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

М.М. Дустмуродов
доцент, Республика Узбекистан
А.А. Файзиев
к.ф-м.н, профессор, Ташкентский экономический и педагогический университет, Республика Узбекистан

https://doi.org/10.5281/zenodo.18303875

АННОТАЦИЯ

Почти в каждой области встречаются явления , которое интересно и важно изучать в их развитии и изменении во времени. Временным(динамическим) рядом называется последовательность значений некоторого показателя во времени. Наблюдения над некоторыми явлениями, характер которых меняется во времени, ведут упорядоченной последовательности, которую называют временным рядом. Цели изучения временных рядов, стремиться предсказать будущее на основании знания прошлого и управлять процессом, описать характерные особенности ряда на основании ограниченного количества информации. В статье, методом статистического анализа временных рядов, изучена статистическая закономерность рядов динамики $\overline{y_t}$ −средней урожайность пшеницы Сурхандарьинской областе в республике Узбекистан (по материалам ЦСУ РУз за 2008-2024 годы). Построены, с 95% ной гарантией точечные и интервальные оценки для средней урожайность пшеницы, определены явные виды трендов и прогнозирована урожайность пшеницы для

КЛЮЧЕВЫЕ СЛОВА

дискретный, динамический, ряд, пшеница, тренд, сезонность, компонента, линейный, наименьший, нормальный, гипотеза, автокорреляция, асимметрия, эксцесса.

последующих лет. С помощью статистических критериев Дарбина -Уотсона установлено, что средняя урожайность пшеницы имеет автокорреляционная зависимость. Следовательно, урожайность пшеницы Сурхандарьинской области в этом году, зависит от урожайности пшеницы прошлых и последующих лет.

Использованные методы обработки и анализ динамических рядов после апробации могут быть использованы в исследованиях магистров, научных работников.

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 изменении 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($[1-5]$).

Материалы и методы. В данной работе используя методы статистического анализа временных рядов построены точечные и интервальные оценки для средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан, определены явные виды трендов и

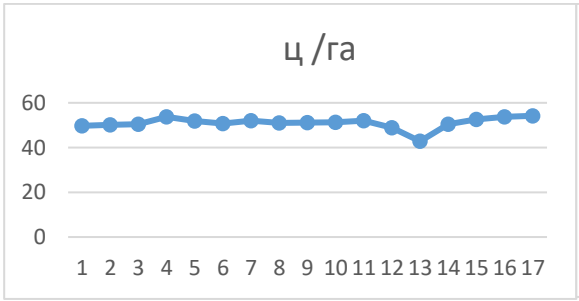
прогнозирован урожай для последующих лет, проверено различные статистические гипотезы.

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

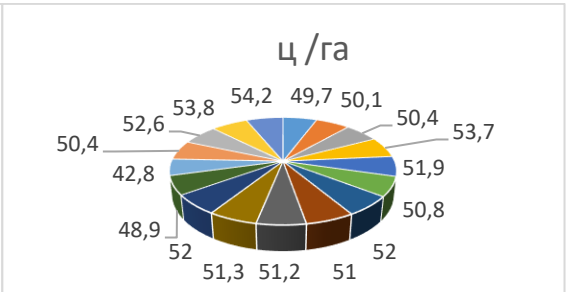
Изучению и анализу динамических рядов посвящены работы: Андерсона[1], Кендала[2], Тихомирова [3], Сулайманова [4] , Файзиева [5] и другие.

Результаты и их обсуждение. Предположим, что статистическая закономерность рядов динамики $\overline{y_t}$ −средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан за период наблюдений, 2008-2024 годы образует дискретный временный ряд. Используя выше изложенных методы статистического анализа временных рядов, построим точечные и интервальные оценки для статистическая закономерность рядов динамики $\overline{y_t}$ −средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан, определим явные вид тренда и прогнозируем средней урожайность пшеницы для последующих лет, проверим различные статистические гипотезы.

