Distribution of the Product of k Maxima from Uniform Ordered Samples

Akram Mohammed Al-Abood
Department of Mathematics and Computer Applications, College of Science,
Al-Nahrain University, Baghdad – Iraq.

Abstract

In this paper, we consider the distribution formulation of the productof*k* maxima of independent uniformly distributed samples over the interval (0,1), where the random variables of the *k* samples are arranged individually in ascending order of magnitude. Farthermore the distribution moments about origin are constructed.

Keyword: Order statistics, distribution theory, distribution of two maxima, generalization and moments.

1. Introduction

In recent years, order statistics come to play an important role in statistical inference theory partly because some of their properties does not depend on the distribution from which the random sample is obtained.

When a set of random variables (abbreviate as r.v's) $X_1, X_2, \ldots, X_n$ is arranged in increasing order of magnitude are represented by the r.v's$Y_1<Y_2<\ldots<Y_n$, where Y_i (i=1,2,…,n) is called the i^{th} order statistics. In most applications the unordered set $\{X_i\}$ represent a random sample drown from a single population in which they are statistically independent and identically distributed [1]. A survey of applications show that the most common use of order statistics are Y_1 and Y_n. For example $Y_1=\min(X_1, X_2, \ldots, X_n)$ is useful in statistical study of breaking strength of a chain which depends on the weakest line [2] and the life of multicomponent system where the system terminated as soon as the first failure occur [3], while $Y_n=\max(X_1, X_2, \ldots, X_n)$ is of interest in statistical study of floods and other meteorological phenomena [4].

2. Joint Distribution of k Sample Maximum

Consider that we have k independent samples each of size n are drown from a continuous distribution having probability density function (p.d.f) $f(x)$ and cumulative distribution function (c.d.f) $F(x)$. If the k maximum values of the samples are denoted by $Y_{n1}, Y_{n2}, \ldots, Y_{nk}$. Then from order statistics theory [5], the p.d.f of the maximum Y_{ni} for the i^{th} sample (i=1,2,…,n) is

$$g_i(y_{ni}) = n[F(y_{ni})]^{n-1} f(y_{ni}) \;,\; -\infty < y_{ni} < \infty \quad \ldots\ldots\ldots\ldots\ldots(2\text{-}1)$$

Since the k samples are independent, then the joint p.d.f of the k samples maxima Y_{n1}, Y_{n2}, …, Y_{nk} is:

$$g(y_{n1}, y_{n2}, \ldots, y_{nk}) = \prod_{i=1}^{k} g_i(y_{ni})$$

$$n^k \left[\prod_{i=1}^{k} F(y_{ni})\right]^{n-1} \prod_{i=1}^{k} f(y_{ni}) \;,\; -\infty < y_{ni} < \infty \quad \ldots\ldots\ldots\ldots\ldots(2\text{-}2)$$

3.Distribution of the Product of Two Maximum from Uniform Samples

For r.v$X \sim U(0,1)$, the p.d.f and c.d.f are

$$f(x) = 1 \;,\; 0 < x < 1 \quad \ldots\ldots\ldots\ldots(3\text{-}1)$$
$$= 0 \;,\; e.w.$$

and

$$F(x) = \begin{cases} 0 & , \; x \le 0 \\ x & , \; 0 < x < 1 \\ 1 & , \; x \ge 1 \end{cases} \quad \ldots\ldots\ldots\ldots(3\text{-}2)$$

We note that if $v_1=y_{n1}$ then by using (2-1), we have the p.d.f of v_1 is

$$h(v_1) = n v_1^{n-1} \;,\; 0 < v_1 < 1 \quad \ldots\ldots\ldots\ldots(3\text{-}3)$$
$$= 0 \;,\; e.w.$$

Now, let us consider the product of two maxima as $V_2=Y_{n1}.Y_{n2}$ and set $V_1=Y_{n1}$. Use of (2-2) with k=2, we have the joint p.d.f of Y_{n1} and Y_{n2} is

$$g(y_{n1},y_{n2}) = n^2 (y_{n1}.y_{n2})^{n-1} \quad ,0<y_{ni}<1 \quad ,i=1,2$$
$$= 0 \quad ,e.w.$$

