

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

имени М. В. Ломоносова

ЭКОНОМИЧЕСКИЙ ФАКУЛЬТЕТ

НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

«ТЕСТЫ НА СТАЦИОНАРНОСТЬ ДЛЯ ВРЕМЕННЫХ РЯДОВ ПРИ МАЛОМ ОБЪЁМЕ  
ВЫБОРКИ»

Выполнил студент группы э201:

Кабанов Илья Дмитриевич

Научный руководитель:

к ф-м н, доцент Артамонов Дмитрий Вячеславович

Москва, 2023

## Содержание

Введение.....	3
Теория структуры временных рядов.....	5
Тесты Дики-Фуллера.....	8
Проведение эмпирического исследования.....	11
Заключение.....	20
Список литературы.....	21
Приложения.....	22

## Введение

Довольно часто в своей исследовательской и аналитической работе экономисты встречаются с данными временных рядов, такие как ВВП, совокупное потребление, уровень безработицы, инфляция и многие другие. В таких рядах особенно важна взаимосвязь между наблюдениями, соответствующими разным периодам времени, тогда исследователь может использовать динамические свойства ряда для того, чтобы формулировать прогнозы на следующие промежутки времени. Одной из ключевых задач в изучении структуры временного ряда  $\{Y_t\}$  является тестирование его слабой стационарности, когда у него существуют моменты первого и второго порядка и они не меняются при сдвиге, то есть при любых  $t, s, t + h, s + h \in T$

$$EY_{t+h} = EY_t, E(Y_{t+h}Y_{s+h}) = E(Y_tY_s) \quad (1.1)$$

( $T$  – временные периоды на которых определен временной ряд  $\{Y_t\}$ )

В дальнейших рассуждениях под стационарностью мы будем иметь в виду именно слабую. Какой содержательный вопрос стоит под стационарностью ряда? Если процесс является таковым, это значит, что некоторый показатель, например, уровень реального выпуска за год, всегда возвращается к своему среднему показателю спустя какое-то время, и внешние шоки вызывают лишь временные отклонения от этого среднего значения, не приводящие к долгосрочным последствиям. Если временной ряд обладает тренд-стационарностью, это говорит о его устойчивости именно относительно тренда, к которому он также со временем возвращается при воздействии внешних шоков. Если две предыдущие гипотезы отвергаются, это значит, что внешние шоки имеют долгосрочные эффекты на показатель. К аналогичному выводу пришли Нельсон и Пlossер, анализируя рост фактического и потенциального ВВП в одной из своих статей (Nelson, Plosser, 1982). Они заключили, что внешние шоки приводят к переходу фактического ВВП с одного тренда на другой, то есть в результате тестирования исследователи отвергли гипотезу о наличии тренд-стационарности в ряду фактического ВВП. На основе этих рассуждений были выведены теории реального делового цикла, по которым основным источником колебаний деловой активности являются шоки производственной функции, приводящие к росту всего тренда (потенциального ВВП).

Для тестирования стационарности чаще всего используются DF-тест и KPSS-тест, в которых при нулевой гипотезе теоретическое распределение расчетной статистики является неклассическим, что подробно описано в основополагающей статье Дики и Фуллера (Dickey, Fuller, 1979) и Квятковского (Kwiatkowski et al., 1992). В данной работе

эмпирически проверяются результаты первой из вышеуказанных статей для разных длин временного ряда, то есть в ходе своей работы я стараюсь ответить на вопрос: «Насколько эмпирические результаты подтверждают теоретические обоснования Дики и Фуллера в их статье 1979 года?». В качестве инструмента для данной эмпирической проверки я собираюсь использовать метод бустрепирования - классический подход для корректировки оценок, критических точек при малых объемах выборки. Его использование и осуществление я подробно опишу в основной части работы, а в заключении я постараюсь сделать некоторые выводы из сравнения результатов полученных мною и Дики и Фуллером.

## Теория структуры временных рядов

Для начала припомним некоторые теоретические обоснования тестирования стационарности временного ряда, которые приведены в книге М. Вербика «Путеводитель по современной эконометрике». Одним из типичных примеров временного ряда является процесс авторегрессии порядка 1 (АР(1)):

$$Y_t = (\mu - \theta\mu) + \theta Y_{t-1} + \varepsilon_t, \text{ где } \mu = E\{Y_t\} \text{ и } \varepsilon_t \sim iid(0, \sigma^2) \quad (2.1)$$

что, сделав замену  $y_t = Y_t - \mu$ , можно переписать как

$$y_t = \theta y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2) \quad (2.2)$$

Частным случаем авторегрессии является процесс случайного блуждания:

$$y_t = y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2) \quad (2.3)$$

Одним из эквивалентных определений стационарности ряда  $\{Y_t\}$  является система условий

$$\text{Для любого } t \in T \quad \left\{ \begin{array}{l} EY_t = \mu < \infty \\ V\{Y_t\} = E\{(Y_t - \mu)^2\} = \gamma_0 < \infty \\ cov\{Y_t, Y_{t-k}\} = E\{(Y_t - \mu)(Y_{t-k} - \mu)\} = \gamma_k, \quad k = 1, 2, 3, \dots \end{array} \right\} \quad (2.4)$$

Необходимым условием стационарности ряда (2.2) является равенство дисперсий в разные моменты времени:  $V\{y_t\} = V\{y_{t-1}\}$ , из чего получаем:  $V\{y_t\} = \frac{\sigma^2}{1-\theta^2}$ .

То есть наложить условие равенства дисперсий можно только в случае, когда  $|\theta| < 1$ .

Также  $cov(y_s, y_t) = \theta^{|s-t|} \frac{\sigma^2}{1-\theta^2}$ , что зависит только от  $(s-t)$  и отражает стационарность процесса.

Получили вывод: процесс (АР(1)) является стационарным только когда  $|\theta| < 1$ , но если это верно, то в таком случае выполняются все условия из системы (2.4) и процесс АР(1) является стационарным по определению. Значит, условие  $|\theta| < 1$  является необходимым и достаточным.

