On the Variances of the Sample Maximum of Order Statistics from A Discrete Uniform Distribution

Kesikli Düzgün Dağılımdaki Sıra İstatistiklerinin Örnek Maksimumunun Varyansları Üzerine

Sinan ÇALIK,[a]
Nurhan HALİSDEMİR,[a]
Mehmet GÜRCAN,[a]
Ayşe METİN[a]

[a]Department of Statistics,
Fırat University Faculty of Medicine,
Elazığ

Geliş Tarihi/*Received:* 03.12.2009
Kabul Tarihi/*Accepted:* 10.02.2010

Yazışma Adresi/*Correspondence:*
Sinan ÇALIK
Fırat University Faculty of Medicine,
Department of Statistics, Elazığ,
TÜRKİYE/TURKEY
scalik@firat.edu.tr

ABSTRACT In this study, formulae for the moments of the sample maximum of order statistics from a discrete uniform distribution are presented in closed form. Illustrative example of the variances of the sample maximum of order statistics from a discrete uniform distribution is given. The current study obtains algebraic expressions for n up to 20 for the variances of the sample maximum of order statistics from a discrete uniform distribution. Using the obtained algebraic expressions, anyone can compute these variances based on the N and n values. Further studies may focus on a software program estimating these variances found in this study.

Key Words: Variance; sample maximum; uniform distribution; order statistics

ÖZET Bu çalışmada kesikli düzgün dağılımdaki sıra istatistiklerinin örnek maksimum momentleri için formüller açık bir şekilde sunuldu. Kesikli düzgün dağılımdaki sıra istatistiklerinin örnek maksimum varyanslarının tanımlayıcı örnekleri verildi. Şimdiki çalışmada kesikli düzgün dağılımdaki sıra istatistiklerinin örnek maksimumunun varyansları için n=20' ye kadar cebirsel ifadeleri elde edildi. Elde edilmiş cebirsel ifadeler kullanılarak N ve n değerlerine bağlı olarak bu varyanslar bilgisayarla hesaplanabilir. Daha ileri çalışmalar bu çalışmada bulunan varyansları hesaplayan yazılım programları üzerine odaklanabilir.

Anahtar Kelimeler: Varyans; örnek maksimumu; düzgün dağılım; sıra istatistikleri

Turkiye Klinikleri J Biostat 2010;2(2):54-7

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a discrete distributions with probability mass function $p(x)$ $(x = 0,1,2,\ldots)$ and cumulative distribution function $P(x)$. Let $X_{1:n} \le X_{2:n} \le \ldots \le X_{n.n}$ be the order statistics obtained from above random sample by arranging the observations in increasing order of magnitude. Let $E\left(X^k_{r:n}\right)$ denote by $\mu^{(k)}_{r:n}$ $\left(1 \le r \le n,\, k \ge 1\right)$. For convenience, $\mu_{r:n}$ for $\mu^{(1)}_{r:n}$ and $\sigma^2_{r:n}$ for variance of $X_{r:n}$ will also be used.

The first two moments of order statistics from discrete distributions were proved by Khatri.[1] Several recurrence relations and identities available for single and product moments order statistics in a sample size n from an arbitrary continuous distribution were extended for the discrete case by Balakrishnan.[2] The review paper by Nagaraja[3] lucidly accounts all the developments on discrete order statistics. The first two moments were also proved by Arnold et al.[4] by a different technique. The *m*th raw moments of order statistics from discrete distributions were proved and a relation between the moments of sample maximum of order statistics from a discrete uniform distribution and the sum *S*(*N-1,n*) as given in (5) were obtained by Calik et al.[5] On the variances of the distribution of the sample range of order statistics from a discrete uniform distribution were obtained by Calik.[6]

In this paper, the variances of the sample maximum of order statistics from a discrete uniform distribution are obtained by using the sum S(N-1,n) as given in (5). For n up to 20, algebraic expressions for the variances are obtained.

