

UDC 519.26

СТАТИСТИЧЕСКИЙ АНАЛИЗ И ПРОГНОЗИРОВАНИЕ ДИНАМИКИ ЗАГОТОВЛЕННОГО ХЛОПКА СИРДАРЬИНСКОЙ ОБЛАСТИ В РЕСПУБЛИКЕ УЗБЕКИСТАН

ФАЙЗИЕВ А. А.

ТАШКЕНТСКИЙ ЭКОНОМИЧЕСКИЙ И ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ, ТАШКЕНТ
fayziev.axtam@bk.ru

Аннотация

Упорядоченных количества случайных явлений, меняющихся во времени, образует временной ряд. В статье, методом статистического анализа временных рядов, изучена статистическая закономерность \bar{y}_t -среднего объема динамики заготовленного хлопка Сирдарьинской области Республика Узбекистан (по материалам ЦСУ РУз за 2003-2024 годы). Построены, с 95%ной гарантией точечные и интервальные оценки для среднего объема динамики заготовленного хлопка Сирдарьинской области, определена явные виды трендов и прогнозирована ожидаемого объема хлопка в области для последующих лет. С помощью статистических критериев Дарбина-Уотсона установлено, что средней объем хлопка в области имеет автокорреляционную зависимость.

Ключевые слова: дискретный, хлопка, динамический, ряд, объем, тренд, сезонность, компонента, линейный, наименьший, нормальный.

Введение. Почти в каждой области встречаются явления, которые важно изучать в их развитии и изменении во времени. Можно, например, стремиться предсказать будущее на основании знания прошлого, управлять процессом, описать характерные особенности ряда на основании ограниченного количества информации. При обработке временных рядов опираются на разработанные математической статистикой, настоящему времени статистика располагает разнообразными методами анализа временных рядов от самых элементарных до весьма сложных ([1, 2, 3, 4]).

Материалы и методы. В настоящей работе, проведена обработка и анализ среднего объема динамики заготовленного хлопка Сирдарьинской области, Республика Узбекистан за период наблюдений 2003-2024 годы, как $\{y_t, t \in T\}$ дискретный стационарный временный ряд.

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

В данной работе использована методы обработки и анализа временных рядов, такие, как методы определения тренда, проверки нормальности и случайности, а также проверка автокорреляции, метод скользящей средней, метод конечных разностей, метод наименьших квадратов, статистическая критерия Дарбина – Уотсона и другие методы.

Используя методы статистического анализа временных рядов построены точечные и интервальные оценки для средней объем хлопка области, определены явные виды трендов и прогнозирована средней объем хлопка области для последующих лет, проверено различные статистические гипотезы.

Изучению и анализу динамических рядов посвящены работы: Андерсона[1], Кендалла[2], Файзиева [3] и другие.

Результаты и их обсуждение. Предположим, что средней объем хлопка Сирдарьинской области Республика Узбекистан за период наблюдений, 2003-2024 годы образует стационарный дискретный временный ряд. Используя выше изложенных методы статистического анализа временных рядов, построим точечные и интервальные оценки для средней объем хлопка области, определим явные вид тренда и прогнозирована объем хлопка для последующих лет, проверим различные статистические гипотезы.

На рисунки-1, с помоши опытных данных (таблица-1, столба-3) геометрически изображены \bar{y}_t – динамика средней объем заготовленного хлопка в Сирдарьинской области виде а) точечная график, б) гистограмма с) круговая диаграмма :

