Negative Binomial Randomization

1/28/02

Let $\mathbb{S}^n$ be the product space $\{0,1,\dots\}\times\cdots\times\{0,1,\dots\}$ and let $\mathbb{S}_t^n$ be the set of all vectors $(s_1,\dots,s_n)$ in $\mathbb{S}^n$ such that $s_1+\dots+s_n=t$. Define $\left(X_{1,t},\dots,X_{n,t}\right)$ to be that random vector such that

$$P\left(X_{1,t}=x_1,\dots,X_{n,t}=x_n\right)=\begin{cases}\dfrac{\binom{x_1+m_1-1}{m_1-1}\cdots\binom{x_n+m_n-1}{m_n-1}}{\binom{t+M-1}{M-1}} & (x_1,\dots,x_n)\in\mathbb{S}_t^n\\ 0 & \text{else}\end{cases}$$

where $M=m_1+\cdots+m_n$.

To see that this is a valid probability distribution we consider a row of M distinguishable (numbered) urns, of which there are m_j urns of color j, $m_1+\cdots+m_n=M$. It follows that

$$\sum_{\mathbb{S}_t^n}\binom{x_1+m_1-1}{m_1-1}\cdots\binom{x_n+m_n-1}{m_n-1}$$

equals the total number of ways of distributing t identical balls into $m_1+\cdots+m_n=M$ distinguishable urns. But of course this also equals $\binom{t+M-1}{M-1}$, therefore this is a valid probability distribution.

We will consider four models leading to the above probability distribution function.

(1) Polya-Eggenberger Urn Model

Suppose an urn contains m_j balls of color j, $j=1,\dots,n$. At each drawing a single ball is removed from the urn at random and returned with another ball of the same color. This is referred to as the Polya-Eggenberger Urn Model. Let $X_{j,t}$ equal the times a ball of color j was selected in the first t draws. Then

$$P\left(X_{1,t}=x_1,\dots,X_{n,t}=x_n\right)=\begin{cases}\dfrac{\binom{x_1+m_1-1}{m_1-1}\cdots\binom{x_n+m_n-1}{m_n-1}}{\binom{t+M-1}{M-1}} & (x_1,\dots,x_n)\in\mathbb{S}_t^n\\ 0 & \text{else}\end{cases}$$

with $M=m_1+\cdots+m_n$.

(2) Exceedances/Placements

Let $V_1, \dots, V_{M-1}$ be a random sample of size $M-1$ from a continuous distribution F. Let $Z_1, \dots, Z_t$ be a second random sample of size t from this same continuous distribution F.

For integers $1 \le r_1 < \cdots < r_{n-1} \le M-1$, define the random variables

$$\begin{aligned} X_{1,t} &= \#\, Z_i\text{'s in } \left(-\infty, V_{(r_1)}\right) \\ X_{2,t} &= \#\, Z_i\text{'s in } \left(V_{(r_1)}, V_{(r_2)}\right) \\ &\vdots \\ X_{n-1,t} &= \#\, Z_i\text{'s in } \left(V_{(r_{n-2})}, V_{(r_{n-1})}\right) \\ X_{n,t} &= \#\, Z_i\text{'s in } \left(V_{(r_{n-1})}, \infty\right) \end{aligned}$$

and define the constants

$$\begin{aligned} m_1 &= r_1 \\ m_2 &= r_2 - r_1 \\ &\vdots \\ m_{n-1} &= r_{n-1} - r_{n-2} \\ m_n &= M - r_{n-1}. \end{aligned}$$

Then,

$$P\left(X_{1,t} = x_1, \dots, X_{n,t} = x_n\right) = \begin{cases} \dfrac{\binom{x_1 + m_1 - 1}{m_1 - 1} \cdots \binom{x_n + m_n - 1}{m_n - 1}}{\binom{t + M - 1}{M - 1}} & (x_1, \dots, x_n) \in \mathbb{S}_t^n \\ 0 & \text{else} \end{cases}$$

with $M = m_1 + \cdots + m_n$.

(3) Multivariate Negative Hypergeometric

Suppose an urn contains $t + M - 1$ balls of which $M - 1$ are white and the remaining t are black. Assume that we continue to draw out balls from this urn at random and without replacement until all $M - 1$ white balls have been selected.

For integers $1 \le r_1 < \cdots < r_{n-1} \le M-1$, define the random variables

$$\begin{aligned} X_{1,t} &= \#\text{ black balls selected prior to the } (r_1)^{st} \text{ white ball} \\ X_{2,t} &= \#\text{ black balls selected between the} (r_1)^{st} \text{and the } (r_2)^{nd} \text{ white balls} \\ &\vdots \end{aligned}$$

$X_{n-1,t}$ = # black balls selected between the $(r_{n-2})^{th}$ and the $(r_{n-1})^{th}$ white balls
$X_{n,t}$ = # black balls selected after the $(r_{n-1})^{th}$ white ball

and define the constants

$$\begin{aligned} m_1 &= r_1 \\ m_2 &= r_2 - r_1 \\ &\vdots \\ m_{n-1} &= r_{n-1} - r_{n-2} \\ m_n &= M - r_{n-1}. \end{aligned}$$

Then,

$$P(X_{1,t} = x_1, \dots, X_{n,t} = x_n) = \begin{cases} \dfrac{\binom{x_1 + m_1 - 1}{m_1 - 1} \cdots \binom{x_n + m_n - 1}{m_n - 1}}{\binom{t + M - 1}{M - 1}} & (x_1, \dots, x_n) \in \mathbb{S}_t^n \\ 0 & \text{else} \end{cases}$$

with $M = m_1 + \cdots + m_n$.

(4) Multivariate Bose-Einstein Allocation

Consider a row of M distinguishable (numbered) urns, of which there are m_j urns of color j, $m_1 + \cdots + m_n = M$. Suppose that t identical balls are distributed among the M urns according to the Bose-Einstein allocation law. That is, all $\binom{t + M - 1}{M - 1}$ allocations are assumed to be equally likely to have occurred.

Let $X_{j,t}$ equal the balls distributed among the m_j urns of color j. Then

$$P(X_{1,t} = x_1, \dots, X_{n,t} = x_n) = \begin{cases} \dfrac{\binom{x_1 + m_1 - 1}{m_1 - 1} \cdots \binom{x_n + m_n - 1}{m_n - 1}}{\binom{t + M - 1}{M - 1}} & (x_1, \dots, x_n) \in \mathbb{S}_t^n \\ 0 & \text{else.} \end{cases}$$

Model 4 conveniently assumes all possible allocations to be equally likely. This raises the question of how one can actually sequentially distribute indistinguishable balls into distinguishable urns so that all possible allocations are in fact equally likely.

Consider a set of n distinguishable urns and let $s_{i,j-1}$ represent the number of (identical) balls in Urn i after $j - 1$ balls have been sequentially distributed among the n urns and let $p_{i,j}$

represent the probability that the j^{th} ball to be sequentially distributed will go in Urn i.

If we define

$$p_{i,j} = \frac{s_{i,j-1} + 1}{n + j - 1}$$

then it is not difficult to show (see Ijiri and Simon, "Some Distributions Associated with Bose-Einstein Statistics", Proc. Nat. Acad. Sci. USA, Vol 72, No. 5, May 1975, pages 1654-1657, for a nice inductive proof) that all possible allocations of m balls are equally likely to occur for each $m = 1,2,\ldots$.

Now define $Y_1,\ldots,Y_n$ to be independent negative binomial random variables such that $Y_i \sim$ Negative Binomial(m_i, p).

i.e.

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

Theorem 1.

For $t \geq 0$,

$$\mathrm{E}_t\big(g(X_1,\ldots,X_n)\big) = \frac{(-1)^t}{\binom{M+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^M \mathrm{E}\big(g(Y_1,\ldots,Y_n)\big)\right)\Bigg|_{p=1}.$$

Corollary

Let $\mathcal{A} \subset \mathbb{S}^n$ and define $\mathcal{A}_t = \mathcal{A} \cap \mathbb{S}_t^n$. Then for $t \geq 0$,

$$P\Big(\big(X_{1,t},\ldots,X_{n,t}\big) \in \mathcal{A}_t\Big) = \frac{(-1)^t}{\binom{M+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^M P\big((Y_1,\ldots,Y_n) \in \mathcal{A}\big)\right)\Bigg|_{p=1} \quad \blacksquare$$

Waiting Times

In keeping with the notation of our model for exceedances/placements we take $V_1,\ldots,V_{M-1}$ to be a random sample of size $M-1$ from a continuous distribution F.