BASIC DISTRIBUTIONAL RESULTS

Let $F_{r:n}(x)\,(r=1,2,\ldots,n)$ denote the cumulative distribution function (cdf) of $X_{r:n}$. Then it is easy to see that

$$\begin{aligned} F_{r:n}(x) &= \Pr\{X_{r:n} \le x\} \\ &= \sum_{i=r}^{n}\binom{n}{i}[P(x)]^{i}[1-P(x)]^{n-i} \\ &= \mathrm{I}_{P(x)}(r,n-r+1) \end{aligned} \tag{1}$$

where $\mathrm{I}_P(a,b)$ is the incomplete beta function defined by

$$\mathrm{I}_P(a,b) = \{B(a,b)\}^{-1}\int^{P} t^{a-1}(1-t)^{b-1}dt \quad , \quad (a,b>0)$$

For discrete population, the probability mass function (pmf) of may be obtained from (1) by differencing as

$$\begin{aligned} f_{r:n}(x) &= F_{r:n}(\mathrm{x}) - F_{r:n}(x-1) \\ &= \frac{1}{B(r,n-r+1)}\int_{P(x-1)}^{P(x)} t^{r-1}(1-t)^{n-r}dt \quad . \end{aligned} \tag{2}$$

Hence the mth raw moment of $X_{r:n}$ can be immediately written down as

$$\mu_{r:n}^{(m)} = \sum_{x=0}^{\infty} x^{m} f_{r:n}(x)$$

where $f_{r:n}(x)$ is as given in (2). These moments can, however, be simplified and written in a convenient form.[1,4,7]

For example, we have for $1 \le r \le n$

$$\mu_{r:n} = \sum_{x=0}^{\infty}[1-F_{r:n}(x)] \tag{3}$$

and

$$\mu_{r:n}^{(2)} = 2\sum_{x=0}^{\infty} x[1-F_{r:n}(x)] + \mu_{r:n} \,. \tag{4}$$

ORDER STATISTICS FROM A DISCRETE UNIFORM DISTRIBUTION

Let the population random variable X be discrete uniform with support $\mathrm{B} = \{1,2,\ldots,N\}$. We then write, X is discrete uniform $[1\,,N]$. Note that its pmf is given by $p(x) = \frac{1}{N}$, and its cdf is $P(x) = \frac{x}{N}$, for $x \in \mathrm{B}$. Consequently the cdf of the r-th order statistics is given by

$$F_{r:n}(x) = \sum_{i=r}^{n}\binom{n}{i}\left(\frac{x}{N}\right)^{i}\left(1-\frac{x}{N}\right)^{n-i} \quad , \qquad x \in \mathrm{B}$$

SPECIAL SUMS

In the theory of nonparametric statistics, particularly when we deal with rank sums, we often need for the sums of powers of the first n positive integers; namely, expression for

$$S(N-1,n) = 1^{n} + 2^{n} + \ldots + (N-1)^{n} = \sum_{x=1}^{N-1} x^{n} \tag{5}$$

for $n = 0,1,2,\cdots$ The following theorem, we provide a convenient way of obtaining these sums.

Theorem 1.1.

$$\sum_{n=0}^{k-1}\binom{k}{n} S(N-1,n) = N^{k} - 1$$

for any positive.[8]

A disadvantage of this theorem is that we have to find the sums $S(N-1,n)$ one at a time, first for n=0, then n=1, then n=2 and so forth. For instance, for k=1 we get

$$\binom{1}{0} S(N-1,0) = N-1$$

and, hence, $S(N-1,0)=1^0+2^0+\ldots+(N-1)^0=N-1$. Similarly, for k=2 we get

$$\binom{2}{0}S(N-1,0)+\binom{2}{1}S(N-1,1)=N^2-1$$

$$N-1+2S(N-1,1)=N^2-1$$

and, hence, $S(N-1,1)=1^1+2^1+\ldots+(N-1)^1=\frac{1}{2}(N-1)N$.

Using the same technique, we can find the sums $S(N-1,2)=\frac{1}{6}(N-1)N(2N-1)$, $S(N-1,3)=\frac{1}{4}(N-1)^2N^2$ and so on.

THE MOMENTS OF THE SAMPLE MAXIMUM OF ORDER STATISTICS

In general, (3) and (4) moments are not easy to evaluate analytically. Sometimes the moments of sample extremes are tractable. Let us see what happens in the case of discrete uniform distribution.

