Open Journal of Statistics, 2016, 6, 274-285
Published Online April 2016 in SciRes. http://www.scirp.org/journal/ojs
http://dx.doi.org/10.4236/ojs.2016.62023

Distribution of the Maximum and Minimum of a Random Number of Bounded Random Variables

Jie Hao[1], Anant Godbole[2]

[1]Department of Mathematics, Kennesaw State University, Marietta, USA
[2]Department of Mathematics and Statistics, East Tennessee State University, Johnson City, USA
Email: jhao2@kennesaw.edu, godbolea@etsu.edu

Received 28 January 2016; accepted 23 April 2016; published 26 April 2016

Abstract

We study a new family of random variables that each arise as the distribution of the maximum or minimum of a random number N of i.i.d. random variables $X_1, X_2, \cdots, X_N$, each distributed as a variable X with support on [0, 1]. The general scheme is first outlined, and several special cases are studied in detail. Wherever appropriate, we find estimates of the parameter θ in the one-parameter family in question.

Keywords

Maximum and Minimum, Random Number of i.i.d. Variables, Statistical Inference

1. Introduction

Consider a sequence $X_1, X_2, \cdots$ of i.i.d. random variables with support on $[0,1]$ and having distribution function F. For any fixed n, the distributions of

$$Y = \max_{1 \le i \le n} X_i$$

and

$$Z = \min_{1 \le i \le n} X_i$$

have been well studied; in fact it is shown in elementary texts that $F_Y(x) = F^n(x)$ and $F_Z(x) = 1 - (1 - F(x))^n$. But what if we have a situation where the number N of X_i's is random, and we are instead considering the

How to cite this paper: Hao, J. and Godbole, A. (2016) Distribution of the Maximum and Minimum of a Random Number of Bounded Random Variables. *Open Journal of Statistics*, **6**, 274-285. http://dx.doi.org/10.4236/ojs.2016.62023

extrema

$$Y = \max_{1 \le i \le N} X_i \tag{1}$$

and

$$Z = \min_{1 \le i \le N} X_i \tag{2}$$

of a random number of i.i.d. random variables? Now the *sum* S of a random number of i.i.d. variables, defined as

$$S = \sum_{i=1}^{N} X_i$$

satisfies, according to Wald's Lemma [1], the equation

$$\mathbb{E}(S) = \mathbb{E}(N)\mathbb{E}(X),$$

provided that N is independent of the sequence $\{X_i\}$ and assuming that the means of X and N exist.

The purpose of this paper is to show that the distributions in (1) and (2) can be studied in many canonical cases, even if N and $\{X_i\}_{i=1}^{\infty}$ are correlated. The main deviation from the papers [2] [3] and [4], where similar questions are studied, is that the variable X is concentrated on the interval $[0,1]$—unlike the above references, where X has lifetime-like distributions on $[0,\infty)$. Even then, we find that many new and interesting distributions arise, none of them to be found, e.g., in [5] or [6] via the "extreme values of a random number of i.i.d. variables" connection. See, however, Remarks 1 and 2 in Section 3. In another deviation from the theory of extremes of random sequences (see, e.g., [7]), we find that the tail behavior of the extreme distributions is not relevant due to the fact that the distributions have compact support. We next cite three examples where our methods might be useful. First, we might be interested in the strongest earthquake in a given region in a given year. The number of earthquakes in a year, N, is usually modeled using a Poisson distribution, and, ignoring aftershocks and similarly correlated events, the intensities of the earthquakes can be considered to be i.i.d. random variables in $[a,b]$ whose distribution can be modeled using, e.g., the data set maintained by Caltech at [8]. Second, many "small world" phenomena have recently been modeled by power law distributions, also sometimes termed discrete Pareto or Zipf distributions. See, for example, the body of work by Chung and her co-authors [9] [10], and the references therein, where vertex degrees $d(v)$ in "internet-like graphs" G (e.g., the vertices of G are individual webpages, and there is an edge between v_1 and v_2 if one of the webpages has a link to the other) are shown to be modeled by

$$\mathbb{P}\big(d(v) = n\big) = \frac{\left[\zeta(k)\right]^{-1}}{n^k}$$

for some constant $k > 1$, where $\zeta(\cdot)$ is the Riemann Zeta function

$$\zeta(k) = \sum_{n=1}^{\infty} \frac{1}{n^k}.$$

Thus if the vertices v in a large internet graph have some bounded i.i.d. property X_i, then the maximum and minimum values of X_i for the neighbors of a randomly chosen vertex can be modeled using the methods of this paper. Third, we note that N and the X_i may be correlated, as in the CSUG example (studied systematically in Section 3) where $X_i \sim U[0,1]$ and $N = \inf\{n \ge 2 : X_n > (1-\theta)\}$ follows the geometric distribution $\text{Geo}(\theta)$. This is an example of a situation where we might be modeling the maximum load that a device might have carried before it breaks down due to an excessive weight or current. It is also feasible in this case that the parameter θ might be unknown.