Let $Z_1, \dots, Z_t$ be a second random sample of size t from this same continuous distribution F.

For integers $1 \le \alpha_1 < \cdots < \alpha_{n-1} \le M-1$, define the random variables

$$\begin{aligned} X_{1,t} &= \#\, Z_i\text{'s in } \left(-\infty, V_{(\alpha_1)}\right) \\ X_{2,t} &= \#\, Z_i\text{'s in } \left(V_{(\alpha_1)}, V_{(\alpha_2)}\right) \\ &\vdots \\ X_{n-1,t} &= \#\, Z_i\text{'s in } \left(V_{(\alpha_{n-2})}, V_{(\alpha_{n-1})}\right) \\ X_{n,t} &= \#\, Z_i\text{'s in } \left(V_{(\alpha_{n-1})}, \infty\right) \end{aligned}$$

When $X_{i,t} = q_i$ we will say interval i has reached its *quota*. Let $W_{r:Q}$ represent the waiting time (i.e. the smallest value of t) until exactly r different intervals have reached their quota.

Theorem 2.

For $M \ge k+1$

$$\mathrm{E}\left(W_{r:Q}^{[k]}\right) = \frac{k(M-1)!}{(M-k-1)!}\int_0^1 p^{-k-1}(1-p)^{k-1} P\left((Y_1, \dots, Y_n) \in \mathbb{A}_{Q:r}\right) dp$$

and for $M \le k$

$$\mathrm{E}\left(W_{r:Q}^{[k]}\right) = \frac{(-1)^{k-M} k}{(M-1)!}\frac{d^{(k-M)}}{dp^{(k-M)}}\left(\frac{(1-p)^{k-1}}{p^M} P\left((Y_1, \dots, Y_n) \in \mathbb{A}_{Q:r}\right)\right)\Bigg|_{p=0}$$

where

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

(2) $\mathcal{A}_j$ is the event that $Y_j < q_j$.

(3) Y_1,…,Y_n are independent negative binomial random variables such that

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

for $j = 1$,…,n, where

$$
\begin{aligned}
m_1 &= \alpha_1 \\
m_2 &= \alpha_2 - \alpha_1 \\
&\vdots \\
m_{n-1} &= \alpha_{n-1} - \alpha_{n-2} \\
m_n &= M - \alpha_{n-1}.
\end{aligned}
$$

■

Theorem 3. $<$**Truncated Negative Binomial Randomization**$>$

Applications

Problem 1.

(a) Let $C_k = \sum_{i=1}^{n} \mathrm{I}_{\{k\}}(X_i) = \#\, X_i\text{'}s$ which equal k. Show that

$$
P(C_1 = c_1, \dots, C_t = c_t) = \begin{cases} \dfrac{\Psi_C(m_1, \dots, m_n)}{\binom{t+M-1}{M-1} c_0!\, c_1! \cdots c_t!} & \sum_{i=0}^{t} c_i = n \text{ and} \sum_{i=0}^{t} i c_i = t \\ 0 & \text{else} \end{cases}
$$

where

$$
\Psi_C(m_1, \dots, m_n) = \sum_\pi \binom{m_{\pi(1)} - 1}{m_{\pi(1)} - 1} \cdots \binom{m_{\pi(c_0)} - 1}{m_{\pi(c_0)} - 1} \times \binom{m_{\pi(c_0+1)} + 0}{m_{\pi(c_0+1)} - 1} \cdots \binom{m_{\pi(c_1+c_0)} + 0}{m_{\pi(c_1+c_0)} - 1} \times \cdots \times \binom{m_{\pi(c_{t-1}+\dots+c_0+1)} + (t-1)}{m_{\pi(c_{t-1}+\dots+c_0+1)} - 1} \cdots \binom{m_{\pi(c_t+\dots+c_0)} + (t-1)}{m_{\pi(c_t+\dots+c_0)} - 1}
$$

and the sum is over all permutations π of $\{1, \dots, n\}$.

(b) If $m_1 = \dots = m_n = m$, then $M = mn$,

$$
\Psi_C(m, \dots, m) = n! \binom{m-1}{m-1}^{c_0} \binom{m+0}{m-1}^{c_1} \cdots \binom{m+(t-1)}{m-1}^{c_t}
$$

and

$$P(C_1 = c_1, \dots, C_t = c_t)$$

$$= \begin{cases} \dfrac{n! \binom{m-1}{m-1}^{c_0} \binom{m+0}{m-1}^{c_1} \cdots \binom{m+(t-1)}{m-1}^{c_t}}{\binom{t+mn-1}{mn-1} c_0!\, c_1! \cdots c_t!} & \sum_{i=0}^{t} c_i = n \text{ and} \sum_{i=0}^{t} i c_i = t \\ 0 & \text{else.} \end{cases}$$

(c) If $m_1 = \cdots = m_n = 1$, then $M = n$, $\Psi_C(m, \dots, m) = n!\, 1^{c_0} 1^{c_1} \cdots 1^{c_t}$ and

$$P(C_1 = c_1, \dots, C_t = c_t) = \begin{cases} \dfrac{n!}{\binom{t+n-1}{n-1} c_0!\, c_1! \cdots c_t!} & \sum_{i=0}^{t} c_i = n \text{ and} \sum_{i=0}^{t} i c_i = t \\ 0 & \text{else.} \end{cases}$$

(See Wilks, Mathematical Statistics, page 445.)

<need to supply details on Wilk's book>

Problem 2.

In the notation of our Waiting Times model, let

$$n = 2, q_1 = \infty, q_2 = 1, m_1 = \alpha_1 = j, m_2 = M - \alpha_1 = M - j, r = 1$$

so that

$$W_{1:Q} = \text{waiting time for } 1^{st} \text{ observation to exceed } V_{(j)}.$$

Show

$$\mathrm{E}(W_{1:Q}) = \frac{M-1}{M-j-1}$$

in agreement with Wenocur, page 41 of her dissertation (she takes $N = M - 1$).

<need to supply reference for Wenocur's dissertation>

Problem 3.

Suppose an urn contains 1 ball each of colors $1, \dots, n$. At each drawing a single ball is removed from the urn at random and returned with another ball of the same color. Let T_r equal the number of draws required to observe r of the n different colored balls in the urn at least once. Let Z equal the number of draws required to observe 1 of the n different colored balls in the urn twice.

Notation:

$$a^{(k)} = a(a+1)\cdots(a+k-1)$$

$$a_{(k)} = a(a-1)\cdots(a-k+1)$$

(a) Show that $\mathrm{E}\left(T_r^{(k)}\right)$, the k^{th} ascending factorial moment of T_r, equals

$$\frac{(n-1)_{(k)}(k+r-1)_{(k)}}{(n-r)_{(k)}}$$

for $n \geq k+r$. In particular, we observe that

$$\mathrm{E}(T_r) = \frac{r(n-1)}{n-r}.$$

(b) Show that $\mathrm{E}\left((T_r+n-1)_{(k)}\right)$, the k^{th} descending factorial moment of T_r+n-1, equals

$$\frac{(n-1)_{(k)}n_{(k)}}{(n-r)_{(k)}}$$

for $n \geq k+r$ in agreement with Holst and Hüsler, "Sequential Urn Schemes and Birth Processes", Advances in Applied Probability, Vol. 17, 1985, pages 257-279.

(c) Show that

$$\mathrm{E}\left(Z^{(k)}\right) = k! \sum_{j=0}^{n} \binom{k+j-1}{j} \frac{n_{(j)}}{n^{(j)}}$$

provided we define $n_{(0)} = n^{(0)} = 1$.

The special case $k=1$ simplifies to

$$\mathrm{E}(Z) = \sum_{j=0}^{n} \frac{n_{(j)}}{n^{(j)}} = 1 + \frac{4^{n-1}}{\binom{2n-1}{n}}$$

in agreement with Blom and Holst, "Embedding Procedures for Discrete Problems in Probability", Mathematical Scientist, Vol. 16, 1981, pages 29-40.

