Лабораторная работа №6 (Вариант 4)

Гудыма Алексей

Анализ и моделирование сезонных колебаний
временного ряда

Ряд x4. Имеются поквартальные данные по розничному товарообороту региона (млн. долларов).

Выполнение работы

Сезонная компонента (S) отражает повторяемость экономических процессов в течение периода времени, не превышающего года (год, квартал, месяц, неделя, сутки). Чаще всего причиной их возникновения считаются природно-климатические условия сезонов года и ритмы человеческой активности. Например, объём продажи товаров в разные времена года, увеличение цен на сахар в период летних заготовок. Иногда причины сезонных колебаний имеют социальный характер, например, увеличение закупок в предпраздничный период, увеличение платежей в конце квартала.

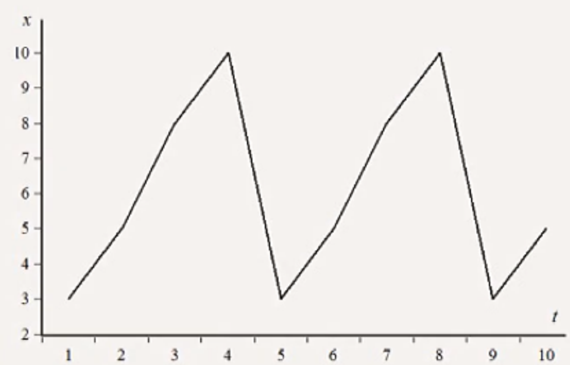
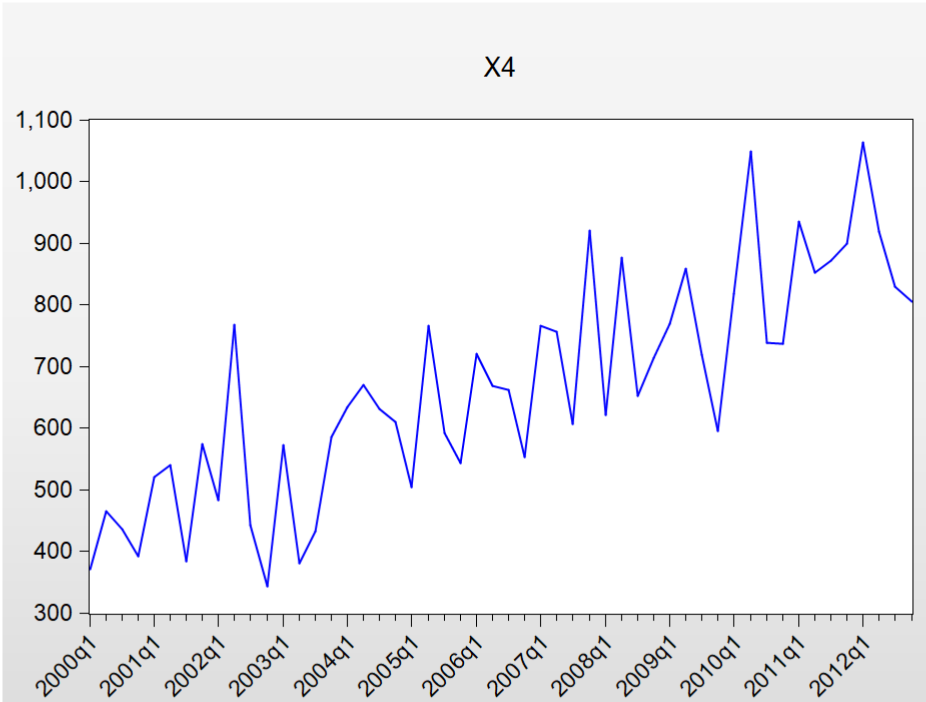


Рисунок 1 – Сезонные колебания

1. **Построить график временного ряда. Сделать вывод о структуре ряда.**



Визуальный анализ графика приводит к следующим выводам:

- присутствует линейный возрастающий тренд;