Here is our general set-up: Suppose $X_1, X_2, \cdots$ are i.i.d. random variables following a continuous distribution on $[0,1]$ with probability density and distribution functions given by $f(x)$ and $F(x)$ respectively. N is a random variable following a discrete distribution on $\{1,2,\cdots\}$ with probability mass function given by $\mathbb{P}(N = n) = p(n)$, $n = 1,2,\cdots$. Let Y and Z be given by (1) and (2) respectively. Then the p.d.f.'s g of Y and Z are derived as follows: Since

$$\mathbb{P}(Y \le y | N = n) = [F(y)]^n,$$

we see that

$$g(y|N = n) = n[F(y)]^{n-1} f(y),$$

and consequently, the marginal p.d.f. of Y is

$$\begin{aligned} g(y) &= \sum_{n=1}^{\infty} g(y|N = n)\mathbb{P}(N = n) \\ &= f(y)\sum_{n=1}^{\infty} n[F(y)]^{n-1} p(n). \end{aligned} \tag{3}$$

In a similar fashion, the p.d.f. of Z can be shown to be

$$g(z) = f(z)\sum_{n=1}^{\infty} n[1-F(z)]^{n-1} p(n); \tag{4}$$

what is remarkable is that the sums in (3) and (4) will be shown to assume simple tractable forms in a variety of cases.

We want to point out that some of our distributions have been studied before but not using this motivation. For example, the Marshall-Olkin distributions [11] give a new method of adding a parameter to a distribution. Also, other distributions such as the beta and Kumaraswamy [12] distributions *can* be used to model continuous bounded data, but these do not apply to our set-up. See also Remark 2 in Section 3.

Our paper is organized as follows. Section 1 provided a summary and *motivation* for studying the distributions in the fashion we do. In Section 2, we study the case of $X \sim U[0,1]$ and $N \sim \text{Geo}(\theta)$. We call this the Standard Uniform Geometric model. The graphs of $g(y)$ and $g(z)$ can be seen in **Figure 1** and **Figure 2** respectively. The CSUG (Correlated Standard Uniform Model) is studied in Section 3. The graphs of $g(y)$ and $g(z)$ in the CSUG model are plotted in **Figure 3** and **Figure 4** respectively. Parameter estimation is done in Section 4. Section 5 is devoted to a summary of a variety of other models.

2. Standard Uniform Geometric (SUG) Model

Since $f(x)=1$, $0 \le x \le 1$, and $\mathbb{P}(N = n) = \theta(1-\theta)^{n-1}$ $(n = 1, 2, \cdots)$ for some $\theta \in (0,1)$, we have from (3) that the p.d.f. of Y in the SUG model is given by

$$\begin{aligned} g(y) &= \sum_{n=1}^{\infty} \theta(1-\theta)^{n-1} \times ny^{n-1} \\ &= \frac{\theta}{[1-(1-\theta)y]^2}. \end{aligned} \tag{5}$$

Similarly, (4) gives that

$$\begin{aligned} g(z) &= \sum_{n=1}^{\infty} \theta(1-\theta)^{n-1} \times n(1-z)^{n-1} \\ &= \frac{\theta}{[1-(1-\theta)(1-z)]^2} \\ &= \frac{\theta}{[\theta+(1-\theta)z]^2}. \end{aligned} \tag{6}$$

Proposition 2.1. *If the random variable Y has the "SUG maximum distribution" (5) and $k \in \mathbb{N}$, then*

$$\mathbb{E}(Y^k) = \frac{\theta}{(1-\theta)^{k+1}} \sum_{j=0}^{k} \binom{k}{j} \int_{\theta}^{1} (-u)^{j-2}\, \mathrm{d}u.$$

Proof.

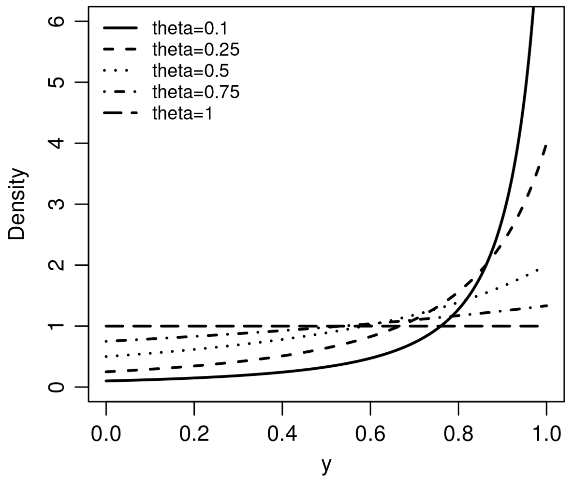

Figure 1. Plot of the SUG maximum density for some values of θ (see Equation (5)).