The special case $k = 2$ simplifies to

$$\mathrm{E}\left(Z^{(2)}\right) = \sum_{j=0}^{n} (j+1) \frac{n_{(j)}}{n^{(j)}} = 2 + n + 3 \cdot \frac{4^{n-1}}{\binom{2n-1}{n}}.$$

Problem 4.

$$P\big(\text{no run of } r+1 \text{ or more sparse cells } (0 \text{ or } 1 \text{ ball in cell})\big)$$

$$= \frac{1}{\binom{M+t-1}{t}} \sum_{i=0}^{n} \sum_{j=0}^{n-i+1} \sum_{\phi=0}^{i} \sum_{s=0}^{n-i} \sum_{\varphi=0}^{s} (-1)^{i-\phi+u+\varphi+j} \left(\binom{n-(r+1)j}{i-(r+1)j} \binom{n-i+1}{j} \binom{i}{\phi} \right.$$

$$\left. \times \binom{n-i}{s} \binom{s}{\varphi} \binom{m(i+s-n) + (i-\phi+s-\varphi)}{t} (m+1)^{\phi+\varphi} m^{i-\phi+s-\varphi} \right)$$

provided $t \leq m(i+s-n) + (i-\phi+s-\varphi)$.

Problem 5.

Suppose t identical balls are distributed among $n+1$ distinguishable urns. The first n urns each contain m (distinguishable) cells or compartments and the last urn contains s (distinguishable) cells or compartments. There is no limit on the number of the number of balls that can go into any of the $mn+s$ cells. Assume that all $\binom{(mn+s)+t-1}{t}$ possible allocations of the t balls into the $mn+s$ distinguishable cells are equally likely to have occurred.

Let the random variable R equal the number of empty urns among the first n urns.

(a) Show that

$$P(R = r) = \frac{\binom{n}{r}}{\binom{mn+s+t-1}{t}} \sum_{i=0}^{n-r} (-1)^{n-r-i} \binom{n-r}{i} \binom{mi+s+t-1}{t}$$

$$= \frac{n_{(n-r)}}{(mn+s)^{(t)}} \left| G(t, n-r; -m, -s) \right|$$

in agreement with Charalambides, Ch. A. and Koutras, M., "On the Differences of the Generalized Factorials at an Arbitrary Point and Their Combinatorial Applications", Discrete Mathematics, Vol. 47, 1983, pages 183 - 201, where $G(t, n; m, s)$ are the *Gould-Hopper* defined by

$$G(t, n; m, s) = \frac{1}{n!} \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} (jm + s)_{(t)}$$

in Gould, H.W. and Hopper, A.T., "Operational Formulas Connected with Two Generalizations of Hermite Polynomials", Duke Mathematics Journal, Vol 29, 1962, pages 51 - 63.

(b) Show that for $0 \le u \le n$

$$\mathrm{E}\left(R_{(u)}\right) = \frac{n_{(u)} (m(n-u)+s)^{(t)}}{(mn+s)^{(t)}}.$$

(c) Suppose balls are distributed until k of the first n urns are occupied (by at least one ball). Let T equal the number of balls required. Show that

$$P(T = t) = \frac{m\, n_{(k)}}{(mn+s)^{(t)}} \left| G(t-1, k-1; -m, -s) \right|$$

in agreement with Charalambides, Ch. A. and Koutras, M., "On the Differences of the Generalized Factorials at an Arbitrary Point and Their Combinatorial Applications", Discrete Mathematics, Vol. 47, 1983, pages 183 - 201.

Problem 6.

The r^{th} descending factorial moment of the number of colors drawn k times after t balls have been drawn from an urn with m balls each of n different colors according to the Polya urn model equals

$$\frac{\binom{k+m-1}{m-1}^r \binom{m(n-r)+(t-kr)-1}{t-kr}}{\binom{mn+t-1}{t}} n_{(r)}$$

provided $n \geq r$ and $t \geq kr$ and equals 0 otherwise.

Proof (Theorem 1)

$$P(Y_1 + \cdots + Y_n = t) = P\big((Y_1, \dots, Y_n) \in \mathbb{S}_t^n\big)$$

$$= \sum_{\mathbb{S}_t^n} \prod_{i=1}^{n} \binom{y_i + m_i - 1}{m_i - 1} p^{m_i} (1-p)^{y_i}$$

$$= \sum_{\mathbb{S}_t^n} \binom{y_1 + m_1 - 1}{m_1 - 1} \cdots \binom{y_n + m_n - 1}{m_n - 1} p^{m_1 + \cdots + m_n} (1-p)^{y_1 + \cdots + y_n}$$

$$= p^{(m_1 + \cdots + m_n)} (1-p)^t \left(\sum_{\mathbb{S}_t^n} \binom{y_1 + m_1 - 1}{m_1 - 1} \cdots \binom{y_n + m_n - 1}{m_n - 1} \right)$$

$$= p^{(m_1 + \cdots + m_n)} (1-p)^t \binom{t + (m_1 + \cdots + m_n) - 1}{(m_1 + \cdots + m_n) - 1}$$

$$= p^M (1-p)^t \binom{t + M - 1}{M - 1}.$$

That is, $Y_1 + \cdots + Y_n \sim$ Negative Binomial(M, p). Thus,

$$P\big((Y_1, \dots, Y_n) \in \mathcal{A} \big| Y_1 + \cdots + Y_n = t\big)$$

$$= \frac{P\big((Y_1, \dots, Y_n) \in \mathcal{A} \text{ and } Y_1 + \cdots + Y_n = t\big)}{P(Y_1 + \cdots + Y_n = t)}$$

$$= \frac{P\big((Y_1, \dots, Y_n) \in \mathcal{A}_t\big)}{P(Y_1 + \cdots + Y_n = t)}$$

$$= \frac{\sum_{\mathcal{A}_t} \binom{y_1 + m_1 - 1}{m_1 - 1} \cdots \binom{y_n + m_n - 1}{m_n - 1} p^M (1-p)^{y_1 + \cdots + y_n}}{p^M (1-p)^t \binom{t + M - 1}{M - 1}}$$

$$= \frac{\sum_{\mathcal{A}_t} \binom{y_1 + m_1 - 1}{m_1 - 1} \cdots \binom{y_n + m_n - 1}{m_n - 1} p^M (1-p)^t}{p^M (1-p)^t \binom{t + M - 1}{M - 1}}$$

$$= \sum_{\mathcal{A}_t} \frac{\binom{y_1 + m_1 - 1}{m_1 - 1} \cdots \binom{y_n + m_n - 1}{m_n - 1}}{\binom{t + M - 1}{M - 1}}$$

$$= P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right).$$

Therefore,

$$P\big((Y_1, \dots, Y_n) \in \mathcal{A}\big)$$

$$= \sum_{t=0}^{\infty} P\left((Y_1, \dots, Y_n) \in \mathcal{A} \middle| \sum_{i=1}^{n} Y_i = t\right) P\left(\sum_{i=1}^{n} Y_i = t\right)$$

$$= \sum_{t=0}^{\infty} P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right) P\left(\sum_{i=1}^{n} Y_i = t\right)$$

$$= \sum_{t=0}^{\infty} P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right) p^M (1-p)^t \binom{t + M - 1}{M - 1}$$

and

$$\left(\frac{1}{p}\right)^M P\big((Y_1, \dots, Y_n) \in \mathcal{A}\big) = \sum_{t=0}^{\infty} P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right) \binom{t + M - 1}{M - 1} (1-p)^t.$$

It follows that

$$\frac{d^r}{dp^r}\left(\left(\frac{1}{p}\right)^M P\big((Y_1, \dots, Y_n) \in \mathcal{A}\big)\right)\Bigg|_{p=1}$$

$$= \frac{d^r}{dp^r}\left(\sum_{t=0}^{\infty} P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right) \binom{t + M - 1}{M - 1} (1-p)^t\right)\Bigg|_{p=1}$$

$$= \sum_{t=0}^{\infty} P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right) \binom{t + M - 1}{M - 1} \left(\frac{d^r}{dp^r}(1-p)^t\Big|_{p=1}\right)$$

$$= \sum_{t=0}^{\infty} P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right) \binom{t + M - 1}{M - 1} \left(r!\,(-1)^r \mathrm{I}_{\{r\}}(t)\right)$$

$$= P\left(\left(X_{1,r}, \dots, X_{n,r}\right) \in \mathcal{A}_r\right)\binom{r+M-1}{M-1} r!\,(-1)^r \mathrm{I}_{\{0,1,\dots\}}(r).$$

Thus for $r \geq 0$,

$$P\left(\left(X_{1,r}, \dots, X_{n,r}\right) \in \mathcal{A}_r\right) = \frac{1}{\binom{r+M-1}{M-1} r!\,(-1)^r} \frac{d^r}{dp^r}\left(\left(\frac{1}{p}\right)^M P((Y_1, \dots, Y_n) \in \mathcal{A})\right)\Bigg|_{p=1}.$$

Proof (Theorem 2)

Define $N_{(q_1,\dots,q_n)}(t) \equiv N_Q(t)$ to be the number of intervals that have not reached their quota after t additional random variables from F have been located to their appropriate interval.

