МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Московский государственный университет экономики, статистики и информатики
Московский международный институт эконометрики, информатики, финансов и права

Садовникова Н.А.
Шмойлова Р.А.

Анализ временных рядов и прогнозирование

Москва 2001

УДК – 311
ББК – 60.6
С – 143

Садовникова Н.А., Шмойлова Р.А. Анализ временных рядов и прогнозирование. Учебное пособие./ Московский государственный университет экономики, статистики и информатики – М., 2001 г., 67 с.

Содержание

Введение

Современный этап социально-экономического развития страны выдвигает на первый план задачу оценки состояния и перспектив развития субъектов рыночных отношений на различных иерархических уровнях управления с целью выбора оптимальных управленческих, организационно-правовых и производственно-хозяйственных решений, направленных на повышение эффективности и деловой активности их функционирования и взаимодействия как в границах внутренней, так и внешней среды. В этой связи возрастает роль методологии статистического оценивания состояния, основных тенденций и перспектив развития субъектов рыночных отношений – организационно-правовых структур вне зависимости от отраслевой принадлежности, форм собственности и внутренней структурной градации, а также возрастает роль анализа числовой информации, представленной в формах статистической, бухгалтерской и других видах отчетности деятельности фирм, коммерческих банков, страховых компаний, различных сегментов финансового и других рынков с целью определения перспектив их развития и путей принятия наиболее эффективных решений и направлений дальнейшей деятельности.

Учебное пособие «Анализ временных рядов и прогнозирование» включает в себя комплексную методологию статистического анализа, моделирования и прогнозирования информации, представленной временными рядами социально-экономических явлений и процессов. В пособии нашло отражение обобщение отечественного и зарубежного опыта использования математико-статистических методов изучения и прогнозирования социально-экономических явлений и процессов.

Применение рассмотренной в учебном пособии методологии анализа и прогнозирования на основе временных рядов имеет достаточно широкое прикладное значение и может использоваться при решении таких конкретных задач исследования реальных социально-экономических явлений и процессов как моделирование и прогнозирование технико-экономических показателей отрасли, фирмы, предприятия; деловой активности и эффективности функционирования организационно-правовой структуры (фирмы, коммерческого банка, страховой компании и так далее); конкурентоспособности; технического и социального развития; кадровых ресурсов и кадровой политики; финансовой устойчивости и финансового состояния фирмы; рынка жилья; мотивов поведения потребителей; товарной структуры, сегментов рынка; рекламы в системе маркетинговых коммуникаций; потребности и управления персоналом; внешней и внутренней предпринимательской среды; ликвидности, доходности, кредитоспособности, эффективности использования капитала, показателей платежеспособности, оборотного капитала, финансовых результатов в отраслях сферы товарного обращения, банковских структурах, страховых компаниях, акционерных, малых и других форм организации предприятий; систем имущественного, подотраслей личного и государственного социального страхования; финансовой устойчивости и деловой активности сегментов фондового рынка, рынка ценных бумаг, биржевых структур; надежности и стабильности, эффективности и деловой активности, конкурентоспособности банковских структур и кредитной политики банковских структур.

Конечной целью создания данного учебного пособия является формирование у будущих специалистов глубоких теоретических знаний методологии анализа временных рядов, статистического моделирования и прогнозирования, и практических навыков по экономико-статистическому анализу состояния и прогнозирования конкретных социально-экономических явлений и процессов на основе построения адекватных, и, в достаточной степени аппроксимирующих реальные явления и процессы, прогностических моделей, на основе которых возможна выработка конкретных предложений, рекомендаций и путей их прикладного использования на макро- и микроуровнях.

Раздел 1. Моделирование динамики.

1.1 Временные ряды как объект статистического прогнозирования.

Основной задачей, решаемой при проведении любого статистического исследования, является определение объективных закономерностей развития социально-экономических явлений и процессов на основе анализа динамической информации.

Статистические модели, построенные на основе временных рядов социально-экономических показателей, позволяют применять математико-статистические методы для описания закономерностей развития объектов экономики как в прошлом, так и в будущем.

Используемые для целей и задач прогнозирования временные ряды экономических показателей обладают целым рядом особенностей.

Временной ряд есть последовательность, в которой каждое значение содержит в себе прошлое для последующих состояний. Любая попытка предвидеть будущее без исследования динамических рядов прошлого является малообоснованной, ненаучной и ошибочной. Поэтому для получения достаточно точных и надежных прогнозов, необходимо подробно изучить настоящее состояние явления или процесса.

Исследование скорости и интенсивности развития временных рядов часто не позволяет сразу определить основную тенденцию поступательного движения изучаемого явления. Это зависит от того, что уровни временного ряда со временем меняются, колеблются, но эта колеблемость не одинакова и может быть вызвана следующими причинами:

- влиянием общих факторов, определяющих главное направление, основную тенденцию развития явления;
- влиянием факторов общего характера, действующих периодически, сезонных колебаний и т.д.;
- влиянием специфических факторов, каждый из которых действует в разных направлениях, и их действие несущественно с точки зрения результатов развития явления, случайных колебаний.

Тип связи между компонентами временного ряда можно определить по нормальному распределению отклонений эмпирических значений уровней временного ряда от теоретических, полученных по уравнению тренда.

