

**ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ МЕТОДОВ ПРОГНОЗИРОВАНИЯ
ДИНАМИКИ УРОЖАЙНОСТИ ОВОЩЕЙ В ТАШКЕНТСКОЙ ОБЛАСТИ
РЕСПУБЛИКИ УЗБЕКИСТАН**

<https://doi.org/10.5281/zenodo.10071815>

А.А.Файзиев

к.ф-м.н, доцент

Н.А.Ашурметова

к.э.н., доцент

Н.Н.Мусаева

ассистент

А.Мынбаева

студентка

*Ташкентский государственный аграрный университет
Республика Узбекистан*

Аннотация

Наблюдения над некоторыми явлениями, характер которых меняется во времени в упорядоченной последовательности, называют временным рядом. В статье методом статистического анализа временных рядов, изучена статистическая закономерность рядов динамики \bar{y}_t –средней урожайности овощей в Ташкентской области Республики Узбекистан за 2000-2020 годы. С 95%- ной гарантией построены точечные и интервальные оценки для средней урожайности овощей, определены явные виды трендов и спрогнозирована урожайность в области для последующих лет. С помощью статистических критериев Дарбина-Уотсона установлено, что средняя урожайность овощей в области имеет автокорреляционную зависимость. Следовательно, урожайность овощей области в следующем году зависит от урожайности прошлых и последующих лет.

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

Ключевые слова

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

Введение

В обеспечении продовольственной безопасности в сфере потребления плодоовощной продукции определенную роль играют овощи, поскольку являются наиболее употребляемыми в течение года продуктами питания, содержащими биологически активные вещества (витамины, минеральные вещества, органические кислоты, клетчатку и пектиновые вещества), необходимые для нормального развития организма. Наличие значительных возможностей выращивания овощных культур в Узбекистане, связанных, прежде всего, с природно-климатическими условиями, создает предпосылки для развития овощеводства в стране. Так, по официальным данным Агентства статистики при Президенте Республики Узбекистан в 2022 году произведено 11163,0 тыс. т овощей, или 102,9 %, по сравнению с 2021 годом и 422,1 % по отношению к 2000 году (2644,7 тыс. тонн). Под овощными культурами занято 233,7 тысячи гектаров, из которых 7175 га это площади закрытого грунта.

Вместе с тем, в республике имеются неиспользованные резервы в развитии овощеводства, которые связаны, в первую очередь, с повышением урожайности овощных культур.

Почти в каждой области, в том числе в овощеводстве, 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 во многом опираются на разработанные математической статистикой методы для динамических рядов распределения. К 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 урожайности овощей в Ташкентской области Республики Узбекистан за период наблюдений 2000-2020 годы, взятые как дискретный временной ряд. Изучение урожайности сельскохозяйственных процессов, как дискретный динамический ряд и прогнозирование их урожайности на основании фактических данных, играют важную роль при определении экономической эффективности деятельности сельскохозяйственных предприятий, в частности хозяйств, выращивающих овощные культуры.

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 использованы 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, критерий Дарбина – Уотсона и другие методы.

