$O\Pi$ «Политология», 2019-20

Математика и статистика, часть 2

Ранги наблюдений. (28.05.2020)

А. А. Макаров, А. А. Тамбовцева, Н. А. Василёнок

Ранг — порядковый номер наблюдения в вариационном ряду, то есть в выборке, упорядоченной по возрастанию. Будем обозначать ранг буквой r, r_i — ранг i-того наблюдения в выборке.

Возможны два случая:

- выборка не содержит повторяющихся значений;
- выборка содержит повторяющиеся значения.

В выборке нет повторяющихся значений

Если в выборке нет повторящихся значений, ранг наблюдения – просто его порядковый номер в выборке, упорядоченной по возрастанию.

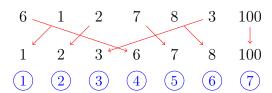
Пример 1. Дана выборка из 7 наблюдений:

Запишем вариационный ряд:

Подпишем номера наблюдений:

Запишем ранги: $r_1 = 4$, $r_2 = 1$, $r_3 = 2$, $r_4 = 5$, $r_5 = 6$, $r_6 = 3$, $r_7 = 7$.

Внимание: ранги определяются для наблюдений в исходной выборке. Например, здесь r_1 — это ранг первого наблюдения в выборке, то есть порядковый номер «шестерки» в вариационном ряду, равный 4.



Аналогично для остальных наблюдений.

В выборке есть повторяющиеся значения

Если в выборке есть повторяющиеся значения, то возникает необходимость считать средний ранг.

Пример 2. Дана выборка из 7 наблюдений:

6 1 2 7 8 2 100

Запишем вариационный ряд:

1 2 2 6 7 8 100

Для неповторяющихся значений ранги определяются обычным образом (точно так же, как в примере 1):

Для повторяющихся значений считается средний ранг. В данном случае у повторяющихся «двоек» порядковые номера в вариационном ряду (ранги) — это 2 и 3. Посчитаем средний ранг — среднее арифметическое этих чисел:

$$\frac{2+3}{2} = 2.5$$

Следовательно:

Запишем ранги: $r_1 = 4$, $r_2 = 1$, $r_3 = 2.5$, $r_4 = 5$, $r_5 = 6$, $r_6 = 2.5$, $r_7 = 7$.

Важно: дробные ранги – это нормально.

Пример 3. Дана выборка из 7 наблюдений:

6 1 7 7 8 7 100

Запишем вариационный ряд:

1 6 7 7 7 8 100

Сначала определим ранги неповторяющихся значений:

Порядковые номера повторяющихся «семерок» – это 3, 4, 5. Посчитаем средний ранг:

$$\frac{3+4+5}{3} = 4$$

Получаем:

Запишем ранги: $r_1=2,\,r_2=1,\,r_3=4,\,r_4=4,\,r_5=6,\,r_6=4,\,r_7=7$

Важно: то, что некоторых «промежуточных» чисел среди рангов нет (например, есть ранги, равные 2 и 4, но нет ранга, равного 3) – это тоже нормально.