When X is a discrete uniform $[1,N]$ random variable in the case of the sample maximum, (3) yields

$$\mu_{n:n}=\sum_{x=0}^{N}\left[1-F_{n:n}(x)\right]$$

$$=\sum_{x=0}^{N-1}\left[1-\left(\frac{x}{N}\right)^n\right] \qquad (6)$$

$$=N-\frac{S(N-1,n)}{N^n}.$$

Further from (4),

$$\mu_{n:n}^{(2)}=2\sum_{x=1}^{N-1}x\left[1-\left(\frac{x}{N}\right)^n\right]+\mu_{n:n}$$

$$=N(N-1)-\frac{2S(N-1,n+1)}{N^n}+\mu_{n:n} \qquad (7)$$

$$=N^2-\frac{2S(N-1,n+1)+S(N-1,n)}{N^n}$$

and hence variance of the sample maximum we obtain

$$\sigma_{n:n}^2=\mu^{(2)}{}_{n:n}-\mu_{n:n}^2$$

$$=N^2-\frac{2S(N-1,n+1)+S(N-1,n)}{N^n}-\left[N-\frac{S(N-1,n)}{N^n}\right]^2.$$

For n up to 20, algebraic expressions for the $\sigma_{n:n}^2$ are obtained, see Table 1.

Example 1: Using (6) and (7), we can conclude for example that

$$\mu_{2:2}=N-\frac{S(N-1,2)}{N^2}=N-\frac{\frac{1}{6}N(N-1)(2N-1)}{N^2}=\frac{4N^2+3N-1}{6N}$$

and

$$\mu_{2:2}^{(2)}=N^2-\frac{S(N-1,3)}{N^2}=N-\frac{\frac{1}{4}(N-1)^2N^2}{N^2}=\frac{3N^3+3N^2+1}{6N}.$$

When n=2, using the values of $\mu_{2:2}$ and $\mu_{2:2}^{(2)}$ and, we obtain

$$\sigma_{2:2}^2=\frac{2N^4-N^2-1}{36N^2}=\frac{(2N^2+1)(N^2-1)}{36N^2}.$$

CONCLUSION

The current study presents algebraic expressions for n up to 10 for the variances of the sample maximum of order statistics from a discrete uniform distribution. Using the obtained algebraic expressions, these variances can be computed. As it is understood from Table 1, different values can be obtained for N and n. For instance, for N=100, using the value n=3 in the Table 1, we obtain $\sigma_{3:3}^2$=374.95833375 Further studies may focus on a software program estimating these variances found in the current study.

TABLE 1: The variances of the sample maximum of order statistics from discrete uniform distribution.