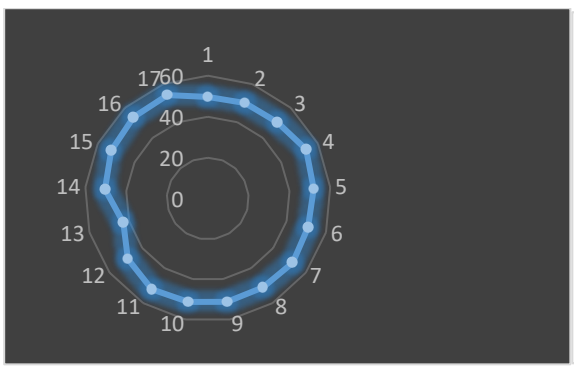
На рисунки-1, с помощи опытных данных (таблица-1, стольба-3) геометрически изо-бражены $\overline{y_t}$ − средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан виде а) точечная график, b) круговая диаграмма, c) область диограммы, d) гистограмма:

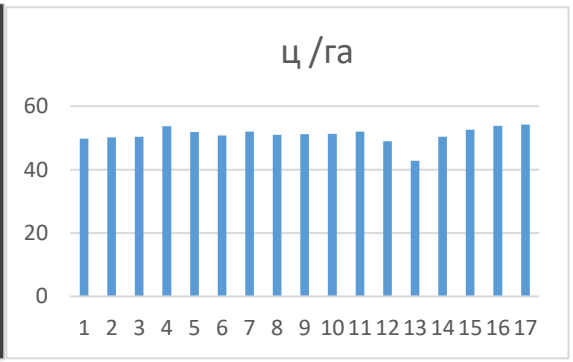


а)



б)





c) d)

Рис.1.

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

$$y(t) = a_1 t + a_0 . \qquad (1)$$

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

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

Решая уравнение (1) находим:

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

Составим таблицу для вычисления тренда временного ряда.

Таблица -1

N п/п	Годы наблюдений	y_t – ц /га	t	t^2	$y_t\, t$	$y_t\, t^2$
1	2008	49,7	-8	64	-397,6	3180,8
2	2009	50,1	-7	49	-350,7	2454,9
3	2010	50,4	-6	36	-302,4	1814,4
4	2011	53,7	-5	25	-268,5	1342,5
5	2012	51,9	-4	16	-207,6	830,4
6	2013	50,8	-3	9	-152,4	457,2
7	2014	52	-2	4	-104	208
8	2015	51	-1	1	-51	51
9	2016	51,2	0	0	0	0
10	2017	51,3	1	1	51,3	51,3
11	2018	52	2	4	104	208
12	2019	48,9	3	9	146,7	440,1
13	2020	42,8	4	16	171,2	684,8
14	2021	50,4	5	25	252	1260
15	2022	52,6	6	36	315,6	1893,6

16	2023	53,8	7	49	376,6	2636,2
17	2024	54,2	8	64	433,6	3468,8
Сумма		866,8	0	408	16,8	20982

Используя, вычисления по таблице 1, имеем:

$$\sum y_t = 866{,}8 \qquad a_0 = \frac{1}{\mathrm{T}}\sum y_t = \frac{866{,}8}{17} = 50{,}99 = 51, \qquad a_1 = \frac{1}{\sum t^2}\sum y_t t = \frac{16{,}8}{408} = 0{,}04.$$

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

$$y(t) = 0{,}04\, t + 51 \qquad (2)$$

В частности, подставляя в уравнение (2) значение t = 2 находим ожидаемые средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан 2026 году, будет 51,08 ц/га.

С помощью статистических критериев $([1]-[5])$ установлено, что в уравнении (2) $y(t) = a_1 t + a_0$ основная гипотеза $\boldsymbol{H_0: a_1 = 0}$ отвергается и принимается альтерна-тивная гипотеза $\boldsymbol{H_1: a_1 \neq 0}$ с уровень значимости $\alpha = 0{,}05$. Следовательно, средней урожайность пшеницы Сурхандарьинской области в Республике Узбекистан имеет линейный тренд.

Многих задачах наблюдения, выборка наблюдения является статистически независи-мы, то во временных рядах они, как правило, зависимы, и характер этой зависимости может определяться положением наблюдений в последовательности. Автокорреляцией представляет собой корреляционную зависимость между последующими и предшествующими членами временного ряда.

