Geometric Randomization

Form 1.

Let $\mathbb{S}^n$ be the product space $\{1,2,\ldots\} \times \cdots \times \{1,2,\ldots\}$ and let $\mathbb{S}^n_t$ be the set of all vectors $(s_1, s_2, \ldots, s_n)$ in $\mathbb{S}^n$ such that $s_1 + s_2 + \ldots + s_n = t$.

Define $(X_{1,t}, X_{2,t}, \ldots, X_{n,t})$ to be that random vector which is equally likely to be any value in $\mathbb{S}^n_t$ and define $Y_1, Y_2, \ldots, Y_n$ to be iid geometric random variables on $y \in \{1,2,\ldots\}$ with parameter p,

i.e.

$$P(Y = y) = p(1-p)^{y-1} \quad y \in \{1,2,\ldots\} \text{ and } 0 \le p \le 1.$$

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

$$P\Big((X_{1,t}, X_{2,t}, \ldots, X_{n,t}) \in \mathcal{A}_t\Big)$$

$$= \frac{(-1)^t}{\binom{t-1}{n-1}\, t!} \cdot \sum_{j=0}^{\infty} \binom{n+j-1}{j} \frac{d^t}{dp^t} \left((1-p)^{n+j}\, P\Big((Y_1, \ldots, Y_n) \in \mathcal{A}\Big) \right) \Bigg|_{p=1}.$$

Form 2.

In many problems it is convenient to define $\mathbb{S}^n$ to be the product space $\{0,1,\ldots\} \times \cdots \times \{0,1,\ldots\}$ and let $\mathbb{S}^n_t$ be the set of all vectors $(s_1, s_2, \ldots, s_n)$ in $\mathbb{S}^n$ such that $s_1 + s_2 + \ldots + s_n = t$.

As before define $(X_{1,t}, X_{2,t}, \ldots, X_{n,t})$ to be that random vector which is equally likely to be any value in $\mathbb{S}^n_t$ and define $Y_1, Y_2, \ldots, Y_n$ to be iid geometric random variables on $\{0,1,\ldots\}$ with parameter p,

i.e.

$$P(Y = y) = p(1-p)^{y} \quad y \in \{0,1,\ldots\} \text{ and } 0 \le p \le 1.$$

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

$$P\Big((X_{1,t}, X_{2,t}, \ldots, X_{n,t}) \in \mathcal{A}_t\Big)$$

$$= \frac{(-1)^t}{\binom{n+t-1}{t}\, t!} \cdot \sum_{j=0}^{\infty} \binom{n+j-1}{j} \frac{d^t}{dp^t} \left((1-p)^{j}\, P\Big((Y_1, \ldots, Y_n) \in \mathcal{A}\Big) \right) \Bigg|_{p=1}.$$

Proof (**Form 1.**)

$$\mathrm{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_n = t) = \mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathbb{S}_t^n \Big)$$

$$= \sum_{\mathbb{S}_t^n} p^n (1-p)^{(\mathrm{s}_1 + \ldots + \mathrm{s}_n) - n}$$

$$= \sum_{\mathbb{S}_t^n} p^n (1-p)^{t-n} = \mathrm{N}(\mathbb{S}_t^n)\, p^n (1-p)^{t-n}.$$

However by the problem of multisets, we have that

$$\mathrm{N}(\mathbb{S}_t^n) = \binom{n + (t-n) - 1}{t-n} = \binom{t-1}{n-1}.$$

Therefore,

$$\mathrm{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_n = t) = \binom{t-1}{n-1} p^n (1-p)^{t-n}.$$

Thus,

$$\mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathcal{A} \,\Big|\, \mathrm{Y}_1 + \ldots + \mathrm{Y}_n = t \Big)$$

$$= \frac{\mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathcal{A} \text{ and } \mathrm{Y}_1 + \ldots + \mathrm{Y}_n = t \Big)}{\mathrm{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_n = t)}$$

$$= \frac{\mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathcal{A}_t \Big)}{\mathrm{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_n = t)}$$

$$= \frac{\sum_{\mathcal{A}_t} p^n (1-p)^{(\mathrm{s}_1 + \ldots + \mathrm{s}_n) - n}}{\binom{t-1}{n-1} p^n (1-p)^{t-n}}$$

