PROCEEDINGS OF THE YEREVAN STATE UNIVERSITY

Physical and Mathematical Sciences 2012, № 2, p. 29–34

Mathematics

RANDOM CONTRACTION SCHEMES FOR EXTREMAL ORDER STATISTICS

V. K. SAGHATELYAN*

Chair of Actuarial Mathematics YSU, Armenia

In this paper a few new types of random contraction and dilation schemes are considered. In particular, four one-sided additive schemes and two two-sided additive schemes are represented. All schemes has been built for sample extremes and exponential random variables. Characterizations of distributions for all those schemes are obtained.

***Keywords*:** order statistics, sample extreme, random contraction, random dilation, characterization of distribution, two-sided contraction.

Introduction. Equality by distribution of the form $X \overset{d}{=} YU$, where U is a random variable, whose distribution is concentrated on $(0,1)$ (mostly power distribution or standard uniform), X and Y are arbitrary random variables, are called random contraction scheme. On the other hand, equality of the form $X \overset{d}{=} YW$, where W is a random variable, whose distribution is concentrated on $(1,\infty)$, is called random dilation scheme. When considering such schemes a question arises to describe those distributions under which the above equations holds. One can use such schemes to switch between differently distributed random variables. The discussion on this topic can be found in [1–4]. During the past few years similar equations have been considered for order statistics (see [5–7]) and mathematical records (see [8]).

In this paper we will discuss several types of special random contraction and dilation schemes constructed for extremal order statistics, in particular for sample extremes. We will also introduce two-sided schemes, where we have contraction on the one side and dilation on the other side.

Characterizations of Distributions Via Random Dilation Schemes. Consider a sequence of independent, identically distributed (i.i.d.) random variables $Y, Y_1, Y_2, \ldots$ with common distribution function (d.f.) $F(x)$. Let W be another random variable independent of $Y, Y_1, Y_2, \ldots$ with standard exponential distribution $F_W(x) = \max\{0, 1-e^{-x}\}$. For some fixed $k = 1,2,\ldots$ and $n = 2,3,\ldots$ let us take

* E-mail: v.saghatelyan@ysu.am

sample minimums $Y_{1,k}=\min\{Y_1,...,Y_k\}$, $Y_{1,n}=\min\{Y_1,...,Y_n\}$. Without loss of generality we can assume that $k<n$ and, hence, $Y_{1,n}<Y_{1,k}$.

Consider the following dilation scheme for introduced quantities

$$Y_{1,k} \overset{d}{=} Y_{1,n}+W\ . \tag{1}$$

Theorem 1. For some fixed $1\le k<n$ equality by distribution (1) holds, iff

$$F(x)=1-\left[\left(e^{C-x}\right)^{\frac{k-n}{k}}+1\right]^{\frac{1}{k-n}},\ -\infty<x<\infty\ , \tag{2}$$

where C is an arbitrary constant.

Remark. We define the inverse function of d.f. $F(x)$ as $G(s)=\inf\{x:F(x)\ge s\}$, so, the continuity of $F(x)$ is enough for existence of the inverse function.

Proof of Theorem 1. Using the principle of convolution, we can rewrite (1) as

$$F_{1,k}(x)=\int_0^\infty F_{1,n}(x-u)f_W(u)du\ , \tag{3}$$

where $F_{1,k}$ and $F_{1,n}$ are d.f.-s of $Y_{1,k}$ and $Y_{1,n}$ respectively, and f_W is the probability density function of W. Making change of variable $x-u=t$ in (3), we get

$$F_{1,k}(x)=e^{-x}\int_{-\infty}^{x} F_{1,n}(t)e^t dt\ . \tag{4}$$

Taking into account that $F_{1,k}(x)=1-(1-F(x))^k$ and $F_{1,n}(x)=1-(1-F(x))^n$ (see, for example [9]), we can rewrite (4) as follow

$$\left[1-(1-F(x))^k\right]e^x=\int_{-\infty}^{x}\left[1-(1-F(t))^n\right]e^t dt\ . \tag{5}$$

