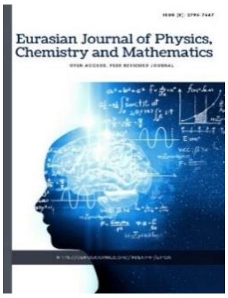

Ordered Statistics And Their Importance In Finding The Distribution Of Variation Series Terms

Gulnoz Sayfulloyeva [1]	[1)] Navoi State University, Faculty of Mathematics and Informatics, Department of Mathematics, Doctor of Philosophy (PhD) in Physics and Mathematics E-mail: sayfullayevagulnoz@gmail.com
Shokirova Durdona [2]	[2)] Teacher of Navoi State University, Faculty of Mathematics and Informatics, Department of Mathematics, shdurdona7810@gmail.com
Ibragimova Dildora [3,]	[3)] Teacher of Navoi State University, Faculty of Mathematics and Informatics, Department of Mathematics, ibragimovadildora13.05@gmail.com
Bozorboyeva Dildora [4]	[4)] Master's student of Navoi State University, dildorabozorboyeva22@gmail.com

ABSTRACT

Ordinal statistics play an important role in mathematical statistics. In many cases, ordinal statistics are used as estimates in problems of estimating unknown parameters. In order to check the properties of these estimates and to assess their error, it is necessary to know the distributions of ordinal statistics. Therefore, the study of ordinal statistics and their distributions is of great importance. This qualification thesis is devoted to these issues.

Keywords: Random Variables, Unordered Variables, Ordered Variables, Mathematical Expectation, Variance

Ordinal statistics play an important role in mathematical statistics. In many cases, ordinal statistics are used as estimates in problems of estimating unknown parameters. In order to check the properties of these estimates and to assess their error, it is necessary to know the distributions of ordinal statistics. Therefore, the study of ordinal statistics and their distributions is of great importance. This qualification thesis is devoted to these issues.

Suppose $(\Omega,\ \mathcal{A},\ P)$there is a space of probabilities, and $(\Omega$ —the measure (probability) defined in the sigma algebra consisting of all subsets of $P-\mathcal{A}$the space of elementary events$\mathcal{A}-\Omega$

$$\xi: (\Omega, \mathcal{A}) \to (x, \mathcal{B})$$

-the ξ - practical values of the random variable in n independent experiments

$$X_1, X_2, \dots, X_n \tag{1}$$

Let (1) be a sample of size n and

$$X^{(n)} = (X_1, X_2, \dots, X_n) \tag{2}$$

It is defined as (2)-the numerical values of the random vector

$$x^{(n)} = (x_1, x_2, \dots, x_n) \tag{3}$$

We mark with

$$x_1, x_2, x_3, \dots, x_n \tag{4}$$

If we arrange the numbers in ascending order,

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

We get a simple variational series, where

$$x_{(1)} = \min(x_1, x_2, \dots, x_n)$$

$$x_{(2)} = \min(\min(x_1, x_{i-1}, x_{i+1}, \dots, x_n) \qquad (1.5)$$

$$\dots\dots\dots\dots\dots\dots\dots\dots i = \overline{1,n}$$

$$x_{(n)} = \max(x_1, x_2, \dots, x_n)$$

$x_{(k)}$ $k = 1,2,\dots,n$ statistics are called ordinal statistics.

If we move from observations (2) to random variables (3) $X_{(k)}$, we find that $x_{(k)}$

$$X_{(1)} \leq X_{(2)} \leq, \dots, X_{(n)} \qquad (6)$$

called variational series .

In addition to simple variational series, grouped and interval variational series can also be constructed.