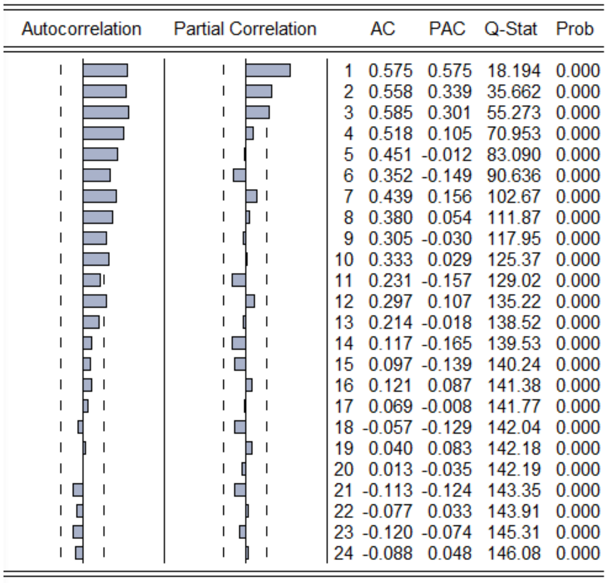
- присутствуют сезонные колебания периодичностью в 4 квартала;

- наблюдается приблизительно равная амплитуда колебаний;

- присутствует случайная компонента.

Объем объемах продаж свежей рыбы сельскохозяйственным предприятием в весенний и летний периоды времени (II и III кварталы) выше, чем в зимний и осенний (I и IV кварталы).

2. **Построить коррелограмму (ВАКФ и ВЧАКФ) данного ВР. Охарактеризовать структуру ряда.**



Date: 05/12/21 Time: 12:25
Sample: 2000Q1 2012Q4
Included observations: 52

	AC	PAC	Q-Stat	Prob
1	0.575	0.575	18.194	0.000
2	0.558	0.339	35.662	0.000
3	0.585	0.301	55.273	0.000
4	0.518	0.105	70.953	0.000
5	0.451	-0.012	83.090	0.000
6	0.352	-0.149	90.636	0.000
7	0.439	0.156	102.67	0.000
8	0.380	0.054	111.87	0.000
9	0.305	-0.030	117.95	0.000
10	0.333	0.029	125.37	0.000
11	0.231	-0.157	129.02	0.000
12	0.297	0.107	135.22	0.000
13	0.214	-0.018	138.52	0.000
14	0.117	-0.165	139.53	0.000
15	0.097	-0.139	140.24	0.000
16	0.121	0.087	141.38	0.000
17	0.069	-0.008	141.77	0.000
18	-0.057	-0.129	142.04	0.000
19	0.040	0.083	142.18	0.000
20	0.013	-0.035	142.19	0.000
21	-0.113	-0.124	143.35	0.000
22	-0.077	0.033	143.91	0.000
23	-0.120	-0.074	145.31	0.000
24	-0.088	0.048	146.08	0.000

Коррелограмма подчеркивает нестационарность исследуемого ряда. Наличие сезонной компоненты можно проследить по волнообразным всплескам на коррелограмме.

3. Сделать предположение о целесообразности построения аддитивной или мультипликативной модели временного ряда.

Общий вид ***аддитивной модели*** с трендом следующий:

$$x_t = T_t + S_t + \varepsilon_t,$$

где x_t – уровень временного ряда в момент времени t; T_t – теоретический уровень ряда согласно тенденции; S_t – сезонная составляющая, измеренная в тех же единицах, что и уровень ряда; ε_t – случайная компонента, измеренная в тех же единицах, что и уровень ряда.

В ***мультипликативной модели*** уровень временного ряда рассматривается как произведение его компонент:

$$x_t = T_t S_t \varepsilon_t,$$

где x_t – фактический уровень временного ряда в момент времени t; T_t – теоретический уровень ряда согласно тенденции; S_t – коэффициент сезонности; ε_t – коэффициент случайной компоненты.