Define N_Q^B as the number of events amongst $\mathcal{A}_1,\dots,\mathcal{A}_n$ that occur.

It follows that

$$W_{r:Q} > t \Leftrightarrow N_Q(t) > n - r$$

However, it follows from

$$\mathrm{E}\left(g(Y_1, \dots, Y_n)\right) = \sum_{t=0}^{\infty} \mathrm{E}\left(g(Y_1, \dots, Y_n)\Bigg|\sum_{i=1}^{n} Y_i = t\right) P\left(\sum_{i=1}^{n} Y_i = t\right)$$

that

$$P\left(N_Q^{NB} > n - r\right) = \sum_{t=0}^{\infty} P\left(N_Q(t) > n - r\right)\binom{M+t-1}{t} p^M (1-p)^t.$$

Thus for $M \geq k + 1$,

$$\int_0^1 p^{-k-1}(1-p)^{k-1}\left(P\left(N_Q^{NB} > n - r\right)\right) dp$$

$$= \int_0^1 p^{-k-1}(1-p)^{k-1}\left(\sum_{t=0}^{\infty} P\left(N_Q(t) > n - r\right)\binom{M+t-1}{t} p^M (1-p)^t\right) dp$$

$$= \int_0^1 p^{-k-1}(1-p)^{k-1}\left(\sum_{t=0}^{\infty} P\left(W_{r:Q} > t\right)\binom{M+t-1}{t} p^M (1-p)^t\right) dp$$

$$= \sum_{t=0}^{\infty} P\big(W_{r:Q} > t\big) \binom{M+t-1}{t} \left(\int_0^1 p^{M-k-1} (1-p)^{t+k-1} dp \right)$$

$$= \sum_{t=0}^{\infty} P\big(W_{r:Q} > t\big) \binom{M+t-1}{t} \left(\frac{(M-k-1)!\,(t+k-1)!}{(M+t-1)!} \right)$$

$$= \frac{(M-k-1)!}{(M-1)!} \sum_{t=0}^{\infty} P\big(W_{r:Q} > t\big) \frac{(t+k-1)!}{t!}$$

$$= \frac{(M-k-1)!}{(M-1)!} \sum_{t=0}^{\infty} P\big(W_{r:Q} > t\big) (t+k-1)_{[k-1]}$$

$$= \frac{(M-k-1)!}{(M-1)!} \sum_{t=k-1}^{\infty} P\big(W_{r:Q} + k - 1 > t\big) t_{[k-1]}$$

$$= \frac{(M-k-1)!}{(M-1)!} \frac{1}{k} \mathrm{E}\left(\big(W_{r:Q} + k - 1\big)_{[k]} \right).$$

But

$$\big(W_{r:Q} + k - 1\big)_{[k]} \equiv W_{r:Q}^{[k]}$$

hence

$$\int_0^1 p^{-k-1} (1-p)^{k-1} \left(P\big(N_Q^{NB} > n - r\big) \right) dp = \frac{(M-k-1)!}{(M-1)!} \frac{1}{k} \mathrm{E}\left(W_{r:Q}^{[k]} \right)$$

and

$$\mathrm{E}\left(W_{r:Q}^{[k]} \right) = \frac{k(M-1)!}{(M-k-1)!} \int_0^1 p^{-k-1} (1-p)^{k-1} \left(P\big(N_Q^{NB} > n - r\big) \right) dp$$

where

$$P\big(N_Q^{NB} > n - r\big)$$

$$= P\big(\text{at least } n - r + 1 \text{ types do } \underline{\text{not}} \text{ obtain their quota} \,\big|\, \text{Negative Binomial model}\big).$$

In the case $M < k+1$, we have

$$P\left(N_Q^{NB} > n-r\right) = \sum_{t=0}^{\infty} P\left(W_{r:Q} > t\right) \binom{M+t-1}{t} p^M (1-p)^t.$$

Therefore,

$$\frac{(1-p)^{k-1}(M-1)!}{p^M} P\left(N_Q^{NB} > n-r\right) = \sum_{t=0}^{\infty} P\left(W_{r:Q} > t\right) \frac{(M+t-1)!}{t!} (1-p)^{t+k-1}$$

and

$$\frac{d^{(k-M)}}{dp^{(k-M)}} \left(\frac{(1-p)^{k-1}(M-1)!}{p^M} P\left(N_Q^{NB} > n-r\right)\Big|_{p=0} \right)$$

$$= \frac{d^{(k-M)}}{dp^{(k-M)}} \left(\sum\nolimits_{t=0}^{\infty} P\left(W_{r:Q} > t\right) \frac{(M+t-1)!}{t!} (1-p)^{t+k-1} \Big|_{p=0} \right)$$

$$= \sum_{t=0}^{\infty} P\left(W_{r:Q} > t\right) \frac{(t+k-1)!}{t!} (-1)^{k-M} \left((1-p)^{t+M-1}|_{p=0} \right)$$

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

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

$$= (-1)^{k-M} \sum_{t=k-1}^{\infty} P\left(W_{r:Q} + k - 1 > t\right) t_{[k-1]}$$

$$= (-1)^{k-M} \frac{1}{k} \mathrm{E}\left(\left(W_{r:Q} + k - 1 \right)_{[k]} \right)$$

$$= (-1)^{k-M} \frac{1}{k} \mathrm{E}\left(W_{r:Q}^{[k]} \right).$$

Hence,

$$\mathrm{E}\left(W_{r:Q}^{[k]}\right) = \frac{d^{(k-M)}}{dp^{(k-M)}}\left((-1)^{k-M}\frac{k(1-p)^{k-1}(M-1)!}{p^M}P\left(N_Q^{NB} > n-r\right)\right)\Bigg|_{p=0}.$$

Solutions

Problem 1(a)

Let $\mathcal{A} \subset \mathbb{S}^n$ and define $\mathcal{A}_t = \mathcal{A} \cap \mathbb{S}_t^n$. Then for $t \geq 0$,

$$P\left(\left(X_{1,t}, \dots, X_{n,t}\right) \in \mathcal{A}_t\right) = \frac{(-1)^t}{\binom{t+M-1}{M-1}t!}\frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^M P\left((Y_1, \dots, Y_n) \in \mathcal{A}\right)\right)\Bigg|_{p=1}.$$

So,

$$P(C_1 = c_1, \dots, C_t = c_t)$$

$$= \frac{(-1)^t}{\binom{t+M-1}{M-1}t!\,(-1)^t}\frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^M p^M(1-p)^{\sum_{i=0}^t ic_i}\frac{\Psi_C(m_1, \dots, m_n)}{c_0!\,c_1!\cdots c_t!}\right)\Bigg|_{p=1}$$

$$= \frac{\Psi_C(m_1, \dots, m_n)(-1)^t}{\binom{t+M-1}{M-1}t!\,c_0!\,c_1!\cdots c_t!}\frac{d^t}{dp^t}\left((1-p)^{\sum_{i=0}^t ic_i}\right)\Bigg|_{p=1}$$

$$= \frac{\Psi_C(m_1, \dots, m_n)(-1)^t}{\binom{t+M-1}{M-1}t!\,c_0!\,c_1!\cdots c_t!}(-1)^t t!\, I\left\{\sum_{i=0}^t ic_i = t\right\}$$

$$= \frac{\Psi_C(m_1, \dots, m_n)}{\binom{t+M-1}{M-1}c_0!\,c_1!\cdots c_t!} I\left\{\sum_{i=0}^t ic_i = t\right\}$$

where

$$\Psi_C(m_1, \dots, m_n) = \sum_{\pi}\binom{m_{\pi_{(1)}}-1}{m_{\pi_{(1)}}-1}\cdots\binom{m_{\pi_{(c_0)}}-1}{m_{\pi_{(c_0)}}-1}\times\binom{m_{\pi_{(c_0+1)}}+0}{m_{\pi_{(c_0+1)}}-1}\cdots\binom{m_{\pi_{(c_1+c_0)}}+0}{m_{\pi_{(c_1+c_0)}}-1}$$

$$\times\cdots\times\binom{m_{\pi_{(c_{t-1}+\cdots+c_0+1)}}+(t-1)}{m_{\pi_{(c_{t-1}+\cdots+c_0+1)}}-1}\cdots\binom{m_{\pi_{(c_t+\cdots+c_0)}}+(t-1)}{m_{\pi_{(c_t+\cdots+c_0)}}-1}$$

and the sum is over all permutations π of $\{1,\dots,n\}$.