n	$\sigma^2_{n:n}$
1	$(1/12)(N^2-1)$
2	$(1/36)N^{-2}(2N^2+1)(N^2-1)$
3	$(1/240)N^{-2}(3N^2-1)(N^2-1)$
4	$(1/900)N^{-6}(24N^6-21N^4-19N^2+1)(N^2-1)$
5	$(1/1008)N^{-6}(20N^6-36N^4-15N^2+7)(N^2-1)$
6	$(1/1764)N^{-10}(78N^4-13N^2+6N^6-78N^8+27N^{10}+1)(N^2-1)$
7	$(1/2880)N^{-10}(213N^4-120N^2+95N^6-145N^8+35N^{10}+20)(N^2-1)$
8	$(1/8100)N^{-14}(541N^4-111N^2-1499N^6+715N^8+575N^{10}-445N^{12}+80N^{14}+9)(N^2-1)$
9	$(1/13200)N^{-14}(106N^{14}-772N^{12}+1571N^{10}+911N^8-5821N^6+4389N^4-1663N^2+297)(N^2-1)$
10	$(1/4356)N^{-18}(1444N^4-305N^2-3572N^6+5690N^8-3622N^{10}-25N^{12}+767N^{14}-267N^{16}+30N^{18}+25)(N^2-1)$
11	$(1/252050)N^{-18}(641036N^4-254800N^2-980525N^6+958375N^8-549707N^{10}-42400N^{12}+63420N^{14}$ $-18880N^{16}+1540N^{18}+45500)(N^2-1)$
12	$(1/7452900)N^{-22}(27342319N^4-5810619N^2-66497141N^6+99659850N^8-104252950N^{10}+63345590N^{12}-13921600N^{14}$ $-3195885N^{16}+2359665N^{18}-487725N^{20}+37800N^{22}+477481)(N^2-1)$
13	$(1/776400)N^{-22}(9519202N^4-2688569N^2-9968838N^6+10132308N^8-7482327N^{10}$ $+3051214N^{12}-399508N^{14}-146255N^{16}+70535N^{18}-11620N^{20}+780N^{22}+477461)(N^2-1)$
14	$(1/16200)N^{-26}(1260092N^4-268170N^2-3053308N^6+4530710N^8-4587230N^{10}+3428080N^{12}-1785320N^{14}+515160N^{16}$ $-36300N^{18}-23235N^{20}+7965N^{22}-1107N^{24}+63N^{26}+22050)(N^2-1)$
15	$(1/135840)N^{-26}(206610740N^4-83880600N^2-310671880N^6+313890720N^8-231867390N^{10}+130845829N^{12}$ $-50143851N^{14}+10539065N^{16}-299395N^{18}-442985N^{20}+115925N^{22}-13905N^{24}+875N^{26}+14994000)(N^2-1)$
16	$(1/260100)N^{-30}(747264681N^4-159086511N^2-1809106919N^6+2678543765N^8-2696299995N^{10}+1979897485N^{12}-$ $1118061475N^{14}+482514435N^{16}-141257045N^{18}+22029315N^{20}+260815N^{22}-868325N^{24}+182275N^{26}$ $-18325N^{28}+800N^{30}+13082689)(N^2-1)$
17	$(1/14309200)N^{-30}$ ([illegible]N^4 [illegible]N^2 [illegible]N^6 [illegible]N^8 [illegible]N^{10} [illegible]N^{12} [illegible]N^{14} [illegible]N^{16} [illegible]N^{18} [illegible]N^{20} [illegible]N^{22} $-20441505N^{24}+3525095N^{26}-307300N^{28}+11900N^{30}+15659978733)(N^2-1)$
18	$(1/15920100)N^{-34}(2747513343304N^4-584973244385N^2-6650236179176N^6+9841058725580N^8$ $-9893272597420N^{10}+7240369373720N^{12}-4049986762360N^{14}+1797830096370N^{16}$ $-640853634750N^{18}+173454706215N^{20}-31448870985N^{22}+2743431621N^{24}+221240901N^{26}$ $-103814025N^{28}+15024135N^{30}-1147335N^{32}+39690N^{34}+48107842225)(N^2-1)$
19	$(1/7781600)N^{-34}$ ([illegible]N^4 [illegible]N^2 [illegible]N^6 [illegible]N^8 [illegible]N^{10} [illegible]N^{12} [illegible]N^{14} [illegible]N^{16} [illegible]N^{18} [illegible]N^{20} [illegible]N^{22} [illegible]N^{24} [illegible]N^{26} $-67739672N^{28}+8365588N^{30}-564564N^{32}+17556N^{34}+2116745057900)(N^2-1)$
20	$(1/48024900)N^{-38}(767878316243179N^4-163492513442139N^2-1858517941834721N^6+2749888811281060N^8$ $-2763602815808140N^{10}+2020971570441980N^{12}-1128280580815540N^{14}+498261362201550N^{16}-179007120908130N^{18}$ $+53175980731170N^{20}-12792741283080N^{22}+2322004704030N^{24}-277071999270N^{26}+12621287250N^{28}$ $+2257541550N^{30}-548675325N^{32}+58605525N^{34}-3521925N^{36}+99000N^{38}+13445649582561)(N^2-1)$

REFERENCES

1. Khatri CG. Distribution of order statistics for discrete case. Ann Inst Statist Math 1962;14(1):167-71.
2. Balakrishnan N. Order statistics from discrete distributions. Commun. Statist.-Theor Meth 1986;15(3):657-75.
3. Nagaraja HN. Order statistics from discrete distributions (with discussion). Statistics 1992;23(3):45-7.
4. Arnold BC, Balakrishnan.N, Nagaraja HN. A First Course in Order Statistics. 1st ed. New York: John Wiley and Sons; 1992. p.42-5.
5. Calik S, Gungor M, Colak C. On the moments of order statistics from discrete distributions. Pak J Statist 2010;26(2):471-26.
6. Calik S. On the variances of the distribution of the sample range of order statistics from a discrete uniform distribution. J Math Stat 2005;1(3):180-3.
7. David HA. Expected values and moments. Order Statistics. 2nd ed. New York: John Wiley and Sons; 1970. p.25-6.
8. Freund JE, Walpole RE. Special Sums. Mathematical Statistics. 1st ed. London: Prentice- Hall International Inc; 1962.