$$= \frac{\sum_{\mathcal{A}_t} p^n (1-p)^{t-n}}{\binom{t-1}{n-1} p^n (1-p)^{t-n}} = \frac{\mathrm{N}(\mathcal{A}_t)\, p^n (1-p)^{t-n}}{\binom{t-1}{n-1} p^n (1-p)^{t-n}}$$

$$= \frac{\mathrm{N}(\mathcal{A}_t)}{\binom{t-1}{n-1}} = \mathrm{P}\Big((\mathrm{X}_{1,t}, \mathrm{X}_{2,t}, \ldots, \mathrm{X}_{n,t}) \in \mathcal{A}_t \Big).$$

Therefore,

$$\mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathcal{A} \Big) = \sum_{t=n}^{\infty} \mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathcal{A} \,\Big|\, \sum \mathrm{Y}_i = t \Big) \, \mathrm{P}\Big(\sum \mathrm{Y}_i = t \Big)$$

$$= \sum_{t=n}^{\infty} \mathrm{P}\Big((\mathrm{X}_{1,t}, \mathrm{X}_{2,t}, \ldots, \mathrm{X}_{n,t}) \in \mathcal{A}_t \Big) \binom{t-1}{n-1} p^n (1-p)^{t-n}$$

and

$$\left(\frac{1-p}{p} \right)^n \mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathcal{A} \Big)$$

$$= \sum_{t=n}^{\infty} \mathrm{P}\Big((\mathrm{X}_{1,t}, \mathrm{X}_{2,t}, \ldots, \mathrm{X}_{n,t}) \in \mathcal{A}_t \Big) \binom{t-1}{n-1} (1-p)^t.$$

It follows that

$$\frac{\mathrm{d}^r}{\mathrm{d}p^r} \left(\left(\frac{1-p}{p} \right)^n \mathrm{P}\Big((\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n) \in \mathcal{A} \Big) \right) \Bigg|_{p=1}$$

$$= \frac{\mathrm{d}^r}{\mathrm{d}p^r} \left(\sum_{t=n}^{\infty} \mathrm{P}\Big((\mathrm{X}_{1,t}, \mathrm{X}_{2,t}, \ldots, \mathrm{X}_{n,t}) \in \mathcal{A}_t \Big) \binom{t-1}{n-1} (1-p)^t \right) \Bigg|_{p=1}$$

$$= \sum_{t=n}^{\infty} \mathrm{P}\Big((\mathrm{X}_{1,t}, \mathrm{X}_{2,t}, \ldots, \mathrm{X}_{n,t}) \in \mathcal{A}_t \Big) \binom{t-1}{n-1} \left(\frac{\mathrm{d}^r}{\mathrm{d}p^r} (1-p)^t \Bigg|_{p=1} \right)$$

$$= \sum_{t=n}^{\infty} \mathrm{P}\Big((\mathrm{X}_{1,t}, \mathrm{X}_{2,t}, \ldots, \mathrm{X}_{n,t}) \in \mathcal{A}_t \Big) \binom{t-1}{n-1} \Big(r!\,(-1)^r \, \mathrm{I}_{\{r\}}(t) \Big)$$

$$= \mathrm{P}\Big((\mathrm{X}_{1,r}, \mathrm{X}_{2,r}, \ldots, \mathrm{X}_{n,r}) \in \mathcal{A}_r \Big) \binom{r-1}{n-1} r!\,(-1)^r \, \mathrm{I}_{\{n,n+1,\ldots\}}(r).$$

Thus for $r \geq n$,

$$\mathrm{P}\Big((\mathrm{X}_{1,r}, \mathrm{X}_{2,r}, \ldots, \mathrm{X}_{n,r}) \in \mathcal{A}_r \Big)$$

$$= \frac{1}{\binom{r-1}{n-1} r!(-1)^r} \cdot \frac{d^r}{dp^r}\left(\left(\frac{1-p}{p}\right)^n P\Big((Y_1, Y_2, \ldots, Y_n) \in \mathcal{A}\Big)\right)\Bigg|_{p=1}.$$

However,

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

and

$$\frac{1}{(-1)^r} = (-1)^r.$$

Thus for $r \geq n$,

$$P\Big((X_{1,r}, X_{2,r}, \ldots, X_{n,r}) \in \mathcal{A}_r\Big)$$

$$= \frac{(-1)^r}{\binom{r-1}{n-1} r!} \cdot \sum_{j=0}^{\infty}\binom{n+j-1}{j}\frac{d^r}{dp^r}\left((1-p)^{n+j} P\Big((Y_1, \ldots, Y_n) \in \mathcal{A}\Big)\right)\Bigg|_{p=1}. \qquad \square$$

The proof of Form 2 is nearly identical to the proof of Form 1 and is therefore omitted.