Проверяем наличие автокорреляции для рассматриваемого процесса, т.е.

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

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

Для дальнейшей исследование вычислим следующие конечные разности по наблюденным данным(таблица- 2):

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

Составим таблица для вычисления конечные разности.

Таблица -2

1	2	3	4	5	6	7	8	9	9

N п/п	Годы наблюдений	y_t – ц /га	Y_t^2	ΔY_t	ΔY_t^2	$\Delta^2 Y_t$	$\Delta^2 Y_t^2$	$\Delta^3 Y_t$	$\Delta^3 Y_t^2$
1	2008	49,7	2470,09						
2	2009	50,1	2510,01	0,4	0,16				
3	2010	50,4	2540,16	0,3	0,09	-0,1	0,01		
4	2011	53,7	2883,69	3,3	10,89	3	9	3,1	9,61
5	2012	51,9	2693,61	-1,8	3,24	-5,1	26,01	-8,1	65,61
6	2013	50,8	2580,64	-1,1	1,21	0,7	0,49	5,8	33,64
7	2014	52	2704	1,2	1,44	2,3	5,29	1,6	2,56
8	2015	51	2601	-1	1	-2,2	4,84	-4,5	20,25
9	2016	51,2	2621,44	0,2	0,04	1,2	1,44	3,4	11,56
10	2017	51,3	2631,69	0,1	0,01	-0,1	0,01	-1,3	1,69
11	2018	52	2704	0,7	0,49	0,6	0,36	0,7	0,49
12	2019	48,9	2391,21	-3,1	9,61	-3,8	14,44	-4,4	19,36
13	2020	42,8	1831,84	-6,1	37,21	-3	9	0,8	0,64
14	2021	50,4	2540,16	7,6	57,76	13,7	187,69	16,7	278,89
15	2022	52,6	2766,76	2,2	4,84	-5,4	29,16	-19,1	364,81
16	2023	53,8	2894,44	1,2	1,44	-1	1	4,4	19,36
17	2024	54,2	2937,64	0,4	0,16	-0,8	0,64	0,2	0,04
Сумма		866,8	44302,38	4,5	129,59	0	289,38	-0,7	828,51

По таблице 2 вычисляя

$$V_k = \frac{\sum_{t=k}^{T} \left(\Delta^k y_t\right)^2}{(T-k)C_{2k}^k} \qquad (3)$$

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

К расчету данных для определения показателей автокорреляционной связи.

Таблица -3

1	2	3	4	5	6	7	8
N п/п	Годы наблюдений	y_t – ц /га	$Y_t \cdot Y_{t+1}$	$Y_t \cdot Y_{t+2}$	$Y_t \cdot Y_{t+3}$	$Y_t \cdot Y_{t+4}$	$Y_t \cdot Y_{t+5}$
1	2008	49,7					
2	2009	50,1	2489,97				
3	2010	50,4	2525,04	2504,88			
4	2011	53,7	2706,48	2690,37	2668,89		
5	2012	51,9	2787,03	2615,76	2600,19	2579,43	
6	2013	50,8	2636,52	2727,96	2560,32	2545,08	2524,76
7	2014	52	2641,6	2698,8	2792,4	2620,8	2605,2
8	2015	51	2652	2590,8	2646,9	2738,7	2570,4
9	2016	51,2	2611,2	2662,4	2600,96	2657,28	2749,44
10	2017	51,3	2626,56	2616,3	2667,6	2606,04	2662,47
11	2018	52	2667,6	2662,4	2652	2704	2641,6
12	2019	48,9	2542,8	2508,57	2503,68	2493,9	2542,8
13	2020	42,8	2092,92	2225,6	2195,64	2191,36	2182,8
14	2021	50,4	2157,12	2464,56	2620,8	2585,52	2580,48
15	2022	52,6	2651,04	2251,28	2572,14	2735,2	2698,38
16	2023	53,8	2829,88	2711,52	2302,64	2630,82	2797,6
17	2024	54,2	2915,96	2850,92	2731,68	2319,76	2650,38
Сумма		866,8	41533,72	38782,12	36115,84	33407,89	31206,31