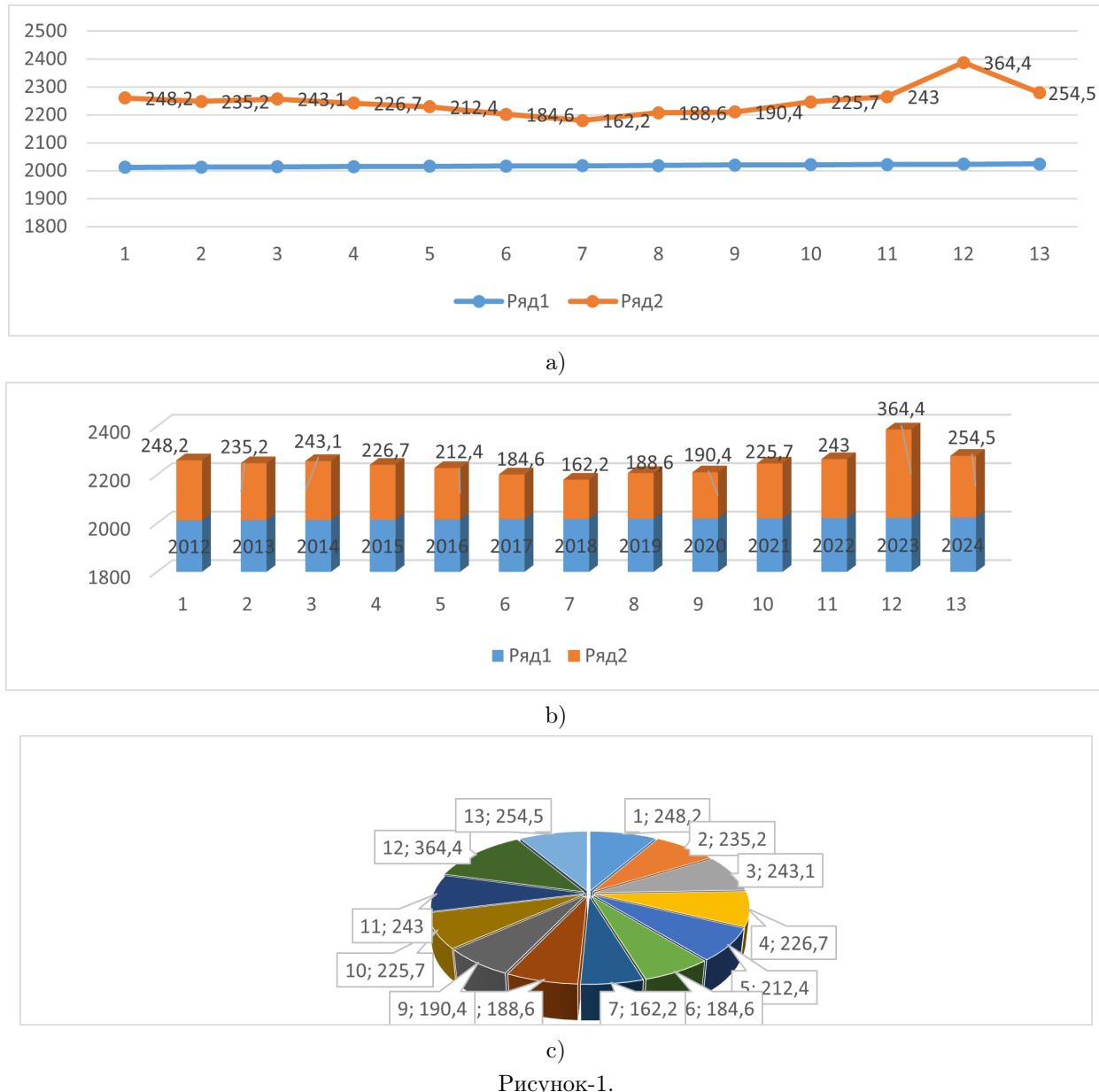


Рисунок-1.

Геометрическое изображение наблюденных данных, система координат дают основание в первом приближении, предполагать гипотезу что, трендовая часть процесса имеет линейную зависимость вида $y(t) = a_1 t + a_0$. Где неизвестные параметры определяются методом наименьших квадратов т.е. на основании опытных данных, решая следующую систему нормальных уравнений:

$$\begin{cases} a_0 T + a_1 \sum t = \sum y_t \\ a_0 \sum t + a_1 \sum t^2 = \sum y_t t \end{cases} \quad (1)$$

Решая систему уравнение (1) и используя, вычисления по таблице-1, имеем

$$\sum y_t = 2978,6 \text{ тысяча тонна}, \quad a_0 = \frac{1}{T} \sum y_t = \frac{2978,6}{13} = 229,12$$

$$a_1 = \frac{1}{\sum t^2} \sum y_t t = \frac{639}{182} = 3,51 \text{ тысяча тонна.}$$

Отсюда, находим уравнение линейного тренда (тенденция) заготовленного объема хлопка в Сирдарьинской области [1, 2, 3]:

$$y(t) = 3,51 t + 229,12 \quad (2)$$

Подставляя в уравнение (2) значение $t = 1$ находим ожидаемые объема хлопка Сирдарьинской области в 2025 году, будет в среднем 232,63 тысяча тонна.