В случае нормальности распределения абсолютных отклонений связь является аддитивной, а относительных — мультипликативной.
Основные компоненты могут воздействовать на величину уровней временного ряда по-разному:

- если факторы, образующие эти компоненты, мультипликативные, то значения уровней временных рядов являются произведением этих компонентов:

$$\bar{y}_t = T \cdot C \cdot \varepsilon;$$

- если факторы аддитивные, то значения уровней временных рядов являются суммой компонентов:

$$\bar{y}_t = T + C + \varepsilon;$$

- если факторы временного ряда выражены комбинированно, то значения уровней являются или произведением, или суммой компонентов:

$$\bar{y}_t = T \cdot C + \varepsilon,$$

где T — тенденция;
C — сезонный компонент;
ε — случайный компонент.

Тип связи между компонентами можно также определить по динамике отклонений эмпирических значений уровней временного ряда от теоретических, полученных по уравнению тренда. Если абсолютные отклонения имеют тенденцию к росту, а относительные варьируют приблизительно на одинаковом уровне, то это свидетельствует о мультипликативной связи тренда и сезонного компонента.

На практике выделить компоненты сложно, так как отдельные последующие значения временных рядов зависят от предыдущих. Поэтому неверно допускать, что факторы, влияющие на колебания уровней, независимы. Кроме того, статистическая совокупность, изучаемая в течение длительного периода, перестает быть такой же самой совокупностью, так как могут измениться основные факторы, влияющие на ее формирование.

Экономическим явлениям свойственны элементы вероятностного характера. Наличие случайного в социально-экономических явлениях объясняется сложным переплетением параметров экономической системы, влиянием на них большого числа взаимосвязанных факторов, действующих в разных направлениях. Это ведет к вариации показателей уровней временного ряда.

Ввиду концепции о наличии вероятностных элементов в динамике процессов, уровни временного ряда могут рассматриваться как сумма детерминированного и случайного компонентов.

Детерминированный компонент выражается некоторой функцией и определяется уравнением основной тенденции или тренда.

Проявление случайного компонента оценивается с некоторой вероятностью.

Отклонения фактических уровней временного ряда от тренда рассматривается как стационарный случайный процесс.

Выявление основной тенденции развития — это один из методов анализа и обобщения временных рядов. Он позволяет выразить особенности изменения явления во времени.

Поэтому следует различать понятия:

- основная тенденция;
- тренд;
- закон развития явления.

Тренд *— некоторая аналитическая функция, которая связывает единым ''законом движения'' все последовательные уровни временного ряда.* Тренд описывает общую тенденцию на базе лишь одного фактора – фактора времени (t). Следовательно не полностью описывает характер тенденции развития и не может рассматриваться как закон развития явления.

Закон развития явления *— выражает сущность, природу явления, не поддающуюся описанием тренда.*

При изучении временных рядов возникают следующие проблемы:

- временной ряд — это числовые последовательности образования уровней во времени (только в одном направлении);
- временной ряд экономических показателей как правило содержит долговременную или краткосрочную тенденции развития, связанные с преодолением случайных колебаний;
- временные ряды могут быть подвержены регулярным колебаниям, связанным с сезонностью, ритмичностью и другими периодическими колебаниями;
- во временных рядах может наблюдаться связь следующих с предыдущими уровнями, то есть автокорреляция;
- при анализе развития взаимосвязанных временных рядов может возникнуть отставание одних рядов от других, выражаемое на основе временного шага;
- развитие социально-экономических явлений происходит непрерывно;
- действие большого числа факторов на развитие экономического явления во временных рядах выступает в виде обобщенного действия одного фактора времени;
- инерционность развития явления, то есть определяется степень сохранения темпов развития, направления развития, колеблемости уровня ряда. Инерционность не исключает наличие в динамике скачков;
- масштаб системы и иерархия характеристик. Чем выше масштаб системы ''предприятие — экономика в целом'', тем выше устойчивость и меньше колеблемость.

Использование особенностей временных рядов позволяет более точно строить по ним модель развития, отображающую процесс изменения явления во времени.

При разложении рядов динамики на отдельные компоненты следует принимать во внимание, что компоненты исходного временного ряда, по существу не наблюдаемы и являются только теоретическими величинами, абстракциями. Но несмотря на это, такой подход к разбиению фактических уровней временных рядов может оказаться довольно полезным для решения проблем анализа и прогнозирования на базе временных рядов.

Следует отметить, что уровни временного ряда не всегда являются составляющими всех трех компонентов одновременно. Единственным компонентом, который встречается во временных рядах является случайный компонент, который может быть представлен в сочетании с определенной тенденцией или с какими-то периодическими колебаниями. Чаще встречаются временные ряды, в которых можно установить тенденцию и случайный компонент, особенно при использовании годовых данных, где влияние сезонности не отражается.

Аналитически данное положение можно выразить уравнением вида:

$$\bar{y}_t = f(t) + \varepsilon(t), \qquad (1.1)$$

где $f(t)$ — некоторая функция времени, описывающая тенденцию исходного временного ряда, называется трендом;

$\varepsilon(t)$ — случайная величина (случайный компонент).