В мультипликативной модели $s_t = T_t S_t$ представляет собой тренд с учетом сезонной волны. Используя величину s_t, мультипликативную модель можно представить как

$$x_t = T_t \frac{s_t}{T_t} \frac{x_t}{s_t},$$

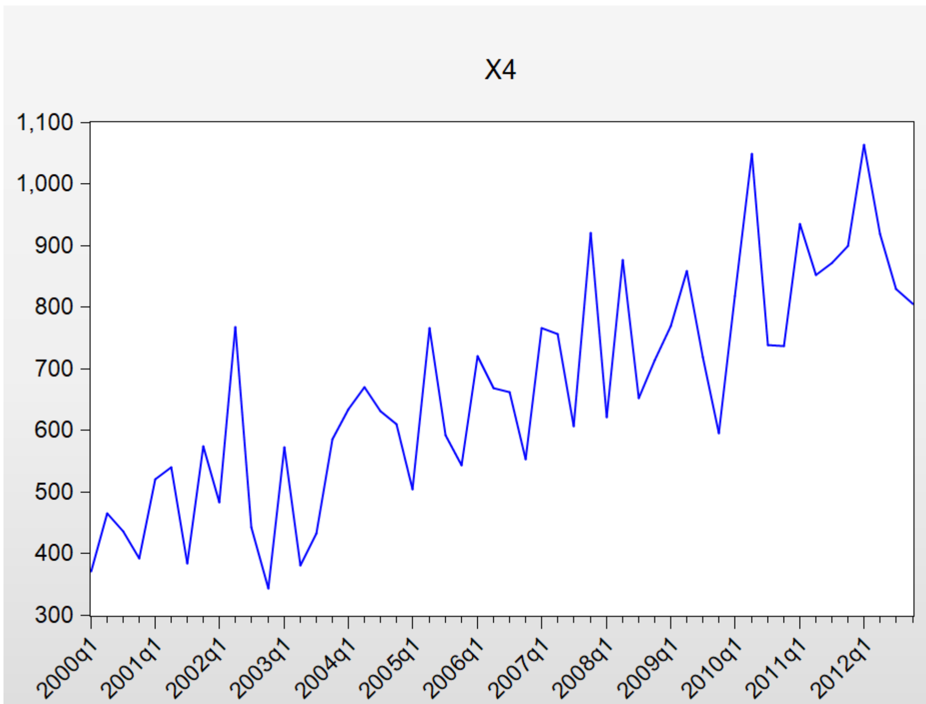
где $\frac{s_t}{T_t} = S_t$, $\frac{x_t}{s_t} = \varepsilon_t$. Тогда, очевидно, отличие мультипликативной модели от аддитивной состоит в том, что *в мультипликативной модели сезонная и случайная составляющие определены в виде относительных величин (коэффициентов)*, а *в аддитивной модели – в виде абсолютных величин (тысяч рублей, тонн, человек и т. п.)*. В мультипликативной модели сезонность выражена в процентах. Так, если коэффициент сезонности для I квартала примет значение 1,3 или 130 %, то при повышающейся тенденции во временном ряду прирост в 30 % будет для I квартала каждого года представлять увеличивающуюся сезонную волну.

Выбор одной из двух моделей (аддитивной или мультипликативной) осуществляется на основе анализа структуры сезонных колебаний.

Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель временного ряда, в которой значения сезонной компоненты предполагаются постоянными для различных циклов.

Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель временного ряда, которая ставит уровни ряда в зависимость от значений сезонной компоненты.

Амплитуда колебаний – разность между наибольшим и наименьшим отклонениями наблюдений от воображаемой линии тренда.



По графику исследуемого ряда видим приблизительно равную амплитуду колебаний временного ряда около воображаемой линии тренда.

4. В зависимости от выводов, сделанных в п. 3, построить аддитивную или мультипликативную модель данного ВР, применить фиктивные переменные.