Here is the construction of these variational series:

A. Grouped variational series

$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ In a simple variational series, each of the variants can be repeated several times. Let's say$x_{(1)}$ $n_1$times, ...,$x_{(r)}$. Let's repeat n_r the option $n_k$times, ..., $x_{(n)}$times. Let's write them accordingly as follows.

$$x_i: \quad x'_{(1)} < x'_2 < \dots < x'_{(k)}$$

$$n_i: \quad n_1 \quad n_2 \quad n_k$$

This $x'_{(1)} < x'_2 < \dots < x'_{(k)}$ creates a variational series. $\sum_{i=1}^{n} n_i = n$

Here n is the number of experiments, n_i −the frequency of the ordinal statistic, i.e. the number of repetitions.

B. Interval variation series

To construct an interval variational series, it is first $R = x_{(n)} - x_{(1)}$found.

Here is R −the sample size. So the experimental results (sample) Rwill be related to a cross section with length . We divide this cross section $h$into equal parts with length .

$$h = \frac{R}{k-1}, \quad k = \sqrt{n}$$

n_i −We denote the number of observations falling into each interval by .

There are various methods that can be used to accomplish this task. Here are the most commonly used ones. It is recommended to divide the time into the following intervals:

$$\begin{cases} a_0 = x_{min} - \dfrac{h}{2} \\ a_1 = a_0 + h \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_k = a_{k-1} + h \end{cases}$$

$a_k - x_{max}$ we stop when or $x_{(n)}$is equal to or greater than . We have the following intervals.

$$(a_0;\ a_1) \quad (a_1;\ a_2) \dots\dots \quad (a_{i-1};\ a_i) \dots\dots. \quad (a_{k-1};\ a_k)$$

the midpoints of these intervals x_i^*as:

$$x_{(1)}^* = \frac{a_0 + a_1}{2} \qquad x_{(2)}^* = \frac{a_1 + a_2}{2} \dots\dots. x_{(k)}^* = \frac{a_{k-1} + a_k}{2}$$

$x_{(1)}^*$, $x_{(2)}^*, \dots\dots, x_{(k)}^*$s form the following variational series.

$$x_{(1)}^* < \quad x_{(2)}^* < \cdots.., < x_{(k)}^*$$

Now we can create a table based on the data obtained.

In the table$\nu_i = \frac{h_i}{n}$ $i = \overline{1,k}$ $x_{(i)}^*$ −relative frequency of ordinal statistics

$a_{i-1}-a_i$	$x^*_{(i)}$	n_i	ν_i
a_0-a_1 a_1-a_2 .... $a_{k-1}-a_k$	$x^*_{(1)}$ $x^*_{(2)}$ .... $x^*_{(k)}$	n_1 n_2 ..... n_k	$\frac{n_1}{n}$ $\frac{n_2}{n}$ ...... $\frac{n_k}{n}$
		$\sum_{i=1}^{k} n_i = n$	$\sum_{i=1}^{k} \nu_i = 1$

Now we will calculate the distributions of ordinal statistics. Let us assume

$$x_1, x_2, \dots, x_n$$

n independent, identically distributed r.v.s , and let their distribution functions $F(x)$be. $F_r(x)$ $(r = 1, 2, \dots, n$)d.f. represents the ordinal statistic $X_r$through . Then

$$X_{(n)} = \max(x_1, x_2, \dots, x_n) \tag{7}$$

The d.f. of the largest order statistic is calculated using the following formula.

$$F_{(n)}(x) = F_{max}(x) = P\{X_{(n)} \le x\} = P\{max(x_1, x_2, \dots, x_n) \le x\}$$

the independent $X_{1,}$ ...., $X_n$terms "x"to be less than the maximum, all of its terms, that is, each of the variants, "x"must be less than . From the independence of the terms, we have the following equality.

$$P\{\max(x_1, x_2, \dots, x_n) \le x\} = P\{max(x_1, x_2, \dots, x_n) \le x\}=$$
$$= P\{X_1 \le x\} \cdot P\{X_2 \le x\} \cdot \dots \ P\{X_{(n)} \le x\}=$$
$$= \prod_{i=1}^{n} P\{X_i \le x\} = \underbrace{F(x) \dots F(x)}_{n\ ta} = F^n(x) = P^n\{X_i \le x\}.$$

So, $X_{(n)}$ —the distribution function of such an order statistic is

$$F_{(n)}(x) = P\{X_{(n)} \le x\} = F^n(x) \tag{8}$$

Similarly, $X_{(1)}$ —the d.f. of the smallest order statistic can be calculated.