Applications

(I) How many ways are there to select j objects from n distinct objects arranged in a circle such that there is no group of k contiguous objects within the j objects selected ? (We will say k objects are contiguous if they form an uninterrupted string.)

This problem was solved in " A direct argument for Kaplansky's theorem on a cyclic arrangement and its generalization ", F. K. Hwang and Y.C. Yao, *Operations Research Letters*, 10, 1991, 241 - 243. They used an argument relating circular to linear arrangements to show that there are a total of

$$\frac{n}{n-j} \sum_{i=0}^{n-j} (-1)^i \binom{n-ki-1}{n-j-1} \binom{n-j}{i}$$

such selections.

(II) How many ways are there to select j objects from n distinct objects arranged in a circle such that there are exactly v groups of k or more contiguous objects ? Show that there are

$$\frac{n}{n-j} \cdot \binom{n-j}{v} \cdot \sum_{i=0}^{n-j-v} (-1)^i \binom{n-k(i+v)-1}{n-j-1} \binom{n-j-v}{i}$$

such selections. (Clearly (I) is the special case of $v = 0$.)

(III) How many ways are there to select j objects from n distinct objects arranged in a circle such that the j objects selected form exactly v groups of contiguous objects ? Show that there are

$$\frac{n}{n-j} \binom{n-j}{v} \binom{j-1}{v-1}$$

such selections. (This is the special case of (II) with $k = 1$.)

(IV) Suppose we randomly arrange m X's and n Y's in a line. Find P(longest run of X's $\leq k$).

(V) Gibrat's Law of Proportionality or Polya Sampling.

(VI) The number of ways of selecting k balls from a cycle of length n with exactly p adjacent selected balls having exactly t unselected balls between them.

The number of ways of selecting k balls from a line of length n with exactly p adjacent selected balls having exactly t unselected balls between them.

[B.S. El-Desouky, "On Selecting k Balls from an n - Line Without Unit Separation", *Indian Journal of Pure and Applied Mathematics*, 19, (2), February 1988, 145 - 148.]

We will show that each of these problems follow directly from the Geometric Randomization Theorem.

Solution (I)

Suppose we arrange n different colored balls in a circle. We pick r of these n balls at random so that the probability of any particular sample is

$$\frac{1}{\binom{n}{r}}.$$

One way to pick r balls at random is to pick all r at once. Alternatively, we could pick 1 ball at random and then pick $r-1$ balls at once from the remaining $n-1$ balls. Both methods assign equal probability to all possible samples.

Consider the latter sampling scheme and suppose that after we pick the first ball we move clockwise from that ball and attach the labels $1, 2, \ldots, n-1$ to the balls as we go around the circle. After we put the labels on the balls we pick the remaining $r-1$ balls. Clearly attaching the labels does not change the fact that all possible samples are equally likely to occur.

Let $X_1 < X_2 < \ldots < X_{r-1}$ be the labels on the $r-1$ balls drawn at this second phase of our sample. We note that $(X_1, X_2, \ldots, X_{r-1})$ is equally likely to be any one of the

$$\binom{n-1}{r-1}$$

possible samples from $\{1, 2, \ldots, n-1\}$.

Let

$$\begin{aligned} S_1 &= X_1 \\ S_2 &= X_2 - X_1 \\ &\vdots \\ S_{r-1} &= X_{r-1} - X_{r-2} \\ S_r &= n - X_{r-1}. \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} X_1 &= S_1 \\ X_2 &= S_1 + S_2 \\ &\vdots \\ X_{r-1} &= S_1 + S_2 + \ldots + S_{r-1} \end{aligned}$$

We see that $S_1 - 1, S_2 - 1, \ldots, S_r - 1$ count the number of unselected objects between each of the selected objects and that $n = S_1 + S_2 + \ldots + S_{r-1} + S_r$.