Аддитивная модель сезонности
как регрессионная модель с фиктивными переменными

Аддитивная модель уровней временного ряда при наличии тенденции и сезонности может быть построена как модель регрессии с включением в нее фактора времени t и фиктивных переменных d. Количество фиктивных переменных в такой модели должно быть на единицу меньше числа моментов (периодов) времени внутри одного цикла колебаний. Например, при моделировании поквартальных данных модель должна включать четыре независимые переменные – фактор времени и три фиктивных переменных. Каждая фиктивная переменная отражает сезонную компоненту временного ряда для какого-либо одного периода. Она равна единице для данного периода и нулю для всех остальных периодов. Так, при квартальном разрезе информации модель примет вид:

$$x_t = \beta_0 + \beta_1 t + \gamma_1 d_1 + \gamma_2 d_2 + \gamma_3 d_3 + \varepsilon_t. \qquad (1)$$

Фактор времени t в этой модели позволит учесть влияние тенденции. Сезонный фактор представлен фиктивными переменными. Переменные d_1, d_2, d_3 соответствуют учету сезонного фактора в I, II и III кварталах соответственно, то есть

$$d_1 = \begin{cases} 1 \text{ – для I квартала,} \\ 0 \text{ – для остальных;} \end{cases} \quad d_2 = \begin{cases} 1 \text{ – для II квартала,} \\ 0 \text{ – для остальных;} \end{cases}$$

$$d_3 = \begin{cases} 1 \text{ – для III квартала,} \\ 0 \text{ – для остальных.} \end{cases}$$

$$x_t = \beta_0 + \beta_1 t + \gamma_1 d_1 + \gamma_2 d_2 + \gamma_3 d_3 + \varepsilon_t. \qquad (1)$$

Так как фиктивные переменные d принимают только значения 1 и 0, то практически мы имеем модель тенденции для каждого квартала:

$$x_t = (\beta_0 + \gamma_1) + \beta_1 t + \varepsilon_t \text{ – для I квартала,}$$

$$x_t = (\beta_0 + \gamma_2) + \beta_1 t + \varepsilon_t \text{ – для II квартала,}$$

$$x_t = (\beta_0 + \gamma_3) + \beta_1 t + \varepsilon_t \text{ – для III квартала,}$$

$$x_t = \beta_0 + \beta_1 t + \varepsilon_t \text{ – для IV квартала.}$$

Иными словами, параметры γ_1, γ_2, γ_3 при фиктивных переменных отражают изменение уровня ряда соответствующего квартала под воздействием сезонности по сравнению с IV кварталом. Параметр β_1 в модели (1) характеризует среднее абсолютное изменение уровней ряда под воздействием тенденций. В сущности, рассмотренная модель есть аналог аддитивной модели временного ряда, поскольку фактический уровень временного ряда – это сумма трендовой, сезонной и случайной компонент.

Применим фиктивные переменные для моделирования сезонности

Dependent Variable: X4
Method: Least Squares
Date: 05/12/21 Time: 12:32
Sample: 2000Q1 2012Q4
Included observations: 52

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	359.1002	34.97121	10.26845	0.0000
T	9.901620	0.853210	11.60514	0.0000
D1	68.46484	36.20870	1.890840	0.0648
D2	119.4476	36.15840	3.303455	0.0018
D3	-11.49411	36.12819	-0.318148	0.7518

R-squared	0.758552	Mean dependent var	665.5977
Adjusted R-squared	0.738004	S.D. dependent var	179.9012
S.E. of regression	92.08348	Akaike info criterion	11.97448
Sum squared resid	398530.3	Schwarz criterion	12.16210
Log likelihood	-306.3365	Hannan-Quinn criter.	12.04641
F-statistic	36.91481	Durbin-Watson stat	2.458789
Prob(F-statistic)	0.000000		