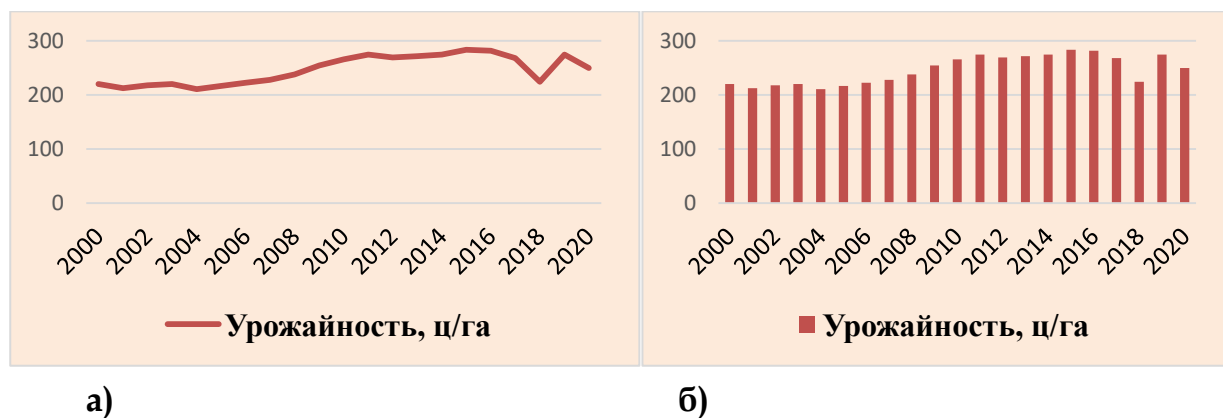
Используя методы статистического анализа временных рядов построены точечные и интервальные оценки для средней урожайности овощей области, определены явные виды трендов и спрогнозирован урожай для последующих лет, проверены различные статистические гипотезы.

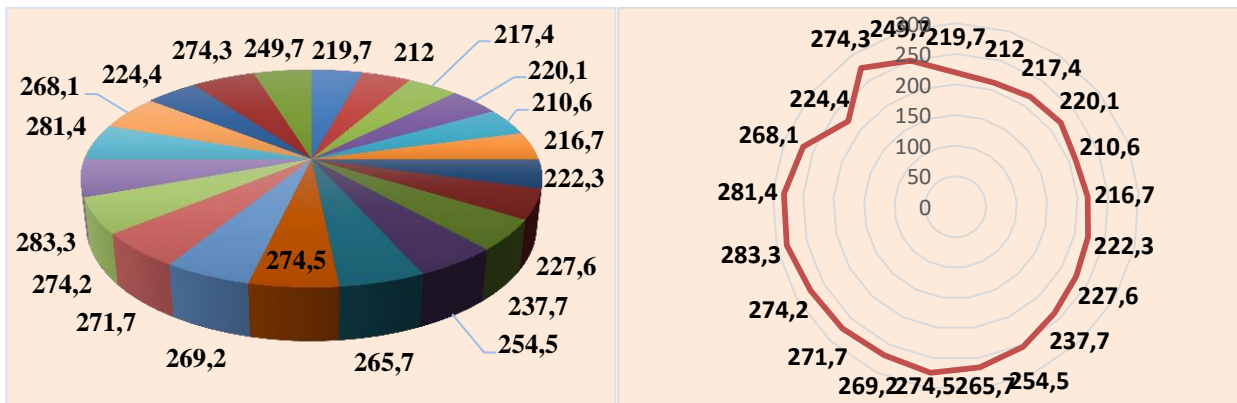
Изучению и анализу динамических рядов посвящены работы Т.Андерсона [1], М.Кендала [2], Н.Тихомирова [2], Б.Сулайманова [3], А.Файзиева [6,7] и др.

Результаты и их обсуждение

Предположим, что урожайность овощей в Ташкентской области Республики Узбекистан за период наблюдений 2000-2020 годы образует дискретный временный ряд. Используя вышеизложенные методы статистического анализа временных рядов, построим точечные и интервальные оценки для средней урожайности овощей области, определим явный вид тренда и спрогнозируем урожай овощей для последующих лет, проверим различные статистические гипотезы.

На рисунке 1, с помощью реальных данных (таблица 1, столб 3) геометрически изображены \bar{y}_t – динамика средней урожайности овощей в Ташкентской области в виде: а) точечного графика, б) гистограммы, с) круговой диаграммы, д) лепестковой диаграммы.