Далее, перепишем процесс (2.2) через оператор сдвига  $L$  ( $Ly_t = y_{t-1}$ ):

$$y_t = \theta Ly_t + \varepsilon_t \quad (2.5)$$

или

$$(1 - \theta L)y_t = \varepsilon_t \quad (2.6)$$

Можно относиться к последнему выражению так: при применении фильтра  $(1 - \theta L)$  к процессу случайного блуждания получается процесс белого шума. В таком случае, можно определить обращение  $(1 - \theta L)$ , определяемое так, что должно выполняться равенство

$$(1 - \theta L)(1 - \theta L)^{-1} = 1. \quad (2.7)$$

Получаем следующий вид для обращения  $(1 - \theta L)$ :

$$(1 - \theta L)^{-1} = \sum_{j=0}^{\infty} \theta^j L^j \text{ при условии, что } |\theta| < 1 \quad (2.8)$$

Этот результат схож с тем, что числовой ряд  $\sum_{j=0}^{\infty} \theta^j$  равен  $(1 - \theta)^{-1}$  тогда и только тогда, когда он сходится и  $|\theta| < 1$ .

Значит, обратимость  $(1 - \theta L)$  тоже является необходимым и достаточным условием стационарности процесса  $AP(1)$ .

Рассмотрим обобщенный случай процесса авторегрессии порядка  $p$  ( $AP(p)$ ):

$$y_t = \theta_1 y_{t-1} + \dots + \theta_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2) \quad (2.9)$$

или

$$y_t = \theta_1 L y_t + \dots + \theta_p L^p y_t + \varepsilon_t \quad (2.10)$$

или

$$\theta(L) y_t = \varepsilon_t, \text{ где } \theta(L) = 1 - \theta_1 L - \dots - \theta_p L^p \quad (2.11)$$

Многочлен  $\theta(L)$  в отсутствии случая комплексных корней (который игнорируется) можно представить в виде  $\theta(L) = (1 - \varphi_1 L) * \dots * (1 - \varphi_p L)$ . Заметим, что весь многочлен  $\theta(L)$  обратим тогда и только тогда, когда каждая из скобок обратима.

Для удобства работы с  $\varphi_i$  вводится характеристическое уравнение относительно  $z$ :

$$(1 - \varphi_1 z) * \dots * (1 - \varphi_p z) = 0 \quad (2.12)$$

Если какое-то решение этого уравнения  $z_i$  по модулю меньше 1, то  $|\varphi_i| > 1$  и многочлен является необратимым. Решение  $z$ , равное единицы, называется единичным корнем.

Аналогично случаю  $AP(1)$ , процесс  $AP(p)$  стационарен если и только если, многочлен  $\theta(L)$  является обратимым, а это равносильно ситуации, когда все скобки обратимы и все  $|\varphi_i| < 1$ . Если в характеристическом уравнении присутствует единичный корень, значит, хотя бы какая-то одна из  $\varphi_i$  равна 1, хотя бы одна из скобок не обратима и весь многочлен не

обратим, а, значит, процесс авторегрессии не стационарен. Поэтому важно тестировать именно наличие единичного корня.

Заметим, что для определения наличия единичного корня вовсе не нужно решать характеристическое уравнение, достаточно проверить:  $\sum_{j=1}^p \theta_j = 1$ . Если это так, то полином  $\theta(z)$ , посчитанный в точке 1, равен 0.

Повторим еще раз, в случае модели процесса авторегрессии порядка  $p$  (AR( $p$ )), временной ряд является не стационарным, когда  $\sum_{j=1}^p \theta_j = 1$ . Именно такая нулевая гипотеза и ставится в тесте Дики-Фуллера (не стационарность), против гипотезы  $\sum_{j=1}^p \theta_j < 1$  (стационарность).

## Тесты Дики-Фуллера

Я собираюсь рассмотреть три ключевые модификации DF-теста для  $AR(1)$ , которые Дики и Фуллер привели в их основной статье (Dickey, Fuller, 1979):

- Без константы и без тренда
- С константой, но без тренда
- С константой и трендом

### 1) Без константы и без тренда

Исследователь предполагает следующую структуру временного ряда:

$$Y_t = \theta Y_{t-1} + \varepsilon_t, \varepsilon_t \sim iid(0, \sigma^2) \quad (3.1)$$

Тогда он переходит к уравнению в первой разности  $\Delta Y_t = (\theta - 1)Y_{t-1} + \varepsilon_t$ , и формулирует гипотезы

$H_0: \theta - 1 = 0$  (не стационарность)

$H_1: \theta - 1 < 0$  (стационарность)

Оценивает парную регрессию  $\Delta Y_t$  на  $Y_{t-1}$  без константы обычным методом наименьших квадратов, и получаем оцененное уравнение:

$$\Delta \hat{Y}_t = (\hat{\theta} - 1)Y_{t-1} \quad (3.2)$$

И проверяет коэффициент при переменной на значимость: считает расчетную t-статистику  $t = \frac{\hat{\theta} - 1}{se(\hat{\theta})}$ , которая при нулевой гипотезе  $\theta = 1$  не имеет стандартного распределения

Стьюдента, в том числе и асимптотически. Причиной этого является не стационарность процесса, которая нарушает предпосылки модели МНК со стохастическими регрессорами и делает МНК-оценки смещенными и несостоятельными. В частности, распределение t-статистик имеет правостороннюю асимметрию, поэтому критические значения будут меньше, чем при аппроксимации Стьюдентом или нормальным распределением. В таком случае нулевая гипотеза отбрасывается на 5%-м уровне значимости, если расчетное значение меньше 5%-го квартиля теоретического распределения. (который, например, при  $T=25$  равен -1,95)



## 2) С константой и без тренда

Исследователь предполагает следующую структуру временного ряда:

$$Y_t = \delta + \theta Y_{t-1} + \varepsilon_t, \varepsilon_t \sim iid(0, \sigma^2), \delta = (1 - \theta)\mu, \mu - \text{среднее значение ряда} \quad (3.3)$$

Тогда он переходит к уравнению в первой разности  $\Delta Y_t = \delta + (\theta - 1)Y_{t-1} + \varepsilon_t$ , и формулирует гипотезы.