Функция $f(t)$ определяет общую тенденцию развития изучаемого явления. Поэтому прежде чем приступать к моделированию и прогнозированию социально-экономических явлений и процессов необходимо проверить гипотезу о наличии тенденции в исходном временном ряду.

1.2 Моделирование тенденции.

Анализ и моделирование тенденции временного ряда целесообразно начинать с выявления наличия тенденции в целом. Для этой цели наиболее эффективны и дают хорошие результаты такие методы как кумулятивный Т-критерий и критерий знаков разностей отклонений Валлиса и Мура.

Кумулятивный Т-критерий позволяет определить наличие не только самой тенденции, но и ее математического выражения — тренда. Выдвигается основная гипотеза (Но:) об отсутствии тенденции в исходном временном ряду.

Расчетное значение критерия определяется как отношение накопленной суммы квадратов отклонений эмпирических значений уровней временного ряда от их среднего значения ($\sum Z_n^2$) и самих отклонений по формуле:

$$T_p = \frac{\sum\limits_n Z_n^2}{\sigma_y^2}, \qquad (1.2)$$

где Z_n — накопленный итог отклонений эмпирических значений от среднего уровня исходного временного ряда;

σ_y^2 —общая сумма квадратов отклонений, определяемая по формуле: $\sigma_y^2 = \sum\limits_n y_t^2 - \left(\bar{y}\right)^2 \cdot n$.

Если анализируется достаточно длинный временной ряд, то для расчета значений критерия можно использовать нормированное отклонение:

$$t_p = \frac{T_p - \left(\frac{n+1}{6}\right)}{\sqrt{(n-2)\frac{2n-1}{90}}}. \qquad (1.3)$$

Расчетные значения кумулятивного Т-критерия и t_p сравниваются с критическими при заданном уровне значимости α. Если расчетное значение T_p или t_p превышает критическое (табличное) значение критерия ($T_{кр}$), то гипотеза об отсутствии тренда отвергается, следовательно в исходном временном ряду существует тенденция, описываемая трендом. В противном случае, если $T_p < T_{кр}$ или $t_p < t_{кр}$, признается отсутствие тенденции в ряду динамики.

Фазочастотный критерий знаков разностей Валлиса и Мура также позволяет определить наличие основной тенденции в целом во временном ряду и относится к группе непараметрических методов оценивания тенденции.

Выдвигается основная гипотеза (H_0 :) о том, что знаки разностей между (y_{i+1}-y_i) образуют случайную последовательность. При этом формируется понятие **фазы** как последовательности одинаковых знаков разностей.

Если знаки при определении числа фаз (без 1-ой и последней) образуют случайную последовательность, то расчетное значение критерия определяется по формуле:

$$t_p = \frac{\left|h - \frac{2n-7}{3}\right| - 0{,}5}{\sqrt{\frac{16n-29}{90}}}, \qquad (1.4)$$

где t_p — фазочастотный критерий разностей;

n — число уровней ряда, которое распределено нормально;

h — число фаз.

Если расчетное значение (t_p) критерия превысит критическое (табличное) $t_{кр}$ при заданном уровне значимости α, то гипотеза H_0 : о том, что знаки разностей между последовательными уровнями временного ряда образуют случайную последовательность, отвергается. Следовательно уровни исходного временного ряда не образуют случайную последовательность, а имеют определенную закономерность в их изменении, а следовательно, и в ряду существует тенденция.

После того, как в ряду динамики подтверждено наличие тенденции, важно определить ее вид.

Во временных рядах социально-экономических явлений может наблюдаться тенденция трех видов:

- тенденция среднего уровня;
- тенденция дисперсии;
- тенденция автокорреляции.

Тенденция среднего уровня может быть выражена с помощью графического метода. *Аналитически тенденция выражается с помощью некоторой математической функции f(t), вокруг которой варьируют эмпирические значения исходного временного ряда изучаемого социально-экономического явления.* При этом теоретические значения, то есть значения, полученные по трендовым моделям в отдельные моменты времени, являются математическими ожиданиями временного ряда.

Тенденция дисперсии *представляет собой тенденцию изменения отклонений эмпирических значений уровней временного ряда от теоретических, полученных по уравнению тренда.*

Тенденция автокорреляции *выражает тенденцию изменения корреляционной связи между отдельными, последовательными уровнями временного ряда.*

Проверка на наличие тенденции среднего уровня и дисперсии может быть произведена методом сравнения средних уровней временного ряда и методом Фостера-Стюарта.

Метод сравнения средних уровней временного ряда предполагает, что исходный временной ряд разбивается на две приблизительно равные части по числу членов ряда, каждая из которых рассматривается как самостоятельная, независимая выборочная совокупность, имеющая нормальное распределение. При этом решаются две задачи.