с)

д)

Рис. 1. Графическое отображение динамики средней урожайности овощей в Ташкентской области

Геометрическое выражение наблюдаемых данных, система координат дают основание в первом приближении предполагать гипотезу, что трендовая часть процесса (общее направление развития процесса) имеет линейную зависимость вида $y(t) = a_1t + a_0$. При этом, неизвестные параметры определяются методом наименьших квадратов, т.е. на основании фактических данных, решая следующую систему нормальных уравнений:

$$\begin{cases} a_0T + a_1 \sum t = \sum y_t \\ a_0 \sum t + a_1 \sum t^2 = \sum y_t t \end{cases} \quad (1)$$

Таблица 1

Расчет данных для определения тренда временного ряда

1	2	3	4	5	6	7
N п/п	Годы наблюдений	y_t -ц/га	t	t^2	$y_t t$	$y_t t^2$
1	2000	219,7	-10	100	-2197	21970
2	2001	212	-9	81	-1908	17172
3	2002	217,4	-8	64	-1739,2	13913,6
4	2003	220,1	-7	49	-1540,7	10784,9
5	2004	210,6	-6	36	-1263,6	7581,6
6	2005	216,7	-5	25	-1083,5	5417,5
7	2006	222,3	-4	16	-889,2	3556,8

8	2007	227,6	-3	9	-682,8	2048,4
9	2008	237,7	-2	4	-475,4	950,8
10	2009	254,5	-1	1	-254,5	254,5
11	2010	265,7	0	0	0	0
12	2011	274,5	1	1	274,5	274,5
13	2012	269,2	2	4	538,4	1076,8
14	2013	271,7	3	9	815,1	2445,3
15	2014	274,2	4	16	1096,8	4387,2
16	2015	283,3	5	25	1416,5	7082,5
17	2016	281,4	6	36	1688,4	10130,4
18	2017	268,1	7	49	1876,7	13136,9
19	2018	224,4	8	64	1795,2	14361,6
20	2019	274,3	9	81	2468,7	22218,3
21	2020	249,7	10	100	2497	24970
Сумма		5175,1	0	770	2433,4	183733,6

Решив систему уравнений (1) и, используя вычисления по данным таблицы 1, имеем:

$$\sum y_t = 5175,1 \frac{\text{ц}}{\text{га}}, \quad a_0 = \frac{1}{T} \sum y_t = \frac{5175,1}{21} = 246,43 \text{ ц/га},$$

$$a_1 = \frac{1}{\sum t^2} \sum y_t t = \frac{2433,4}{770} = 3,16 \text{ ц/га}.$$

Отсюда находим уравнение линейного тренда (тенденция) урожайности овощей в Ташкентской области:

$$y(t) = 3,16 t + 246,43 \quad (2)$$

В частности, подставляя в уравнение (2) значение $t = 3$, находим ожидаемую урожайность овощей в Ташкентской области в 2023 году, которая составит в среднем 255,91 ц/га.

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 уровнем значимости $\alpha = 0,05$. Следовательно, урожайность овощей области имеет линейный тренд.

Если во многих задачах наблюдения, выборка наблюдения является статистически независимой, то во временных рядах они, как правило, зависимы, и характер этой зависимости может определяться положением наблюдений в последовательности. Автокорреляция представляет собой

корреляционную зависимость между последующими и предшествующими членами временного ряда.

Проверяем наличие автокорреляции средней урожайности овощей в области, 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 дальнейшего исследования нам необходимо вычислить следующие конечные разности по наблюдаемым данным. Обозначим $\Delta Y_t = Y_{t+1} - Y_t$, $\Delta^2 Y_t = \Delta Y_{t+1} - \Delta Y_t$, $\Delta^3 Y_t = \Delta^2 Y_{t+1} - \Delta^2 Y_t$, $\Delta Y_t = Y_{t+1} - Y_t$, $\Delta^2 Y_t = \Delta Y_{t+1} - \Delta Y_t$, $\Delta^3 Y_t = \Delta^2 Y_{t+1} - \Delta^2 Y_t$ как конечные разности (таблица 2).