The functions $v_1=y_{n1}$ and $v_2=y_{n1}.y_{n2}$ define one-to-one transformation that maps the space $A=\{(y_{n1},y_{n2}): 0<y_{ni}<1, i=1,2\}$ onto the space $B=\{(v_1,v_2): 0<v_2<v_1<1\}$ with inverses $y_{n1}=v_1$ and $y_{n2}=\frac{v_2}{v_1}$ and the Jacobian of the transformation is:

$$J=\frac{\partial(y_{n1},y_{n2})}{\partial(v_1,v_2)}=\begin{vmatrix} \frac{\partial y_{n1}}{\partial v_1} & \frac{\partial y_{n1}}{\partial v_2} \\ \frac{\partial y_{n2}}{\partial v_1} & \frac{\partial y_{n2}}{\partial v_2} \end{vmatrix}$$

$$=\begin{vmatrix} 1 & 0 \\ \frac{-v_2}{v_1^2} & \frac{1}{v_1} \end{vmatrix}=\frac{1}{v_1}$$

Then the joint p.d.f of V_1 and V_2 is:

$$h(v_1,v_2)=g(v_1,\frac{v_2}{v_1})|J|$$
$$=\frac{n^2 v_2^{n-1}}{v_1}, \quad 0<v_2<v_1<1$$
$$=0 \quad , \quad e.w.$$

and the marginal p.d.f of V_2 is:

$$h_2(v_2)=\int_{v_1} h(v_1,v_2)\,dv_1 = n^2 v_2^{n-1}\int_{v_1=v_2}^{1}\frac{1}{v_1}dv_1$$
$$=n^2 v_2^{n-1}\ln v_1\Big|_{v_2}^{1}$$
$$=-n^2 v_2^{n-1}\ln v_2 \quad ,0<v_2<1 \qquad \text{........ (3-4)}$$
$$=0 \quad ,e.w.$$

4. Distribution of the Product of k Maxima from Uniform Samples

In this section, we shall consider the generalization to the distribution of the product of k sample maxima from uniform distribution by formulating the following theorem.

Theorem (1):

Let $Y_{n1},Y_{n2},\ldots,Y_{nk}$ be k r.v's representing the maximum of k independent samples of same size n from uniform distribution. If $V_k=\prod_{i=1}^{k}Y_{ni}$, the V_k has p.d.f

$$h_k(v_k)=\frac{(-1)^{k-1}}{\Gamma(k)}\, n^k v_k^{n-1}[\ln v_k]^{k-1} \quad ,0<v_k<1$$
$$=0, \qquad e.w.$$

Proof:

Using (2-2) with p.d.f and c.d.f given by (3-1) and (3-2), we have the joint p.d.f of $Y_{n1},Y_{n2},\ldots,Y_{nk}$ is

$$g(y_{n1},y_{n2},\ldots,y_{nk})=n^k\left(\prod_{i=1}^{k}y_{ni}\right)^{n-1} \quad ,0<y_{ni}<1 \quad ,i=1,2,\ldots,k$$
$$=0, \qquad e.w.$$

with $V_k=\prod_{i=1}^{k}Y_{ni}$, set $V_1=Y_{n1}$, $V_2=Y_{n1}.Y_{n2}$, …, $V_{k-1}=\prod_{i=1}^{k-1}Y_{ni}$,

so that the functions $v_1=y_{n1}$, $v_2=y_{n1}.y_{n2}$, …, $v_{k-1}=\prod_{i=1}^{k-1}y_{ni}$, $v_k=\prod_{i=1}^{k}y_{ni}$ define one-to-one transformation that maps the space:

$A=\{(y_{n1},y_{n2},\ldots,y_{nk}) : 0 < y_{ni} < 1,\ i = 1,2,\ldots,k\}$

onto the space:

$B=\{(v_1,v_2,\ldots,v_k) \ : \ 0<v_k<v_{k-1}<\ldots<v_2<v_1<1\}$

with inverses $y_{n1}=v_1$, $y_{n2}=\frac{v_2}{v_1}$, $y_{n3}=\frac{v_3}{v_2}$, …, $y_{nk-1}=\frac{v_{k-1}}{v_{k-2}}$, $y_{nk}=\frac{v_k}{v_{k-1}}$ and the Jacobian of this transformation is:

$$J=\frac{\partial(y_{n1},y_{n2},\ldots,y_{nk})}{\partial(v_1,v_2,\ldots,v_k)}$$

$$=\begin{vmatrix} \frac{\partial y_{n1}}{\partial v_1} & \frac{\partial y_{n1}}{\partial v_2} & \cdots & \frac{\partial y_{n1}}{\partial v_k} \\ \frac{\partial y_{n2}}{\partial v_1} & \frac{\partial y_{n2}}{\partial v_2} & \cdots & \frac{\partial y_{n2}}{\partial v_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{nk-1}}{\partial v_1} & \frac{\partial y_{nk-1}}{\partial v_2} & \cdots & \frac{\partial y_{nk-1}}{\partial v_k} \\ \frac{\partial y_{nk}}{\partial v_1} & \frac{\partial y_{nk}}{\partial v_2} & \cdots & \frac{\partial y_{nk}}{\partial v_k} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \frac{-v_2}{v_1^2} & \frac{1}{v_1} & 0 & 0 & \cdots & 0 \\ 0 & \frac{-v_3}{v_2^2} & \frac{1}{v_2} & 0 & \cdots & 0 \\ \vdots & & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{-v_{k-1}}{v_{k-2}^2} & \frac{1}{v_{k-2}} & 0 \\ 0 & 0 & \cdots & 0 & \frac{-v_k}{v_{k-1}^2} & \frac{1}{v_k} \end{vmatrix}$$

$$= \frac{1}{v_1 v_2 \ldots v_k}$$

Then the joint p.d.f of $V_1, V_2, \ldots, V_k$ is

$$h(v_1, v_2, \ldots, v_k) = g\left(v_1, \frac{v_2}{v_1}, \frac{v_3}{v_2}, \ldots, \frac{v_{k-1}}{v_{k-2}}, \frac{v_k}{v_{k-1}}\right)|J|$$

$$= n^k \frac{v_k^{n-1}}{v_1 v_2 \ldots v_k} \qquad , 0 < v_k < v_{k-1} < \ldots < v_2 < v_1 < 1$$

$$= 0 \qquad , e.w$$

The marginal p.d.f of V_k is

$$h_k(v_k) = \int_{v_1}\int_{v_2} \cdots \int_{v_{k-1}} h(v_1, v_2, \ldots, v_k)\, dv_1\, dv_2 \ldots dv_{k-1}$$

$$= n^k v_k^{n-1} \int_{v_{k-1}=v_k}^{1} \cdots \int_{v_2=v_3}^{1} \int_{v_1=v_2}^{1} \frac{dv_1\, dv_2 \ldots dv_{k-1}}{v_1 v_2 \ldots v_{k-1}}$$

Since

$$\int_{v_1=v_2}^{1} \frac{dv_1}{v_1} = \ln v_1 \Big|_{v_2}^{1} = -\ln v_2 = (-1)^1 \frac{\ln v_2}{1!}$$

also,

$$\int_{v_2=v_3}^{1} \frac{-\ln v_2}{v_2} dv_2 = -\frac{(\ln v_2)^2}{2}\Bigg|_{v_3}^{1} = \frac{(\ln v_3)^2}{2}$$

$$= (-1)^2 \frac{(\ln v_3)^2}{2!}$$

and

$$\int_{v_3=v_4}^{1} \frac{(\ln v_3)^2}{2} \frac{dv_3}{v_3} = \frac{1}{2}\frac{(\ln v_3)^3}{3}\Bigg|_{v_4}^{1} = -\frac{(\ln v_4)^3}{6}$$

$$= (-1)^3 \frac{(\ln v_4)^3}{3!}$$

If successive integration are carried out in this manner for $V_4, V_5, \ldots, V_{k-1}$, we will have

