

Probability Distributions Motivated by Permutations and Featuring Telescoping Series

Jie Hao and Anant Godbole, ETSU

Abstract

In this poster, we study the simplest type of infinite telescoping sum, namely

$$\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} (q_n - q_{n+1}) = 1,$$

due to our assumptions that $1 = q_1 \geq q_2 \geq q_3 \dots \geq 0$ and $\lim_{n \rightarrow \infty} q_n = 0$. Now the above clearly defines a probability distribution on $\mathbf{Z}^+$, but making unmotivated choices such as $q_n = \frac{1}{3n^4 - 2n^2}$ would serve little purpose, however, so let us state our objective: We wish to propose three probability distributions that each

(i) are similar in form to familiar discrete distributions;

(ii) arise from the cycle or pattern containment structure of random permutations; and

(iii) lead to open-ended sets of questions.

Let us go one step further, by asking

(iv) that there be events $A_n, n \geq 1$ in some probability space so that $q_n = P(A_n)$ with $A_{n+1} \subseteq A_n$ so that, as we shall see, $p_n, n \geq 1$ provides a model for *first occurrence distributions* of certain events.

The viewer of this poster should note that discrete models, other than those provided by random permutations, could equally well have been used as launching pads for our discrete distributions, and s/he is invited to come up with well-motivated first occurrence distributions along the lines of the ones in this paper.

Poisson, etc.

ZETA: The zeta distribution with parameter $k > 1$ has associated probability mass function (pmf) given by

$$f(n) = \frac{1}{\zeta(k)} \cdot \frac{1}{n^k}, n = 1, 2, \dots,$$

where $\zeta(k) = \sum_{n=1}^{\infty} \frac{1}{n^k}$ is the Riemann zeta function.

POISSON: The Poisson distribution with parameter $\lambda > 0$, denoted by $Po(\lambda)$, is the distribution of a random variable X with pmf

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, \dots$$

and the so-called moment generating function (m.g.f.) of X is

$$M_X(t) = E(e^{tX}) = \exp\{\lambda(e^t - 1)\}.$$

GEOMETRIC: The geometric random variable counts the number of independent Bernoulli trials with success probability p that need to be conducted in order to get the first success. We have, for $x \geq 1$,

$$f(x) = (1-p)^{x-1} \cdot p; \quad E(X) = \frac{1}{p}; \quad \text{Var}(X) = \frac{1-p}{p^2}.$$

GENERALIZED TELESOPING POISSON: This is a toy example. The probability that a random permutation on $[n]$ has its first ascent at positions $k, k+1$ is given, for $1 \leq k \leq n-1$, by $\frac{k}{(k+1)!}$. To see this, choose any one of the $k+1$ elements in positions 1 through $k+1$, except for the smallest, to occupy the $k+1$ st position, and then arrange the other elements in a monotone decreasing fashion. But, to reformulate this argument via telescoping series, we let A_k be the event that the permutation restricted to the first k integers is monotone decreasing, note that $A_{k+1} \subseteq A_k$ and that the probability $f(k)$ that the first ascent is at $k, k+1$ satisfies, for $k \leq n-1$,

$$f(k) = P(A_k \cap A_{k+1}^C) = \frac{1}{k!} - \frac{1}{(k+1)!} = \frac{k}{(k+1)!},$$

as before. The chance that the first ascent is at position n is, of course, $\frac{1}{n!}$. Letting $n \rightarrow \infty$, we get the discrete distribution on $\mathbf{Z}^+$ given by

$$f(x) = \frac{x}{(x+1)!}, x = 1, 2, \dots, \quad (1)$$

which is quite similar to the unit Poisson distribution on $0, 1, \dots$ with mass function $f(x) = e^{-1}/x!$. If X is the corresponding first ascent random variable generated, e.g., by a sequence $\{X_n\}_{n=1}^{\infty}$ of i.i.d. uniform random variables, then

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x \times \frac{x}{(x+1)!} \\ &= e - 1, \end{aligned}$$

and

$$\begin{aligned} \text{Var}(X) &= E(X^2) - (E(X))^2 \\ &= (e+1) - (e-1)^2 \\ &= e(3-e), \end{aligned}$$

More generally, the m.g.f. of X is given by

$$\begin{aligned} E(e^{tX}) &= \sum_{x=1}^{\infty} e^{tx} \times \frac{x}{(x+1)!} \\ &= e^{-t}(1 - e^{e^t} + e^{t+e^t}), \end{aligned}$$

compared to the m.g.f. $\exp\{e^t - 1\}$ of the unit Poisson r.v. A θ analog has also been studied by us.

