РАСЧЕТ ВЫБРАННЫХ ХАРАКТЕРИСТИК СТАТИСТИЧЕСКОГО РАСПРЕДЕЛЕНИЯ УРОЖАЙНОСТИ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР

***Мамбетов Атамурат Баймуратович**, ассистент кафедры «Информационные технологии, математики, физики и химии» Каракалпакский институт сельского хозяйства и агротехнологии.*
***Губенова Галия Турганбаевна**, ассистент кафедры «Информационные технологии, математики, физики и химии» Каракалпакский институт сельского хозяйства и агротехнологии.*
***Алламуратова Венера Жумамуратовна**, ассистент Каракалпакский госуниверситет.*

Аннотация: Статья носит практический характер и посвящена оценке и расчету значимых средних, средних геометрических, среднеквадратичных и других количественных показателей растений на основе выборочных данных. Показана значимость формул для расчета выборочной дисперсии, выборочного среднеквадратического отклонения, моды, медианы и коэффициентов вариации в раскрытии фундаментальности исследуемой случайной величины.

Ключевые слова: случайная величина, выборочные данные, выборочная дисперсия, выборочное среднеквадратичное отклонение, мода, медиана, коэффициент вариации.

Вводный раздел: При решении практических задач в области сельского хозяйства необходимо оценивать важные числовые характеристики изучаемого числа на основе выборочных данных. Определение средней урожайности хлопчатника, пшеницы и риса в текущем году, среднегодового прироста культурных деревьев приводит к расчету выбранных характеристик статистического распределения[2].

Учитывая статистическое распределение X –количества символов,

X	x_1	x_2	…	x_k
N	n_1	n_2	…	n_k

размер выборки в этом $\sum_{i=1}^{k} n_i = N$;

Цена, определяемая одним числовым значением, называется точечной статистической ценой. Например, выборочное среднее, выборочная дисперсия,

ISSN 2181-3187

скорректированная выборочная дисперсия, среднеквадратичное отклонение, центральный момент: $\bar{x}_t, D_t, S_T^2, б_T, m_k$ и другие статистические точечные оценки.

Теоретическая часть: Выборочное значение: это статистическая оценка математического ожидания в теории вероятностей, равного среднему арифметическому значению $\bar{x}_t$ выборочного набора.

Если все наблюдаемые значения X в выборке размера n равны $x_1, x_2, \dots, x_n$ в противном случае мы вычисляем выборочное среднее, используя приведенную ниже формулу.

$$\bar{x}_T = \frac{1}{n}\sum x_j = \frac{1}{n}(x_1 + x_2 + \dots + x_k)\,; \quad (1)$$

Если значения $x_1, x_2, \dots, x_k$ количества символов X имеют абсолютные частоты $n_1, n_2, \dots, n_k$ раз соответственно, мы вычисляем выбранное среднее значение по следующей формуле.

$$\bar{x}_T = \frac{1}{n}\sum_{j=1}^{k} x_j n_j = \frac{1}{n}(x_1 n_1 + x_2 n_2 + \dots + x_k n_k); \quad (2)$$

Расчет различных средних значений. В математической статистике, в зависимости от задачи, могут быть рассчитаны различные средние значения, такие как среднее арифметическое, среднее геометрическое, среднее квадратичное и другие средние значения. Мы получаем все это из общей формулы среднего значения.

$$v_k = \left(\frac{1}{n}\sum_{i=1}^{k} x_i^k n_i\right)^{\frac{1}{k}}; \quad (3)$$

а) Из формулы (2) получаем формулу (3) для значения k=1.

$$\bar{x}_T = v_1 = \frac{1}{n}\sum_{i=1}^{k} x_i n_i\,;$$

б) (3) - если мы находим k=-1 из формулы, мы вычисляем следующее среднее значение гармоники.

$$\bar{x}_T = v_{-1} = \frac{n}{\sum \frac{n_i}{x_i}}; \quad (4)$$

в) (3) - даем следующее среднеквадратичное значение при k=2 по формуле.

$$\bar{x}_T = v_2 = \left[\frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i\right]^{\frac{1}{2}};$$

всегда выполняется неравенство для этих выбранных характеристик.

$$\bar{x}_{гарм} < \bar{x}_{арифм} < \bar{x}_{квадр};$$

ISSN 2181-3187

Если статистическое распределение задано в виде вектора вариации интервала, то в этом случае среднее значение $i-$ интервала принимается за x_i при вычислении выборочного среднего значения, а статистические условия изучаются, как указано выше, для полученного дискретного статистического распределения и рассчитываются характеристики выборки.

