

СТАТИСТИЧЕСКИЙ АНАЛИЗ И ПРОГНОЗИРОВАНИЕ ДИНАМИКА УРОЖАЙНОСТИ ПШЕНИЦЫ В ХОРЕЗМСКОЙ ОБЛАСТИ РЕСПУБЛИКЕ УЗБЕКИСТАН

А.А. Файзиев, С.К. Уринбаев

Ташкентский экономический и педагогический институт

Аннотация. В статье, методом статистического анализа временных рядов, изучена статистическая закономерность рядов динамики \bar{y}_t – средней урожайности пшеницы Хорезмской области в республике Узбекистан (по материалам ЦСУ РУз за 2012-2024 годы). С помощью статистических критериев Дарбина-Уотсона установлено, что средняя урожайность пшеницы имеет автокорреляционную зависимость.

Ключевые слова: дискретный, динамический, ряд, пшеница, тренд, сезонность, компонента, линейный, наименьший, нормальный, гипотеза, автокорреляция, эксцесса.

Annotatsiya. Ushbu maqolada O'zbekiston Respublikasi Xorazm viloyatida o'rtacha bug'doy hosildorligining (y_t) vaqt qatorlarining statistik muntazamligini o'rganish uchun vaqt qatorlari statistik tahlili qo'llaniladi (O'zbekiston Respublikasi Markaziy statistika boshqarmasining 2012-2024 yillardagi ma'lumotlari asosida). Durbin-Uotson statistik testidan foydalanib, o'rtacha bug'doy hosildorligi avtokorrelyatsiyani namoyon qilishi aniqlandi.

Kalit so'zlar: diskret, dinamik, qator, bug'doy, trend, mavsumiylik, komponent, chiziqli, eng kichik, normal, gipoteza, avtokorrelyatsiya, ortiqcha.

Annotation. In the article, using the method of statistical analysis of time series, the statistical pattern of the dynamics series $\{y_t, t \in T\}$ of the average wheat yield in the Khorezm region of the Republic of Uzbekistan (based on the data of the Central Statistical Bureau of the Republic of Uzbekistan for 2012-2024) is studied. Using the statistical criteria of the Darbin-Watson, it has been established that the average wheat yield has an autocorrelation dependence.

Keywords: discrete, dynamic, series, wheat, trend, seasonality, component, linear, least, normal, hypothesis, autocorrelation, asymmetry, excess.

Введение. При обработке временных рядов методы во многом опираются на разработанные математической статистикой методы для рядов распределения. К настоящему времени статистика располагает разнообразными методами анализа временных рядов [1,2,3,4,5].

Можно выделить три основные задачи исследования временных рядов. Первая из них заключается в описании изменения соответствующего показателя

во времени и выявлении тех или иных свойств исследуемого ряда. Для этого прибегают к разнообразным способам: расчету обобщающего показателя изменения уровней во времени и среднего темпа роста; применению различных сглаживающих фильтров, уменьшающих колебания уровней во времени и позволяющих более четко представить тенденции развития; подбору кривых, характеризующих эту тенденцию; выделению сезонных и иных периодических и случайных колебаний; измерению зависимости между членами ряда (автокорреляции). Второй задачей анализа является объяснение механизма изменения уровней ряда, для ее решения обычно прибегают к регрессионному анализу. В третьих описание изменений временного ряда и объяснения механизма формирования рядов часто используются для статистического прогнозирования, которое в большинстве случаев сводится к экстраполяции обнаруженных тенденций развития. Перечисленные задачи решаются с помощью различных методов [1-5].

Анализ литературы. В данной работе используя методы статистического анализа временных рядов построены точечные и интервальные оценки для средней урожайности пшеницы Хорезмский области в республике Узбекистан, определены явные виды трендов и прогнозирован урожай для последующих лет.

В общем случае временной ряд $\{y_t, t \in T\}$ состоит из четырех составляющих: тренд; колебания относительно тренда; эффект сезонности; случайная компонента.

Изучению и анализу динамических рядов посвящены работы: Андерсона[1], Кендала[2], Тихомирова [3], Сулайманова [4], Файзиева [5] и другие.