Используя таблицу -3, формулы (4) из литературы [1,2,3,4,5] определяются значения коэффициентов автокорреляции R_L при $L = 1,2,3,4,5$:

$$R_L = \frac{\sum_{t=1}^{N-L} Y_t Y_{t+L} - \frac{\sum_{t=1}^{N-L} Y_t \sum_{t=L+1}^{N} Y_t}{N-L}}{\sqrt{\left[\sum_{t=1}^{N-L} Y_t^2 - \frac{\left(\sum_{t=1}^{N-L} Y_t\right)^2}{N-L}\right]\left[\sum_{t=L+1}^{N} Y_t^2 - \frac{\left(\sum_{t=L+1}^{N} Y_t\right)^2}{N-L}\right]}} \qquad (4)$$

Отличие значения R_L от нуля, даёт основание полагать, что между средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан имеется существенная автокорреляционная зависимость.

С другой стороны проверим гипотезу, существования автокорреляционная зависи-мость между средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан с помощью критерия Дарбина – Уотсона :

$$d = \sum_{t=1}^{T-1} (Y_{t+1} - Y_t)^2 / \sum_{t=1}^{T} Y_t^2 \ . \qquad (5)$$

Вычисляется по формуле (5) $d_{\text{наб}} = 0{,}19$ и сравнивается с $d_{\text{крит}} = 1{,}22$ табличным значением [1], [2], [5]. Поскольку $d_{\text{наб}} = 0{,}19 < d_{\text{крит}} = 1{,}22$. Следовательно, критерия Дарбина – Уотсона 95% тоже гарантией доказывает, что средней урожайность пшеницы Сурхандарьинской 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, $\overline{y_t}$ – средней урожайность пшеницы Сурхандарьинской области в республике Узбекистан ([1] – [5]):

$$\text{H}_0 : P(\overline{y_t} < x) = \Phi_{\text{a},\sigma}(\text{x}), \qquad \text{H}_1 : P(\overline{y_t} < x) \neq \Phi_{\text{a},\sigma}(\text{x})$$

принимается уровень значимости $\alpha = 0{,}05$ (см. таблица – 4).

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

$$\overline{Y}_{T+i} - t(T-2;\alpha)\overline{\sigma}_y \le a_0 + a_1(T+i) \le \overline{Y}_{T+i} + t(T-2;\alpha)\overline{\sigma}_y \qquad (6)$$

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

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

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]):

$$|A_s| < 1{,}5\sigma_1\ , \quad \left|E_k + \frac{\sigma}{T+1}\right| < 1{,}5\sigma_2 \qquad (7)$$

тогда можно предположить, что случайная величина с достоверностью 95% имеет примерно нормальное распределение. Здесь асимметрия, эксцесса, исправленная выборочная дисперсия и другие числовые характеристики вычисляется по формулы:

$$A_s = \frac{m_3}{S_T^{\,3}}, \quad E_R = \frac{m_4}{S_T^{\,4}} - 3, \quad S_T^{\,2} = \frac{1}{(n-1)}\sum_{i=1}^{k}(x_i - \bar{x}_T)^2 n_i$$

$$m_3 = \frac{1}{(n-1)}\sum_{i=1}^{k}(x_i - \bar{x}_T)^3 n_i \quad m_4 = \frac{1}{(n-1)}\sum_{i=1}^{k}(x_i - \bar{x}_T)^4 n_i$$

$$\sigma_1 = \sqrt{\frac{6(T-2)}{(T+1)(T+3)}} \quad \sigma_2 = \sqrt{\frac{24T(T-2)(T-3)}{(T+1)^2(T+3)(T+5)}} \qquad (8)$$

Вычисление показывает, что условия (7) выполняется. Следовательно, тогда можно предположить, что случайная величина $\bar{y}_\text{T}$ − имеет примерно нормальное распределение с достоверностью 95% (см. Таблица-4).

Используя формула (6), построим с вероятностью 0.95 интервальная оценка для $\bar{y}_t$ −средней урожай пшеницы Сурхандарьинской областе:

(49,69; 52,31) ц/га

Теперь на основании выборочных данных, используя пакет программу x7.2019 и Excel ЭВМ [4,5,6], вычислим числовые характеристики y_t − средней урожайность пшеницы (таблица - 4).