Note, that equality (5) and continuity of $F(x)$ ensures the differentiability of $F(x)$. So, differentiating both sides of (5), we obtain the following equality

$$e^x k(1-F(x))^{k-1}F'(x)+e^x\left(1-(1-F(x))^k\right)=\left(1-(1-F(x))^n\right)e^x,$$

consequently,

$$kF'(x)=\left(1-F(x)\right)\left(1-(1-F(x))^{n-k}\right). \tag{6}$$

As $F(G(x))=x$, so, $F'(G(x))=\dfrac{1}{G'(x)}$. Hence, substituting $G(x)$ instead of x in (6), we get $G'(x)=\dfrac{k}{(1-x)\left(1-(1-x)^{n-k}\right)}$, $0<x<1$. Integration of both sides of the last equality yields

$$G(x)=-k\int\frac{1}{(1-x)\left(1-(1-x)^{n-k}\right)}d(1-x)=-k\left(\ln(1-x)+\frac{1}{k-n}\ln\left(1-(1-x)^{n-k}\right)\right)+C,$$

where C is an arbitrary constant. Thus for function $G(x)$, $0<x<1$, we obtain

$$G(x) = \ln \frac{(1-x)^{-k} e^C}{(1-(1-x)^{n-k})^{\frac{k}{k-n}}}. \tag{7}$$

Note that since $k-n<0$, from (7) it follows that $G(0+) = -\infty$, $G(1-) = \infty$, viz the support of initial d.f. $F(x)$ is $(-\infty, \infty)$. Returning to d.f. $F(x)$, we can rewrite (7) as $x = \ln \frac{(1-F(x))^{-k} e^C}{(1-(1-F(x))^{n-k})^{\frac{k}{k-n}}}$, thus $\frac{(1-(1-F(x))^{n-k})^{\frac{k}{k-n}}}{(1-F(x))^{-k}} = e^{C-x}$ and finally $(1-F(x))^{k-n} = (e^{C-x})^{\frac{k-n}{k}} + 1 = C_1 e^{-\frac{k-n}{k}x} + 1$, where C_1 is an arbitrary positive constant. So, we have proved the first part of the Theorem. To prove the second part, one can just substitute d.f. (2) into the equality for distributions (3).

Theorem 1 provides a similar result for sample maximums. Consider a new sequence of random variables, constructed in the following manner $X = -Y,\ X_1 = -Y_1,\ X_2 = -Y_2, \ldots$. It is obvious, that $X,\ X_1,\ X_2, \ldots$ will be i.i.d. with common continuous d.f. $H(x) = 1 - F(-x)$, $-\infty < x < \infty$. Also, one can easily obtain that for some $m = 1, 2, \ldots$ $X_{m,m}$ and $-Y_{1,m}$ have the same distribution. Indeed, $P\{X_{m,m} < x\} = H^m(x) = (1-F(-x))^m$ (see, for example, [9]). On the other hand, $P\{-Y_{1,m} < x\} = 1 - P\{Y_{1,m} < -x\} = (1-F(-x))^m$. Further, by changing the sign in equation (1), we get $-Y_{1,k} \overset{d}{=} -Y_{1,n} - W$ which is the same as $X_{k,k} \overset{d}{=} X_{n,n} - W$ which, in turn, is a random contraction scheme. Hence, we formulate the following corollary.

Corollary 1. Equality by distribution $X_{k,k} \overset{d}{=} X_{n,n} - W$ holds, iff $H(x) = ((\tilde{C} e^x)^{\frac{k-n}{k}} + 1)^{\frac{1}{k-n}}$ for all $-\infty < x < \infty$ with $k < n$ and arbitrary positive constant $\tilde{C}$.

To prove the sufficiency of Corollary 1, one just need to substitute (2) in the expression of d.f. $H(x)$.

Random Contraction Schemes. Here we keep all definitions and notations from the previous section. For the same sequence $Y, Y_1, Y_2, \ldots$ of i.i.d. random variables and W, consider an equality by distribution

$$Y_{1,k} - W \overset{d}{=} Y_{1,n}. \tag{8}$$

Theorem 2. The equality by distribution (8) for some fixed $1 \le k < n$ holds, iff the d.f. $F(x)$, $-\infty < x < C$, is given by

$$F(x) = 1 - \left[1 - (e^{x-C})^{\frac{n-k}{n}}\right]^{\frac{1}{n-k}}, \tag{9}$$

where C is some constant.