К расчету данных для определения тренда временного ряда.

Таблица-1

N п/п	Годы наблюдения	y_t – Тысяча тонна	t	t^2	$y_t t$	$y_t t^2$
1	2012	248,2	-6	36	-1 489	8 935
2	2013	235,2	-5	25	-1 176	5 880
3	2014	243,1	-4	16	-972	3 890
4	2015	226,7	-3	9	-680	2 040
5	2016	212,4	-2	4	-425	850
6	2017	184,6	-1	1	-185	185
7	2018	162,2	0	0	0	0
8	2019	188,6	1	1	189	189
9	2020	190,4	2	4	381	762
10	2021	225,7	3	9	677	2 031
11	2022	243	4	16	970	3 882
12	2023	364,4	5	25	1822	9110
13	2024	254,5	6	36	1 527	9 162
Сумма		2 978,6	0	182	639	46 914

С помощью статистических критериев ([1] – [3]) 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 уровнем значимости $\alpha = 0,05$. Следовательно, объем хлопка в Сирдарьинской области имеет линейный тренд.

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

Для дальнейшей исследование нам необходима вычислить следующие конечные разности по опытным данным. Обозначим,

$$\Delta Y_t = Y_{t+1} - Y_t, \Delta^2 Y_t = \Delta Y_{t+1} - \Delta Y_t, \Delta^3 Y_t = \Delta^2 Y_{t+1} - \Delta^2 Y_t \text{ конечные разности (таблица -2).}$$

Составим таблица для расчету конечные разности.

Таблица -2

N п/п	Годы наблю- дения	y_t	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	2012	248,2	61 603,2						
2	2013	235,2	55 319,0	-13,0	169,0				
3	2014	243,1	59 097,6	7,9	62,4	20,9	436,8		
4	2015	226,7	51 392,9	-16,4	269,0	-24,3	590,5	-45,2	2 043,0
5	2016	212,4	45 113,8	-14,3	204,5	2,1	4,4	26,4	697,0
6	2017	184,6	34 077,2	-27,8	772,8	-13,5	182,3	-15,6	243,4
7	2018	162,2	26 308,8	-22,4	501,8	5,4	29,2	18,9	357,2
8	2019	188,6	35 570,0	26,4	697,0	48,8	2 381,4	43,4	1 883,6
9	2020	190,4	36 252,2	1,8	3,2	-24,6	605,2	-73,4	5 387,6
10	2021	225,7	50 940,5	35,3	1 246,1	33,5	1 122,3	58,1	3 375,6
11	2022	243	58 854,8	16,9	285,6	-18,4	338,6	-51,9	2 693,6
12	2023	364,4	132 787,4	121,8	14 835,2	104,9	11 004,0	123,3	15 202,9
13	2024	254,5	64 770,3	-109,9	12 078,0	-231,7	53 684,9	-336,6	113 299
	Сумма	2 978,6	712 087,5	6,3	31 124,6	-96,9	70 379,4	-252,6	145 183

По таблице- 2 вычислим

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

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

Для определения коэффициентов автокорреляции R_L при $L = 1, 2, 3, 4, 5$ (где: $L_{\text{лаг}}$, временной сдвиг), т.е. промежуток времени отставания одного явления от другого, связанного с ним) составим таблица-3 : Таблица-3

N п/п	Годы наблю- дения	y_t - тысяча тонна	$Y_t \cdot Y_{t+1}$	$Y_t \cdot Y_{t+2}$	$Y_t \cdot Y_{t+3}$	$Y_t \cdot Y_{t+4}$	$Y_t \cdot Y_{t+5}$
1	2012	248,2					
2	2013	235,2	58 376,6				
3	2014	243,1	57 177,1	60 337,4			
4	2015	226,7	55 110,8	53 319,8	60 337,4		
5	2016	212,4	48 151,1	51 634,4	53 319,8	52 717,7	
6	2017	184,6	39 209,0	41 848,8	51 634,4	43 417,9	45 817,7
7	2018	162,2	29 942,1	34 451,3	41 848,8	39 430,8	38 149,4
8	2019	188,6	30 590,9	34 815,6	34 451,3	42 755,6	45 848,7
9	2020	190,4	35 909,4	30 882,9	34 815,6	40 441,0	43 163,7
10	2021	225,7	42 973,3	42 567,0	30 882,9	41 664,2	47 938,7
11	2022	243	54 754,8	46 191,0	42 567,0	39 349,7	44 784,0
12	2023	364,4	88 403,4	82 245,1	46 191,0	68 725,8	59 105,7
13	2024	254,5	92 739,8	61 741,7	82 245,1	48 456,8	47 998,7
	Сумма	2 978,6	633 338,5	540 035,1	478 293,4	416 959,6	372 806,5