Выборочная дисперсия: выборочное среднее не дает полной информации о статистическом распределении. Существуют случайные величины с разными распределениями, но с одинаковым математическим ожиданием. Кроме того, необходимо знать распределение значений наблюдаемого варианта x_i вокруг его выборочного среднего значения $\bar{x}_T$. Например, имеет практическое значение знать, насколько годовой доход рабочих на предприятии отличается от среднемесячного дохода рабочего[1].

Выборочная дисперсия характеризует распределение значений наблюдаемых вариантов вокруг выборочного среднего значения $\bar{x}_T$. Если размер равен n, а значения $x_1, x_2, \dots, x_k$ имеют абсолютные частоты $n_1, n_2, \dots, n_k$ соответственно, мы вычисляем выборочную дисперсию по следующей формуле:

$$D_T = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x}_T)^2 n_i; \qquad (5)$$

Если значения $x_1, x_2, \dots, x_n$ символа различаются, мы вычисляем выборочную дисперсию по следующей формуле:

$$D_T = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x}_T)^2; \qquad (6)$$

эту формулу легко использовать, когда значения вариантов представляют собой небольшие числа.

Среднеквадратичное отклонение выборки: Квадратный корень выборочной дисперсии называется среднеквадратичным отклонением выборки и обозначается $б_T$.

$$б_T = \sqrt{D_T} = \sqrt{\frac{1}{n}\sum(x_i - \bar{x}_T)^2 n_i}\,; \qquad (7)$$

Мода: значение варианта с наибольшей частотой статистического распределения и обозначается M_0.

Медиана: это значение варианта, которое делит статистическое распределение на две равные части, и мы анализируем его в соответствии со следующим:

$$M_e = \begin{cases} x_{m+1}, & \text{если } k = 2m+1, \\ \dfrac{(x_m + x_{m+1})}{2}, & \text{если } k = 2m\,. \end{cases} \qquad (8)$$

ISSN 2181-3187

Коэффициент вариации: мы используем коэффициент вариации при сравнении значений разных выборок, распределение расположения вокруг среднего значения:

$$v_T = \frac{б_T}{\bar{x}_T} 100\%; \quad (9)$$

Все характеристики выборки являются случайными величинами, которые являются функцией выборки. Их значения меняется в зависимости от размера выборки.

Функцию $\Theta_n^* = \Theta_n^*(x_1, x_2, \dots, x_n)$, сформированную из набора выборок как -ϴ, назовем статистической оценкой неизвестного параметра теоретического распределения.

Чем меньше разница между Θ_n^*- ϴ, тем точнее значение Θ_n^* для параметра ϴ.

Практическая часть: Теперь рассчитаем характеристики приведенной выше выборки и сделаем теоретические и практические выводы об исследуемой случайной величине.

Пример: чтобы оценить среднее количество картошек, выращенных на одно растение, мы случайным образом выбираем 50 растений с поля. Количество картофеля в каждом из этих растений следующее:

6, 7, 5, 8, 3, 7, 9, 5, 8, 7, 4, 6, 8, 7, 5, 8, 10, 6, 7, 8, 9, 7, 8, 6, 9, 6, 7, 5, 10,9, 7, 8, 6, 11, 7, 5, 4, 6, 7, 8, 10, 6, 7, 8, 11, 9, 7, 8, 10, 12.

1) На основании этих данных представляем дискретный вектор вариации и полигон выборочной совокупности[3].

2) Находим важные числовые характеристики построенного статистического распределения: выборочное среднее, выборочную дисперсию, среднеквадратичное отклонение, моду, медиану и коэффициент вариации[2].

Решение: 1) Записываем значения X, количество картофелин, снятых с каждого из 50 выкопанных кустов, в порядке возрастания, то есть составляем вариационный ряд:

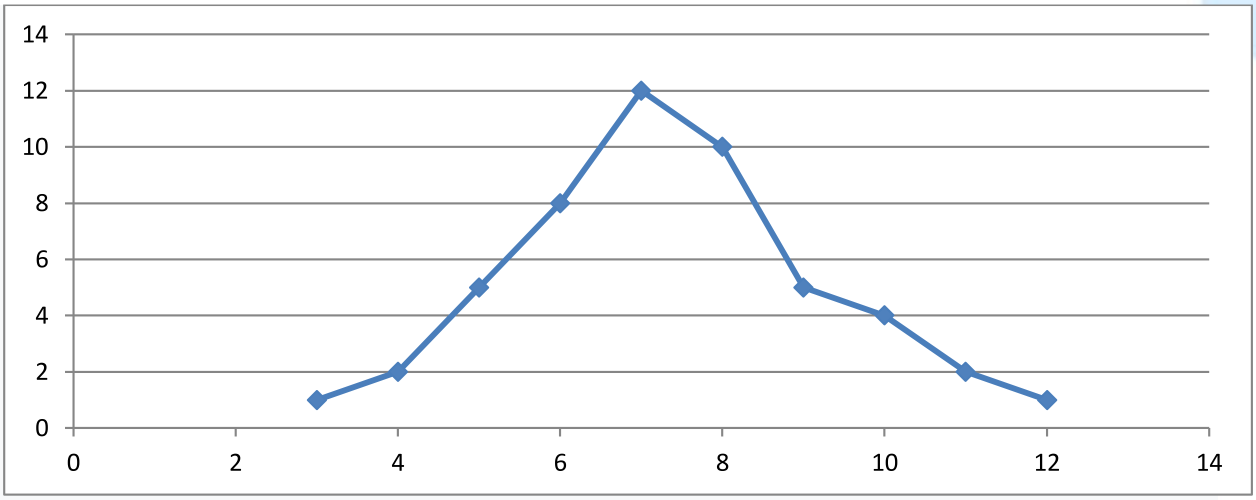
3, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 9, 9, 9, 9, 9, 10, 10, 10, 10, 11, 11, 12.

Статистическое распределение выборки создается с помощью построенной ряд вариаций:

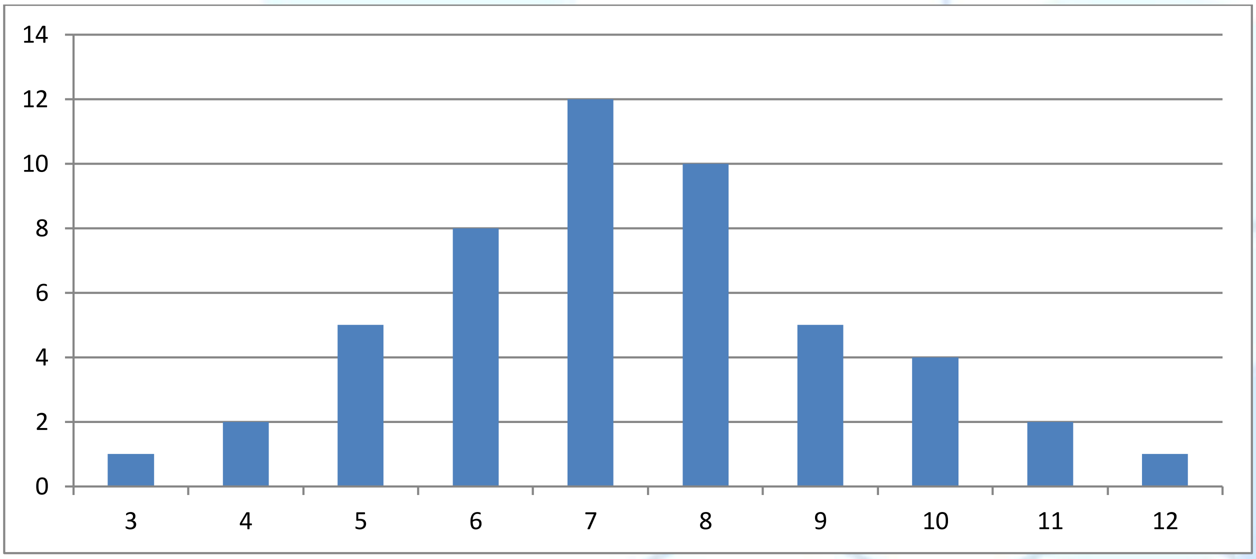
x_i	3	4	5	6	7	8	9	10	11	12
n_i	1	2	5	8	12	10	5	4	2	1

В этом случае объем выборки составляет $n = 50$.