Note:

$$P\big((Y_1,\dots,Y_n)\in\mathcal{A}\big)=P(\text{exactly } c_k \text{ of the } n\ Y_i\text{'s}=k, k=0,\dots,t)$$

$$=p^M(1-p)^{\sum_{i=0}^t ic_i}\frac{1}{c_0!\,c_1!\cdots c_t!}\sum_{\pi}\binom{m_{\pi(1)}-1}{m_{\pi(1)}-1}\cdots\binom{m_{\pi(c_0)}-1}{m_{\pi(c_0)}-1}$$
$$\times\binom{m_{\pi(c_0+1)}+0}{m_{\pi(c_0+1)}-1}\cdots\binom{m_{\pi(c_1+c_0)}+0}{m_{\pi(c_1+c_0)}-1}$$
$$\times\cdots\times\binom{m_{\pi(c_{t-1}+\cdots+c_0+1)}+(t-1)}{m_{\pi(c_{t-1}+\cdots+c_0+1)}-1}\cdots\binom{m_{\pi(c_t+\cdots+c_0)}+(t-1)}{m_{\pi(c_t+\cdots+c_0)}-1}$$

$$=p^M(1-p)^{\sum_{i=0}^t ic_i}\frac{\Psi_C(m_1,\dots,m_n)}{c_0!\,c_1!\cdots c_t!}$$

where $M=m_1+\cdots+m_n$ and the sum is over all $n!$ permutations π of $\{1,2,\dots,n\}$.

Problem 2

$$\mathrm{E}\big(W_{r:Q}\big)=(M-1)\int_0^1 p^{-2}P\left((Y_1,\dots,Y_n)\in\mathbb{A}_{Q:r}\right)dp$$

$$=(M-1)\int_0^1 p^{-2}P(Y_1<\infty,Y_2<1)dp$$

$$=(M-1)\int_0^1 p^{-2}P(Y_2=0)dp$$

$$=(M-1)\int_0^1 p^{-2}\binom{0+m_2-1}{m_2-1}p^{m_2}(1-p)^0dp$$

$$=(M-1)\int_0^1 p^{M-j-2}dp=\frac{M-1}{M-j-1}$$

which agrees with Wenocur, page 41 of her dissertation (she takes $N=M-1$).

Problem 3

(a) By Theorem 2, with $q_1 = \cdots = q_n = 1$, $m_1 = \cdots = m_n = 1$, and for $n \geq k+1$, we have

$$\mathrm{E}\left(T_r^{(k)}\right) = \frac{k(n-1)!}{(n-k-1)!}\int_0^1 p^{-k-1}(1-p)^{k-1}P\big(\text{at least } n-r+1 \text{ of the } n\ Y_j\text{'s} = 0\big)dp$$

$$= \frac{k(n-1)!}{(n-k-1)!}\int_0^1 p^{-k-1}(1-p)^{k-1}\left(\sum_{i=n-r+1}^{n}(-1)^{i-(n-r+1)}\binom{i-1}{(n-r+1)-1}\times\binom{n}{i}\big(P(Y_1=0)\big)^i dp\right)$$

$$= \frac{k(n-1)!}{(n-k-1)!}\sum_{i=n-r+1}^{n}(-1)^{i-(n-r+1)}\binom{i-1}{(n-r+1)-1}\binom{n}{i}\left(\int_0^1 p^{i-k-1}(1-p)^{k-1}dp\right)$$

$$= \frac{k(n-1)!}{(n-k-1)!}\sum_{i=n-r+1}^{n}(-1)^{i-(n-r+1)}\binom{i-1}{(n-r+1)-1}\binom{n}{i}\frac{(i-k-1)!\,(k-1)!}{(i-1)!}$$

(provided $n \geq k+r$)

$$= \frac{k!\,(n-1)!\,(n-r-k)!}{(n-r)!\,(n-k-1)!}\sum_{i=n-r+1}^{n}(-1)^{i-n+r-1}\binom{i-k-1}{n-r-k}\binom{n}{i}$$

$$= \frac{k!\,(n-1)!\,(n-r-k)!}{(n-r)!\,(n-k-1)!}\binom{k+r-1}{k}$$

$$= \frac{(n-1)_{(k)}(k+r-1)_{(k)}}{(n-r)_{(k)}}.$$

(b) We can use the result in (a) along with the identity

$$(x+c)_{(k)} = \sum_{j=0}^{k}\frac{k!}{j!}\binom{c-j}{k-j}x^{(j)}$$

to show that

$$\mathsf{E}\big((T_r+n-1)_{(k)}\big)=\sum_{j=0}^{k}\frac{k!}{j!}\binom{(n-1)-j}{k-j}\mathsf{E}\left(T_r^{(j)}\right)$$

$$=\sum_{j=0}^{k}\frac{k!}{j!}\binom{(n-1)-j}{k-j}\left(\frac{(n-1)_{(j)}(j+r-1)_{(j)}}{(n-r)_{(j)}}\right)$$

$$=\frac{(n-1)_{(k)}k!}{(n-r)_{(k)}}\sum_{j=0}^{k}\binom{r-1+j}{j}\binom{n-r-j}{n-r-k}$$

$$=\frac{(n-1)_{(k)}k!}{(n-r)_{(k)}}\binom{n}{k}$$

$$=\frac{(n-1)_{(k)}n_{(k)}}{(n-r)_{(k)}}.$$

(c) By Theorem 2, with $q_1=\cdots=q_n=2$, $m_1=\cdots=m_n=1$,

$$\mathsf{E}\big(Z^{[k]}\big)=\mathsf{E}\left(W_{1:Q}^{[k]}\right)$$

$$=\frac{k(n-1)!}{(n-k-1)!}\int_0^1 p^{-k-1}(1-p)^{k-1}P\left((Y_1,\dots,Y_n)\in\mathbb{A}_{Q:1}\right)dp$$

$$=\frac{k(n-1)!}{(n-k-1)!}\int_0^1 p^{-k-1}(1-p)^{k-1}P(Y_1<2,\dots,Y_n<2)dp$$

$$=\frac{k(n-1)!}{(n-k-1)!}\int_0^1 p^{-k-1}(1-p)^{k-1}\big(p+p(1-p)\big)^n dp$$

$$=\frac{k(n-1)!}{(n-k-1)!}\int_0^1 p^{-k-1}(1-p)^{k-1}\sum_{j=0}^{n}\binom{n}{j}p^{n-j}\big(p(1-p)\big)^j dp$$

$$=\frac{k(n-1)!}{(n-k-1)!}\sum_{j=0}^{n}\binom{n}{j}\left(\int_0^1 p^{n-k-1}(1-p)^{k+j-1}dp\right)$$

$$= \frac{k(n-1)!}{(n-k-1)!}\sum_{j=0}^{n}\binom{n}{j}\frac{(n-k-1)!\,(k+j-1)!}{(n-1+j)!}$$