H0:  $\theta - 1 = 0$  (не стационарность)

H1:  $\theta - 1 < 0$  (стационарность)

Оценивает парную регрессию  $\Delta Y_t$  на  $Y_{t-1}$  с константой обычным методом наименьших квадратов, и получаем оцененное уравнение:

$$\Delta \hat{Y}_t = \hat{\delta} + (\hat{\theta} - 1)Y_{t-1} \quad (3.4)$$

И проверяет коэффициент при переменной на значимость: считает расчетную t-статистику  $t = \frac{\hat{\theta}-1}{se(\hat{\theta})}$ , которая при нулевой гипотезе  $\theta = 1$  также имеет не стандартное распределение Стьюдента. В таком случае нулевая гипотеза отвергается на 5%-м уровне значимости, если расчетное значение меньше соответствующего 5%-го квантиля теоретического распределения, которое отличен от случая теста без константы. (соответствующий квантиль, например, при T=25 в таком случае равен -3,00)

## 3) С константой и трендом

Исследователь предполагает следующую структуру временного ряда:

$$Y_t = \delta + \theta Y_{t-1} + \gamma t + \varepsilon_t, \varepsilon_t \sim iid(0, \sigma^2), \delta = (1 - \theta)\mu, \mu - \text{среднее значение ряда} \quad (3.5)$$

Тогда он переходит к уравнению в первой разности  $\Delta Y_t = \delta + (\theta - 1)Y_{t-1} + \gamma t + \varepsilon_t$ , и формулирует гипотезы

H0:  $\delta = \gamma = \theta - 1 = 0$  (случайное блуждание)

H1:  $\delta \neq 0, \gamma \neq 0, \theta - 1 < 0$  (тренд-стационарность)

Но чаще вместо этой множественной гипотезы проверяют только  $\theta - 1 = 0$ , подразумевая выполнение остальных двух ( $\delta = \gamma = 0$ ).

Оценивает множественную регрессию  $\Delta Y_t$  на  $Y_{t-1}$ , t и константу обычным методом наименьших квадратов, и получает оцененное уравнение:

$$\Delta \hat{Y}_t = \hat{\delta} + (\hat{\theta} - 1)Y_{t-1} + \hat{\gamma}t \quad (3.6)$$

И проверяет коэффициент при переменной  $Y_{t-1}$  на значимость: считает расчетную t-статистику  $t = \frac{\hat{\theta}-1}{se(\hat{\theta})}$ , которая при нулевой гипотезе  $\theta = 1$  также имеет не стандартное распределение Стьюдента. В таком случае нулевая гипотеза отвергается на 5%-м уровне значимости, если расчетное значение меньше соответствующего 5%-го квантиля теоретического распределения, который отличен от двух предыдущих случаев. (соответствующий квантиль, например, при T=25 в таком случае равен -3,60)

## Проведение эмпирического исследования

В данном разделе я подробно опишу этапы моего эмпирического исследования – проверки критических значений теста Дики-Фуллера при тестировании стационарности для разных объемов выборок. Я собираюсь рассмотреть три ключевые модификации этого теста:

- Без константы и без тренда
- С константой, но без тренда
- С константой и трендом

Для процесса, который должен соответствовать нулевой гипотезе и быть не стационарным, я взял самый простой процесс случайного блуждания.

### 1) Тест без константы и без тренда

Положим  $T = 25$ , сначала я сгенерировал процесс белого шума со стандартным нормальным распределением  $\{\varepsilon_t\}$  с длиной  $T+1$ , потом из него получил процесс случайного блуждания  $\{Y_t\}$  длины  $T+1$  следующим индуктивным способом;

$$Y_t = Y_{t-1} + \varepsilon_t, Y_0 = 0, \quad (4.1)$$

который впоследствии я тестировал на наличие единичного корня через тест Дики-Фуллера.

Перейдем к уравнению в первых разностях, вычитая из обеих частей  $Y_{t-1}$ ,

$$\Delta Y_t = \varepsilon_t, \Delta Y_1 = \varepsilon_0 \quad (4.2)$$

Далее, я оценил парную регрессию  $\Delta Y_t$  на  $Y_{t-1}$  стандартным методом наименьших квадратов без константы:

$$\Delta Y_t = Y_{t-1} + \varepsilon_t,$$

$$\text{где } \Delta Y_t = Y_t - Y_{t-1} \quad (4.3)$$

Получил оцененное уравнение:

$$\Delta \hat{Y}_t = \hat{\beta} Y_{t-1} \quad (4.4)$$

и посчитал расчетную t-статистику  $t = \frac{\hat{\beta}}{se(\hat{\beta})}$  для тестирования значимости коэффициента при переменной. Заметим, что при таком тесте нулевая гипотеза  $\beta = 0$  равносильна гипотезе  $\theta = 1$ .

Затем я таким же способом сгенерировал процесс случайного блуждания еще 999 раз для  $T=25$ , и получил некоторое эмпирическое распределение  $t$ -статистик (Гистограмма 1), которое в пределе должно совпадать с теоретическим распределением Дики-Фуллера. Нулевой гипотезе в этом тесте соответствует отсутствие стационарности, и получил я именно распределение  $t$ -статистики при гипотезе  $H_0$ .



Гистограмма 1. Эмпирическое распределение расчетных  $t$ -статистик для DF-теста без константы и тренда ( $T=25$ )

Как можно увидеть из Гистограмма 1, эмпирическое распределение имеет правостороннюю асимметрию, что верно согласуется с теорией.

В конечном счете, я получил 5%-й квартиль эмпирического распределения (-1.97), который должен приблизительно совпадать с соответствующим критическим значением теста при  $T=25$  и 5%-м уровнем значимости (-1.95). (Fuller, 1976, p. 373)<sup>1</sup>. Такие 5%-е квартили я посчитал при каждом  $T = [25, 50, 75, 100, 150, 250, 350, 500, 600]$  и построил график, как от длины временного ряда зависит критическое значение статистики Дики-Фуллера в двух случаях (График 1): при сгенерированном эмпирическом распределении (parametric\_bootstrap) и при теоретическом распределении (table\_values, значения взяты из той же таблицы). Абсолютно аналогичные шаги я предпринял при расчетах эмпирических критических точек для этого же теста при уровне значимости в 1% (График 2).