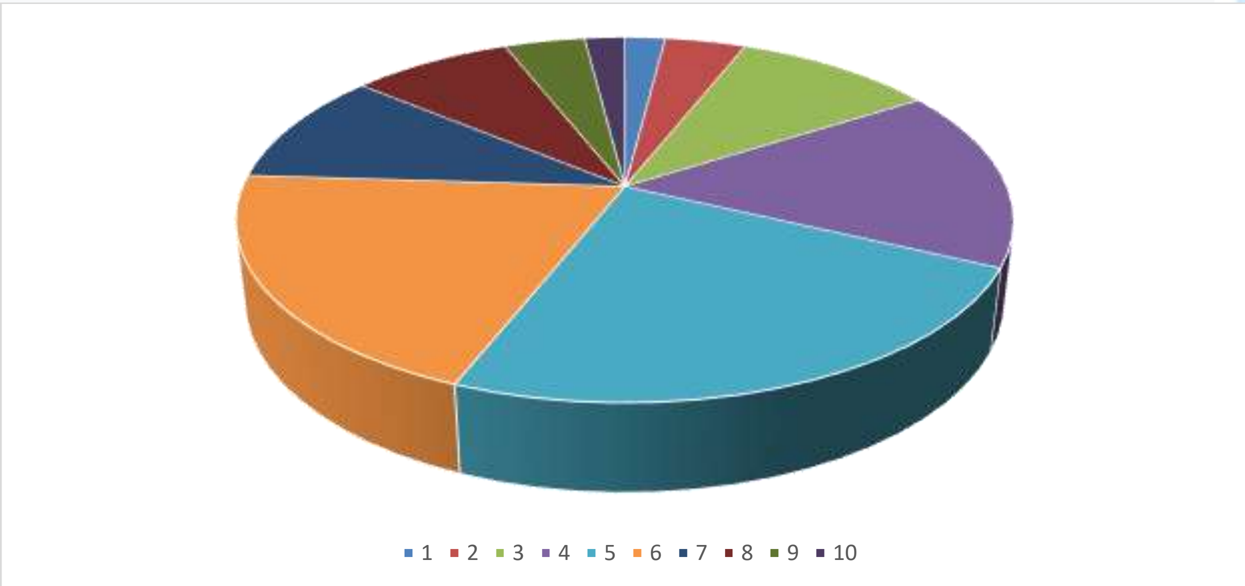
a) Нарисуйте многоугольник[3];



б) Гистограмма[3];



в) Круговая диаграмма [3];



2) Приведем расчет характеристик выборки с использованием статистического распределения:

ISSN 2181-3187

a) Поскольку значения X_i повторялись, вычисляем среднее значение выборки как обычно,

$$\bar{x}_T = \frac{1}{n}(n_1x_1 + n_2x_2 + n_3x_3 + n_4x_4 + n_5x_5 + n_6x_6 + n_7x_7 + n_8x_8 + n_9x_9 + n_{10}x_{10}) = \frac{1}{50}(1*3 + 2*4 + 5*5 + 8*6 + 12*7 + 10*8 + 5*9 + 4*10 + 2*11 + 1*12) = \frac{1}{50}(3 + 8 + 10 + 48 + 84 + 80 + 45 + 40 + 22 + 12) = 7{,}04;$$

б) Мы рассчитываем выборочную дисперсию следующим образом:

$$D_T = \frac{1}{n}\sum_{i=1}^{10}(x_i - \bar{x}_T)^2 n_i = \frac{1}{50}[(3-7{,}04)^2*1 + (4-7{,}04)^2*2 + (5-7{,}04)^2*5 + (6-7{,}04)^2*8 + (7-7{,}04)^2*12 + (8-7{,}04)^2*10 + (9-7{,}04)^2*5 + (10-7{,}04)^2*4 + (11-7{,}04)^2*2 + (12-7{,}04)^2*1 = 3{,}67;$$

в) среднеквадратичное отклонение, $б_T = \sqrt{D_T} = \sqrt{3{,}67} = 1{,}91;$

г) способ статистического распределения; $M_0 = 7$, медиана потому что количество вариантов четное 2k=10, k=5; $M_e = \frac{1}{2}(x_5 + x_6) = \frac{1}{2}(7+8) = 7{,}5;$

крыла вариации $R = x_{max} - x_{min} = 12 - 3 = 9,$

коэффициент вариации $V_k = \frac{б_T}{\bar{x}_T} * 100\% = \frac{1{,}91}{7{,}04} * 100\% = 27{,}2\%.$

Вывод: выборочная совокупность должна быть отделена от основной выборки таким образом, чтобы сохранить основные характеристики основной выборки или получить репрезентативную выборочную совокупность. Для того чтобы выборка была репрезентативной, выборка должна быть случайной и все элементы генеральной совокупности должны иметь одинаковую вероятность попадания в выборку. В противном случае статистические исследования могут привести к неверным выводам.

Список литературы

1. Б.А. Колемаев,О.Б. Староверов, В.Б.Трундаевский “Теория вероятностей и математическая статистика”. М. ВШ, 1991. 400стр.
2. A.A. Fayziev, B.Rajabov, L.Rajabova “Oliy matematika, ehtimollar nazariyasi va matematik statistika”, Tashkent. TashDAU, 2014,306 bet.
3. Гарнаев А.Ю. Использование MS EXCEL и VBA в экономике и финансах. –СПб.: БХВ – Санкт-Петербург, 1999. -336 с., ил.