I.Если временной ряд имеет тенденцию, то средние, вычисленные для каждой совокупности в отдельности, должны существенно, значимо различаться между собой. Если же расхождение незначимо, несущественно и носит случайный характер, то временной ряд не имеет тенденции средней. Таким образом, проверка гипотезы (H_0 :) о наличии тенденции в исследуемом ряду сводится к проверке гипотезы о равенстве средних двух нормально распределенных совокупностей, то есть:

$$H_0: \quad \overline{y_1} = \overline{y_2}$$
$$H_1: \quad \overline{y_1} \neq \overline{y_2}.$$

Гипотеза проверяется на основе t-критерия Стьюдента, расчетное значение которого определяется по формуле:

$$t_p = \frac{\overline{y_1} - \overline{y_2}}{\sqrt{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}} * \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}, \qquad (1.5)$$

где $\overline{y_1}$ и $\overline{y_2}$ — средние уровни временного ряда согласно порядка разбиения;
n_1 и n_2 — число уровней временного ряда, соответственно первой и второй части;
σ_1^2 и σ_2^2 — дисперсия уровней ряда.

Расчетное значение (t_p) критерия сравнивается с его критическим (табличным) значением ($t_{кр}$) при уровне значимости α и числе степеней свободы $\nu = n - 2$.

Если $t_p > t_{кр}$, то гипотеза о равенстве средних уровней двух нормально распределенных совокупностей отвергается, следовательно расхождение между вычисленными средними значимо, существенно и носит неслучайный характер, и, следовательно, во временном ряду существует тенденция средней и существует тренд.

II. Если временной ряд имеет тенденцию, то дисперсии, вычисленные для каждой совокупности в отдельности, должны существенно и значимо различаться между собой. Если же расхождение между ними не значимо, то временной ряд не имеет тенденции дисперсии. Таким образом проверяется гипотеза (H_0 :) об отсутствии тенденции в дисперсиях в исходном временном ряду, которая сводится к проверке гипотезы о равенстве дисперсий двух нормально распределенных совокупностей, то есть :

$$H_0: \sigma_1^2 = \sigma_2^2;$$
$$H_1: \sigma_1^2 \neq \sigma_2^2.$$

Гипотеза проверяется на основе F-критерия Фишера-Снедекора, расчетное значение которого определяется по формуле:

$$F_p = \frac{\sigma_2^2}{\sigma_1^2}, \qquad \text{если } \sigma_2^2 \rangle \sigma_1^2$$

и

$$F_p = \frac{\sigma_1^2}{\sigma_2^2}, \qquad \text{если } \sigma_1^2 \rangle \sigma_2^2. \qquad (1.6)$$

Проверка гипотезы осуществляется на основе сравнения расчетного и критического значений F-критерия, полученного при заданном уровне значимости α и числе степеней свободы ν_1 и ν_2.

Если $\sigma_2^2 \rangle \sigma_1^2$, то $\nu_1 = n_2 - 1$;
$\nu_2 = n_1 - 1$.
Если $\sigma_1^2 \rangle \sigma_2^2$, то $\nu_1 = n_1 - 1$;
$\nu_2 = n_2 - 1$. (1.7)

Гипотеза о равенстве дисперсий двух нормально распределенных совокупностей отвергается, если $F_p > F_{кр}$. Следовательно, расхождение между вычисленными дисперсиями значимо, существенно, носит неслучайный характер и в ряду динамики существует тенденция в дисперсиях и существует тренд.

Следует заметить, что данный метод дает вполне приемлемые результаты лишь в случае рядов с монотонной тенденцией. Если же ряд динамики меняет общее направление развития, то точка поворота тенденции может оказаться близкой к середине ряда, в силу этого средние двух отрезков ряда будут близки и проверка может не показать наличие тенденции.

Метод Фостера-Стюарта основан на двух простых характеристиках S и d.

$$S = \sum S_t,$$
$$d = \sum d_t,$$

где

$$S_t = U_t + l_t,$$
$$d_t = U_t - l_t. \qquad (1.8)$$

Суммирование производится по всем членам ряда. Значения U_t и l_t определяются путем последовательного сравнения уровней.

Если значение уровня ряда превышает по своей величине каждый из предыдущих уровней, то величине Ut присваивается значение 1, в остальных случаях она равна 0. Таким образом,

$$U_t = \begin{cases} 1, \text{ если } y_t > y_{t-1}; y_{t-2}; y_{t-3} \dots y_1; \\ 0 \text{ в остальных случаях.} \end{cases}$$

Наоборот, если значение уровня ряда меньше всех предыдущих, то l_t присваивается значение 1. Таким образом,

$$l_t = \begin{cases} 1, \text{ если } y_t < y_{t-1};\ y_{t-2};\ y_{t-3} \ldots y_1; \\ 0 \text{ в остальных случаях.} \end{cases}$$

Показатели S и d асимптотически нормальные и имеют независимые распределения, но на них влияет порядок расположения уровней во времени. Показатель S применяется для обнаружения тенденции изменения в дисперсиях, d — для обнаружения тенденций в средней. После того, как для исследуемого ряда найдены фактические значения d и S, проверяется гипотеза о том, можно ли считать случайными разности d - 0 и S - μ. Гипотезы можно проверять, применяя t-критерий Стьюдента, то есть:

$$t_d = \frac{d - 0}{\sigma_2}, \qquad (1.9)$$

$$t_s = \frac{S - \mu}{\sigma_1}, \qquad (1.10)$$

где μ — математическое ожидание величины S, определенное для случайного расположения уровней во времени;

σ_1 — средняя квадратическая ошибка величины S;

σ_2 — средняя квадратическая ошибка величины d.

Значения μ, σ_1, σ_2 табулированы.