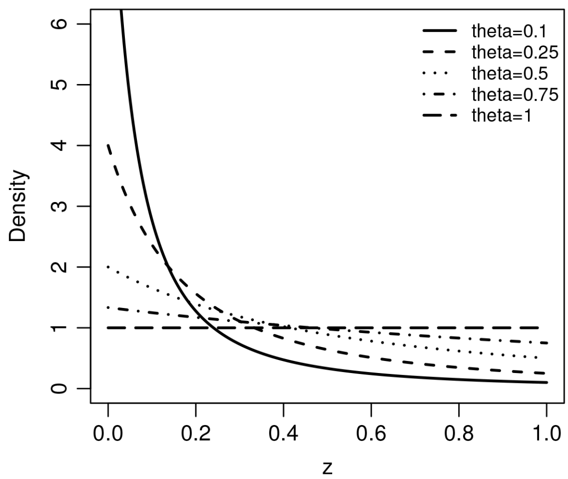

Figure 2. Plot of the SUG minimum density for some values of θ (see Equation (6)).

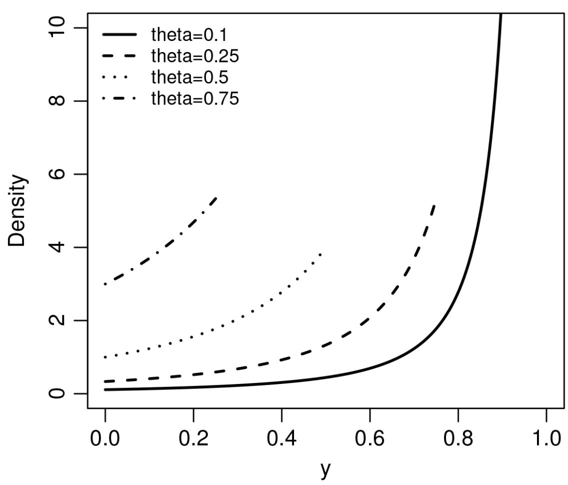

Figure 3. Plot of the CSUG maximum density for some values of θ.

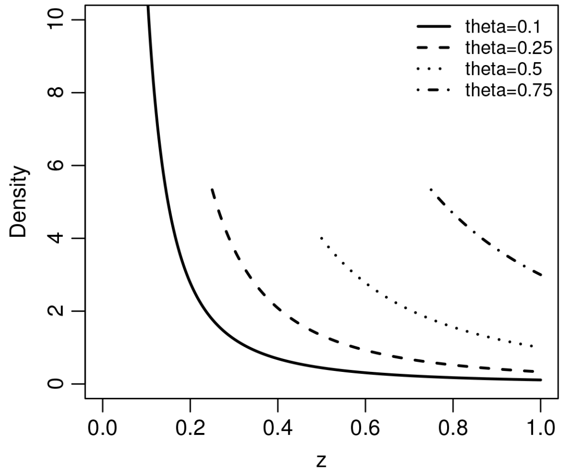

Figure 4. Plot of CSUG minimum density for some values of θ.

$$\begin{aligned}\mathbb{E}\left(Y^k\right) &= \int_0^1 y^k \times \frac{\theta}{\left[1-(1-\theta)y\right]^2}\mathrm{d}y \\ &= \int_1^\theta \left(\frac{1-u}{1-\theta}\right)^k \times \frac{\theta}{u^2} \times \left(-\frac{1}{1-\theta}\right)\mathrm{d}u \\ &= \frac{\theta}{(1-\theta)^{k+1}}\int_\theta^1 \frac{(1-u)^k}{u^2}\mathrm{d}u \\ &= \frac{\theta}{(1-\theta)^{k+1}}\int_\theta^1 \frac{\sum_{j=0}^{k}\binom{k}{j}(-u)^j}{u^2}\mathrm{d}u \\ &= \frac{\theta}{(1-\theta)^{k+1}}\sum_{j=0}^{k}\binom{k}{j}\int_\theta^1 (-u)^{j-2}\,\mathrm{d}u,\end{aligned}$$

as claimed. □

Note. Even though we take the distributions to have support on $[0,1]$, this may be done by changing the survival function in [3], where the same compounding method is used. Specifically we can use the transformation $z = \exp(-\lambda y)$ in the proofs of [3].

Proposition 2.2. *The random variable Y has mean and variance given, respectively, by*

$$\mathbb{E}(Y) = \frac{\theta\left(\ln\theta + \frac{1}{\theta} - 1\right)}{(1-\theta)^2} \quad \textit{and} \quad \mathbb{V}(Y) = \frac{\theta^3 - 2\theta^2 - \theta^2\ln^2\theta + \theta}{(1-\theta)^4}.$$

Proof. Using Proposition 2.1, we can directly compute the mean and variance by setting $k = 1, 2$, and using the fact that $\mathbb{V}(W) = \mathbb{E}(W^2) - \left[\mathbb{E}(W)\right]^2$ for any random variable *W*. (This proof could equally well have been based on calculating the moments of $1-(1-\theta)Y$ and then recovering the values of $\mathbb{E}(W)$ and $\mathbb{V}(Y)$. The same is true of other proofs in the paper.) □