$$\int_{v_i=v_{i+1}}^{1} (-1)^{i+1} \frac{(\ln v_i)^2}{(i-1)!} \frac{dv_i}{v_i} = (-1)^i \frac{(\ln v_{i+1})^i}{i!},$$

$$i = 1, 2, \ldots, k-1$$

and since we have k-1 integrals, then

$$h_k(v_k) = n^k v_k^{n-1} (-1)^{k-1} \frac{(\ln v_k)^{k-1}}{(k-1)!}$$

$$= \frac{(-1)^{k-1}}{\Gamma(k)} n^k v_k^{n-1} (\ln v_k)^{k-1}, 0 < v_k < 1$$

$$= 0 \qquad , e.w.$$

.... (4-1)

Note: when k=1,2equation (4-1) leads to (3-3) and (3-4).

5. The Central Moments of the Distribution

The r^th central moment about origin defined as

$$E[v_k^r] = \int_{v_k} v_k^r h_k(v_k) dv_k$$

$$= \int_0^1 v_k^r \frac{(-1)^{k-1}}{\Gamma(k)} n^k v_k^{n-1} (\ln v_k)^{k-1} dv_k$$

$$= \frac{n^k}{(n+r)^k} \int_0^1 \frac{(-1)^{k-1}}{\Gamma(k)} (n+r)^k v_k^{(n+r)-1} (\ln v_k)^{k-1} dv_k$$

.... (5-1)

The integral side of equation (5-1) is unity with n replaced by $n+r$

$$E[v_k^r] = \left(\frac{n}{n+r}\right)^k \qquad , r = 1, 2, 3, \ldots \quad \text{..............(5-2)}$$

For r=1,2, we have from equation (5-1)

$$m_k = E[v_k] = \left(\frac{n}{n+1}\right)^k \text{ and } E[v_k^2] = \left(\frac{n}{n+2}\right)^k$$

.... (5-3)

$$s_k = \operatorname{var}(v_k) = E[v_k^2] - m_k^2$$

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

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

$$= \frac{[n(n+1)^2]^k - [n^2(n+2)]^k}{[(n+1)^2(n+2)]^k} \quad \text{........... (5-4)}$$

Conclusions

1. The c.d.f of the p.d.f given by equation (4-1) has the form:

$$H(v) = \Pr(V_k \leq v) = \int_0^v h(t)dt$$

$$= \frac{(-1)^{k-1}}{\Gamma(k)} n^k \int_0^v t^{n-1} (\ln t)^{k-1} dt$$

Unfortunately a close form expression for the integral side does not exist. But numerical approximations can be used for tabulating the values of $H(v)$ for some specific values of n, k and x.

2. Generating r.v's from the p.d.f of equation (4-1) can be made easily by two methods of Monte Carlo simulation as follows:
   i) Generating k samples from U(0,1), and the product of k maxima lead to a r.v follow equation (4-1).
   ii) Acceptance-rejection Monte Carlo method could be used by making use of the inequality $\ln x < x$ for $0 < x < 1$.
3. The driven probability distribution might be find applications in certain fields of discipline.

References

[1] Herbert, A. D., Order Statistics, third edition, Wiley Series in Probability and Statistics, 2003.

[2] Gunbel,E. J., Statistical Theory of Breaking Strength and Fatigue, Bull Inc., 38(3), 375-393,1961.

[3] Coben, A. C. 1966, Life Testing and Early Failure, Technometrics 8, 157, 539-545, 1966.

[4] Zidck, J. V., Francis, P. D., Navln and Richard Lockhart, Statistics of Extreme an Alternative Method with Applications to Bridge Design Code, Technometrics, 21, 185-192, 1979

[5] Barry, C. D., N. Balakrshnan and H. N. Najaraja, A first course in Order Statistics, Springer, New York, 2008.

الخلاصة

في هذا البحث توصلنا الى صياغة توزيع حاصل ضرب القيم العظمى التي تعود الى k من العينات المستقلة من التوزيع المتجانس بالفترة (0,1) حيث ان المتغيرات العشوائية للعينات قد رتبت تصاعدياً في المقدار كل على حده. بالاضافة الى ايجاد الصيغة العامه لعزوم التوزيع حول نقطة الاصل.