Если $t_d > t_{кр}$ (α; $\nu = n - 1$), то гипотеза об отсутствии тенденции в средней отвергается, следовательно в исходном временном ряду существует тренд.

Если $t_s > t_{кр}$ (α; $\nu = n - 1$), то гипотеза об отсутствии тенденции в дисперсиях отвергается, следовательно существует тенденция дисперсии и существует тренд. После того, как установлено наличие тенденции во временном ряду необходимо ее описать, то есть определить тип протекания процесса, имеющего место в данном явлении, направление роста и изменение, проходящие в нем.

Можно выделить следующие ***типы процессов:***

I. По возрастанию или убыванию уровней ряда:

- монотонно- возрастающие;
- монотонно- убывающие;
- комбинированные.

II. По наличию насыщения и стремлению к некоторой предельной величине:

- имеющие пределы насыщения;

- не имеющие пределов насыщения.

III. По наличию экстремальных значений и перегибов:

- процессы, имеющие экстремальные значения;
- процессы, имеющие переходы от возрастания к убыванию или наоборот.

Для выявления типа развития могут использоваться различные методы и критерии.

Критерий Кокса-Стюарта позволяет выявить характер тенденции исходного временного ряда. При этом проверяется гипотеза H_0: об отсутствии убывающей или возрастающей тенденции индивидуально при решении каждой конкретной задачи. Этапы реализации критерия Кокса-Стюарта следующие:

1. весь исходный временной ряд делится на три группы уровней;
2. численность 1-ой и 3-ей групп уровней равны и составляют [n:3] уровней каждая (кратны трем).
   Если ‘‘n’’ не делится на ‘‘3’’, то к ряду добавляют ‘‘1’’ или ‘‘2’’;
3. определяются отклонения каждого уровня 3-ей группы от соответствующего уровня 1-ой группы;
4. из суммы знаков ‘‘+’’ (при возрастающем тренде) и ‘‘—’’ (при убывающем тренде) вычитается ожидаемое число знаков [n:6].

Расчетное значение критерия Кокса-Стюарта определяется по формуле:

$$t_p = \frac{\left|S - \frac{n}{6}\right| - 0,5}{\sqrt{n:12}}. \qquad (1.11)$$

Если $t_p > t_{кр}$ при заданном уровне значимости α, то можно утверждать о наличии убывающей или возрастающей тенденции в исходном временном ряду.

Для определения типа развития явления могут быть использованы и следующие методы сглаживания:

1. метод скользящей средней;
2. метод аналитического выравнивания.

Выбор метода зависит от:

а) технических возможностей;
б) умения использовать методы;
в) цели и задач исследования.

Метод скользящих средних используется когда необходимо дать общую картину развития, основанную на механическом повторении одних и тех же действий по увеличению интервала времени.

Если исследование требует подробного аналитического выражения движения во времени, то используется аналитическое выравнивание.

Метод скользящих средних дает оценку среднего уровня за некоторый период времени. Так как средняя образуется за более длительный отрезок времени, она выступает не как средство обобщения единиц совокупности, а как средство их сглаживания. Чем больше интервал времени, к которому относится средняя, тем более плавным будет сглаженный уровень. Использование скользящей средней требует логического обоснования периода, взятого для выявления основной тенденции.

Метод аналитического выравнивания заключается в нахождении аналитической функции, выражающей развитие явления за рассматриваемый период времени. При этом решаются следующие задачи:

а) выбор вида уравнения, отображающего тип развития;

б) анализ схемы сбора фактических данных и определение параметров модели;

в) определение методов преобразования исходных данных с целью сведения сложных уравнений к более простым;

г) выявление степени близости теоретических и фактических данных.

1.3 Выбор формы тренда.

Остановимся подробнее на проблеме выбора математической функции описания основной тенденции развития, то есть выбора подобной реальной динамике формы уравнения.

Для отображения основной тенденции развития явлений во времени или модели этого процесса применяются разные уравнения, полиномы разной степени, экспоненты, логистические кривые и другие функции.

Полиномы имеют следующий вид:

полином первой степени $\overline{y_t} = a_0 + a_1 t$,

полином второй степени $\overline{y}_t = a_0 + a_1 t + a_2 t^2$, (1.12)

полином n-й степени $\overline{y}_t = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n$.

Для полиномиальных моделей характерно отсутствие прямой связи между абсолютными приростами и приростами уровней рядов динамики.

Предполагаемой функцией, отражающей процесс роста явления, может быть и экспонента $\overline{y}_t = a_0 \cdot a_1^t$ или $\overline{y}_t = a_0 \cdot a_1^{b_1 t + b_2 t^2}$. Экспоненты характеризуют прирост, зависящий от величины основания функции. Прологарифмировав левую и правую части, найдем $\lg \overline{y}_t = \lg a_0 + t \cdot \lg a_1$;

$\lg \overline{y}_t = \lg a_0 + \left(b_1 t + b_2 t^2\right) \cdot \lg a_1$, то есть логарифмические кривые. После замены lg a_0 = c_0 и lg a_1 = c_1 получим уравнение $\lg \overline{y}_t = c_0 + c_1 t$, из которого видно, что логарифм ординаты линейно зависит от t.