Proof. As $F_{1,n}(x) = 1 - (1-F(x))^n$ and $F_{1,k}(x) = 1 - (1-F(x))^k$, so, we can rewrite (8) in the following form $1 - (1-F(x))^n = \int_0^\infty \left[1 - (1-F(x+u))^k\right] f_W(u)\,du$, and, by denoting $x + u = t$, we get

$$e^{-x}\left[1-(1-F(x))^n\right]=\int\limits_x^\infty\left[1-(1-F(t))^k\right]e^{-t}dt. \qquad (10)$$

Since $F(x)$ is a differentiable function, we can differentiate both sides of (10) to obtain $\left[1-(1-F(x))^n\right]'=\left[1-(1-F(x))^n\right]-\left[1-(1-F(x))^k\right]$, from which we get $F'(x)=\dfrac{(1-F(x))^k-(1-F(x))^n}{n(1-F(x))^{n-1}}$. Switching to the inverse function $G(x)=F^{-1}(x)$ and taking into account that $F'(G(x))=\dfrac{1}{G'(x)}$, we obtain $G'(x)=\dfrac{n}{(1-x)((1-x)^{k-n}-1)}$ for $0<x<1$. Integrating both sides of last equation, we get following representation for inverse function

$$G(x)=\ln\left[1-(1-x)^{n-k}\right]^{\frac{n}{n-k}}+C, \qquad (11)$$

where C is an arbitrary constant. It is easily seen, that the equality (11) can be written in following manner

$$x=\ln\left[1-(1-F(x))^{n-k}\right]^{\frac{n}{n-k}}+C,$$

which can be easily transformed to (9).

Similarly to the previous scheme, we can formulate a corollary from Theorem 2 for random dilation scheme $X_{k,k}+W\overset{d}{=}X_{n,n}$.

Corollary 2. Equality by distribution $X_{k,k}+W\overset{d}{=}X_{n,n}$ holds for some fixed $1\le k<n$, iff $H(x)=\left[1-\left(e^{-x-C}\right)^{\frac{n-k}{n}}\right]^{\frac{1}{n-k}}$, $C<x<\infty$, where C is an arbitrary constant.

Some Generalizations of the Proven Results. First we want to take more general random variable instead of W. Let W_α be an exponential random variable, independent of $Y,Y_1,Y_2,\dots$ with a density function $f_{W_\alpha}(x)=\alpha e^{-\alpha x}$. Then the following is true.

Theorem 3. Equality by distribution $Y_{1,k}\overset{d}{=}Y_{1,n}+W_\alpha$ for some fixed $1\le k<n$ and $\alpha>0$ holds, iff the d.f. $F(x)$ is given by

$$F(x)=1-\left[(e^{\alpha(C-x)})^{\frac{k-n}{k}}+1\right]^{\frac{1}{k-n}},\ -\infty<x<\infty,$$

where C is some constant.

Proof of this Theorem is similar to the proof of Theorem 1, the only difference is the α parameter appearing in expressions. Moreover, we can similarly formulate analogous result for the scheme given in Theorem 2.

The second type of generalization is the consideration of two-sided random contraction/dilation scheme. Keeping the notations, for some positive α and β

consider the following equality $Y_{1,k} - W_\alpha \overset{d}{=} Y_{1,n} + W_\beta$. Rewriting this equality in terms of convolution of distributions, we get

$$\int_0^\infty F_{1,k}(x+u) f_{W_\alpha}(u)du = \int_0^\infty F_{1,n}(x-u) f_{W_\beta}(u)du .$$

Considering that $f_{W_\alpha}(x) = \alpha e^{-\alpha x}$, $f_{W_\beta}(x) = \beta e^{-\beta x}$ and making change of variable $x+u=t$ and $x-u=t$ respectively on the left and right sides of the last equality, we obtain $\dfrac{\alpha}{\beta} e^{(\alpha+\beta)x} \int_x^\infty F_{1,k}(t) e^{-\alpha t} dt = \int_{-\infty}^x F_{1,n}(t) e^{\beta t} dt$. Differentiation of both sides leads to

$$\frac{\alpha}{\beta} e^{(\alpha+\beta)x} (\alpha+\beta) \int_x^\infty F_{1,k}(t) e^{-\alpha t} dt - \frac{\alpha}{\beta} e^{(\alpha+\beta)x} F_{1,k}(x) e^{-\alpha x} = F_{1,n}(x) e^{\beta x} ,$$

and so we have

$$\int_x^\infty F_{1,k}(t) e^{-\alpha t} dt = \frac{\beta}{\alpha(\beta+\alpha)} F_{1,n}(x) e^{-\alpha x} + \frac{1}{\alpha+\beta} F_{1,k}(x) e^{-\alpha x} .$$