It follows that

$$\begin{aligned} &P\Big(S_1 = s_1, S_2 = s_2, \ldots, S_r = s_r \Big) \\ &\quad = P\Big(X_1 = s_1, X_2 = s_1 + s_2, \ldots, X_{r-1} = s_1 + s_2 + \ldots + s_{r-1} \Big) \\ &\quad = \frac{1}{\binom{n-1}{r-1}} \quad \text{for all } (s_1, s_2, \ldots, s_r) \in \mathbb{S}_n^r. \end{aligned}$$

Now we can show how to use this result in connection with our theorem on Geometric Randomization to solve the stated problem.

We note that

there is no group of k contiguous objects among the j objects selected

$\Leftrightarrow$

the gap between each of the $n-j$ objects <u>not</u> selected contains less than k objects.

Therefore,

N(samples of size j such that there is no group of k contiguous objects)

$=$ N(samples of size j such that the gap between each of the $n-j$ objects <u>not</u> selected contains less than k objects)

$=$ N(samples of size $n-j$ such that the gap between each of the $n-j$ objects selected contains less than k objects)

[Selecting j out of n objects $\equiv$ not selecting $n-j$ out of n objects.]

$$= \mathrm{N}\Big(\mathrm{S}_1 - 1 < k, \mathrm{S}_2 - 1 < k, \ldots, \mathrm{S}_{n-j} - 1 < k \Big)$$

$$= \mathrm{N}\Big(\mathrm{S}_1 \leq k, \mathrm{S}_2 \leq k, \ldots, \mathrm{S}_{n-j} \leq k \Big)$$

$$= \binom{n}{n-j} \mathrm{P}\Big(\mathrm{S}_1 \leq k, \mathrm{S}_2 \leq k, \ldots, \mathrm{S}_{n-j} \leq k \Big)$$

$$= \binom{n}{n-j} \mathrm{P}\Big((\mathrm{S}_1, \mathrm{S}_2, \ldots, \mathrm{S}_{n-j}) \in \mathcal{A}_n \Big)$$

where $\mathcal{A}_n = \Big\{ (\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_{n-j}) \Big| \, 1 \leq \mathrm{s}_1 \leq k, \ldots, 1 \leq \mathrm{s}_{n-j} \leq k \text{ and } \mathrm{s}_1 + \ldots + \mathrm{s}_{n-j} = n \Big\}$.

Now we can apply the Geometric Randomization Theorem to find $\mathrm{P}\Big((\mathrm{S}_1, \mathrm{S}_2, \ldots, \mathrm{S}_{n-j}) \in \mathcal{A}_n \Big)$. By this theorem,

$$\mathrm{P}\Big((\mathrm{S}_1, \mathrm{S}_2, \ldots, \mathrm{S}_{n-j}) \in \mathcal{A}_n \Big)$$

$$= \frac{(-1)^n}{\binom{n-1}{n-j-1} n!} \cdot \sum_{u=0}^{\infty} \binom{n-j+u-1}{u} \frac{\mathrm{d}^n}{\mathrm{d}p^n} \Bigg((1-p)^{n-j+u} \, \mathrm{P}\Big((\mathrm{Y}_1, \ldots, \mathrm{Y}_{n-j}) \in \mathcal{A} \Big) \Bigg) \Bigg|_{p=1}$$

where $\mathcal{A}$ is <u>any</u> set such that $\mathcal{A} \subset \mathbb{S}^{n-j}$ and $\mathcal{A}_n = \mathcal{A} \cap \mathbb{S}_n^{n-j}$ and $\mathrm{Y}_1, \ldots, \mathrm{Y}_{n-j}$ are iid Geometric random variables.