Используя таблицу-3, формулы (4) из литературы [1, 2, 3] определяются значения коэффициентов автокорреляции R_L $L = 1, 2, 3, 4, 5$ (где: $L_{\text{лаг}}$, временной сдвиг):

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

Отличие (4) значения R_L от нуля, даёт основание полагать, что между годовая объема хлопка имеется существенная автокорреляционная зависимость.

Проверим статистическую гипотезу, о существования автокорреляционная зависимость между годовая объема хлопка с помощью критерия Дарбина – Уотсона :

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

Используя таблица-2 , по формуле (5) вычислим $d_{\text{наб}} = \frac{31124,6}{712087,5} = 0,044$. Сравнивая их с табличным значением(см. [3]-стр. 194, приложение-9) $d_{\text{крит}} = 1,06$ установим, что $d_{\text{наб}} = 0,044 < d_{\text{крит}} = 1,06$. Следовательно, критерия Дарбина – Уотсона 95% тоже гарантией доказывает, что средние объем заготовленного хлопка в Сирдарьинской области имеет автокорреляционную зависимость

$$Y_t = \rho Y_{t-1} + \epsilon_t, \quad \rho = \text{Cov}(Y_t, Y_{t+1}) = M[(Y_t - \bar{y}_t)(Y_{t+1} - \bar{y}_t)].$$

Следовательно, заготовленная объем хлопка области в этом году, зависит от объема прошлых и последующих лет.

Проверка статистическая гипотезы (6) о нормальности \bar{y}_t – среднего объема заготовленного хлопка в Сирдарьинской области ([1] – [3]) :

$$H_0 : P(\bar{y}_t < x) = \Phi_{,\sigma}(x), \quad H_1 : P(\bar{y}_t < x) \neq \Phi_{,\sigma}(x) \quad (6)$$

принимается уровень значимости $\alpha = 0,05$ (Таблица-4).

Тогда с помощью следующие формулы построим интервальные оценка для \bar{y}_t – среднего объема заготовленного хлопка в Сирдарьинской области:

$$\bar{Y}_{T+i} - t(T-2; \alpha) \bar{\sigma}_y \leq a_0 + a_1(T+i) \leq \bar{Y}_{T+i} + t(T-2; \alpha) \bar{\sigma}_y \quad (7)$$

Значение $t_{\text{крит.}} = t(T-2; \alpha)$ определяется по таблице Стьюдента (см. [3]-стр. 190, приложение-3). По формула (6) построим с вероятностью 0.95 интервальная оценка для \bar{y}_t -среднего объема хлопка в Сирдарьинской области :

$$(199,14 ; 259,16) \text{ тысяча тонна.}$$

На основании выборочных данных, используя пакет программы x7.2019 и Excel ЭВМ [3,4,5], вычислим числовые характеристики y_t – среднего объема заготовленного хлопка в Сирдарьинской области (таблица - 4).

Оценка основных параметров динамического ряда.

Таблица - 4

Выборочные характеристики	Статистические оценки для выборочные характеристики
Средний объем ежегодно заготовленного хлопка в Сирдарьинская области \bar{y}_T	229, 12 тысяча тонна
Дисперсия	2469,371
Среднее квадратичное отклонение σ_T	49,693
Коэффициент вариации v (%)	7, 07 %
Асимметрия ζ	1,593
Эксцесса E_{K_ζ}	2,309
Ошибка среднего значения \bar{y} , m_y	$m_y = \frac{\sigma}{\sqrt{n}} = 13, 765$
Предельная ошибка t'	$m'_y = tm = 2, 18 \cdot 13, 765 = 30, 01$
Интервальная оценка (95%) $\bar{y}_T \pm tm$ для объема хлопка	$\bar{y}_T \pm tm = 229, 08 \pm 30, 01$ (199,14 ; 259,16) тысяча тонна.
Проверка статистической гипотезы $H_0 : P(\bar{y}_t < x) = \Phi_{,\sigma}(x)$ $H_1 : P(\bar{y}_t < x) \neq \Phi_{,\sigma}(x)$	95% гарантий гипотезы H_0 принимается