Вторая функция после логарифмирования дает уравнение логарифмической параболы $\lg \overline{y}_t = a_0 + b_2 t + b_2 t^2$, в котором темп прироста линейно зависит от времени.

Надо помнить, что практика моделирования свидетельствует о том, что выбор тех или иных кривых всегда оказывается под воздействием представлений о желаемой форме кривой, и что на координатном поле, отображающем расстояние точек, можно построить бесконечное множество кривых. При этом необходимо отражать особенности процесса. Свойства процесса должны соответствовать свойствам функций, используемых для построения моделей.

Надо иметь в виду, что отдельные уравнения выражают определенный тип динамики.

Монотонное возрастание или убывание процесса характеризуют функции:

- линейная;
- параболическая;
- степенная;
- экспоненциальная простая (показательная) и производная от нее логарифмическая линейная;
- сложная экспоненциальная и производная от нее логарифмическая парабола;
- гиперболическая (главным образом убывающих процессов);
- комбинация их видов.

Для моделирования динамических рядов, которые характеризуются стремлением к некоторой предельной величине, насыщением, применяются логистические функции.

Логистическую функцию часто записывают в следующем виде:

$$\overline{y}_t = \frac{k}{1 + e^{-a_0 t}} \text{ или } \overline{y}_t = \frac{k}{1 + 10^{a_0 + a_1 t}}, \qquad (1.13)$$

где **e**— основание натуральных логарифмов.

Логистическая кривая симметрична относительно точки перегиба и при $t = -\infty$ стремится к нулю, а при $t = +\infty$ стремится к некоторой постоянной величине, к которой кривая асимптотически приближается. Если найти вторую производную от y_t по t и приравнять ее к нулю, то для логистической кривой, выражаемой через е, место положения точки перегиба кривой равно: $t = \ln a_1 : a_0$; $\overline{y}_t = n{:}2$.

Тип процессов, характеризующихся наличием экстремальных значений, описывается кривой Гомперца, имеющей следующее выражение:

$$\overline{y}_t = k \cdot a_0^{a_1 t}. \qquad (1.14)$$

Возможны четыре варианта этой кривой. Для экономистов наибольшее значение имеет кривая, у которой $\lg a_0 < 0$ и $a_1 < 1$. Развитие уровня такой кривой имеет следующие этапы: если коэффициент a_1 меньше единицы при отрицательном значении $\lg a_0$, то на первом этапе прирост кривой незначителен. Он медленно увеличивается по мере роста t, но на следующем этапе прирост увеличивается быстрее, а затем, после точки перегиба, прирост начинает уменьшаться; подойдя к линии асимптоты, прирост кривой опять незначителен.

Прологарифмировав функцию Гомперца, получим

$$\lg \overline{y}_t = \lg k + a_1^t \cdot \lg a_0.$$

Следовательно, после логарифмирования получим модифицированную экспоненту. Введя в модифицированной экспоненте величину, обратную y_t, получим сходство с кривой Гомперца. Однако различие состоит в том, что изменение во времени первых разностей кривой Гомперца асимметрично, а у логистической кривой их изменение во времени имеет симметричный вид, напоминающий нормальное распределение.

При выборе формы тренда наряду с теоретическим анализом закономерностей развития изучаемого явления используются эмпирические методы, такие как:

- метод разностного исчисления;
- метод дисперсионного анализа;
- расчет и анализ средней квадратической ошибки;
- критерий наименьшей суммы квадратов отклонений эмпирических и теоретических значений уровней временного ряда.

Метод разностного исчисления предполагает:

- если первые разности равны $\Delta_i^{(1)} = y_i - y_{i-1}$ (цепные абсолютные приросты), а вторые разности $\Delta_i^{(2)} = \Delta_i^{(1)} - \Delta_{i-1}^{(1)}$ равны нулю, то наилучшим образом тенденция описывается уравнением линейного тренда;
- если вторые разности $\Delta_i^{(2)} = \Delta_i^{(1)} - \Delta_{i-1}^{(1)}$ (цепные абсолютные приросты из первых абсолютных приростов цепных) равны, а третьи разности $\Delta_i^{(3)}$ равны нулю, то парабола второго порядка наилучшим образом опишет тенденцию исходного временного ряда и так далее.

Дисперсионный метод анализа основывается на сравнении дисперсий.

Суть метода в следующем: общая дисперсия временного ряда делится на две части:

- вариация вследствие тенденции $V_{f(t)}$;
- случайная вариация V_ε.