Методология исследования. Предположим, что статистическая закономерность рядов динамики \bar{y}_t –средней урожайности пшеницы Хорезмский области в республике Узбекистан за период наблюдений, 2012-2024 годы образует дискретный временной ряд. Используя выше изложенных методы статистического анализа временных рядов, построим точечные и интервальные оценки для статистическая закономерность рядов динамики \bar{y}_t –средней урожайности пшеницы Хорезмский области в республике Узбекистан, определим явные вид тренда и прогнозируем средней урожайности пшеницы для последующих лет, проверим различные статистические гипотезы. На рисунки-1, с помощью опытных данных (таблица-1, столбец-3) геометрически изображены \bar{y}_t – средней урожайности пшеницы Хорезмский области в республике Узбекистан виде а) точечная график, б) гистограмма, в) круговая диаграмма, д) область диаграммы:

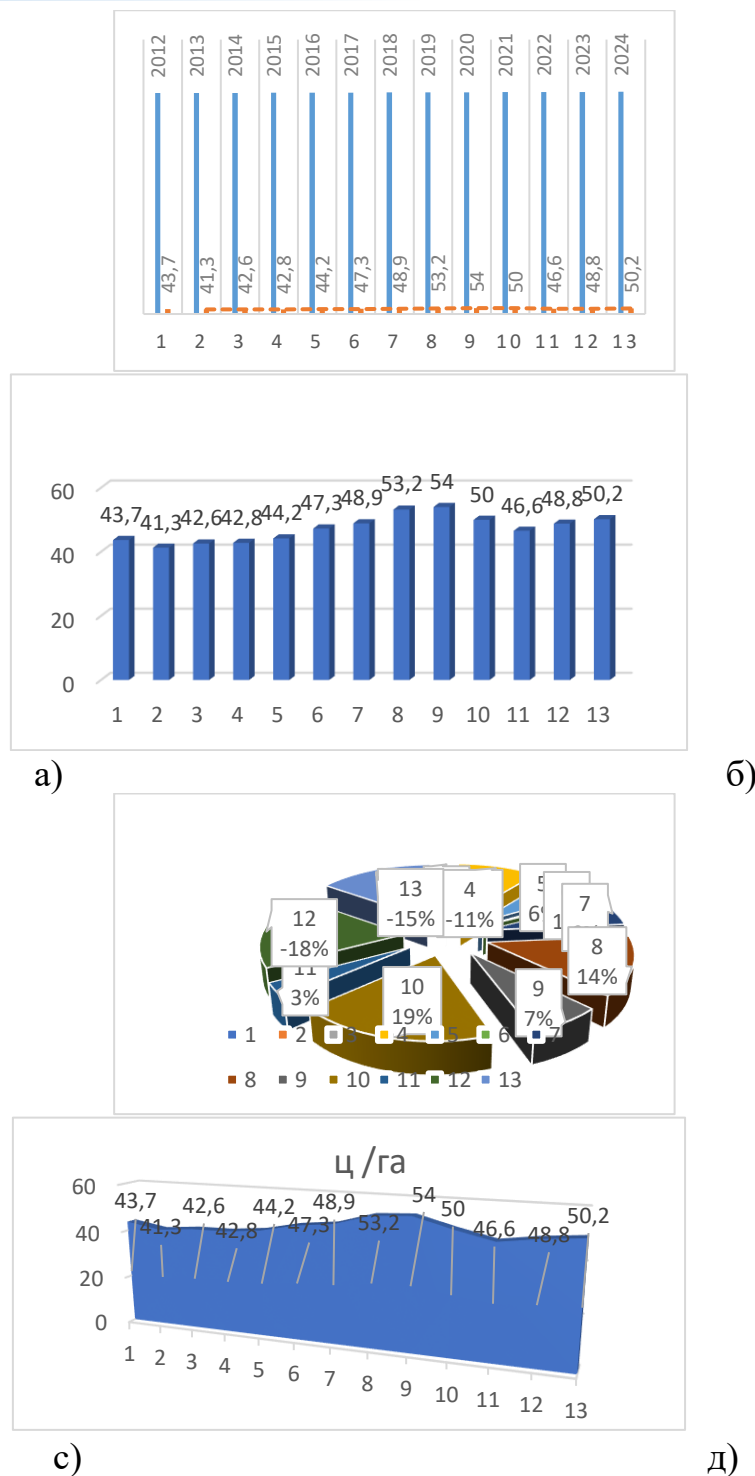


Рис.1.