Полученная модель имеет вид

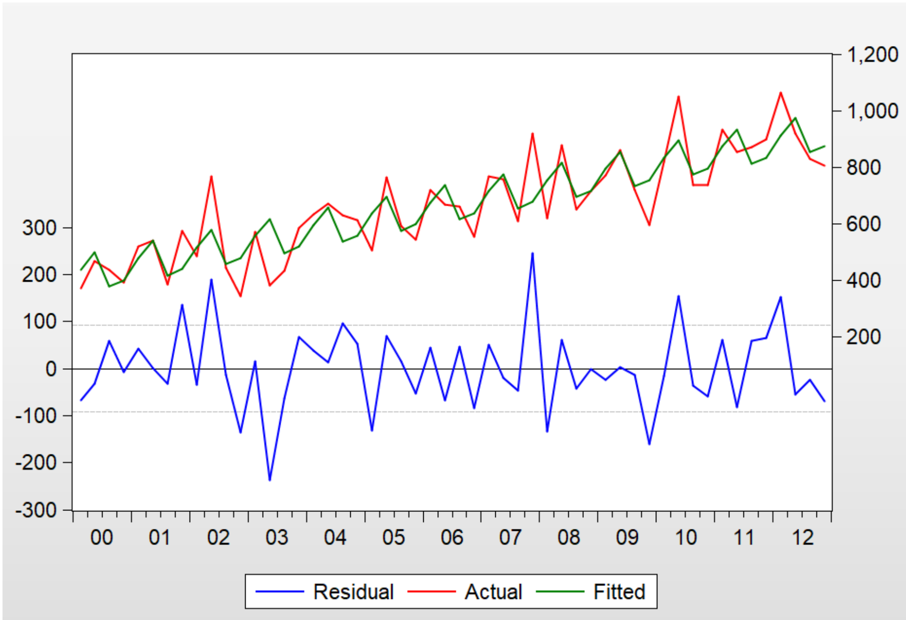
$$X4 = 359.1002 + 9.9016*T + 68.4648*D1 + 119.4476*D2 -11.49411*D3$$

5. Оценить общее качество построенной модели.

Dependent Variable: X4
Method: Least Squares
Date: 05/12/21 Time: 12:32
Sample: 2000Q1 2012Q4
Included observations: 52

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	359.1002	34.97121	10.26845	0.0000
T	9.901620	0.853210	11.60514	0.0000
D1	68.46484	36.20870	1.890840	0.0648
D2	119.4476	36.15840	3.303455	0.0018
D3	-11.49411	36.12819	-0.318148	0.7518

R-squared	0.758552	Mean dependent var	665.5977
Adjusted R-squared	0.738004	S.D. dependent var	179.9012
S.E. of regression	92.08348	Akaike info criterion	11.97448
Sum squared resid	398530.3	Schwarz criterion	12.16210
Log likelihood	-306.3365	Hannan-Quinn criter.	12.04641
F-statistic	36.91481	Durbin-Watson stat	2.458789
Prob(F-statistic)	0.000000		



Качество уравнения высокое, т.к. коэффициент детерминации R^2=0.75, модель дает хорошую подгонку к исходным данным. Уравнение является значимым по F-статистике, все его коэффициенты значимы по соответствующим t-статистикам.

Статистика Дарбина – Уотсона = 2.4, выводам по статистическим характеристикам качества модели можно доверять.

6. Исследовать остатки на выполнение характеристических свойств белого шума.

Гауссовским белым шумом (*white noise*) называют *стационарный* случайный процесс $\varepsilon(t)$, который удовлетворяет характеристическим свойствам:

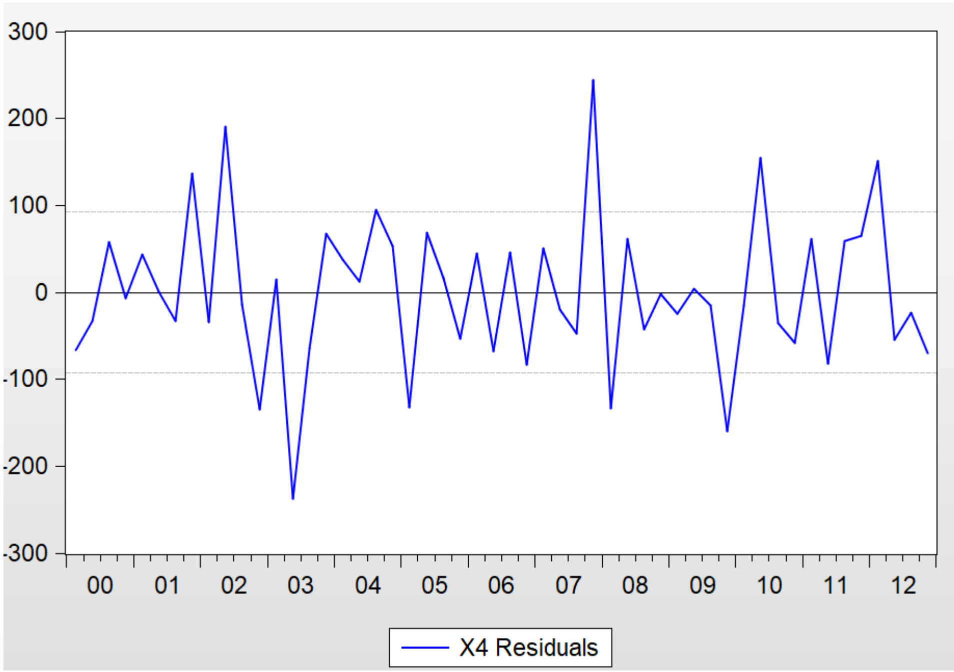
(wn1). $M(\varepsilon_t) = 0$ – среднее значение равно нулю.