$F_{(1)}(x) = P\{X_{(1)} \le x\} = P\{\min(x_1 \dots x_n) \le x\} = 1 - P\{\min(x_1 \dots x_n) > x\} = (x_1, x_2, \dots, x_n)$ x_i because the are independent.}

$$1 - P\{X_1 > x, \ X_2 > x, \dots X_n > x_n\}=$$
$$= 1 - P\{X_1 > x\} \cdot P\{X_2 > x\} \cdot \dots P\{X_n > x\}=$$
$$= 1 - [1 - P\{X_1 \le x\}] \cdot [1 - P\{X_2 \le x\}] \dots [1 - P\{X_n \le x\}] =$$
$$= 1 - \prod_{i=1}^{n} [1 - P\{X_i \le x\}] = 1 - [1 - F(x)]^n$$

So, the minimum term of the sample d.f. has the following form.

$F_{(1)}(x) = P\{X_{(1)} \le x\} = P\{x_i$ at least one of $x$them $r$is equal to or less than}in a Bernoulli scheme $\mathcal{A} = \{x_i \le x\}$, then for n trials, the occurrence of this event at least r times $X_{(r)} \le x$is { }. Therefore, $\mu_n$if we consider the number of events in a Bernoulli scheme in n trials, then

$$\{\mu_n \ge r\} = \{X_{(r)} < x\} \tag{10}$$

according to the property of binomial probabilities

$$P\{r \le \mu_n\} = \sum_{k=r}^{n} P\{\mu_n = k\}$$

because

$$P\{\mu_n \ge r\} = P\{X_{(r)} < x\} = \textstyle\sum_{i=r}^{n} C_n^k p^k q^{n-k} (1.11)$$

Then, according to (10) and (11), $k$the probability that the number of events occurring in the Bernoulli scheme is not less than

$$P_n(k) + P_n(k+1) + \cdots + P_n(n)$$

Therefore, $X_{(r)} - r -$the form of the d.f. of the ordinal statistics is calculated as follows.

$$P\{X_{(r)} \le x\} = \sum_{i=r}^{n} C_n^i F^i(x) \cdot [1 - F(x)]^{n-i} \quad (12)$$

Here, $i -$the binomial probability of -had being $x_1, x_2, \dots, x_n$exactly equal to or less than $i$the value obtained from x. Therefore, we can write (12) as follows.

$$F_{(r)} = E_{p(x)}(n, r) \quad (13)$$

In this case, we know that the function is often Egiven in tabular form in sources. In addition, the well-known relationship between binomial sums and incomplete beta functions leads to the following.

$$F_{(x)}(x) = I_{p(x)}(r, n - r + 1)$$

here $I_p(a,\ b)$is known according to (12).

Formulas (8), (9), (12) are valid for both continuous and discrete r.v..

a) continuous case:

Now we $p(x)$ take the continuous r.v., which is the density of the distribution.$X_i -$

If $f_{(r)}(x)$ - $X_{(r)}$. If r.v. is the density function, then the following equality follows from (1.14).

$$f_{(r)}(x) = \frac{1}{B(r,\ n-r+1)} \frac{d}{dx} \int_0^{p(x)} t^{r-1}(1-t)^{n-r} dt =$$

$$= \frac{1}{B(r,\ n-r+1)} P^{r-1}[1 - P(x)]^{n-r}\, P(x)(1.15)$$

Since this formula is important in science, we will look at it in another way.

$x < X_{(r)} < x + \delta x$. The event can be implemented in the following way.