<sup>1</sup> Таблица 2 теоретических критических значений представлена в Приложении

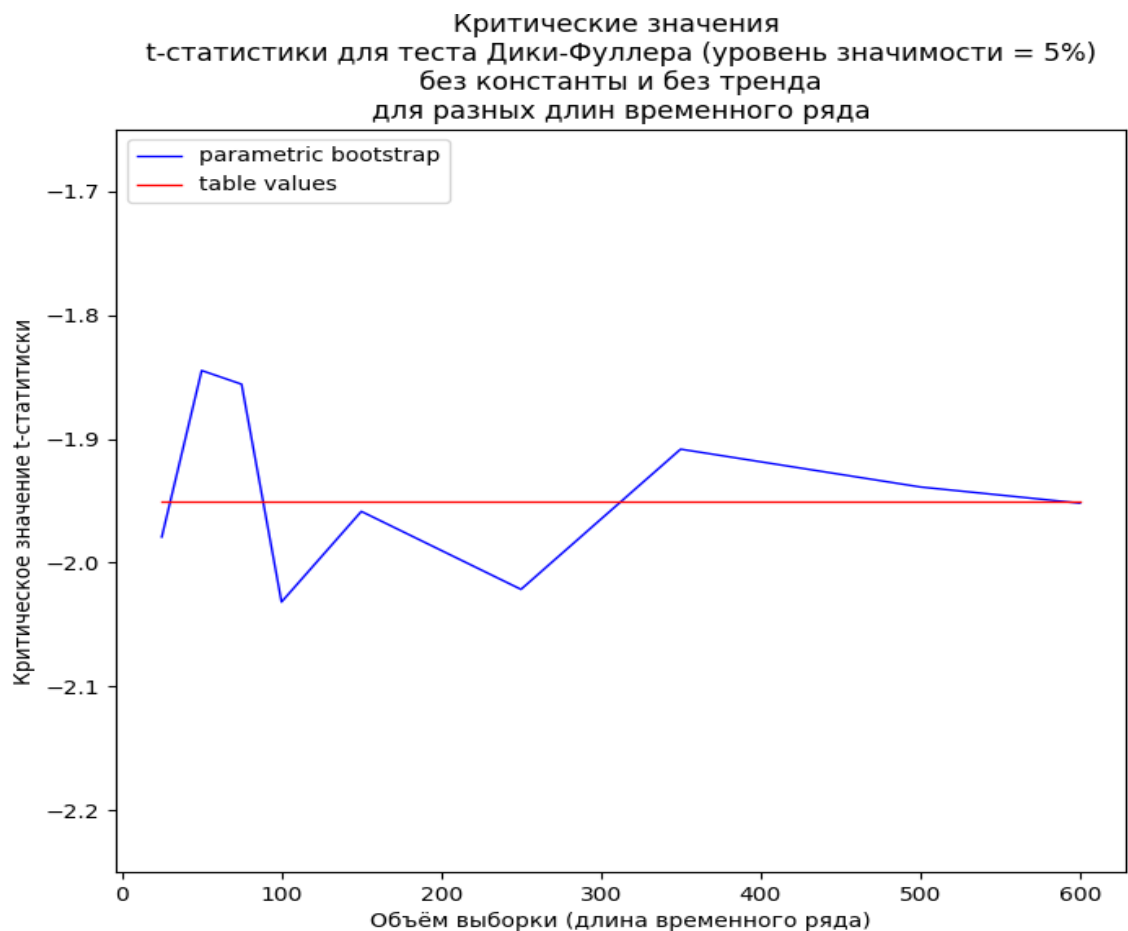


График 1. Критические значения  $t$ -статистик для DF-теста (уровень значимости = 5%) без константы и тренда для разных длин временного ряда

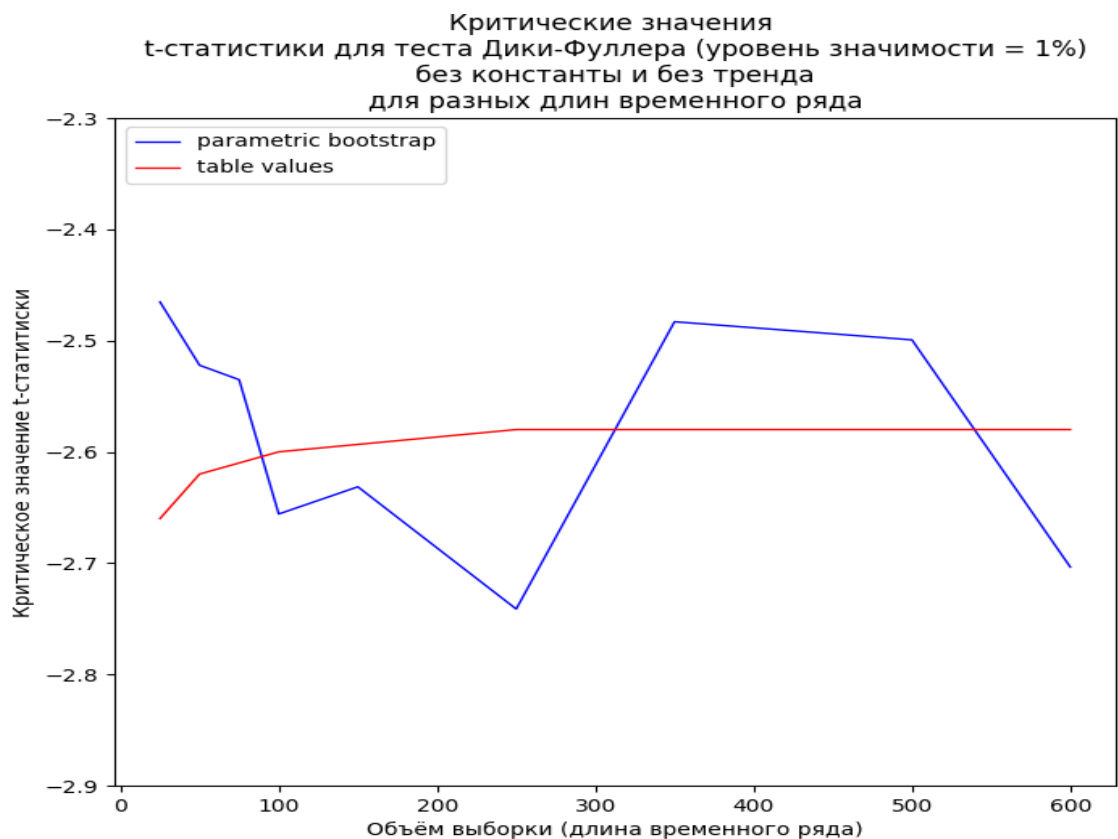


График 2. Критические значения  $t$ -статистик для DF-теста (уровень значимости = 1%) без константы и тренда для разных длин временного ряда