(wn2). $D(\varepsilon_t) = \sigma_\varepsilon^2 > 0$ – гомоскедастичность (*тест Уайта и др.*).

(wn3). $\rho_k = \rho(k) = corr(\varepsilon_t, \varepsilon_{t-k}) = 0$ при $k \neq 0$ – отсутствие автокорреляции (*коррелограмма, тест Льюинга – Бокса*).

(wn4). $\varepsilon(t)$ имеет нормальное распределение (*тест Жака – Бера*).

(wn1) График остатков



Визуальный анализ графика остатков показывает, что в среднем остатки колеблются около нуля в горизонтальной полосе приблизительно постоянной ширины.

Sample Mean = 665.5977
Sample Std. Dev. = 179.9012

Method	Value	Probability
t-statistic	26.67960	1.0000

Видим, что P=1>0,05, значит с вероятностью 95% среднее значение остатков значимо не отличается от нуля. Характеристическое условие (wn1) для белого шума выполняется.

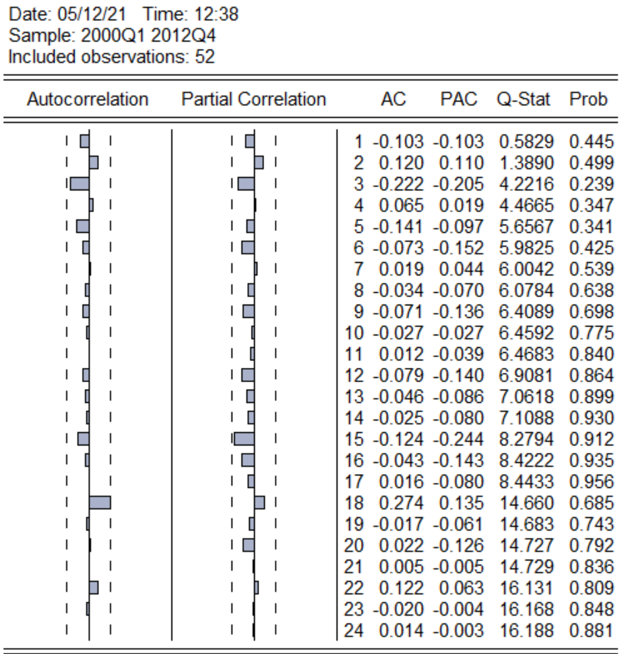
Heteroskedasticity Test: White

F-statistic	0.790016	Prob. F(8,43)	0.6141
Obs*R-squared	6.663537	Prob. Chi-Square(8)	0.5733
Scaled explained SS	7.553530	Prob. Chi-Square(8)	0.4782

(wn2)

Так как P=0.6141>0.05, согласно тесту Уайта, остатки гомоскедастичны. Характеристическое условие (wn2) для белого шума выполняется.