It is easy to see that

$$\mathrm{P}(\mathrm{Y}_1 > k) = \sum_{r=k+1}^{\infty} p\,(1-p)^{r-1} = p\,(1-p)^k \sum_{r=0}^{\infty}(1-p)^r$$

$$= p\,(1-p)^k \frac{1}{1-(1-p)} = (1-p)^k.$$

Therefore,

$$\mathrm{P}\Big((\mathrm{Y}_1,\ldots,\mathrm{Y}_{n-j}) \in \mathcal{A}\Big) = \mathrm{P}\Big(\mathrm{Y}_1 \le k,\ldots,\mathrm{Y}_{n-j} \le k\Big)$$

$$= \Big(1 - \mathrm{P}(\mathrm{Y}_1 > k)\Big)^{n-j} = \Big(1 - (1-p)^k\Big)^{n-j}.$$

Thus,

$$\frac{\mathrm{d}^n}{\mathrm{d}p^n}\left((1-p)^{n-j+u}\,\mathrm{P}\Big((\mathrm{Y}_1,\ldots,\mathrm{Y}_{n-j}) \in \mathcal{A}\Big)\right)\Bigg|_{p=1}$$

$$= \frac{\mathrm{d}^n}{\mathrm{d}p^n}\left((1-p)^{n-j+u}\,\Big(1 - (1-p)^k\Big)^{n-j}\right)\Bigg|_{p=1}$$

$$= \frac{\mathrm{d}^n}{\mathrm{d}p^n}\left(\sum_{i=0}^{n-j}\binom{n-j}{i}(-1)^i\,(1-p)^{n-j+u+ki}\right)\Bigg|_{p=1}$$

$$= \sum_{i=0}^{n-j}\binom{n-j}{i}(-1)^i\,\frac{\mathrm{d}^n}{\mathrm{d}p^n}\left((1-p)^{n-j+u+ki}\right)\Bigg|_{p=1}$$

$$= \sum_{i=0}^{n-j}\binom{n-j}{i}(-1)^i\Big((-1)^n\,n!\;\mathrm{I}_{\{n-j+u+ki\}}(n)\Big).$$

$$= \sum_{i=0}^{n-j}\binom{n-j}{i}(-1)^i\Big((-1)^n\,n!\;\mathrm{I}_{\{j-ki\}}(u)\Big).$$

Thus,

$$\mathrm{P}\Big((\mathrm{S}_1,\mathrm{S}_2,\ldots,\mathrm{S}_{n-j}) \in \mathcal{A}_n\Big)$$

$$= \frac{(-1)^n}{\binom{n-1}{n-j-1} n!} \cdot \sum_{u=0}^{\infty} \binom{n-j+u-1}{u} \left(\sum_{i=0}^{n-j} \binom{n-j}{i} (-1)^i \Big((-1)^n n! \, \mathrm{I}_{\{j-ki\}}(u) \Big) \right)$$

$$= \frac{(-1)^n}{\binom{n-1}{n-j-1} n!} \cdot \left(\sum_{i=0}^{n-j} \binom{n-j+(j-ki)-1}{j-ki} \binom{n-j}{i} (-1)^i \Big((-1)^n n! \Big) \right)$$

$$= \frac{1}{\binom{n-1}{n-j-1}} \cdot \left(\sum_{i=0}^{n-j} (-1)^i \binom{n-ki-1}{j-ki} \binom{n-j}{i} \right).$$

Thus, we have

N(samples of size j such that there is no group of k contiguous objects)

$$= \binom{n}{n-j} \mathrm{P}\Big((\mathrm{S}_1, \mathrm{S}_2, \dots, \mathrm{S}_{n-j}) \in \mathcal{A}_n \Big)$$

$$= \frac{n}{n-j} \sum_{i=0}^{n-j} (-1)^i \binom{n-ki-1}{n-j-1} \binom{n-j}{i}. \qquad \square$$

Solution (II)

Following the pattern developed in (I), we note that

N(samples of size j such that there are exactly v groups of k or more contiguous objects)

$=$ N(samples of size $n-j$ such that exactly v of the gaps between the $n-j$ objects not selected contain k or more objects)

$$= \binom{n}{n-j} \mathrm{P}\Big((\mathrm{S}_1, \mathrm{S}_2, \dots, \mathrm{S}_{n-j}) \in \mathcal{A}_n \Big)$$

where

$$\mathcal{A}_n = \Big\{ (\mathrm{s}_1, \mathrm{s}_2, \dots, \mathrm{s}_{n-j}) \,\Big|\, \text{exactly } v \text{ of components of } (\mathrm{s}_1, \mathrm{s}_2, \dots, \mathrm{s}_{n-j}) \text{ are values strictly greater than } k \text{ and } \mathrm{s}_1 + \dots + \mathrm{s}_{n-j} = n \Big\}.$$