provided $n \geq k+1$ and $k \geq 1$

$$= \frac{k!}{\binom{2n-1}{n}}\sum_{j=0}^{n}\binom{k+j-1}{j}\binom{2n-1}{n-j}$$

$$= k!\sum_{j=0}^{n}\binom{k+j-1}{j}\frac{n_{(j)}}{n^{(j)}}.$$

In the special case $k = 1$ this simplifies to

$$\frac{1}{\binom{2n-1}{n}}\sum_{j=0}^{n}\binom{2n-1}{n-j}$$

$$= \frac{1}{\binom{2n-1}{n}}\left(2^{2n-2} + \binom{2n-1}{n}\right)$$

using Identity #1.84 of Gould

$$= 1 + \frac{4^{n-1}}{\binom{2n-1}{n}}.$$

In the special case $k = 2$ this simplifies to

$$\frac{2}{\binom{2n-1}{n}}\sum_{j=0}^{n}\binom{j+1}{j}\binom{2n-1}{n-j}$$

$$= \frac{2}{\binom{2n-1}{n}}\left(\sum_{j=0}^{n}(n+j-1)\binom{2n-1}{n-j} - (n-2)\sum_{j=0}^{n}\binom{2n-1}{n-j}\right)$$

$$= \frac{2}{\binom{2n-1}{n}}\left((2n-1)\sum_{j=0}^{n}\binom{2n-2}{n-j} - (n-2)\sum_{j=0}^{n}\binom{2n-1}{n-j}\right)$$

$$= \frac{2}{\binom{2n-1}{n}} \left((2n-1)\left(\binom{2n-2}{n} + \sum_{j=1}^{n} \binom{2n-2}{n-j} \right) - (n-2) \sum_{j=0}^{n} \binom{2n-1}{n-j} \right)$$

$$= \frac{2}{\binom{2n-1}{n}} \left((2n-1)\left(\binom{2n-2}{n} + 2^{2n-3} + \binom{2n-3}{n-1} \right) - (n-2)\left(2^{2n-2} + \binom{2n-1}{n} \right) \right)$$

applying Identities 1.85 and 1.84 of Gould

$$= 2 + n + 3 \cdot \frac{4^{n-1}}{\binom{2n-1}{n}}.$$

Problem 4 (No run of $r+1$ or more sparse cells (0 or 1 ball in a cell))

$P($longest success run of length less than or equal to r in a series of n *iid* Bernoulli trials$)$

$$= \sum_{i=0}^{n} \sum_{j=0}^{n-i+1} (-1)^j \binom{n-(r+1)j}{i-(r+1)j} \binom{n-i+1}{j} p^i (1-p)^{n-i}$$

success ≡ event $Y_j = \{0 \text{ or } 1\}$ where $Y_1, \ldots, Y_n$ are assumed to be independent negative binomial random variables such that $Y_i \sim$Negative Binomial(m, θ).

i.e.

$$P(Y_i = y) = \binom{y+m-1}{m-1} \theta^m (1-\theta)^y \qquad y \in \{0,1,\ldots\} \text{ and } 0 \le \theta \le 1$$

so

$$p = P(Y_i \le 1) = \binom{0+m-1}{m-1} \theta^m (1-\theta)^0 + \binom{1+m-1}{m-1} \theta^m (1-\theta)^1$$

$$= (m+1)\theta^m - m\theta^{m+1} = g(\theta)$$

$$M = mn$$

$$P\left(\left(X_{1,t},\dots,X_{n,t}\right)\in\mathcal{A}_t\right)=\frac{(-1)^t}{\binom{M+t-1}{t}t!}\frac{d^t}{d\theta^t}\left(\left(\frac{1}{\theta}\right)^M P\left((Y_1,\dots,Y_n)\in\mathcal{A}\right)\right)\Bigg|_{\theta=1}$$

$$=\frac{(-1)^t}{\binom{M+t-1}{t}t!}\frac{d^t}{d\theta^t}\left(\left(\frac{1}{\theta}\right)^M\sum_{i=0}^{n}\sum_{j=0}^{n-i+1}(-1)^j\binom{n-(r+1)j}{i-(r+1)j}\binom{n-i+1}{j}\left(g(\theta)\right)^i\left(1-g(\theta)\right)^{n-i}\right)\Bigg|_{\theta=1}$$

$$=\frac{(-1)^t}{\binom{M+t-1}{t}t!}\sum_{i=0}^{n}\sum_{j=0}^{n-i+1}(-1)^j\binom{n-(r+1)j}{i-(r+1)j}\binom{n-i+1}{j}\left(\frac{d^t}{d\theta^t}\left(\left(\frac{1}{\theta}\right)^M\left(g(\theta)\right)^i\left(1-g(\theta)\right)^{n-i}\right)\Bigg|_{\theta=1}\right)$$

$$M = mn$$

$$g(\theta)=(m+1)\theta^m-m\theta^{m+1}$$

$$\frac{d^t}{d\theta^t}\left(\left(\frac{1}{\theta}\right)^M\left(g(\theta)\right)^i\left(1-g(\theta)\right)^{n-i}\right)\Bigg|_{\theta=1}$$

$$=\frac{d^t}{d\theta^t}\left(\left(\frac{1}{\theta}\right)^{mn}\left((m+1)\theta^m-m\theta^{m+1}\right)^i(1-(m+1)\theta^m+m\theta^{m+1})^{n-i}\right)\Bigg|_{\theta=1}$$

$$=\frac{d^t}{d\theta^t}\left(\left(\frac{1}{\theta}\right)^{mn}\sum_{\phi=0}^{i}\sum_{s=0}^{n-i}\sum_{\varphi=0}^{s}(-1)^{i-\phi+\varphi}\binom{i}{\phi}\binom{n-i}{s}\binom{s}{\varphi}\left((m+1)\theta^m\right)^{\phi+\varphi}(m\theta^{m+1})^{i-\phi+s-\varphi}\right)\Bigg|_{\theta=1}$$

$$=\frac{d^t}{d\theta^t}\left(\sum_{\phi=0}^{i}\sum_{s=0}^{n-i}\sum_{\varphi=0}^{s}(-1)^{i-\phi+\varphi}\binom{i}{\phi}\binom{n-i}{s}\binom{s}{\varphi}(m+1)^{\phi+\varphi}m^{i-\phi+s-\varphi}\theta^{m(i+s-n)+(i-\phi+s-\varphi)}\right)\Bigg|_{\theta=1}$$

$$= \frac{d^t}{d\theta^t}\left(\sum_{\phi=0}^{i}\sum_{s=0}^{n-i}\sum_{\varphi=0}^{s}(-1)^{i-\phi+\varphi}\binom{i}{\phi}\binom{n-i}{s}\binom{s}{\varphi}(m+1)^{\phi+\varphi}m^{i-\phi+s-\varphi}\theta^{m(i+s-n)+(i-\phi+s-\varphi)}\right)\Bigg|_{\theta=1}$$

$$= \frac{d^t}{d\theta^t}\sum_{\phi=0}^{i}\sum_{s=0}^{n-i}\sum_{\varphi=0}^{s}\sum_{u=0}^{m(i+s-n)+(i-\phi+s-\varphi)}$$

$$\times\ (-1)^{i-\phi+\varphi+u}\binom{i}{\phi}\binom{n-i}{s}\binom{s}{\varphi}\binom{m(i+s-n)+(i-\phi+s-\varphi)}{u}$$

$$\times (m+1)^{\phi+\varphi}m^{i-\phi+s-\varphi}(1-\theta)^u)\big|_{\theta=1}$$

$$= \sum_{\phi=0}^{i}\sum_{s=0}^{n-i}\sum_{\varphi=0}^{s}\sum_{u=0}^{m(i+s-n)+(i-\phi+s-\varphi)}(-1)^{i-\phi+\varphi+u}\binom{i}{\phi}\binom{n-i}{s}\binom{s}{\varphi}$$
$$\times\binom{m(i+s-n)+(i-\phi+s-\varphi)}{u}(m+1)^{\phi+\varphi}m^{i-\phi+s-\varphi}\left(\frac{d^t}{d\theta^t}((1-\theta)^u)\Bigg|_{\theta=1}\right)$$

$$= \sum_{\phi=0}^{i}\sum_{s=0}^{n-i}\sum_{\varphi=0}^{s}\sum_{u=0}^{m(i+s-n)+(i-\phi+s-\varphi)}(-1)^{i-\phi+\varphi+u}\binom{i}{\phi}\binom{n-i}{s}\binom{s}{\varphi}$$
$$\times\ \binom{m(i+s-n)+(i-\phi+s-\varphi)}{u}(m+1)^{\phi+\varphi}m^{i-\phi+s-\varphi}\left(t!\,(-1)^t\mathrm{I}_{\{u\}}(t)\right)$$

$$= \sum_{\phi=0}^{i}\sum_{s=0}^{n-i}\sum_{\varphi=0}^{s}(-1)^{i-\phi+u}\binom{i}{\phi}\binom{n-i}{s}\binom{s}{\varphi}\binom{m(i+s-n)+(i-\phi+s-\varphi)}{t}t!$$
$$\times (m+1)^{\phi+\varphi}m^{i-\phi+s-\varphi}I\big(t \le m(i+s-n)+(i-\phi+s-\varphi)\big)$$