Differentiating this last equality again, we get

$$F'(x) = \frac{\beta(1-F(x))^k - \beta(1-F(x))^n}{\frac{\beta}{\alpha} n(1-F(x))^{n-1} + k(1-F(x))^{k-1}} .$$

Switching to the inverse function $G(x)$, we can rewrite this equality as $G'(x) = \dfrac{\frac{\beta}{\alpha} n(1-x)^{n-k} + k}{\beta(1-x) - \beta(1-x)^{n-k+1}}$, and by taking the antiderivative on both sides, we obtain $G(x) = \displaystyle\int \frac{\frac{\beta}{\alpha} n(1-x)^{n-k} + k}{\beta(1-x) - \beta(1-x)^{n-k+1}} dx$. Computing this integral, we obtain

$$G(x) = \ln \frac{\left(1-(1-x)^{n-k}\right)^{\frac{n\beta+k\alpha}{\alpha\beta(n-k)}} C_1}{(1-x)^{\frac{k}{\beta}}} , \qquad (12)$$

where C_1 is an arbitrary positive constant.

It should be noted, that initial d.f. $F(x)$ was supposed to be continuous, so, it can be uniquely given by its inverse. Thus, we actually proved the following theorem.

Theorem 4. Equality by distribution $Y_{1,k} - W_\alpha \overset{d}{=} Y_{1,n} + W_\beta$ for some fixed $1 \le k < n$ and $\alpha, \beta > 0$ holds, iff the d.f. $F(x)$ is uniquely defined through its inverse, which is given by (12).

For the two sided random contraction/dilation scheme we can also derive analogous result for sample maximums, as it had been done in Corollary 1. Consider the sequence X, X_1, X_2,... from the previous section. The d.f. of these random variables $H(x)=1-F(-x)$ is continuous, and hence it has uniquely defined inverse. It is obvious that the inverse function of $H(x)$ will be given by $R(x)=H^{-1}(x)=-G(1-x)$. Now we can formulate similar result for two sided random contraction/dilation scheme for sample maximums.

Corollary 3. Equality by distribution $X_{k,k}+W_\alpha \overset{d}{=} X_{n,n}-W_\beta$ for some fixed $1\le k<n$ and $\alpha,\beta>0$ holds, iff the d.f. $H(x)$ is uniquely defined through its inverse, which is given by $R(x)=\ln\left[x^{\frac{k}{\beta}}(1-x^{n-k})^{\frac{n\beta+k\alpha}{\alpha\beta(k-n)}}C_2\right]$, where C_2 is an arbitrary positive constant.

Received 10.02.2012

REFERENCES

1. **Pakes A.G.** // J. Statist. Plann. Infer., 1997, v. 63, p. 285–310.
2. **Pakes A.G., Khattree R.** // Austral. NZ J. Statist., 1992, v. 49, p. 307–322.
3. **Pakes A.G., Navarro J.** // Austral. NZ J. Statist., 2007, v. 49, p. 115–135.
4. **Yeo G.F., Milne R.K.** // Statist. Probab. Lett., 1991, p. 239–242.
5. **Beutner E., Kamps U.** // Commun. Statist. Theory Meth., 2008, v. 37, p. 2185–2201.
6. **Navarro J.** // Commun. Statist. Theory Meth., 2008, v. 37, p. 987–997.
7. **Wesolowsky J., Ahsanullah M.** // Austral. NZ J. Statist., 2004, v. 46, p. 297–303.
8. **Oncel S., Ahsanullah M., Aliev F., Aygun F.** // Statist. Probab. Lett., 2005, v. 73, p. 207–217.
9. **David H.** Order Statistics. Wiley-Interscience, 2003, p. 482.