Proposition 2.3. *If the random variable Z has the "SUG minimum distribution" and* $k \in \mathbb{N}$, *then*

$$\mathbb{E}\left(Z^k\right) = \frac{\theta}{(1-\theta)^{k+1}}\sum_{j=0}^{k}\binom{k}{j}(-\theta)^j\int_\theta^1 u^{k-j-2}\mathrm{d}u.$$

Proof.

$$\begin{aligned}\mathbb{E}\left(Z^k\right) &= \int_0^1 z^k \times \frac{\theta}{\left[\theta+(1-\theta)z\right]^2}\mathrm{d}z \\ &= \int_\theta^1 \left(\frac{u-\theta}{1-\theta}\right)^k \times \frac{\theta}{u^2} \times \frac{1}{1-\theta}\mathrm{d}u \\ &= \frac{\theta}{(1-\theta)^{k+1}}\int_\theta^1 \frac{(u-\theta)^k}{u^2}\mathrm{d}u \\ &= \frac{\theta}{(1-\theta)^{k+1}}\int_\theta^1 \frac{\sum_{j=0}^{k}\binom{k}{j}u^{k-j}(-\theta)^j}{u^2}\mathrm{d}u \\ &= \frac{\theta}{(1-\theta)^{k+1}}\sum_{j=0}^{k}\binom{k}{j}(-\theta)^j\int_\theta^1 u^{k-j-2}\mathrm{d}u,\end{aligned}$$

as asserted. □

Proposition 2.4. *The random variable Z has mean and variance given, respectively, by*

$$\mathbb{E}(Z) = \frac{\theta(\theta-1-\ln\theta)}{(1-\theta)^2} \text{ and } \mathbb{V}(Z) = \frac{\theta^3-2\theta^2-\theta^2\ln^2\theta+\theta}{(1-\theta)^4}.$$

Proof. Using Proposition 2.3, it is easily to compute the mean and variance by setting $k = 1$, $k = 2$. □

The m.g.f.'s of Y, Z are easy to calculate too. Notice that the logarithmic terms above arise due to the contributions of the $j = 1$ and $j = k-1$ terms, and it is precisely these logarithmic terms that make, e.g., method of moments estimates for θ to be intractable in a closed (*i.e.*, non-numerical) form. Similar difficulties arise when analyzing the likelihood function and likelihood ratios.

3. The Correlated Standard Uniform Geometric (CSUG) Model

The Correlated Standard Uniform Geometric (CSUG) model is related to the SUG model, as the name suggests, but X and N are correlated as indicated in Section 1. The CSUG problems arise in two cases. One case is that we conduct standard uniform trials until a variable X_i exceeds $1-\theta$, where θ is the parameter of the correlated geometric variable, and the maximum of $X_1, X_2, \cdots, X_{i-1}$ is what we seek. The maximum is between 0 and $1-\theta$. The other case is where standard uniform trials are conducted until X_i is less than θ, and we are looking for the minimum of $X_1, X_2, \cdots, X_{i-1}$. The minimum is between θ and 1.

Specifically, let $X_1, X_2, \cdots$ be a sequence of standard uniform variables and define

$$N = \inf\left\{n \ge 2 : X_i > 1-\theta\right\},$$

or

$$N = \inf\left\{n \ge 2 : X_i < \theta\right\}.$$

In either case N has probability mass function given by

$$\mathbb{P}(N = n) = \theta(1-\theta)^{n-2}, 0 < \theta < 1, n = 2,3,\cdots; \tag{7}$$

note that this is simply a geometric random variable conditional on the success having occurred at trial 2 or later. Clearly N is dependent on the X sequence.

Proposition 3.1. *Under the CSUG model, the p.d.f. of Y, defined by* (1), *is given by*

$$g(y) = \frac{\theta}{(1-\theta)(1-y)^2}, 0 \le y \le 1-\theta.$$

Proof. The conditional c.d.f. of Y given that $N = n$ is given by

$$\mathbb{P}(Y \le y | N = n) = \left(\frac{y}{1-\theta}\right)^{n-1}, n = 2,3,\cdots.$$

Taking the derivative, we see that the conditional density function is given by

$$g\left(y \middle| N=n\right)=\frac{n-1}{1-\theta}\left(\frac{y}{1-\theta}\right)^{n-2}, n=2,3,\cdots$$

Consequently, the p.d.f. of Y in the CSUG model is given by

$$\begin{aligned} g(y) &= \sum_{n=2}^{\infty}\theta(1-\theta)^{n-2}\times\frac{n-1}{1-\theta}\left(\frac{y}{1-\theta}\right)^{n-2} \\ &= \frac{\theta}{1-\theta}\times\sum_{n=2}^{\infty}(n-1)y^{n-2} \\ &= \frac{\theta}{(1-\theta)(1-y)^2}. \end{aligned}$$