2) Тест с константой, но без тренда

Положим  $T = 25$ , сначала я сгенерировал процесс белого шума со стандартным нормальным распределением  $\{\varepsilon_t\}$  с длиной  $T+1$ , потом из него получил процесса случайного блуждания  $\{y_t\}$  длины  $T+1$  следующим индуктивным способом;

$$Y_t = Y_{t-1} + \varepsilon_t, Y_0 = 0, \quad (4.5)$$

который впоследствии я тестировал на наличие единичного корня через тест Дикки-Фуллера.

Перейдем к уравнению в первых разностях, вычитая из обеих частей  $Y_{t-1}$ ,

$$\Delta Y_t = \varepsilon_t, \Delta Y_1 = \varepsilon_0 \quad (4.6)$$

Далее я оценил парную регрессию с константой стандартным методом наименьших квадратов

$$\begin{aligned} \Delta Y_t &= \delta + Y_{t-1} + \varepsilon_t, \\ \text{где } \Delta Y_t &= Y_t - Y_{t-1} \end{aligned} \quad (4.7)$$

Получил оцененное уравнение:

$$\Delta \hat{Y}_t = \hat{\delta} + \hat{\beta} Y_{t-1} \quad (4.8)$$

и посчитал расчетную t-статистику  $t = \frac{\hat{\beta}}{se(\hat{\beta})}$  для тестирования значимости коэффициента при переменной. Заметим, что при таком тесте нулевая гипотеза  $\beta = 0$  равносильна гипотезе  $\theta = 1$ .

Затем я таким же способом сгенерировал процесс случайного блуждания еще 999 раз для  $T=25$ , и получил некоторое эмпирическое распределение t-статистик (Гистограмма 2), которое в пределе должно совпадать с теоретическим распределением Дики-Фуллера. Нулевой гипотезе в этом тесте соответствует отсутствие стационарности, и получил я именно распределение t-статистики при гипотезе  $H_0$ .



*Гистограмма 2. Эмпирическое распределение расчетных t-статистик для DF-теста с константой, но без тренда (T=25)*

Как можно увидеть из Гистограмма 2, эмпирическое распределение имеет правостороннюю асимметрию, что верно согласуется с теорией.

В конечном счете, я получил 5%-й квантиль эмпирического распределения (-2.96), который должен приблизительно совпадать с соответствующим критическим значением теста при  $T=25$  и 5%-м уровне значимости (-3.00) (Fuller, 1976, p. 373).

Такие 5%-е квантили я посчитал при каждом  $T = [25, 50, 75, 100, 150, 250, 350, 500, 600]$  и построил график, как от длины временного ряда зависит критическое значение статистики Дикки-Фуллера в двух случаях (График 3): при сгенерированном эмпирическом распределении (`parametric_bootstrap`) и при теоретическом распределении (`table_values`, значения взяты из той же таблицы). Абсолютно аналогичные шаги я предпринял при расчетах эмпирических критических точек для этого же теста при уровне значимости в 1% (График 4).

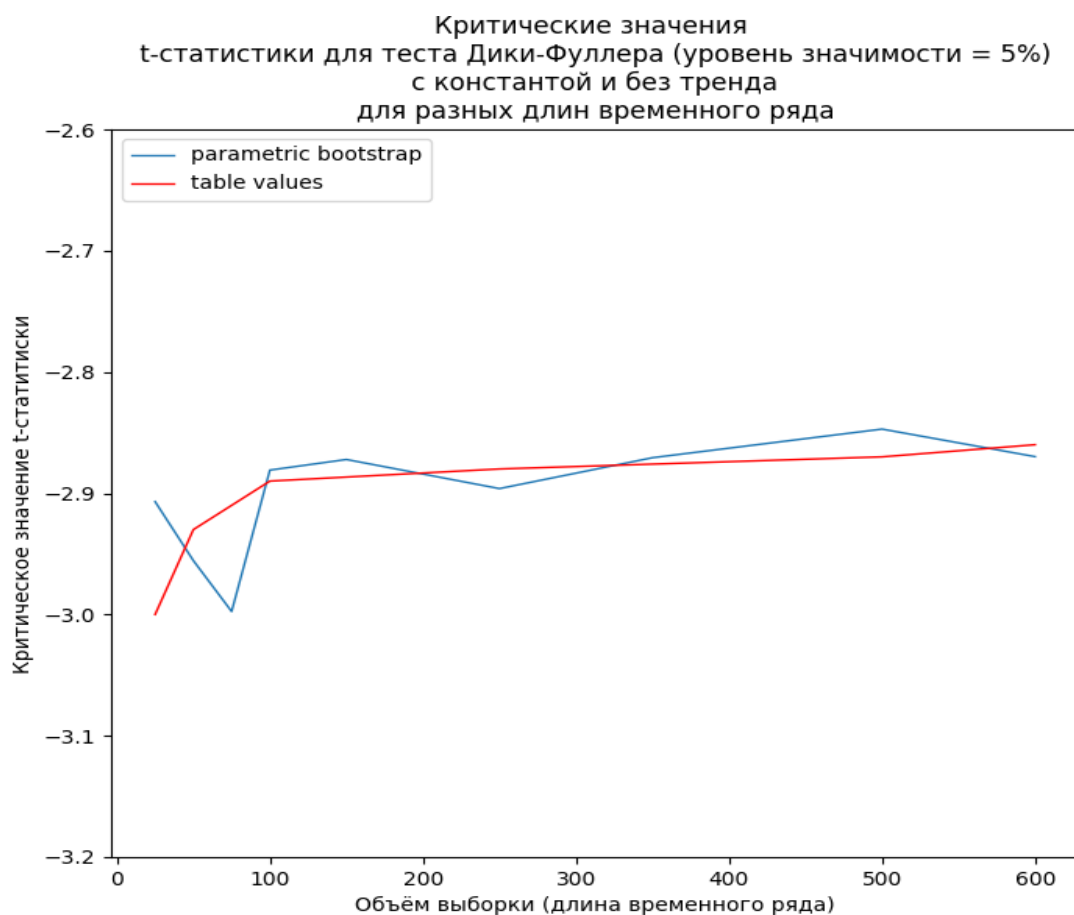


График 3. Критические значения t-статистик для DF-теста (уровень значимости = 5%) с константой и без тренда для разных длин временного ряда.