$r-1$	1	$n-r$
x		$x+\delta x$

1) $x \ge x_i$ for x_i from $r-1$
2) $x < X_i < x + \delta x$ for X_i
3) $X_i > x + \delta x$ for the amount X_i of $n - r$.

n The number of observation methods can be divided into 3 groups, namely

$$\frac{n\ !}{(r-1)! \cdot 1 \cdot (n-r)!} = \frac{1}{B(r,\ n-r+1)}$$

and each of them has the following probability.

$$P^{r-1}(x)[P(x+\delta x) - P(x)][1 - P(x+\delta x)]^{n-r}$$

δx. Taking is small, we obtain the following.

$$P\{x < X_{(r)} \le x + \delta x\} =$$

$$= \frac{1}{B(r,\ n-r+1)} P^{r-1} p(x)\delta x[1 - P(x+\delta x)]^{n-r} + 0(\delta x^2)$$

It means very small relative to the limit and involves $x < X_{(r)} < x + \delta x$ the implementation of a process in which $X_i$more than one of the values $(x,\ x+\delta x)$ falls into the interval.

Dividing both sides of the equation $\delta x$by and δx setting to zero, we obtain formula (15).

b) Discrete case.

If $p(x)$ $x = 0,\ 1,\ 2,\ \dots$.

Let us define $X_{(r)}$ the values. This function $f_{(r)}(x) = P\{X_{(r)} = x\}$is called the distribution of the statistic. From (1.13) and (1.14) the following follows.

$$f_{(r)}(x) = F_{(r)}(x) - F_{(r)}(x-1) = E_{p(x)}(n, r) - E_{p(x-1)}(n, r) = I_{p(r)}(r,\ n-r+1) - I_{p(x)}(r, n-r+1) \quad (16)$$

Although this expression seems convenient for calculation, there is another version of it, attributed to Khatri (1962).

$x < y$ Let's create the following table for.

$r-1-i$	$1+i+t$	$s-r-1-u-t$	$1+j+u$	$n-s-j$
	x		y	

$$f_{rs}(x,y) = \sum_{i=o}^{r-1}\sum_{j=0}^{n-s}\sum_{u,t}\frac{n!}{(r-1-i)!\,(1+i+t)!\,(s-r-1-u-t)!\,(1+j+u)!\,(n-s-j)!}$$
$$\cdot P(x-1)^{r-1-i}[p(x)]^{1+i+t}[P(y-1)-P(x)]^{s-r-1-u-t}\cdot[p(y)]^{1+y+u}\cdot[1-P(y)]^{n-s-j}$$

Here $\sum_{u,t}$ represents the sum over $u+t<s-r-1$ nonnegative and integer values satisfying the inequality $u,\ t$.

$$C_{rs} = \frac{n!}{(n-1)!(s-r-1)!(n-s)!}$$

By introducing notation, we get the following.

$$f_{rs} = (x,y) = C_{rs}\sum_{i=o}^{r-1}\sum_{i=o}^{n-s}\sum_{u,t} C_{r-1}^{i}C_{n-s}^{j}\frac{(s-r-1)!}{(s-r-1-u-t)!\,u!\,t!}$$
$$\cdot[P(x-1)]^{r-1-i}[1-P(y)]^{n-s-j}[P(y-1)-P(x)]^{s-r-1-u-t}$$
$$\cdot[p(x)]^{1+i+t}[p(y)]^{1+j+u}\int_0^1\int_0^1 z^i(1-z)^t z'^j(1-z')^u dzdz'$$

By interchanging the sum and integral signs, and $v = P(y) - z'p(y)$

$\omega = P(x-1) + zp(x)$ If we make the substitutions.

$$f_{rs}(x,y) = C_{rs}\int_{P(x-1)}^{P(x)}\int_{P(y-1)}^{P(y)}\omega^{r-1}\,(v-\omega)^{s-r-1}(1-v)^{n-s}dvd\omega \quad (17)$$

we will have.