This completes the proof. □

Proposition 3.2. *The p.d.f. of Z under the CSUG model is given by*

$$g(z)=\frac{\theta}{(1-\theta)z^2}, \theta\le z\le 1.$$

Proof. The conditional cumulative distribution function of Z given that $N=n$ is given by

$$\mathbb{P}\left(Z\le z \middle| N=n\right)=1-\mathbb{P}\left(Z>z \middle| N=n\right)=1-\left(\frac{1-z}{1-\theta}\right)^{n-1}, n=2,3,\cdots$$

Thus, the conditional density function is given by

$$g\left(z \middle| N=n\right)=\frac{n-1}{1-\theta}\left(\frac{1-z}{1-\theta}\right)^{n-2}, n=2,3,\cdots,$$

which yields the p.d.f. of Z under the CSUG model as

$$\begin{aligned} g(z) &= \sum_{n=2}^{\infty}\theta(1-\theta)^{n-2}\times\frac{n-1}{1-\theta}\left(\frac{1-z}{1-\theta}\right)^{n-2} \\ &= \frac{\theta}{1-\theta}\times\sum_{n=2}^{\infty}(n-1)(1-z)^{n-2} \\ &= \frac{\theta}{(1-\theta)z^2}, \end{aligned}$$

which finishes the proof. □

Proposition 3.3. *If the random variable Y has the "CSUG maximum distribution" and* $k\in\mathbb{N}$, *then*

$$\mathbb{E}\left(Y^k\right)=\frac{\theta}{1-\theta}\sum_{j=0}^{k}\binom{k}{j}\int_{\theta}^{1}(-u)^{j-2}\,\mathrm{d}u.$$

Proof.

$$\begin{aligned} \mathbb{E}\left(Y^k\right) &= \int_0^{1-\theta} y^k\times\frac{\theta}{(1-\theta)(1-y)^2}\,\mathrm{d}y \\ &= \frac{\theta}{1-\theta}\int_1^{\theta}\frac{(1-u)^k}{u^2}(-\mathrm{d}u) \\ &= \frac{\theta}{1-\theta}\int_{\theta}^{1}\frac{\sum_{j=0}^{k}\binom{k}{j}(-u)^j}{u^2}\,\mathrm{d}u \\ &= \frac{\theta}{1-\theta}\sum_{j=0}^{k}\binom{k}{j}\int_{\theta}^{1}(-u)^{j-2}\,\mathrm{d}u, \end{aligned}$$

as claimed. □

Proposition 3.4. *The random variable Y has mean and variance given, respectively, by*

$$\mathbb{E}(Y)=\frac{\theta\ln\theta-\theta+1}{1-\theta}$$

and

$$\mathbb{V}(Y)=\frac{\theta^3-2\theta^2-\theta^2\ln^2\theta+\theta}{(1-\theta)^2}.$$

Proof. Using Proposition 3.3, we can directly compute the mean and variance by setting $k = 1, 2$. For example with $k = 1$ we get

$$\mathbb{E}(Y)=\frac{\theta}{1-\theta}\sum_{j=0}^{1}\int_\theta^1(-u)^{j-2}\,\mathrm{d}u=\frac{\theta}{1-\theta}\int_\theta^1\frac{1}{u^2}-\frac{1}{u}\mathrm{d}u=\frac{\theta\ln\theta-\theta+1}{1-\theta}.$$

Notice that the variance of Y is smaller than that of Y under the SUG model, with an identical numerator term. Also, the expected value is smaller under the CSUG model than in the SUG case. This can be best seen by the inequalities

$$\frac{\theta^3-2\theta^2-\theta^2\ln^2\theta+\theta}{(1-\theta)^2}<\frac{\theta^3-2\theta^2-\theta^2\ln^2\theta+\theta}{(1-\theta)^4}$$

and

$$\frac{\theta\ln\theta-\theta+1}{1-\theta}<\frac{\theta\left(\ln\theta+\frac{1}{\theta}-1\right)}{(1-\theta)^2},$$

valid for $0<\theta<1$. □

Proposition 3.5. *If the random variable Z has the "CSUG Minimum distribution" and* $k\in\mathbb{N}$, *then*

$$\mathbb{E}\left(Z^k\right)=\frac{\theta}{1-\theta}\int_\theta^1 z^{k-2}\mathrm{d}z.$$

Proof. Routine, as before. □

Proposition 3.6. *The random variable Z has mean and variance given, respectively, by*

$$\mathbb{E}(Z)=\frac{-\theta\ln\theta}{1-\theta}$$

and

$$\mathbb{V}(Z)=\frac{\theta^3-2\theta^2-\theta^2\ln^2\theta+\theta}{(1-\theta)^2}.$$

Proof. A special case of Proposition 3.3; note that as in the SUG model, $\mathbb{V}(Y)=\mathbb{V}(Z)$. □

Remark 1. The four distributions of Y and Z under the SUG and the CSUG models can be shown to be affine transformations of the same distribution as seem by the following results (proofs omitted):