Therefore,

$$P\left((X_{1,t},\dots,X_{n,t}) \in \mathcal{A}_t\right)$$

$$= \frac{(-1)^t}{\binom{M+t-1}{t} t!} \sum_{i=0}^{n} \sum_{j=0}^{n-i+1} (-1)^j \binom{n-(r+1)j}{i-(r+1)j} \binom{n-i+1}{j} \times \left(\frac{d^t}{d\theta^t} \left(\left(\frac{1}{\theta}\right)^M \left(g(\theta)\right)^i \left(1-g(\theta)\right)^{n-i} \right)\Big|_{\theta=1} \right)$$

$$= \frac{1}{\binom{M+t-1}{t}} \sum_{i=0}^{n} \sum_{j=0}^{n-i+1} (-1)^j \binom{n-(r+1)j}{i-(r+1)j} \binom{n-i+1}{j}$$

$$\times \left(\sum_{\phi=0}^{i} \sum_{s=0}^{n-i} \sum_{\varphi=0}^{s} (-1)^{i-\phi+u+\varphi} \binom{i}{\phi} \binom{n-i}{s} \binom{s}{\varphi} \binom{m(i+s-n)+(i-\phi+s-\varphi)}{t} (m+1)^{\phi+\varphi} m^{i-\phi+s-\varphi} I\big(t \le m(i+s-n)+(i-\phi+s-\varphi)\big) \right)$$

$$= \frac{1}{\binom{M+t-1}{t}} \sum_{i=0}^{n} \sum_{j=0}^{n-i+1} \sum_{\phi=0}^{i} \sum_{s=0}^{n-i} \sum_{\varphi=0}^{s} (-1)^{i-\phi+u+\varphi+j} \binom{n-(r+1)j}{i-(r+1)j} \times \binom{n-i+1}{j} \binom{i}{\phi} \binom{n-i}{s} \binom{s}{\varphi}$$

$$\times \binom{m(i+s-n)+(i-\phi+s-\varphi)}{t} (m+1)^{\phi+\varphi} m^{i-\phi+s-\varphi} I\big(t \le m(i+s-n)+(i-\phi+s-\varphi)\big)$$

Problem 5

(a)

If we define

$$\mathcal{A}_t = \left\{(a_1, a_2, \dots, a_n, a_{n+1}) \middle| a_1 + \cdots + a_{n+1} = t \text{ and exactly } r \text{ of } (a_1, a_2, \dots, a_n) \text{ equal } 0\right\}$$

$$\mathcal{A} = \left\{(a_1, a_2, \dots, a_n, a_{n+1}) \middle| \text{exactly } r \text{ of } (a_1, a_2, \dots, a_n) \text{ equal } 0\right\}$$

then by Theorem 1,

$P(\text{exactly } r \text{ of the first } n \text{ urns are empty})$

$$= P\left(\left(X_{1,t}, \dots, X_{n,t}, X_{n+1,t}\right) \in \mathcal{A}_t\right)$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s} P\left((Y_1, \dots, Y_n, Y_{n+1}) \in \mathcal{A}\right)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s}\left(\sum_{i=r}^{n} (-1)^{i-r} \binom{i}{r}\binom{n}{i} \left(P(Y_1 = 0)^i\right)\right)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s}\left(\sum_{i=r}^{n} (-1)^{i-r} \binom{i}{r}\binom{n}{i} p^{im}\right)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\sum_{i=r}^{n} (-1)^{i-r} \binom{i}{r}\binom{n}{i} \left(\frac{1}{p}\right)^{m(n-i)+s}\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\sum_{i=r}^{n} (-1)^{i-r} \binom{i}{r}\binom{n}{i} \left(\frac{1}{1-(1-p)}\right)^{m(n-i)+s}\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\sum_{i=r}^{n} \sum_{j=0}^{\infty} (-1)^{i-r} \binom{i}{r}\binom{n}{i} \binom{m(n-i)+s+j-1}{j} (1-p)^j\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \sum_{i=r}^{n} \sum_{j=0}^{\infty} (-1)^{i-r} \binom{i}{r}\binom{n}{i} \binom{m(n-i)+s+j-1}{j} \left(\frac{d^t}{dp^t}\left((1-p)^j\right)\Bigg|_{p=1}\right)$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \sum_{i=r}^{n} \sum_{j=0}^{\infty} (-1)^{i-r} \binom{i}{r}\binom{n}{i} \binom{m(n-i)+s+j-1}{j} t!\,(-1)^t \mathrm{I}_{\{t\}}(j)$$

$$= \frac{1}{\binom{mn+s+t-1}{t}} \sum_{i=r}^{n} (-1)^{i-r} \binom{i}{r}\binom{n}{i} \binom{m(n-i)+s+t-1}{t}$$

$$= \frac{n_{(n-r)}}{(mn+s)^{(t)}}(-1)^t\left(\frac{(-1)^t t!}{(n-r)!}\sum_{i=0}^{n-r}(-1)^{n-r-i}\binom{n-r}{i}\binom{mi+s+t-1}{t}\right)$$

$$= \frac{n_{(n-r)}}{(mn+s)^{(t)}}(-1)^t G(t, n-r, -m, -s)$$

$$= \frac{n_{(n-r)}}{(mn+s)^{(t)}}\left|G(t, n-r, -m, -s)\right|.$$

(b)

It is a standard result of probability distribution theory that for general random variables $X_1, \dots, X_n$

$$\mathsf{E}\left(\left(\mathrm{I}_{\{A\}}(X_1) + \cdots + \mathrm{I}_{\{A\}}(X_n)\right)_{(u)}\right) = u! \sum_{(j_1,\dots,j_u)\in\mathbb{C}_u} \mathrm{P}\left(X_{j_1} \in A, \dots, X_{j_u} \in A\right)$$

where $\mathbb{C}_u$ is the set of all samples of size u drawn without replacement from $\{1, \dots, n\}$, when the order of sampling is considered unimportant and where we define

$$\mathrm{I}_{\{A\}}(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases}$$

Now define $g(a_1, \dots, a_n, a_{n+1}) = \left(\mathrm{I}_{\{0\}}(a_1) + \cdots + \mathrm{I}_{\{0\}}(a_n)\right)_{(u)}$. Then it follows from Theorem 1 that for $t \geq 0$,

$$\mathsf{E}\left(R_{(u)}\right)$$

$$= \mathsf{E}_t\left(g(X_1, \dots, X_n, X_{n+1})\right)$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!}\frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s} \mathsf{E}\left(g(Y_1, \dots, Y_n, Y_{n+1})\right)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!}\frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s} \left(u!\binom{n}{u}\left(P(Y_1 = 0)\right)^u\right)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s} n_{(u)} p^{mu}\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t n_{(u)}}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{1-(1-p)}\right)^{m(n-u)+s}\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t n_{(u)}}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\sum_{j=0}^{\infty} \binom{m(n-u)+s+j-1}{j} (1-p)^j\right)\Bigg|_{p=1}$$

$$0 \le u \le n$$

$$= \frac{(-1)^t n_{(u)}}{\binom{mn+s+t-1}{t} t!} \sum_{j=0}^{\infty} \binom{m(n-u)+s+j-1}{j} \left(\frac{d^t}{dp^t}\left((1-p)^j\right)\Big|_{p=1}\right)$$