Общая вариация определяется как сумма квадратов отклонений эмпирических значений уровней ряда (y_t) от среднего уровня исходного временного ряда ($\bar{y}$), то есть из выражения вида:

$$V_{общ} = \sum_{t=1}^{n} \left(y_t - \bar{y}\right)^2 . \qquad (1.15)$$

Случайная вариация — это сумма квадратов отклонений эмпирических значений уровней (y_t) от теоретических полученных по уравнению тренда ($\bar{y}_t$), и определяется по выражению следующего вида:

$$V_\varepsilon = \sum_{t=1}^{n} \left(y_t - \bar{y}_t\right)^2 . \qquad (1.16)$$

Вариация вследствие тенденции определяется как разность общей и случайной вариаций из выражения вида:

$$V_{f(t)} = V_{общ} - V_\varepsilon. \qquad (1.17)$$

На основе рассмотренных показателей вариации определяются различные виды дисперсии:

- общая дисперсия: $$\sigma^2_{общ} = \frac{V_{общ}}{n-1} = \frac{\sum\left(y_i - \bar{y}\right)^2}{n-1}; \qquad (1.18)$$
- дисперсия случайного компонента: $$\sigma^2_\varepsilon = \frac{V_\varepsilon}{n-k} = \frac{\sum\left(y_i - \bar{y}_t\right)^2}{n-k},$$

  где **k** — число параметров уравнения тренда. (1.19)

- дисперсия тенденции: $\sigma^2_{f(t)} = \frac{V_{f(t)}}{k-1} = \frac{V_{общ} - V_\varepsilon}{k-1}$. (1.20)

Выдвигается и проверяется гипотеза о том, что подходит или не подходит рассматриваемое уравнение тренда для описания тенденции исходного временного ряда.

Гипотеза проверяется на основе F-критерия Фишера-Снедекора, расчетное значение которого определяется по следующей формуле:

$$F_p = \frac{\sigma^2_{f(t)}}{\sigma^2_\varepsilon}. \qquad (1.21)$$

Если $F_p > F_{кр}$ при заданном уровне значимости α и числе степеней свободы ($\nu_1 = k - 1$, $\nu_2 = n - k$), то уравнение тренда подходит для отражения тенденции исходного временного ряда.

Анализ необходимо начинать с более простого уравнения к сложным, пока не подойдет.

Средняя квадратическая ошибка определяется по формуле:

$$\sigma_{ош} = \sqrt{\frac{\sum\left(y_i - \bar{y}_t\right)^2}{n-k-1}}, \qquad (1.22)$$

где **k** — число параметров уравнения.

Чем меньше значение ошибки, тем функция наилучшим образом описывает тенденцию исходного временного ряда.

Критерий наименьшей суммы квадратов отклонений эмпирических уровней от теоретических $\sum\left(y_i - \bar{y}_t\right)^2 \longrightarrow \min$ также предполагает, что наилучшим образом тенденция описывается трендом, которому соответствует наименьшее значение суммы квадратов отклонений.

Отдельно взятый критерий или метод при выборе формы тренда не обеспечивает правильность ее выбора. Необходим обязательно учет специфики объекта исследования, методов прогнозирования и оценки точности и надежности получаемых прогнозов.

После того, как определена форма трендовой модели (уравнения), необходимо проанализировать наличие, характер и закон распределения отклонений эмпирических значений от теоретических, полученных по уравнению тренда.

1.4 Моделирование случайного компонента.

Исследование случайного компонента проводится с целью решения двух основных задач:

1. оценки правильности выбора трендовой модели;
2. оценки стационарности случайного процесса.

При верном выборе формы тренда отклонения от него будут носить случайный характер, что означает, что изменение случайной величины ε_t не связано с изменением t.

Для этого определяются отклонения эмпирических значений от теоретических: $\varepsilon_t = y_t - f(t)$ для каждого уровня исходного временного ряда.

Проверяется гипотеза H_0: о том, что значения случайной величины ε_t случайны и величина ε_t не зависят от времени.

Методами проверки данной гипотезы являются следующие:

- коэффициент корреляции;
- критерий серий, основанный на медиане выборки;
- критерий “восходящих” и “нисходящих” серий;
- критерий min и max.

Наиболее простой сводится к расчету **коэффициента корреляции** между ε_t (отклонениями от тренда) и фактором времени t, и проверке его значимости.

Критерий серий, основанный на медиане выборки.

Этапы реализации метода:

- рассчитываются отклонения эмпирических значений от теоретических, полученных по уравнению тренда: $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$.
- ε_t ранжируются, где $\varepsilon^{(1)}$ — наименьшее значение: $\varepsilon^{(1)}, \varepsilon^{(2)}, ..., \varepsilon^{(n)}$ в порядке возрастания или убывания.
- Определяется медиана отклонений ε_{med}.
- Значения ε_t сравниваются со значением ε_{med} и ставится знак ‘‘+’’ или ‘‘-’’:

$\varepsilon_t > \varepsilon_{med}$ — ‘‘+’’

$\varepsilon_t < \varepsilon_{med}$ — ‘‘-’’

$\varepsilon_t = \varepsilon_{med}$ — пропускается уровень и ставится ‘‘0’’.

Таким образом получается ряд ‘‘+’’ и ‘‘-’’.

- Выдвигается и проверяется следующая основная гипотеза H_0 : Если отклонения от тренда случайны, то их чередование должно быть случайным.
- Последовательность ‘‘+’’ и ‘‘-’’ называется ***серией.***
- Определяется $k_{max}(n)$ — длина наибольшей серии.
- Определяется V(n) — число серий.

Выборка признается случайной, если одновременно выполняются неравенства (α = 0,05):

$$\begin{cases} k_{max}(n) < [3{,}3(\lg n + 1)]; \\ V(n) > \left[\frac{1}{2}\left(n + 1 - 1{,}96\sqrt{n-1}\right)\right]. \end{cases} \quad (1.23)$$

Если хотя бы одно неравенство нарушается, то гипотеза о случайности отклонений уровней временного ряда от тренда отвергается.