Таблица 2

Таблица расчета конечных разностей

	2	3	4	5	6	7	8	9
Годы наблю дений	$Y(t)$ ц/га	Y_t^2	ΔY_t	ΔY_t^2	$\Delta^2 Y_t$	$\Delta^2 Y_t^2$	$\Delta^3 Y_t$	$\Delta^3 Y_t^2$
2000	219,7	48268,09						
2001	212	44944	-7,7	59,29				
2002	217,4	47262,76	5,4	29,16	13,1	171,61		
2003	220,1	48444,01	2,7	7,29	-2,7	7,29	-15,8	249,64
2004	210,6	44352,36	-9,5	90,25	-12,2	148,84	-9,5	90,25
2005	216,7	46958,89	6,1	37,21	15,6	243,36	27,8	772,84
2006	222,3	49417,29	5,6	31,36	-0,5	0,25	-16,1	259,21
2007	227,6	51801,76	5,3	28,09	-0,3	0,09	0,2	0,04
2008	237,7	56501,29	10,1	102,01	4,8	23,04	5,1	26,01
2009	254,5	64770,25	16,8	282,24	6,7	44,89	1,9	3,61
2010	265,7	70596,49	11,2	125,44	-5,6	31,36	-12,3	151,29
2011	274,5	75350,25	8,8	77,44	-2,4	5,76	3,2	10,24
2012	269,2	72468,64	-5,3	28,09	-14,1	198,81	-11,7	136,89
2013	271,7	73820,89	2,5	6,25	7,8	60,84	21,9	479,61
2014	274,2	75185,64	2,5	6,25	0	0	-7,8	60,84
2015	283,3	80258,89	9,1	82,81	6,6	43,56	6,6	43,56
2016	281,4	79185,96	-1,9	3,61	-11	121	-17,6	309,76
2017	268,1	71877,61	-13,3	176,89	-11,4	129,96	-0,4	0,16
2018	224,4	50355,36	-43,7	1909,69	-30,4	924,16	-19	361
2019	274,3	75240,49	49,9	2490,01	93,6	8760,96	124	15376

2020	249,7	62350,09	-24,6	605,16	-74,5	5550,25	-168,1	28257,61
Сумма	5175,1	1289411,01	30	6178,54	-16,9	16466,03	-87,6	46588,56

В таблице 2 вычислены коэффициенты вариации разностей по формуле 3 и установлено, что $V_1 \approx V_2 \approx V_3$.

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

Следовательно, конечные разности первого порядка элиминируют линейную тенденцию.

Таблица 3

Расчет данных для определения показателей автокорреляционной СВЯЗИ

1	2	3	4	4	5	6	7
№ п/п	Годы наблюдений	y_t	$y_t \cdot y_{t+1}$	$y_t \cdot y_{t+2}$	$y_t \cdot y_{t+3}$	$y_t \cdot y_{t+4}$	$y_t \cdot y_{t+5}$
1	2000	219,7					
2	2001	212	46576,4				
3	2002	217,4	46088,8	47762,78			
4	2003	220,1	47849,74	46661,2	48355,97		
5	2004	210,6	46353,06	45784,44	44647,2	46268,82	
6	2005	216,7	45637,02	47695,67	47110,58	45940,4	47608,99
7	2006	222,3	48172,41	46816,38	48928,23	48328,02	47127,6
8	2007	227,6	50595,48	49320,92	47932,56	50094,76	49480,24
9	2008	237,7	54100,52	52840,71	51509,59	50059,62	52317,77
10	2009	254,5	60494,65	57924,2	56575,35	55150,15	53597,7
11	2010	265,7	67620,65	63156,89	60473,32	59065,11	57577,19
12	2011	274,5	72934,65	69860,25	65248,65	62476,2	61021,35
13	2012	269,2	73895,4	71526,44	68511,4	63988,84	61269,92
14	2013	271,7	73141,64	74581,65	72190,69	69147,65	64583,09
15	2014	274,2	74500,14	73814,64	75267,9	72854,94	69783,9
16	2015	283,3	77680,86	76972,61	76264,36	77765,85	75272,81
17	2016	281,4	79720,62	77159,88	76456,38	75752,88	77244,3
18	2017	268,1	75443,34	75952,73	73513,02	72842,77	72172,52
19	2018	224,4	60161,64	63146,16	63572,52	61530,48	60969,48
20	2019	274,3	61552,92	73539,83	77188,02	77709,19	75213,06
21	2020	249,7	68492,71	56032,68	66944,57	70265,58	70740,01
Сумма		5175,1	1231012,65	1170550,06	1120690,31	1059241,26	995979,93