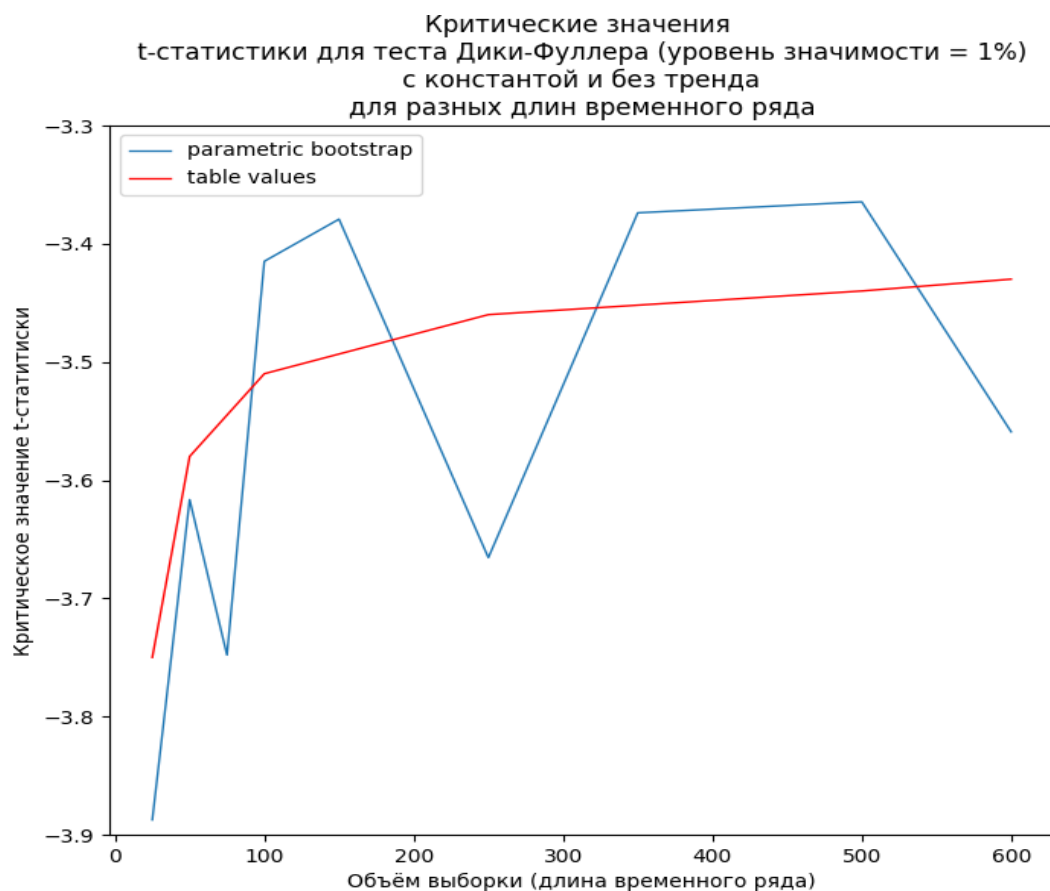


График 4. Критические значения t-статистик для DF-теста (уровень значимости = 1%) с константой и без тренда для разных длин временного ряда.



### 3) Тест с константой и трендом

Положим  $T = 25$ , сначала я сгенерировал процесс белого шума со стандартным нормальным распределением  $\{\varepsilon_t\}$  с длиной  $T+1$ , потом из него получил процесса случайного блуждания  $\{y_t\}$  длины  $T+1$  следующим индуктивным способом;

$$Y_t = Y_{t-1} + \varepsilon_t, Y_0 = 0, \quad (4.9)$$

который впоследствии я тестировал на наличие единичного корня через тест Дикки-Фуллера.

Перейдем к уравнению в первых разностях, вычитая из обеих частей  $y_{t-1}$ ,

$$\Delta Y_t = \varepsilon_t, \Delta Y_1 = \varepsilon_0 \quad (4.10)$$

Далее я оценил множественную регрессию  $\Delta Y_t$  на  $Y_{t-1}$ ,  $t$  и константу стандартным методом наименьших квадратов

$$\begin{aligned} \Delta Y_t &= \delta + (\theta - 1)Y_{t-1} + \gamma t + \varepsilon_t \\ \text{где } \Delta Y_t &= Y_t - Y_{t-1} \end{aligned} \quad (4.11)$$

Получил оцененное уравнение:

$$\Delta \hat{Y}_t = \hat{\delta} + (\hat{\theta} - 1)Y_{t-1} + \hat{\gamma}t \quad (4.12)$$

и посчитал расчетную t-статистику  $t = \frac{\hat{\beta}}{se(\hat{\beta})}$  для тестирования значимости коэффициента при переменной  $Y_{t-1}$ . (Как было описано в предшествующем разделе, обычно проверяется именно эта нулевая гипотеза, подразумевающая выполнение остальных условий  $\delta = \gamma = 0$ ) Заметим, что при таком тесте нулевая гипотеза  $\beta = 0$  равносильна гипотезе  $\theta = 1$ .

Затем я таким же способом сгенерировал процесс случайного блуждания еще 999 раз для  $T=25$ , и получил некоторые эмпирическое распределение t-статистик (Гистограмма 3), которое в пределе должно совпадать с теоретическим распределением Дикки-Фуллера. Нулевой гипотезе в этом тесте соответствует отсутствие стационарности, и получил я именно распределение t-статистики при гипотезе  $H_0$ .