Proposition 3.7. *Changing the variable Y of* (5) *as* $Z=1-Y$ *yields* (6). *Thus the SUG maximum and SUG minimum variables are related by the fact that*

$$Y\sim\mathrm{SUG}(\mathrm{Max})\Rightarrow 1-Y\sim\mathrm{SUG}(\mathrm{Min}).$$

Proposition 3.8. *Changing the variable Y of the CSUG model* (*in Proposition* 3.1) *as* $W=Y/(1-\theta)$ *yields* $g(w)=\theta/(1-(1-\theta)w)^2$, *which equals the pdf of* (5). *Hence*

$$Y\sim\mathrm{CSUG}(\mathrm{Max})\Rightarrow Y/(1-\theta)\sim\mathrm{SUG}(\mathrm{Max}).$$

Proposition 3.9. *Changing the variable Z of the CSUG model* (*in Proposition* 3.2) *as* $T=(Z-\theta)/(1-\theta)$

yields $g(t)=\theta/((1-\theta)t+\theta)^2$ *, which equals the pdf of* (6). *Thus*

$$Y\sim \text{CSUG}(\text{Min})\Rightarrow (Z-\theta)/(1-\theta)\sim \text{SUG}(\text{Min})$$

As a result of these affine transformations, the moment equations (Propositions from 2.1 to 2.4 and from 3.3 to 3.6) can be derived in an easier fashion, though these facts are easier to observe *post facto*.

Remark 2. As stated earlier the distributions of this paper are related to other distributions in the literature, but these do not exploit the extreme value connection as we do. For example, when $\theta=1/2$, (5) reduces to

$$g(y)=\frac{2}{(1+(1-y))^2},0\le y\le 1,$$

which is a special case, with $k = 1$, of the generalized half-logistic distribution [5], eq. 23.83.

Second, the distribution of Z under the CSUG model is a special case of a truncated Pareto distribution, which, for positive a, is defined by

$$f(x)=\frac{a\theta^a x^{-a-1}}{1-(\theta/\xi)^a},\theta\le x\le \xi.$$

Putting $a=1$ and $\xi=1$, we obtain the pdf of Proposition 3.2. This special case of $f(x)=K\cdot x^{-2}$ appears in the 2nd type of Zipf's Law; see Urzúa [13]. The truncated Pareto distribution appears, e.g., in Aban *et al.* [14] and the references therein.

4. Parameter Estimation

The intermingling of polynomial and logarithmic terms makes method of moments estimation difficult in closed form, as in the SUG case. However, if θ is unknown, the maximum likelihood estimate of θ can be found in a satisfying form, both in the CGUG maximum and CSUG minimum cases. Suppose that $Y_1,Y_2,\cdots,Y_n$ form a random sample from the CSUG Maximum distribution with unknown θ. Since the pdf of each observation has the following form:

$$f(y|\theta)=\begin{cases}\dfrac{\theta}{(1-\theta)(1-y)^2}, & \text{for } 0\le y\le 1-\theta\\ 0, & \text{otherwise}\end{cases}$$

the likelihood function is given by

$$\ell(\theta)=\begin{cases}\left(\dfrac{\theta}{1-\theta}\right)^n\dfrac{1}{\prod_{i=1}^{n}(1-y_i)^2}, & \text{for } 0\le y_i\le 1-\theta\ \ (i=1,2,\cdots,n)\\ 0, & \text{otherwise}\end{cases}$$

The MLE of θ is a value of θ, where $\theta\le 1-y_i$ for $i=1,2,\cdots,n$, which maximizes $\dfrac{\theta}{1-\theta}$. Let $\varphi(\theta)=\dfrac{\theta}{1-\theta}$. Since $\varphi'(\theta)\ge 0$, it follows that $\varphi(\theta)$ is a increasing function, which means the MLE is the largest possible value of θ such that $\theta\le 1-y_i$ for $i=1,2,\cdots,n$. Thus, this value should be $1-\max(Y_1,\cdots,Y_n)$, *i.e.*, $\hat{\theta}=1-Y_{(n)}$.

Suppose next that $Z_1,Z_2,\cdots,Z_n$ form a random sample from the CSUG minimum distribution. Since the pdf of each observation has the following form:

$$f(z|\theta)=\begin{cases}\dfrac{\theta}{(1-\theta)z^2}, & \text{for } \theta\le z\le 1\\ 0 & \text{otherwise,}\end{cases}$$

it follows that the likelihood function is given by

$$\ell(\theta)=\begin{cases}\left(\frac{\theta}{1-\theta}\right)^n \frac{1}{\prod_{i=1}^n z_i^2}, & \text{for } \theta \le y_i \le 1 \; (i=1,2,\cdots,n) \\ 0 & \text{otherwise.}\end{cases}$$

As above, it now follows that $\hat{\theta}=Y_{(1)}$. It is not too hard to write down the distribution of the MLE's but we do not do so here.