(wn3)



Date: 05/12/21 Time: 12:38
Sample: 2000Q1 2012Q4
Included observations: 52

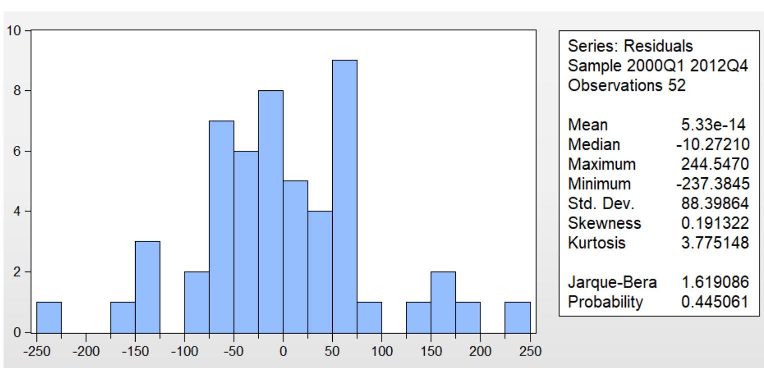
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.103	-0.103	0.5829	0.445
		2	0.120	0.110	1.3890	0.499
		3	-0.222	-0.205	4.2216	0.239
		4	0.065	0.019	4.4665	0.347
		5	-0.141	-0.097	5.6567	0.341
		6	-0.073	-0.152	5.9825	0.425
		7	0.019	0.044	6.0042	0.539
		8	-0.034	-0.070	6.0784	0.638
		9	-0.071	-0.136	6.4089	0.698
		10	-0.027	-0.027	6.4592	0.775
		11	0.012	-0.039	6.4683	0.840
		12	-0.079	-0.140	6.9081	0.864
		13	-0.046	-0.086	7.0618	0.899
		14	-0.025	-0.080	7.1088	0.930
		15	-0.124	-0.244	8.2794	0.912
		16	-0.043	-0.143	8.4222	0.935
		17	0.016	-0.080	8.4433	0.956
		18	0.274	0.135	14.660	0.685
		19	-0.017	-0.061	14.683	0.743
		20	0.022	-0.126	14.727	0.792
		21	0.005	-0.005	14.729	0.836
		22	0.122	0.063	16.131	0.809
		23	-0.020	-0.004	16.168	0.848
		24	0.014	-0.003	16.188	0.881

По коррелограмме делаем вывод, что автокорреляция отсутствует, так как все столбцы коррелограммы лежат внутри доверительной трубки.

Р-значения Q-статистики Льюинга - Бокса больше 0.05. Поэтому автокорреляция остатков отсутствует.

Можем считать, что характеристическое свойство (wn3) белого шума выполняется.

(wn4)



Используем тест Жака-Бера. Значение P=0.45>0,05, значит распределение остатков согласуется с нормальным распределением и характеристическое свойство (wn4) гауссовского белого шума выполняется.

Вывод: Выполнены характеристические свойства wn1-wn4. Можно утверждать, что остатки построенной модели являются гауссовским белым шумом.

7. Сделать экономические выводы на основе полученной модели.

$$x_t = \beta_0 + \beta_1 t + \gamma_1 d_1 + \gamma_2 d_2 + \gamma_3 d_3 + \varepsilon_t. \quad (1)$$

Еще раз вспомним, что параметры γ_1, γ_2, γ_3 при фиктивных переменных отражают изменение уровня ряда соответствующего квартала под воздействием сезонности по сравнению с IV кварталом. Параметр β_1 в модели (1) характеризует среднее абсолютное изменение уровней ряда под воздействием тенденций. В сущности, рассмотренная модель есть аналог аддитивной модели временного ряда, поскольку фактический уровень временного ряда – это сумма трендовой, сезонной и случайной компонент.

Эмпирическое уравнение имеет вид

Полученная модель имеет вид

X4 = 359.1002 + 9.9016*T + 68.4648*D1 + 119.4476*D2 -11.49411*D3

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	359.1002	34.97121	10.26845	0.0000
T	9.901620	0.853210	11.60514	0.0000
D1	68.46484	36.20870	1.890840	0.0648
D2	119.4476	36.15840	3.303455	0.0018
D3	-11.49411	36.12819	-0.318148	0.7518

Проанализируем полученные результаты.

Коэффициент $\beta 0$= 359.1002 статистически значим и характеризует уровень IV квартала нулевого ряда.