Геометрическое изображение наблюденных данных, система координат Декарта дают основание в первом приближении, предполагать гипотезу что, трендовая часть процесса (общее направление развития процесса) имеет линейную зависимость следующего вида:

$$y(t) = a_1 t + a_0 \quad (1)$$

Где неизвестные параметры определяются методом наименьших квадратов(мнк), т.е. на основании опытных данных, решая следующую систему нормальных уравнений:

$$\begin{cases} a_0 T + a_1 \sum t = \sum y_t \\ a_0 \sum t + a_1 \sum t^2 = \sum y_t t \end{cases} \quad (2)$$

Решая уравнение (1) находим:

$$a_0 = \frac{1}{T} \sum y_t, \quad a_1 = \frac{1}{\sum t^2} \sum y_t t. \quad (3)$$

Составим таблица для вычисления тренда временного ряда.

Таблица -1

1	2	3	4	5	6	7
N п/п	Годы наблю- дений	y_t - ц /га	t	t ²	ytt	ytt ²
1	2012	43,7	-6	36	-262,2	1573,2
2	2013	41,3	-5	25	-206,5	1032,5
3	2014	42,6	-4	16	-170,4	681,6
4	2015	42,8	-3	9	-128,4	385,2
5	2016	44,2	-2	4	-88,4	176,8
6	2017	47,3	-1	1	-47,3	47,3
7	2018	48,9	0	0	0	0
8	2019	53,2	1	1	53,2	53,2
9	2020	54	2	4	108	216
10	2021	50	3	9	150	450
11	2022	46,6	4	16	186,4	745,6
12	2023	48,8	5	25	244	1220
13	2024	50,2	6	36	301,2	1807,2
Сумма		613,6	0	182	139,6	8388,6

Используя, вычисления по таблице 1, имеем:

$$\sum y_t = 613,6, \quad a_0 = \frac{1}{T} \sum y_t = \frac{613,6}{13} = 47,2 \quad a_1 = \frac{1}{\sum t^2} \sum y_t t = \frac{139,6}{182} = 0,77.$$

Отсюда, находим уравнение линейного тренда (тенденция) для средней урожайность пшеницы Хорезмский областе в республике Узбекистан [1,2,3,4,5]:

$$y(t) = 0,77 t + 47,2 \quad (2)$$

В частности, подставляя в уравнение (2) значение t=2 находим ожидаемые средней урожайность пшеницы Хорезмский областе в республике Узбекистан 2026 году, будет 48,74 ц/га.

Анализ материала и результаты исследования

С помощью статистических критериев ([1] – [5]) установлено, что в уравнении (2) $y(t) = a_1 t + a_0$ основная гипотеза $H_0: a_1 = 0$ отвергается и принимается альтернативная гипотеза $H_1: a_1 \neq 0$ с уровень значимости $\alpha = 0,05$. Следовательно, средней урожайность пшеницы Хорезмский областе в республике Узбекистан имеет линейный тренд.

Отметим, что автокорреляцией представляет собой корреляционную зависимость между последующими и предшествующими членами временного ряда. Проверяем наличие автокорреляции для рассматриваемого процесса, т.е.

$$y_t = \rho y_{t-1} + \varepsilon_t,$$

$$\text{где } \rho = \text{Cov}(y_t, y_{t+1}) = M[(y_t - \bar{y}_t)(y_{t+1} - \bar{y}_t)].$$

Для дальнейшей исследование вычислим следующие конечные разности по наблюдаемым данным:

$$\Delta Y_t = Y_{t+1} - Y_t, \quad \Delta^2 Y_t = \Delta Y_{t+1} - \Delta Y_t, \quad \Delta^3 Y_t = \Delta^2 Y_{t+1} - \Delta^2 Y_t, \quad \Delta Y_t = Y_{t+1} - Y_t, \quad \Delta^2 Y_t = \Delta Y_{t+1} - \Delta Y_t, \\ \Delta^3 Y_t = \Delta^2 Y_{t+1} - \Delta^2 Y_t, \dots$$

$$V_k = \frac{\sum_{t=k}^T (\Delta^k y_t)^2}{(T-k)C_{2k}^k}$$

Вычисляя

(3)

коэффициенты вариации разностей и установлено, что $V_1 \approx V_2 \approx V_3$. Следовательно, конечные разности первого порядка элиминируют линейную тенденцию.