Оценка основных параметров динамического ряда.

Таблица – 4

Выборочные характеристики	Оценки выборочные характери
Средний урожайность пшеницы $\bar{y}_\text{T}$ – ц/га	51 ц/га
Дисперсия	6,61
Среднее квадратичное	2,57
Коэффициент вариации v (%)	5,01 %
Асимметрия A_ς	-2,0
Эксцесса $E_{K\varsigma}$	2,24
Ошибка среднего значения $\bar{y}_\text{T}$, m_y	$m_y = \frac{\sigma_y}{\sqrt{n}} = 0{,}62$
Предельная ошибка m'_y	$m'_y = t\, m_y = 2{,}12 \cdot 0{,}62 = 1{,}31$
Ошибка среднего квадратичного отклонения σ_T	$m_\sigma = \frac{\sigma}{\sqrt{2n}} = 0{,}44$

Интервальная оценка (95%) $\bar{y}_T \pm tm_y$ для урожайности пшеницы	$\bar{y}_T \pm t\, m_y = 51,0 \pm 1,31$ (49,69; 52,31) ц/га
Проверка статистической гипотезы $H_0 : P(\bar{y}_t < x) = \Phi_{a,\sigma}(x)$ $H_1 : P(\bar{y}_t < x) \neq \Phi_{a,\sigma}(x)$	95% гарантий гипотезы H_0 принимается

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

1. построены точечные $\bar{y}_T = 51$ ц/га и интервальные статистические оценки для $\bar{y}_t -$ средней урожайность пшеницы с 95% гарантией (49,69; 52,31) ц/га ;
2. определены явные виды тренда и установлена её линейность $y(t) = 0,04\, t + 51$;
3. критерием Дарбина – Уотсона установлены, что, средней урожайность пшеницы области имеет автокорреляционную зависимость $Y_t=\rho Y_{t-1}+\varepsilon_t$, где $\rho = \mathrm{Cov}(Y_t, Y_{t+1}) = M[(Y_t - \bar{y}_t)(Y_{t+1} - \bar{y}_t)]$ т.е. в этом году, зависит от урожайности прошлых и последу-ющих лет;
4. Отметим, что в первом полугодии 2025 года в Узбекистане собрано почти 6 млн тонна пшеницы.

ЛИТЕРАТУРА

1. Т.Андерсон “Статистический анализ временных рядов”. –Москва: “МИР”, 1976. – 759 с.
2. М. Кендал, А. Стьюарт “Многомерный статистический анализ и временные ряды.- Москва: “Наука”, 1976. -736 с.
3. Н.П.Тихомиров «Эконометрика».- Москва: «Экзамен», 2003. – 512 с.
4. Б.А.Сулайманов, А.А.Файзиев, Ж.Н. Файзиев "Тажриба маълумотларининг статистик таҳлили".– Тошкент: ТошДАУ, 2015. – 124 с.
5. А.А.Файзиев “Matematik statistika”, O’quv qo’llanma. “Ilm-ziyo-zakovat”, 218-bet, Toshkent – 2022.
6. Б.Абдалимов, А.А.Файзиев “Математика”. Дарслик.“Ilm-ziyo-zakovat”, 518-bet, Toshkent – 2022.

7. V. Vahabov, A.A. Fayziev “Statistical analysis and forecasting of cotton yield dynamics Bukhara region”, Tashkent state transport university. 1 st International Scientifik Conference “Modern Materials Science: Topical Issues, Achievements and Innovations (ISCMMSTIAL-2022)” (Tashkent, Mart 4-5, 2022). 5 –pej. (in Inglish).

8. А.А.Файзиев “Статистический анализ и прогнозирование динамики инвестиции в республике Узбекистан”. “UNIVERSAL JOURNAL OF MATHEMATICAL THEORY END COMPUTER SCIENCES”. In volume # 3 issue#25, june 2025. Researchbib if 10.2 ojs\ pkp international standart serial number.issn 2992-8818. 13-22 pej.