t-1}{m_t-1}$$

$$= \binom{t+(m_1+\ldots+m_t)-1}{(m_1+\ldots+m_t)-1} = \binom{t+M-1}{M-1}$$

$\frac{(t+k-1)!}{t!}$

$= \int_0^1 h(\theta,k) g(n,t) \frac{1}{b_1(\theta)\cdots b_n(\theta)} \theta^t d\theta$

$= \int_0^1 h(\theta,k) \binom{t+M-1}{M-1} \frac{1}{(1-\theta)^{-M}} \theta^t d\theta$

$$= \binom{t+M-1}{M-1} \int_0^1 h(\theta, k)(1-\theta)^M \theta^t d\theta$$

$$= \binom{t+M-1}{M-1} \frac{(M-1)!}{(M-k-1)!} \int_0^1 (1-\theta)^{-k-1} \theta^{k-1} (1-\theta)^M \theta^t d\theta$$

$$= \binom{t+M-1}{M-1} \frac{(M-1)!}{(M-k-1)!} \int_0^1 (1-\theta)^{M-k-1} \theta^{t+k-1} d\theta$$

$$= \binom{t+M-1}{M-1} \frac{(M-1)!}{(M-k-1)!} \left(\frac{(t+k-1)!(M-k-1)!}{(M+t-1)!} \right)$$

$$= \frac{(t+M-1)!\,(t+k-1)!\,(M-k-1)!}{(M-1)!\,t!\,(M+t-1)!} \frac{(M-1)!}{(M-k-1)!}$$

$$= \frac{(M-k-1)!}{(M-1)!} \frac{(t+k-1)!}{t!} \frac{(M-1)!}{(M-k-1)!}$$

$$= \frac{(t+k-1)!}{t!}$$

$$h(\theta, k) = \frac{(M-1)!}{(M-k-1)!} (1-\theta)^{-k-1} \theta^{k-1}$$

where

$$M = m_1 + \ldots + m_n$$