С другой стороны проверим гипотезу, существования автокорреляционная зависимость между средней урожайность пшеницы Хорезмский области в республике Узбекистан с помощью критерия Дарбина – Уотсона :

$$d = \frac{\sum_{t=1}^{T-1} (Y_{t+1} - Y_t)^2}{\sum_{t=1}^T Y_t^2} \quad (4)$$

Вычисляется по формуле (4) $d_{\text{наб}} = \frac{75,11}{29163,4} = 0,03$ и сравнивается с $d_{\text{крит}} = 1,08$ табличным значением [1],[2],[5]. Поскольку $d_{\text{наб}} = 0,03 < d_{\text{крит}} = 1,08$. Следовательно, критерия Дарбина – Уотсона 95% тоже гарантией доказывает, что средней урожайность пшеницы Хорезмский области в республике Узбекистан имеет автокорреляционную зависимость $Y_t = \rho Y_{t-1} + \varepsilon_t$. Следовательно, средней урожайность пшеницы области в этом году, зависит от урожайности прошлых и последующих лет.

Проверка статистическая гипотезы о нормальности, \bar{y}_t – средней урожайность пшеницы Хорезмский области в республике Узбекистан [1-5]:

$$H_0 : P(\bar{y}_t < x) = \Phi_{a,\sigma}(x), \quad H_1 : P(\bar{y}_t < x) \neq \Phi_{a,\sigma}(x)$$

принимается уровень значимости $\alpha = 0,05$.

Тогда с помощью следующие формулы построим интервальные оценка для средней урожайность пшеницы:

$$\bar{Y}_{T+i} - t(T-2; \alpha) \bar{\sigma}_y \leq a_0 + a_1(T+i) \leq \bar{Y}_{T+i} + t(T-2; \alpha) \bar{\sigma}_y \quad (5)$$

$$\bar{\sigma}_y = \bar{\sigma} \left[\frac{\frac{1}{T} + \left(\frac{T-1}{2} + i \right)^2}{\sum_{i=1}^T (t - \bar{t})^2} \right]^{0.5};$$

где

Значение $t_{\text{крит.}} = t(T-2; \alpha)$ определяется по таблице Стьюдента.

Кроме того, гипотезы о нормальности средний урожайность пшеницы можно проверить с помощью параметрических гипотезы. Если для исследуемой случайной величины одновременно выполняются следующие условия([1] – [5]):

$$|A_s| < 1,5\sigma_1, \quad \left| E_k + \frac{\sigma}{T+1} \right| < 1,5\sigma_2 \quad (6)$$

тогда можно предположить, что случайная величина с достоверностью 95% имеет примерно нормальное распределение.

Вычисление показывает, что условия (6) выполняется. Следовательно, тогда можно предположить, что случайная величина \bar{y}_T – имеет примерно нормальное распределение с достоверностью 95% .

Используя формула (5), построим с вероятностью 0.95 интервальная оценка для \bar{y}_t –средней урожай пшеницы Хорезмский области:
(44,74; 49,66) ц/га.

Выводы. На основании выше изложенных статистических анализов, динамики \bar{y}_t –средней урожайность пшеницы Хорезмский области в республике Узбекистан как дискретный временный ряда с вероятностью $\gamma = 0,95$ (таблица-4) можно сделать следующие выводы:

Построены точечные $\bar{y}_T = 47,2$ ц/га и интервальные статистические оценки для \bar{y}_t – средней урожайность пшеницы с 95% гарантией
(44,74; 49,66) ц/га ;

2. Определены явные виды тренда и установлена её линейность

$$y(t) = 0,77 t + 47,2;$$

Критерием Дарбина – Уотсона установлены, что, средней урожайность пшеницы области имеет автокорреляционную зависимость $Y_t = \rho Y_{t-1} + \epsilon_t$, где $\rho = \text{Cov}(Y_t, Y_{t+1}) =$

$M[(Y_t - \bar{y}_t)(Y_{t+1} - \bar{y}_t)]$ т.е. в этом году, зависит от урожайности прошлых и последующих лет;

4. Отметим, что в республике Узбекистан 2025 году, заготовлена более 8,1 млн. тонна зерно.

Литература

Т.Андерсон “Статистический анализ временных рядов”. –Москва: “МИР”, 1976. – 759 с.

М. Кендал, А. Стьюарт “Многомерный статистический анализ и временные ряды.- Москва: “Наука”, 1976. -736 с.

Н.П.Тихомиров «Эконометрика».- Москва: «Экзамен», 2003. – 512 с.

Б.А.Сулайманов, А.А.Файзиев, Ж.Н. Файзиев "Тажриба маълумотларининг статистик таҳлили".– Тошкент: ТошДАУ, 2015. – 124 с.