$$= \frac{(-1)^t n_{(u)}}{\binom{mn+s+t-1}{t} t!} \sum_{j=0}^{\infty} \binom{m(n-u)+s+j-1}{j} (-1)^t t!\, \mathrm{I}_{\{t\}}(j)$$

$$= \frac{n_{(u)}}{\binom{mn+s+t-1}{t}} \binom{m(n-u)+s+t-1}{t}$$

$$= \frac{n_{(u)} (m(n-u)+s)^{(t)}}{(mn+s)^{(t)}}.$$

(c) Suppose balls are distributed until k of the first n urns are occupied (by at least one ball). Let T equal the number of balls required. Show that

$$P(T=t) = \frac{m\, n_{(k)}}{(mn+s)^{(t)}} \left|G(t-1, k-1; -m, -s)\right|$$

we see that $T = t$ if and only if

(1) $k-1$ of the first n urns are occupied by at least one ball after $t-1$ balls have been distributed

and

(2) the t^{th} ball goes into an unoccupied urn among the first n urns

Therefore,

$$P(T = t) = P(B|A)P(A)$$

where

$A: k-1$ of the first n urns are occupied by at least one ball after $t-1$ balls have been distributed

and

$B: t^{th}$ ball goes into an unoccupied urn among the first n urns.

However, from part (a)

$$P(A) = \frac{n_{(k-1)}}{(mn+s)^{(t-1)}} |G(t-1, k-1; -m, -s)|.$$

Furthermore, it follows from our understanding about sequentially distributing balls so as to maintain a Bose-Einstein distribution that the probability that the t^{th} ball goes into any particular unoccupied urn is

$$\frac{0+1}{(mn+s)+(t-1)}.$$

Whence it follows that

$$P(B|A) = m(n-k+1)\left(\frac{0+1}{(mn+s)+(t-1)}\right).$$

Hence,

$$P(T = t) = \left(\frac{m(n-k+1)}{(mn+s)+(t-1)}\right)\left(\frac{n_{(k-1)}}{(mn+s)^{(t-1)}}\left|G(t-1, k-1; -m, -s)\right|\right)$$

$$= \frac{mn_{(k)}}{(mn+s)^{(t)}}\left|G(t-1, k-1; -m, -s)\right|.$$

(d)

If we define

$$\mathcal{A}_t = \left\{(a_1, a_2, \ldots, a_n, a_{n+1}) \middle| a_1 + \cdots + a_{n+1} = t \text{ and at least } r \text{ of } (a_1, a_2, \ldots, a_n) \text{ equal } 0\right\}$$

And

$$\mathcal{A} = \left\{(a_1, a_2, \dots, a_n, a_{n+1}) \middle| \text{at least } r \text{ of } (a_1, a_2, \dots, a_n) \text{ equal } 0\right\}$$

then by Theorem 1,

$$P(\text{at least } r \text{ of the first } n \text{ urns are empty})$$

$$= P\left((X_{1,t}, \dots, X_{n,t}, X_{n+1,t}) \in \mathcal{A}_t\right)$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s} P\big((Y_1, \dots, Y_n, Y_{n+1}) \in \mathcal{A}\big)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s}\left(\sum_{i=r}^{n} (-1)^{i-r}\binom{i-1}{r-1}\binom{n}{i}\left(P(Y_1 = 0)^i\right)\right)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{p}\right)^{mn+s}\left(\sum_{i=r}^{n} (-1)^{i-r}\binom{i-1}{r-1}\binom{n}{i} p^{im}\right)\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\sum_{i=r}^{n} (-1)^{i-r}\binom{i-1}{r-1}\binom{n}{i}\left(\frac{1}{p}\right)^{m(n-i)+s}\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\sum_{i=r}^{n} (-1)^{i-r}\binom{i-1}{r-1}\binom{n}{i}\left(\frac{1}{1-(1-p)}\right)^{m(n-i)+s}\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \frac{d^t}{dp^t}\left(\sum_{i=r}^{n}\sum_{j=0}^{\infty} (-1)^{i-r}\binom{i-1}{r-1}\binom{n}{i}\binom{m(n-i)+s+j-1}{j}(1-p)^j\right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \sum_{i=r}^{n} \sum_{j=0}^{\infty} (-1)^{i-r} \binom{i-1}{r-1} \binom{n}{i} \binom{m(n-i)+s+j-1}{j} \left(\frac{d^t}{dp^t} \left((1 - p)^j\right)\Big|_{p=1} \right)$$

$$= \frac{(-1)^t}{\binom{mn+s+t-1}{t} t!} \sum_{i=r}^{n} \sum_{j=0}^{\infty} (-1)^{i-r} \binom{i-1}{r-1} \binom{n}{i} \binom{m(n-i)+s+j-1}{j} t!\,(-1)^t \mathrm{I}_{\{t\}}(j)$$

$$= \frac{1}{\binom{mn+s+t-1}{t}} \sum_{i=r}^{n} (-1)^{i-r} \binom{i-1}{r-1} \binom{n}{i} \binom{m(n-i)+s+t-1}{t}.$$

Problem 6.

The r^{th} descending factorial moment of the number of colors drawn k times after t balls have been drawn from an urn with m balls each of n different colors according to the Polya urn model equals

$$\frac{\binom{k+m-1}{m-1}^r \binom{m(n-r)+(t-kr)-1}{t-kr}}{\binom{mn+t-1}{t}} n_{(r)}$$

provided $n \geq r$ and $t \geq kr$ and equals 0 otherwise.

$$\mathrm{E}_t\big(g(X_1, \dots, X_n)\big) = \frac{(-1)^t}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{p}\right)^{mn} \mathrm{E}\big(g(Y_1, \dots, Y_n)\big) \right)\Bigg|_{p=1}$$

$$= \frac{(-1)^t}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{p}\right)^{mn} n_{(r)} \left(\binom{k+m-1}{m-1} p^m (1-p)^k \right)^r \right)\Bigg|_{p=1}$$

Theorem 8.10 Binomial Moment of Sum of Indicator Variables

$$= \frac{(-1)^t \binom{k+m-1}{m-1}^r n_{(r)}}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{p}\right)^{m(n-r)} (1-p)^{kr} \right) \Bigg|_{p=1}$$

$$= \frac{(-1)^t \binom{k+m-1}{m-1}^r n_{(r)}}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-(1-p)}\right)^{m(n-r)} (1-p)^{kr} \right) \Bigg|_{p=1}$$

provided $n \geq r$

$$= \frac{(-1)^t \binom{k+m-1}{m-1}^r n_{(r)}}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\sum_{j=0}^{\infty} \binom{m(n-r)+j-1}{j} (1-p)^{j+kr} \right) \Bigg|_{p=1}$$

$$= \frac{(-1)^t \binom{k+m-1}{m-1}^r n_{(r)}}{\binom{mn+t-1}{t} t!} \sum_{j=0}^{\infty} \binom{m(n-r)+j-1}{j} \left(\frac{d^t}{dp^t} (1-p)^{j+kr} \Big|_{p=1} \right)$$

$$= \frac{(-1)^t \binom{k+m-1}{m-1}^r n_{(r)}}{\binom{mn+t-1}{t} t!} \sum_{j=0}^{\infty} \binom{m(n-r)+j-1}{j} \left((-1)^{j+kr} (j+kr)!\, \mathrm{I}_{\{t-kr\}}(j) \right)$$

$$= \frac{\binom{k+m-1}{m-1}^r n_{(r)}}{\binom{mn+t-1}{t}} \binom{m(n-r)+(t-kr)-1}{t-kr}$$

provided $t \geq kr$.

The special case $m = 1, r = 1, k = i, t = j, n = N$

$$\frac{N}{\binom{N+j-1}{j}} \binom{N+j-i-2}{j-i} = \frac{j!}{(j-i)!} \frac{\prod_{p=0}^{j-i-1} (N+p-1)}{\prod_{p=1}^{j-1} (N+p)}$$

$$\mathrm{E}\left(S_{(r)}\right) = n_{(r)} \left(P(Y_1 = k)\right)^r$$

$$= n_{(r)} \left(\binom{k+m-1}{m-1} p^m (1-p)^k \right)^r$$

$$P(Y_i = k) = \binom{k+m-1}{m-1} p^m (1-p)^k.$$