Коэффициент $\beta 1$= 9.901620 свидетельствует о возрастании уровней ряда при элиминировании сезонности, то есть ежеквартально независимо от влияния сезонности розничный товарооборот региона возрастает в среднем на 9.901620 млн. долларов.

Поскольку фактическое значение t-критерия равно 11.60, можно утверждать, что существование тенденции в уровнях ряда установлено надежно.

Влияние сезонной компоненты не в каждом из кварталов статистически значимо по t-критерию. Розничный товарооборот региона в I квартале статистически значим, во втором и третьем статистически незначим, отличается от розничного товарооборота региона в IV квартале.

Параметр γ1 = 68.46 – положительный, значит розничный товарооборот региона в I квартале на 68.46 млн. долларов больше, чем в IV квартале.

Он показывает, что во I квартале объем экспорта региона независимо от влияния тенденции был в среднем ниже, чем в IV квартале на 2.352776 тыс. кг.

Параметр γ2 = 119.44 – положительный, значит объем продаж свежей рыбы региона во II квартале на 119.44 млн. долларов больше, чем в IV квартале.

Параметр γ3 = -11.49 – отрицательный. Он показывает, что в I I I квартале объем экспорта региона независимо от влияния тенденции был в среднем ниже, чем в IV квартале на 11.49 млн. долларов.

8. Рассчитайте прогноз исследуемого показателя на 5 периодов времени вперед.

Рассчитаем прогнозные значения переменной x4.

T	X4
1 квартал 2013	871.90482
2 квартал 2013	899.34768
3 квартал 2013	1064.6748
4 квартал 2013	828.9063
1 квартал 2014	804.7893

9. Определить значения сезонных компонент аддитивной модели. Сделать экономические выводы.

Моделирование сезонности подробно описано в § 6.7 (стр. 239 -253).

Отметим, что параметры $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}_3$ не равны значениям сезонной компоненты, поскольку они характеризуют не уровни сезонности соответствующего квартала, а их отличие от уровней, учитывающих сезонные воздействия в IV квартале.

Вместе с тем, зная параметры $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}_3$, можно определить показатели сезонности. Исходя из содержания параметров при фиктивных переменных, имеем

$$\gamma_1 = S_1 - S_4,\ \gamma_2 = S_2 - S_4,\ \gamma_3 = S_3 - S_4,$$

где S_1, S_2, S_3, S_4 – показатели сезонности соответствующих кварталов. Тогда

$$\gamma_1 + \gamma_2 + \gamma_3 = S_1 + S_2 + S_3 - 3S_4.$$

В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются. В аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна нулю: $\sum_{i=1}^{4} S_i = 0$, тогда можно записать, что $\gamma_1 + \gamma_2 + \gamma_3 = 0 - 4S_4 = -4S_4$, отсюда

$$S_4 = -\frac{1}{4}(\gamma_1 + \gamma_2 + \gamma_3).$$

X4 = 359.1002 + 9.9016*T + 68.4648*D1 + 119.4476*D2 -11.49411*D3

Для нашего примера, используя значения эмпирических коэффициентов модели, получим

$\hat{S4} = -1/4(68.46 + 119.44 - 11.49) = 176{,}41$

Далее определяем сезонные компоненты 1, 2 и 3 кварталов.

$\hat{S1} = \hat{\gamma 1} + \hat{S4} = 244{,}87$

$\hat{S2} = \hat{\gamma 2} + \hat{S4} = 295{,}85$

$\hat{S3} = \hat{\gamma 3} + \hat{S4} = 164{,}92$

Контроль $\sum Si$=0 выполняется.

Рассмотрим аддитивную сезонную модель в общем виде: $xt = Tt + St + \varepsilon t$.

Отсюда очевидны экономические выводы, вытекающие из значений сезонных компонент. Каждый 1 квартал объем экспорта региона выше на 244,87 млн. долларов относительно уровня тренда, 2 – выше на 295,85 млн. долларов, 3 – выше на 164,92 млн. долларов, 4 – выше на 176,41 млн. долларов.