*Гистограмма 3. Эмпирическое распределение расчетных t-статистик для DF-теста с константой и трендом (T= 25)*

Как можно увидеть из Гистограмма 3, эмпирическое распределение имеет правостороннюю асимметрию, что верно согласуется с теорией.

В конечном счете, я получил 5%-й квартиль эмпирического распределения (-3.57), который должен приблизительно совпадать с соответствующим критическим значением теста при  $T=25$  и 5%-м уровнем значимости (-3.6) (Fuller, 1976, p. 373). Такие 5%-е квартили я посчитал при каждом  $T = [25, 50, 75, 100, 150, 250, 350, 500, 600]$  и построил график, как от длины временного ряда зависит критическое значение статистики Дикки-Фуллера в двух случаях (График 5): при сгенерированном эмпирическом распределении (parametric\_bootstrap) и при теоретическом распределении (table\_values, значения взяты из той же таблицы).

Абсолютно аналогичные шаги я предпринял при расчетах эмпирических критических точек для этого же теста при уровне значимости в 1% (График 6).

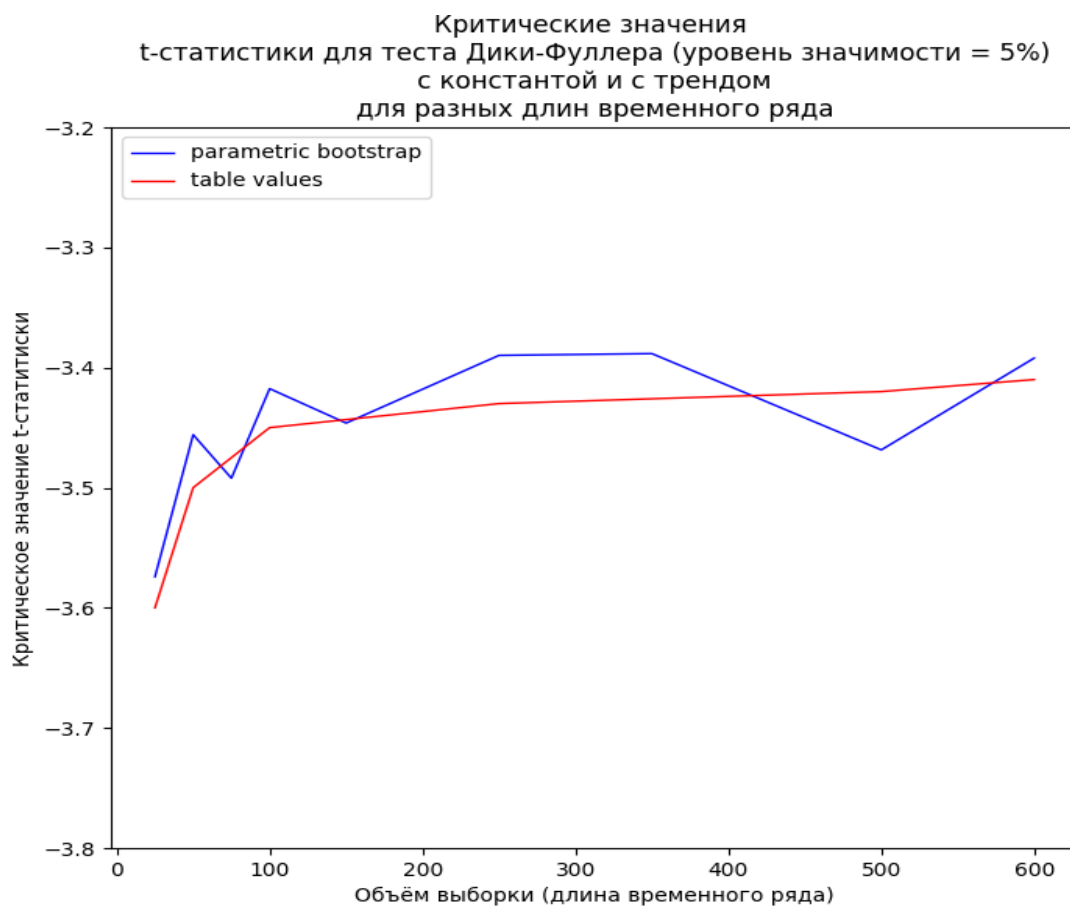


График 5. Критические значения t-статистик для DF-теста (уровень значимости = 5%) с константой и трендом для разных длин временного ряда.