In this case we have

$$\mathrm{P}\Big((\mathrm{Y}_1, \dots, \mathrm{Y}_{n-j}) \in \mathcal{A} \Big) = \mathrm{P}\Big(\text{exactly } v \text{ of the statements } \mathrm{Y}_1 > k, \dots, \mathrm{Y}_{n-j} > k \text{ are true} \Big)$$

$$= \binom{n-j}{v}\Big(\mathrm{P}(\,\mathrm{Y}_1 \,>\, k\,)\Big)^{v}\Big(1 \,-\, \mathrm{P}(\,\mathrm{Y}_1 \,>\, k\,)\Big)^{n-j-v}$$

$$= \binom{n-j}{v}\Big((\,1-p\,)^{k}\Big)^{v}\Big(1 \,-\, (\,1-p\,)^{k}\Big)^{n-j-v}$$

$$= \binom{n-j}{v}\sum_{i=0}^{n-j-v}(\,-\,1\,)^{i}\binom{n-j-v}{i}\,(\,1-p\,)^{k(v+i)}.$$

Thus,

$$\frac{\mathrm{d}^n}{\mathrm{d}p^n}\Bigg((\,1-p\,)^{n-j+u}\,\mathrm{P}\Big((\,\mathrm{Y}_1,\ldots,\mathrm{Y}_{n-j}\,)\,\in\,\mathcal{A}\Big)\Bigg)\Bigg|_{p=1}$$

$$= \frac{\mathrm{d}^n}{\mathrm{d}p^n}\Bigg(\binom{n-j}{v}\sum_{i=0}^{n-j-v}(\,-\,1\,)^{i}\binom{n-j-v}{i}\,(\,1-p\,)^{n-j+u+k(v+i)}\Bigg)\Bigg|_{p=1}$$

$$= \binom{n-j}{v}\sum_{i=0}^{n-j-v}(\,-\,1\,)^{i}\binom{n-j-v}{i}\Big((\,-\,1\,)^{n}\,n\,!\;\mathrm{I}_{\{\,j-k(v+i)\,\}}(\,u\,)\Big).$$

It follows that the only value of u we need consider is $u \,=\, j \,-\, k(i+v)$. Therefore,

$$\mathrm{P}\Big((\,\mathrm{S}_1,\mathrm{S}_2,\ldots,\mathrm{S}_{n-j}\,)\,\in\,\mathcal{A}_n\Big)$$

$$= \frac{(-1)^n}{\binom{n-1}{n-j-1}\,n\,!}\cdot\sum_{u=0}^{\infty}\binom{n-j+u-1}{u}\frac{\mathrm{d}^n}{\mathrm{d}p^n}\Bigg((\,1-p\,)^{n-j+u}\,\mathrm{P}\Big((\,\mathrm{Y}_1,\ldots,\mathrm{Y}_{n-j}\,)\,\in\,\mathcal{A}\Big)\Bigg)\Bigg|_{p=1}$$

$$= \frac{(-1)^{2n}\,n\,!}{\binom{n-1}{n-j-1}\,n\,!}\cdot\Bigg(\sum_{i=0}^{n-j-v}(\,-\,1\,)^{i}\binom{n-j+(j-k\,i-kv)-1}{j-k\,i-kv}\binom{n-j}{v}\binom{n-j-v}{i}\Bigg).$$

After simplifying we have that

$$\binom{n}{n-j}\mathrm{P}\Big((\,\mathrm{S}_1,\mathrm{S}_2,\ldots,\mathrm{S}_{n-j}\,)\,\in\,\mathcal{A}_n\Big)$$

$$= \frac{n}{n-j} \cdot \binom{n-j}{v} \cdot \sum_{i=0}^{n-j-v} (-1)^i \binom{n-k(i+v)-1}{n-j-1} \binom{n-j-v}{i}. \qquad \square$$

Solution (**III**)

Substituting $k = 1$ into the general solution developed in (II) we have that there are

$$\frac{n}{n-j} \cdot \binom{n-j}{v} \cdot \sum_{i=0}^{n-j-v} (-1)^i \binom{n-1(i+v)-1}{n-j-1} \binom{n-j-v}{i}$$

such solutions. In Problem ??? we showed that

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{x-k}{r} = \binom{x-n}{r-n}.$$

It follows from this identity that

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

which establishes the solution.