Используя таблицу 3, формулы из литературы [1,2,3,4,5] определяются значения 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):

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

Чем выше значение коэффициента корреляции, тем больше зависимость между величинами, в нашем случае значение R_L намного выше нуля, это даёт основание полагать, что между урожайностью овощей имеется существенная автокорреляционная зависимость.

С другой стороны, проверим гипотезу существования автокорреляционной зависимости между урожайностью овощей с помощью критерия Дарбина - Уотсона:

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

Вычисленное по формуле (5) значение $d_{\text{наб}} = 0,006$ сравнивается с $d_{\text{крит}} = 1,22$ табличным значением [1], [2].. [5].

Поскольку $d_{\text{наб}} = 0,006 < d_{\text{крит}} = 1,22$, то критерий Дарбина - Уотсона с 95%-й гарантией тоже доказывает, что средние урожайности овощей в области имеют автокорреляционную зависимость $Y_t = \rho Y_{t-1} + \varepsilon_t$. Следовательно, урожайность овощей области в последнем году зависит от урожайности прошлых лет.

Проверка статистической гипотезы нормальности, \bar{y}_t -средний урожай овощей области ([1] - [5]):

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

при этом, принимается уровень значимости $\alpha = 0,05$ (см. таблица 4).

Теперь с помощью следующей формулы построим интервальные оценка для средней урожайности овощей:

$$\bar{Y}_{T+i} - t(T-2; \alpha) \bar{\sigma}_y \leq a_0 + a_1(T+i) \leq \bar{Y}_{T+i} + t(T-2; \alpha) \bar{\sigma}_y \quad (6)$$

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

Значение $t_{\text{крит.}} = t(T-2; \alpha)$ определяется по таблице Стьюдента (см. [6], стр. 181).

По формуле (6) определим интервальную оценку для \bar{y}_t - средней урожайности овощей в Ташкентской области с вероятностью 0.95:

(234,39; 258,47) ц/га.

На основании выборочных данных, используя пакет программы x7.2019 и Excel, вычислим числовые характеристики y_t - для средней урожайности овощей Ташкентской области (таблица 4).

Таблица 4

Оценка основных параметров динамического ряда

Выборочные характеристики	Оценки выборочных характеристик
Средний урожай овощей, ц/га \bar{y}_T	246,43
Дисперсия	704,69
Среднее квадратичное	26,55
Коэффициент вариации v (%)	10,77 %
Асимметрия A_ζ	-0,0095
Экцесса E_{K_ζ}	-1,8162
Ошибка среднего значения \bar{y}_T, m_y	$m_y = \frac{\sigma_y}{\sqrt{n}} = 5,79$
Предельная ошибка m'_y	$m'_y = t m_y = 2,08 \cdot 5,79 = 12,04$
Ошибка среднего квадратичного отклонения σ_T	$m_\sigma = \frac{\sigma}{\sqrt{2n}} = \frac{26,55}{6,48} = 4,09$
Интервальная оценка (95%) $\bar{y}_T \pm t m_y$ для урожайности овощей	$\bar{y}_T \pm t m_y = 246,43 \pm 12,04$ (234,39; 258,47) ц/га
Проверка статистической гипотезы $H_0 : P(\bar{y}_t < x) = \Phi_{a,\sigma}(x)$	95% гарантии гипотезы H_0