Выводы. На основании выше изложенных статистических анализов, динамики ежегодно \bar{y}_t – заготовленного хлопка в Сирдарьинской области республика Узбекистан как дискретный стационарный временный ряд с надежностью $\gamma = 0, 95$ (таблица-4) можно сделать следующие выводы:

- построена точечные 229, тысяча тонна и интервальные (199,14 ; 259,16) тысяча тонна статистические оценки для среднего объема ежегодно \bar{y}_t - заготовленного хлопка в Сирдарьинской области;
- определен виды тренда и установлена её линейность $(t) = 3, 51 t + 229, 12$;
- критерием Дарбина – Уотсона установлены что, средней объем заготовленного хлопка в Сирдарьинской области имеет автокорреляционную зависимость:

$$Y_t = \rho Y_{t-1} + \epsilon_t, \quad \text{где} \quad \rho = \text{Cov}(Y_t, Y_{t+1}) = M[(Y_t - \bar{y}_t)(Y_{t+1} - \bar{y}_t)].$$

т.е. объем заготовленного хлопка в этом году, зависит от объема заготовленных прошлых и последующих лет.

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.
- А.А.Fayziyev “Ehtimollar nazariyasi va matematik statistika”. Darslik, “METODIST NASHRIYOTI”, 218-bet, Toshkent – 2024.
- M.U.Achilov, A.A.Fayziev “The analysis of dynamics of fruits and berry productivity grown in Uzbekistan”, EPRA Internatsional journal of Research and Development (IJRD.Indiya).Volum: 4. Issue: 8. August 2019, pp.5-9. (in English)
- V. Vahabov, A.A. Fayziev “Statistical analysis and forecasting of cotton yield dynamics Bukhara region”, Tashkent state transport university. 1 st International Scientific Conference “Modern Materials Science: Topical Issues, Achievements and Innovations (ISCMSTIAL-2022)” (Tashkent, Mart 4-5, 2022). p.5. (in English).

Annotatsiya

Vaqt o'tishi bilan o'zgarib turadigan tasodifiy hodisalarlarning tartiblangan ketma-ketligi vaqtli (dinamik)qatorni hosil qiladi. Maqolada dinamik qatorlarni statistik tahlil qilish usullariidan foydalanib, O'zbekiston Respublikasi Sirdaryo viloyatida (\bar{y}_t)-o'rtacha yillik terib olingan paxta hajmi dinamikasini statistik qonuniyatlari (O'zSTAT boshqarmasining 2003-2024 yillardagi ma'lumotlari asosida) o'rganilgan. Sirdaryo viloyatida terib olingan paxtaning o'rtacha dinamikasi hajmining nuqta va intervalli hisob-kitoblari 95 % li kafolat bilan tuzilib, tendensiyalarning aniq turlari belgilab olindi, keyingi yillar uchun viloyatda paxta yetishtirishning kutilayotgan hajmi proqnoz qilindi. Durbin-Watson statistik mezonlaridan foydalanib, mintaqadagi paxtaning o'rtacha hajmi avtokorrelyatsion bog'liqlikka ega ekanligi aniqlandi.

Kalit so'zlar: diskret, paxta, dinamik, qator, hajm, trend, mavsumiylik, komponent, chiziqli, eng kichik, normal.

Abstract

An ordered number of random events changing over time forms a time series. In the article, using the method of statistical analysis of time series, the statistical regularity of (\bar{y}_t)-average volume of dynamics of harvested cotton in the Sirdarya region of the Republic of Uzbekistan (based on the materials of the Central Statistical Bureau of the Republic of Uzbekistan for 2003-2024) is studied. Point and interval estimates for the average volume of dynamics of harvested cotton in the Sirdarya region have been constructed with a 95% guarantee, clear types of trends have been determined, and the expected volume of cotton in the region for subsequent years has been predicted. Using the Durbin-Watson statistical criteria, it has been established that the average volume of cotton in the region has an autocorrelation dependence.

Key words: discrete, cotton, dynamic, series, volume, trend, seasonality, component, linear, least, normal.