THE TELESOPING ZETA(2) DISTRIBUTION: There are $(n-1)!$ permutations on $[n]$ that are unicyclic. Now if a permutation on $[n+1]$ is unicyclic, then its *reduction* to $[n]$ is also unicyclic. For example, the permutation

123456
452613,

written in “two line notation,” can be reduced, on eliminating the 6 in both the first and second rows and mapping 4 to 3 directly, as the unicyclic 5-permutation

12345
45231;

a general definition of reduction can now be easily formulated. Let $n \geq 1$ and consider $\pi \in S_{n+1}$. On setting A_n to be the event that the reduced permutation on $[n]$ is unicyclic, we have that $A_{n+1} \subseteq A_n$, and so $P(A_n \cap A_{n+1}^C) = P(A_n) - P(A_{n+1})$. It follows that

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

is the probability that $\pi \in S_{n+1}$ has a unicyclic reduction but we have $\pi(n+1) = n+1$, so that π itself is not unicyclic, and has an n cycle on $[n]$ followed by a 1-cycle on $\{n+1\}$. Given a random permutation in S_n , we let $X = 1, 2, \dots, n$ be the largest $i \leq n$ such that the reduction of π to $[i]$ is unicyclic. We then have

$$p_i = P(X = i) = \frac{1}{i} - \frac{1}{i+1} = \frac{1}{i(i+1)}, \quad i = 1, 2, \dots, n-1, \quad (2)$$

with

$$P(X = n) = \frac{1}{n}.$$

THE TELESOPING GEOMETRIC DISTRIBUTION: The following is a very natural question: in how many permutations, in which the longest increasing subsequence is of length 2, does the first ascent occur in positions $k, k+1$? Catalan [2] proved the k -fold Catalan convolution formula

$$C_{n,k} := \sum_{i_1 + \dots + i_k = n} \prod_{r=1}^k C_{i_r-1} = \frac{k}{2n-k} \binom{2n-k}{n},$$

and in [3] it was proved that there are precisely $C_{n,k}$ permutations on $[n]$ with longest increasing subsequence of size 2 and for which the first ascent occurs at positions $k, k+1$. Thus, for a randomly chosen 123-avoiding permutation on $[n]$, the distribution of the location of first ascent is given by

$$f(k) = \frac{C_{n,k}}{C_n} = k \frac{(2n-k-1)!(n+1)!}{(2n)!(n-k)!}, \quad k = 1, 2, \dots, n,$$

which, for small k and large n , may be approximated by $f(x) = \frac{k}{2^{k+1}}$. Accordingly, let us define the telescoping geometric-like distribution on $\mathbf{Z}^+ = 1, 2, \dots$ by

$$f(x) = \frac{x}{2^{x+1}} = \frac{x+1}{2^x} - \frac{x+2}{2^{x+1}}, x = 1, 2, \dots$$

We see that for a telescoping geometric random variable X ,

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x \times \frac{x}{2^{x+1}} \\ &= 3. \end{aligned}$$

Thus for a random permutation on a large $[n]$, we expect the first ascent to be at position $e-1 \approx 1.718$, whereas this value increases to 3 for a random 123-avoiding permutation. Moreover, motivated by size-biased permutations, we define a telescoping θ -analog of the above distribution defined by

$$f(x) = \frac{(\theta-1)^2 x}{\theta^{x+1}} = \frac{(\theta-1)x+1}{\theta^x} - \frac{(\theta-1)(x+1)+1}{\theta^{x+1}}, x = 1, 2, \dots, \theta > 1,$$

and its properties can be studied, e.g., the MLE of θ $\hat{\theta} = 1 + \frac{2}{\bar{X}-1}$. The MOM estimator is the same. Finally consider the null hypothesis $H_0 : \theta = \theta_0$ versus the alternative hypothesis $H_1 : \theta = \theta_1$. Assume that $\theta_1 > \theta_0$. By the Neyman Pearson theorem, the most powerful test rejects H_0 for small values of $\bar{X}$, where the critical value is determined by the level of significance used.

References

- [1] M. Bóna (2004). *Combinatorics of Permutations*, Chapman and Hall, Boca Raton.
- [2] E. Catalan (1887). Sur les nombres de Segner, *Rend. Circ. Mat. Palermo* **1**, 190–201.
- [3] S. Connolly, Z. Gabor, and A. Godbole (2014+). The location of the first ascent in a 123-avoiding permutation, submitted, available at <http://arxiv.org/pdf/1401.2691.pdf>