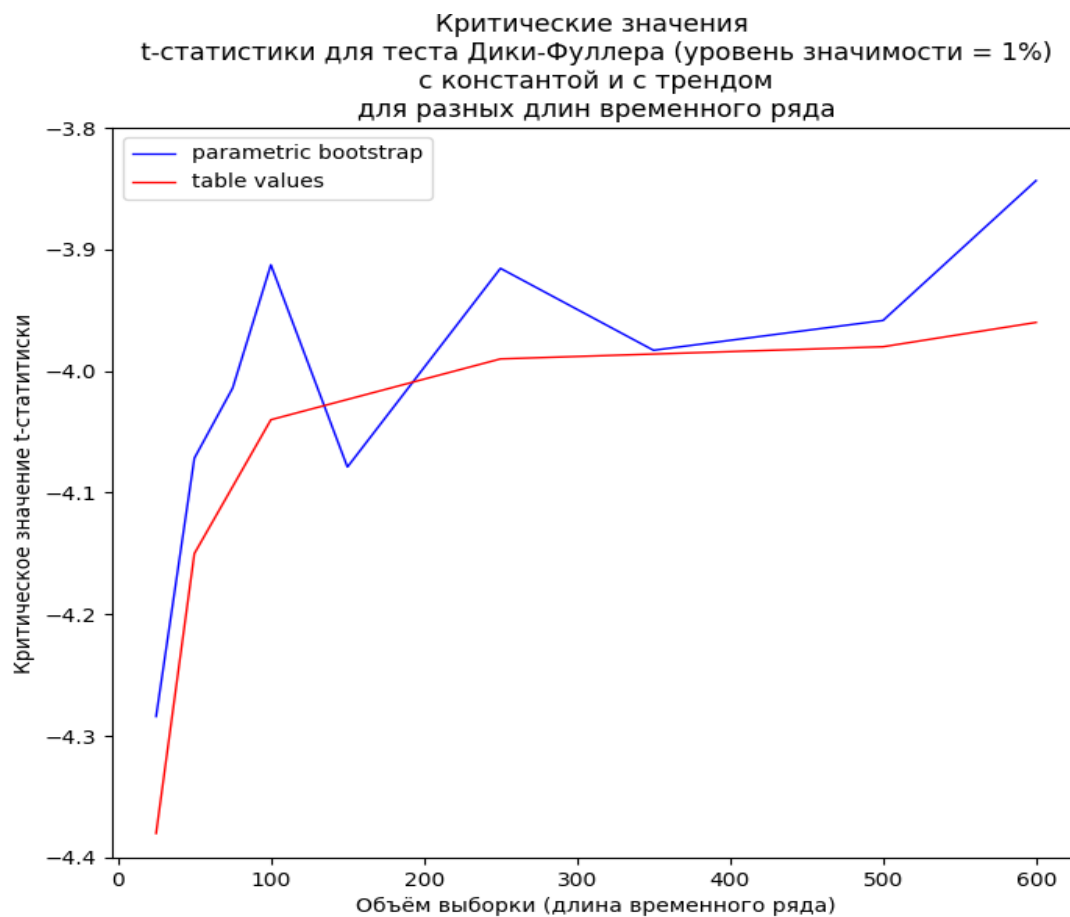


График 6. Критические значения t-статистик для DF-теста (уровень значимости = 1%) с константой и трендом для разных длин временного ряда.

## Заключение

В данном разделе я бы хотел ответить на вопрос, поставленный в самом начале этой работе: «Насколько эмпирические результаты подтверждают теоретические обоснования Дики и Фуллера в их статье 1979 года?». Согласно графикам 1-6, критические точки, полученные эмпирической генерацией случайных блужданий, довольно близки с соответствующими теоретическими значениями. Для того чтобы иметь некоторые численные характеристики этой точности, для каждого из ранее указанных тестов посчитаем модули разностей между теоретическими и эмпирическими критическими значениями. Далее по таким модулям посчитаем максимальное и среднее значение для каждого вида теста и для каждого уровня значимости. Результаты этих вычислений (среднее/максимальное значение) представлены в Таблица 1.

*Таблица 1. Среднее и максимальное значение модулей разности теоретических и эмпирических критических точек.*

Вид теста / Уровень значимости	Без константы и тренда	С константой, но без тренда	С константой и трендом
5 %	(0.046, 0.105)	(0.025, 0.093)	(0.035, 0.059)
1%	(0.096, 0.195)	(0.080, 0.137)	(0.066, 0.127)

Согласно Таблица 1 ни один эмпирический 5%-й квантиль не отклонился от теоретического более, чем 0.105. А в среднем такое отклонение составило около 0,035. Это говорит о довольно высоком совпадении результатов, полученных двумя путями. Для 1%-х тестов точность между эмпирическими и теоретическими значениями, очевидно, будет несколько меньше, что подтверждается и на данном примере. Сгенерированные 1%-е квантили максимально отклонились от табличных значений примерно на 0.2, а их среднее отклонение составило 0.081. Несмотря на то что оба значения говорят о более низкой точности эмпирических 1%-х квантилей по сравнению с 5%-ми, эти расхождения от табличных значений все равно можно считать несущественными. На основе данных рассуждений можно без сомнений сделать вывод о том, что нестандартное распределение t-статистик при тестировании стационарности временных рядов, обоснованное Дики и Фуллером, не только верно с теоретической точки зрения, но и подтверждается приведенными экспериментальными результатами. Эти соображения лишний раз позволяют быть уверенными в своих выводах, прибегая к DF-тесту.

## Список литературы

*M. Verbeek, 2008. "A Guide to Modern Econometrics", 379-397*

*A. Colin Cameron, Pravin K. Trivedi, 2005. "Microeconometrics. Methods and applications", 357-383*

*B. Efron, R. J. Tibshirani, 1988. "An introduction to the Bootstrap", Journal of Economic Literature, Vol. 34, No. 3, pp. 1340-1342*

*S. Anatolyev, 2007. "The basics of bootstrapping", Quantile, No. 3*

*B. E. Hansen, 1995. "Rethinking the univariate approach to unit root testing", Economic Theory, 11, 1148-1171*

*A. Hornok, R. Larsson. "The finite sample distribution of the KPSS test", The Econometrics Journal, Vol. 3, No. 1(2000), pp. 108-121*

*Charles R. Nelson, Charles I. Plosser. "Trends and random walks in macroeconomic time series. Some Evidence and Implications", Journal of Monetary Economics, 10 (1982), 139-162*

*David A. Dickey, Wayne A. Fuller. "Distribution of the Estimators for Autoregressive Time Series With a Unit Root", Journal of the American Statistical Association, Vol. 74, No. 366 (1979), pp. 427-431*

*Denis Kwiatkowski, Peter C. B. Phillips, Peter Schmidt, Yongcheol Shin. "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?", Journal of Econometrics, Nol. 54, 1-3 (1992), pp. 159-178*

## Приложение

Таблица 2. 1%-е и 5%-е критические значения для тестов  
Дики-Фуллера (Fuller, 1976, p. 373)

Объем выборки	Без константы Без тренда		Константа Без тренда		Константа Тренд	
	1%	5%	1%	5%	1%	5%
$T = 25$	-2,66	-1,95	-3,75	-3,00	-4,38	-3,60
$T = 50$	-2,62	-1,95	-3,58	-2,93	-4,15	-3,50
$T = 100$	-2,60	-1,95	-3,51	-2,89	-4,04	-3,45
$T = 250$	-2,58	-1,95	-3,46	-2,88	-3,99	-3,43
$T = 500$	-2,58	-1,95	-3,44	-2,87	-3,98	-3,42
$T = \infty$	-2,58	-1,95	-3,43	-2,86	-3,96	-3,41