5. A Summary of Some Other Models

The general scheme given by (3) and (4) is quite powerful. As another example, suppose (using the example from Section 1) that

$$p(n)=\frac{6}{\pi^2}\frac{1}{n^2}$$

and $X \sim U[0,1]$. Then it is easy to show that

$$g(y)=\frac{6}{\pi^2}\frac{1}{y}\ln\left(\frac{1}{1-y}\right), 0 \le y \le 1,$$

and that $\mathbb{E}(Y)=\frac{6}{\pi^2}$. (The expected value of Y can also be calculated by using the identity $\mathbb{E}(Y)=\mathbb{E}(\mathbb{E}(Y|N))$. In this section, we collect some more results of this type, without proof:

UNIFORM-POISSON MODEL. Here we let $X \sim U[0,1]$ and $p(n)=\frac{e^{-\lambda}\lambda^n}{(1-e^{-\lambda})n!}, n=1,2,\cdots$, so that N follows a left-truncated Poisson distribution.

Proposition 5.1. *Under the Uniform-Poisson model*,

$$g(y)=\frac{\lambda e^{-\lambda}e^{\lambda y}}{1-e^{-\lambda}}; g(z)=\frac{\lambda e^{-\lambda z}}{1-e^{-\lambda}};$$

$$\mathbb{E}(Y)=\frac{1}{1-e^{-\lambda}}-\frac{1}{\lambda}; \mathbb{E}(Z)=\frac{1}{\lambda}-\frac{e^{-\lambda}}{1-e^{-\lambda}};$$

$$\mathbb{V}(Y)=\frac{1}{\lambda^2}+\frac{1}{1-e^{-\lambda}}-\frac{1}{(1-e^{-\lambda})^2}; \mathbb{V}(Z)=\frac{1}{\lambda^2}-\frac{e^{-\lambda}}{\lambda(1-e^{-\lambda})}-\frac{e^{-2\lambda}}{(1-e^{-\lambda})^2};$$

$$M_Y(t)=\mathbb{E}(e^{ty})=\frac{\lambda e^{-\lambda}(e^{t+\lambda}-1)}{(t+\lambda)(1-e^{-\lambda})}; M_Z(t)=\mathbb{E}(e^{tz})=\frac{\lambda(e^{t-\lambda}-1)}{(t-\lambda)(1-e^{-\lambda})}.$$

In some sense, the primary motivation of this paper was to produce extreme value distributions that did not fall into the Beta family (such as $f(y)=nt^{n-1}$ for the maximum of n i.i.d. $U[0,1]$ variables). A wide variety of non-Beta-based distributions may be found in [6]. Can we add extreme value distributions to that collection? In what follows, we use both the Beta families $B(2,2)$ and $B(1/2,1/2)$, the arcsine distribution, and a "Beyond Beta" distribution, the Topp-Leone distribution [15], as "input variables" to make further progress in this direction.

GEOMETRIC-BETA(2, 2) MODEL. Here $X \sim B(2,2)$ and $N \sim \text{Geo}(\theta)$. In this case we get

$$g(y)=\frac{6y(1-y)\theta}{\left[1-(1-\theta)y^2(3-2y)\right]^2}$$

and

$$g(z)=\frac{6z(1-z)\theta}{\left[1-(1-\theta)\left(2z^3-3z^2+1\right)\right]^2}.$$

POISSON-BETA(2, 2) MODEL. Here $X\sim B(2,2)$ and $N\sim \mathrm{Po}_0(\theta)$, the Poisson ($\theta$) distribution left-truncated at 0. In this case we get

$$g(y)=\frac{6\theta y(1-y)\mathrm{e}^{-\theta\left(2y^3-3y^2+1\right)}}{1-\mathrm{e}^{-\theta}}$$

and

$$g(z)=\frac{6\theta z(1-z)\mathrm{e}^{-\theta\left(3z^2-2z^3\right)}}{1-\mathrm{e}^{-\theta}}.$$

GEOMETRIC-ARCSINE MODEL. Here $X\sim B(1/2,1/2)$ and $N\sim \mathrm{Geo}(\theta)$. In this case we get

$$g(y)=\frac{\theta\pi^{-1}\left[y(1-y)\right]^{-1/2}}{\left[1-(1-\theta)\frac{2}{\pi}\arcsin\sqrt{y}\right]^2}$$

and

$$g(z)=\frac{\theta\pi^{-1}\left[z(1-z)\right]^{-1/2}}{\left[1-(1-\theta)\left(1-\frac{2}{\pi}\arcsin\sqrt{z}\right)\right]^2}.$$

POISSON-ARCSINE MODEL. Here $X\sim B(1/2,1/2)$ and $N\sim \mathrm{Po}_0(\theta)$. Here we have

$$g(y)=\frac{\theta\pi^{-1}\left[y(1-y)\right]^{-1/2}\mathrm{e}^{-\theta\left(1-\frac{2}{\pi}\arcsin\sqrt{y}\right)}}{1-\mathrm{e}^{-\theta}}$$

and

$$g(z)=\frac{\theta\pi^{-1}\left[z(1-z)\right]^{-1/2}\mathrm{e}^{-\frac{2\theta\arcsin\sqrt{z}}{\pi}}}{1-\mathrm{e}^{-\theta}}.$$