If $x = y$ so, then

$$f_{rs}(x,x) = C_{rs}\iint \omega^{r-1}(v-\omega)^{s-r-1}(1-v)^{n-s}dvd\omega \quad (18)$$

It will be. Here the domain of integration is

$$P(x-1) \le \omega < v \le P(x)$$

In equality(17), $\omega < v$ the inequality is valid.

So, here we come to the following general conclusion.

$$f_{rs}(x,y) = C_{rs}\iint \omega^{r-1}(v-\omega)^{s-r-1}(1-v)^{n-s}dvd\omega \quad (19)$$

Here, $\omega \le v$ integration is taking place by sector.

Now we find the joint distribution function of two or more order statistics. $X_{(r)}$We defined the joint density of $f_{rs}(x,y)$ and as . $X_{(s)}$

If $x < X_{(r)} \le x + \delta x$ $y < X_{(s)} \le y + \delta y$the event can be represented as a configuration, then the expression corresponding to (14) can be written in the following form.

$r-1$	1	$s-r-1$	1	$n-s$
x		$x+\delta x$ y		$y+\delta y$

All observations smaller than , $r-1$one of which $(x,\ x+\delta x)$falls in the interval. Here, $x \le y$the following follows for .

$$f_{rs}(x,y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}P^{r-1}(x)\cdot[P(y)-P(x)]^{s-r-1}\cdot p(y)[1-P(y)]^{n-s} \quad (20)$$

$x_1 \le x_2 \le \ldots \le x_k$ for $X_{(r_1)}, X_{(r_2)}, \ldots, X_{r(k)}$has the following form.

$$f_{r_1 r_2 \ldots r_k}(x_1 \ldots x_k) = \frac{n!}{(r_1-1)!\,(r_1-r_1-1)!\ldots(n-r_k)}$$

$$\times P^{r_1-1}(x_1)p(x_1)[P(x_2)-P(x_1)]^{r_1-r_2-1}p(x_2)\ldots$$
$$\times [1-P(x_k)]^{n-r_k} \quad (21)$$

If $x_0=-\infty$ $x_{k+1}=+\infty$ $r_{k+1}=n+1$it is found that (21) then the right-hand side can be written in the form (22).

$$n!\,[\prod_{i=1}^{k}P(x_i)]\prod_{i=0}^{k}\{\frac{[\{P(x_{i+1})-P(x_i)]^{r_{i+1}-r_k-1}}{(r_{i+1}-r_i-1)!}\} \quad (22)$$

In particular, n —the joint distribution density of all ordinal statistics has the following form:

$$n!\,p(x_1)p(x_2)\ldots\ldots p(x_n)$$

The final result is not necessarily accurate, x_i but the value of $n!$ has an equally likely ordering. Can be used as a starting point for inferring the joint distribution of continuous-state k —ordinal statistics.

$$F_{rs}(x,y)=\sum_{i=r}^{n}\sum_{\max(0,5-i)}^{n-i}\frac{n!}{i!\,j!\,(n-i-j)!}\times P^{i}(x)[P(y)-P(x)]^{j}\,[1-P(y)]^{n-i-j} \quad (23)$$

$x\geq y$ It follows that in the equation $X_{(r)}\leq x$ for , so $X_{(s)}\leq y$

$$F_{rs}(x,y)=F_s(y)$$

References

1. A.A.Abdushukurov “Ehtimollar nazariyasi”. Ma’ruzalar matni. Toshkent: «Universitet», 2000.

2. A.A.Abdushukurov, T.A.Azlarov, A.A.Djamirzayev “Ehtimollar nazariyasi va matematik statistikadan misol va masalalar to‘plami”.Toshkent 2003y.

3. В.Б.Невзоров "Рекорды математическая теория" выпуск 4 Фазис. Москва 2000
4. Wikipedia electronic encyclopedia.
5. http://www.el.tfi.uz/pdf/enmcoq22.uzl.pdf;
6. http://www.eknigu.com/lib/mathematics/