Выводы. В результате вышеизложенного статистического анализа динамики \bar{y}_t -урожайности овощей в Ташкентской области Республики Узбекистан на основе дискретных временных рядов с надежностью $\gamma = 0,95$ (таблица 4) можно сделать следующие выводы:

1. Построены точечные и интервальные статистические оценки для \bar{y}_t – средней урожайности овощей в Ташкентской области. В частности, средняя урожайность \bar{y}_t – овощей в области составит с 95%-й гарантией (234,39; 258,47) ц/га;

2. Определены явные виды тренда и установлена её линейность $y(t) = 0,54t + 25,27$;

3. Посредством критерия Дарбина – Уотсона установлено, что, средние урожайности овощей в области имеет автокорреляционную зависимость $Y_t = \rho Y_{t-1} + \varepsilon_t$, где $\rho = \text{Cov}(Y_t, Y_{t+1}) = M[(Y_t - \bar{y}_t)(Y_{t+1} - \bar{y}_t)]$, т.е. урожайность овощей в области в последнем году зависит от урожайности прошлых лет.

4. В целом, доказано, что динамика средней урожайности овощей в Ташкентской области Республики Узбекистан образует нестационарный временный ряд.

5. Статистические исследования, основанные на анализе временных рядов, дают возможность понять некоторые явления с точки зрения их зависимости от предыдущих показателей, в частности в нашем случае, такой анализ позволяет проследить интенсивность и закономерность процесса изменения урожайности овощей. В связи с тем, что в сельском хозяйстве урожайность культур выступает одним из главных факторов эффективности производства, выявление и количественная характеристика основной тенденции урожайности необходим при планировании производства сельскохозяйственных продуктов. Вместе с тем, для того, чтобы производитель сельскохозяйственной продукции стремился к постоянному повышению урожайности культур, он должен не только наблюдать общее направление к росту, снижению или стабилизации уровня урожайности с течением времени, но и выявлять зависимость урожайности от определяющих его факторов. К примеру, изучается зависимость урожайности определенной сельскохозяйственной культуры от количества внесенных удобрений (минеральных или органических), средств защиты растений от сорняков и вредителей, правильного и своевременного орошения и т.п.

Таким образом, анализ урожайности овощей в Ташкентской области Республики Узбекистан за период 2000-2020 годы на основе дискретного временного ряда будет иметь логическое продолжение в виде регрессионно – корреляционного анализа, результаты которого будут освещены в следующих публикациях.

ЛИТЕРАТУРА:

1. Андерсон Т. "Статистический анализ временных рядов". - Москва: "МИР", 1976. - 759 с.
2. Кендал М., Стьюарт А. "Многомерный статистический анализ и временные ряды. - Москва: "Наука", 1976. -736 с.
3. Сулайманов Б.А., Файзиев А.А., Файзиев Ж.Н. "Таъриба маълумотларининг статистик таҳлили".- Тошкент: ТошДАУ, 2015. - 124 с.
4. Тихомиров Н.П. «Эконометрика». - Москва: «Экзамен», 2003. - 512 с.
5. Тихонов Э.Е. Методы прогнозирования в условиях рынка: учебное пособие. - Невинномысск, 2006. - 221 с.
6. Файзиев А.А. "Matematik statistika". O'quv qo'llanma. "Ilm-ziyozakovat", 218-bet, Toshkent, 2022.
7. Файзиев А.А., Тургунов Т. "Статистический анализ и прогнозирование динамики урожайности хлопка в Республике Узбекистан" //Журнал.-Бюллетень Института Математики. ISSN 2181-9483, [http: //mib. Mathinst. Uz](http://mib.Mathinst.Uz). Ташкент, 2020. № 1.- С.107-111.
8. Achilov M.U., Fayziev A.A. "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.