GEOMETRIC-TOPP-LEONE MODEL. Here $X\sim TL(a)$ and $N\sim \mathrm{Geo}(\theta)$:

$$g(y)=\frac{2a(1-y)y^{a-1}(2-y)^{a-1}\theta}{\left[1-(1-\theta)y^a(2-y)^a\right]^2}$$

and

$$g(z)=\frac{2a(1-z)z^{a-1}(2-z)^{a-1}\theta}{\left\{1-(1-\theta)\left[1-z^a(2-z)^a\right]\right\}^2}.$$

POISSON-TOPP-LEONE MODEL. $X\sim TL(a)$ and $N\sim \mathrm{Po}_0(\theta)$:

$$g(y)=\frac{2\theta a(1-y)y^{a-1}(2-y)^{a-1}\mathrm{e}^{-\theta\left[1-y^a(2-y)^a\right]}}{1-\mathrm{e}^{-\theta}}$$

and

$$g(z)=\frac{2\theta a(1-z)z^{a-1}(2-z)^{a-1}\mathrm{e}^{-\theta\left[z^a(2-z)^a\right]}}{1-\mathrm{e}^{-\theta}}.$$

6. Conclusion

In this paper we studied a general scheme for the distribution of the maximum or minimum of a random number of i.i.d. random variables with compact support. While some of the distributions obtained through this process have appeared before in the literature, they do not been studied using this approach. Our biggest open problem is to find data sets for which these new distributions are appropriate.

Acknowledgements

The research of AG was supported by NSF Grants 1004624 and 1040928. We thank the referees for their insightful suggestions for improvement.

References

[1] Durrett, R. (1991) Probability: Theory and Examples. Wadsworth and Brooks/Cole, Pacific Grove. http://dx.doi.org/10.4236/am.2012.34054

[2] Louzada, F., Roman, M. and Cancho, V. (2011) The Complementary Exponential Geometric Distribution: Model, Properties, and a Comparison with Its Counterpart. *Computational Statistics and Data Analysis*, **55**, 2516-2524. http://dx.doi.org/10.1016/j.csda.2010.09.030

[3] Louzada, F., Bereta, E. and Franco, M. (2012) On the Distribution of the Minimum or Maximum of a Random Number of i.i.d. Lifetime Random Variables. *Applied Mathematics*, **3**, 350-353.

[4] Morais, A. and Barreto-Souza, W. (2011) A Compound Class of Weibull and Power Series Distributions. *Computational Statistics and Data Analysis*, **55**, 1410-1425.

[5] Johnson, N., Kotz, S. and Balakrishnan, N. (1995) Continuous Univariate Distributions. Vol. 2, Wiley, New York.

[6] Kotz, S. and van Dorp, J. (2004) Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications. World Scientific Publishing Co., Singapore. http://dx.doi.org/10.1142/5720

[7] Leadbetter, M., Lindgren, G. and Rootzén, H. (1983) Extremes and Related Properties of Random Sequences and Processes. Springer Verlag, New York. http://dx.doi.org/10.1007/978-1-4612-5449-2

[8] Earthquake Data Set. http://www.data.scec.org/eq-catalogs/date_mag_loc.php

[9] Chung, F., Lu, L. and Vu, V. (2003) Eigenvalues of Random Power Law Graphs. *Annals of Combinatorics*, **7**, 21-33. http://dx.doi.org/10.1007/s000260300002

[10] Aiello, B., Chung, F. and Lu, L. (2001) A Random Graph Model for Power Law Graphs. *Experimental Mathematics*, **10**, 53-66. http://dx.doi.org/10.1080/10586458.2001.10504428

[11] Marshall, A. and Olkin, I. (1997) A New Method of Adding a Parameter to a Family of Distributions with Applications to the Exponential and Weibull Families. *Biometrika*, **84**, 641-652. http://dx.doi.org/10.1093/biomet/84.3.641

[12] Kumaraswamy, P. (1980) A Generalized Probability Density Function for Double-Bounded Random Processes. *Journal of Hydrology*, **46**, 79-88. http://dx.doi.org/10.1016/0022-1694(80)90036-0

[13] Urzúa, C.M. (2000) A Simple and Efficient Test for Zipf's Law. *Economics Letters*, **66**, 257-260. http://dx.doi.org/10.1016/S0165-1765(99)00215-3

[14] Aban, I.B., Meerschaert, M.M. and Panorska, A.K. (2006) Parameter Estimation for the Truncated Pareto Distribution. *Journal of the American Statistical Association*, **101**, 270-277. http://dx.doi.org/10.1198/016214505000000411

[15] Topp, C.W. and Leone, F.C. (1955) A Family of J-Shaped Frequency Functions. *Journal of the American Statistical Association*, **50**, 209-219. http://dx.doi.org/10.1080/01621459.1955.10501259