Solution (**IV**)

One method of randomly arranging the m X's and n Y's is to consider placing an urn before, after, and between each Y and to then distribute the m X's into these urns in such a manner that all

$$\binom{(n+1)+m-1}{m} = \binom{n+m}{m}$$

distributions are equally likely.

Let $X_{j,m}$ = the number of X's that are put into the j^{th} urn, $j = 1, 2, \ldots, n+1$. It follows that $(X_{1,m}, X_{2,m}, \ldots, X_{n+1,m})$ is a random vector which is equally likely to be any value in $\mathbb{S}_m^{n+1}$.

Therefore,

$$\begin{aligned} P(\text{longest run of X's} \le k) &= P(X_{1,m} \le k, \ldots, X_{n+1,m} \le k) \\ &= P\Big((X_{1,m}, X_{2,m}, \ldots, X_{n+1,m}) \in \mathcal{A}_m \Big) \end{aligned}$$

where $\mathcal{A}_m = \Big\{ (\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_{n+1}) \Big| \; 0 \le \mathrm{s}_1 \le k, \ldots, 0 \le \mathrm{s}_{n+1} \le k \text{ and } \mathrm{s}_1 + \ldots + \mathrm{s}_{n+1} = m \Big\}$.

Define $\mathcal{A} = \Big\{ (\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_{n+1}) \Big| \; 0 \le \mathrm{s}_1 \le k, \ldots, 0 \le \mathrm{s}_{n+1} \le k \Big\}$. Then by the Geometric Randomization Theorem,

P(longest run of X's $\le k$)

$$= \mathrm{P}\Big((\mathrm{X}_{1,m}, \mathrm{X}_{2,m}, \ldots, \mathrm{X}_{n+1,m}) \in \mathcal{A}_m \Big)$$

$$= \frac{(-1)^m}{\binom{n+m}{m} m!} \cdot \sum_{j=0}^{\infty} \binom{n+j}{j} \frac{\mathrm{d}^m}{\mathrm{d}p^m} \left((1-p)^j \, \mathrm{P}\Big((\mathrm{Y}_1, \ldots, \mathrm{Y}_{n+1}) \in \mathcal{A} \Big) \right) \Bigg|_{p=1}$$

$$= \frac{(-1)^m}{\binom{n+m}{m} m!} \cdot \sum_{j=0}^{\infty} \binom{n+j}{j} \frac{\mathrm{d}^m}{\mathrm{d}p^m} \left((1-p)^j \Big(\mathrm{P}(\mathrm{Y}_1 \le k) \Big)^{n+1} \right) \Bigg|_{p=1}$$

$$= \frac{(-1)^m}{\binom{n+m}{m} m!} \cdot \sum_{j=0}^{\infty} \binom{n+j}{j} \frac{\mathrm{d}^m}{\mathrm{d}p^m} \left((1-p)^j \Big(1 - (1-p)^{(k+1)} \Big)^{n+1} \right) \Bigg|_{p=1}$$

$$= \frac{(-1)^m}{\binom{n+m}{m} m!} \cdot \sum_{j=0}^{\infty} \binom{n+j}{j} \frac{\mathrm{d}^m}{\mathrm{d}p^m} \left(\sum_{i=0}^{n+1} (-1)^i \binom{n+1}{i} (1-p)^{j+(k+1)i} \right) \Bigg|_{p=1}$$

$$= \frac{(-1)^m}{\binom{n+m}{m} m!} \cdot \sum_{j=0}^{\infty} \binom{n+j}{j} \left(\sum_{i=0}^{n+1} (-1)^{i+m} \, m! \binom{n+1}{i} \mathrm{I}_{\{(k+1)i+j\}}(m) \right)$$

$$= \frac{1}{\binom{n+m}{m}} \cdot \sum_{i=0}^{n+1} \sum_{j=0}^{\infty} (-1)^i \binom{n+j}{j} \binom{n+1}{i} \mathrm{I}_{\{m-(k+1)i\}}(j)$$

$$= \frac{1}{\binom{n+m}{m}} \cdot \sum_{i=0}^{n+1} (-1)^i \binom{n+m-(k+1)i}{m-(k+1)i} \binom{n+1}{i}. \qquad \square$$