Критерий ‘‘восходящих’’ и ‘‘нисходящих’’ серий.

Этапы реализации метода:

- Последовательно сравниваются каждое следующее значение ε_{t+1} с предыдущим и ставится знак ‘‘+’’ или ‘‘-’’:

  $\varepsilon_{i+1} > \varepsilon_i$ — ‘‘+’’

  $\varepsilon_{i+1} < \varepsilon_i$ — ‘‘-’’

  $\varepsilon_{i+1} = \varepsilon_i$ — учитывается только одно наблюдение (другие опускаются).
- Определяется $k_{max}(n)$ — длина наибольшей серии.
- Определяется V(n) — общее число серий.
- Выдвигается и проверяется гипотеза H_0 : о случайности выборки и подтверждается, если выполняются следующие неравенства (α = 0,05) ;

$$\begin{cases} V(n) > \left[\frac{1}{3}(2n-1) - 1{,}96\sqrt{\frac{16n-29}{90}}\right]; \\ k_{max}(n) \le k_0(n), \end{cases} \quad (1.24)$$

где $k_0(n)$ — число подряд идущих ‘‘+’’ или ‘‘-’’ в самой длинной серии.

$k_0(n)$ определяется следующим образом:

N	$k_0(n)$
$n \le 26$	5
$26 < n \le 153$	6
$153 < n \le 1170$	7

Если хотя бы одно из неравенств не выполняется, то гипотеза о случайном характере отклонений уровней временного ряда от тренда отвергается.

Критерий min и max : ‘‘пиков’’ и ‘‘ям’’.

Сущность метода заключается в том, что:

- уровень ряда является максимальным, если:

$$\begin{cases} 1; & y_{i-1} < y_i > y_{i+1} \\ 0 & \text{в других случаях;} \end{cases} \qquad (1.25)$$

- уровень временного ряда является минимальным, если:

$$\begin{cases} 1; & y_{i-1} > y_i < y_{i+1} \\ 0 & \text{в других случаях;} \end{cases} \qquad (1.26)$$

y_i является ***точкой поворота*** в ряду.

В случайном ряду математическое ожидание числа точек поворота (p) и дисперсии определяется по формулам:

$$\bar{p} = \frac{2}{3}(N-2); \qquad p = \sum_{i=1}^{n-2} X_i$$

$$\sigma_p^2 = \frac{16N-29}{90}. \qquad (1.27)$$

Критерием случайности с вероятностью 0,95 является неравенство:

$$p > \left[\frac{2}{3}(N-2) - 1{,}96\sqrt{\frac{16N-29}{90}}\right]. \qquad (1.28)$$

Основным условием анализа исходных значений уровней временного ряда и отклонений эмпирических значений от выравненных по тренду является нормальность закона распределения этих значений. Часто при прогнозировании на основе авторегрессионных моделей необходимо проверять гипотезу о том, что отклонения от тренда подчиняются закону нормального распределения.

При нормальном распределении коэффициенты асимметрии A_s и эксцесса E_x должны быть равны нулю. Так как отклонения от тренда — это выборка из некоторой генеральной совокупности, то можно определить только выборочные характеристики A_s и E_x:

$$A_s = \frac{\frac{1}{n}\sum\left(y_t - \bar{y}_t\right)^3}{\left(\sqrt{\frac{\sum\left(y_t - \bar{y}_t\right)^2}{n}}\right)^3} = \frac{\frac{1}{n}\sum\varepsilon_t^3}{\sqrt{\left(\frac{1}{n}\sum\varepsilon_t^2\right)^3}}$$

$$E_x = \frac{\frac{1}{n}\sum\left(y_t - \bar{y}_t\right)^4}{\left(\sqrt{\frac{\sum\left(y_t - \bar{y}_t\right)^2}{n}}\right)^4} - 3 = \frac{\frac{1}{n}\sum\varepsilon_t^4}{\left(\frac{1}{n}\sum\varepsilon_t^2\right)^2} - 3 \qquad (1.29)$$

Если $\Sigma\, y_t = 0$, то формулы преобразуются :

$$A_s = \frac{\frac{1}{n}\sum y_t^3}{\left(\sqrt{\frac{1}{n}\sum y_t^2}\right)^3}$$

$$E_x = \frac{\frac{1}{n}\sum y_t^4}{\frac{1}{n^2}\left(\sqrt{\sum y_t^2}\right)^4} - 3 \qquad (1.30)$$

Уровни ряда являются нормально распределенными, если одновременно выполняются следующие условия:

$$\begin{cases} \left|A_s\right| < 1{,}5 \cdot \sigma_{A_s} \\ \left|E_x - \frac{6}{n+1}\right| < 1{,}5 \cdot \sigma_{E_x}, \end{cases} \qquad (1.31)$$

где σ_{A_s} и σ_{E_x} — среднеквадратические ошибки коэффициентов A_s и E_x и определяются по формулам:

$$\sigma_{A_s} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}}$$

$$\sigma_{E_x} = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}} \qquad (1.32)$$

Если выполняется хотя бы одно неравенство:

$$\begin{cases} \left|A_s\right| \geq 2 \cdot \sigma_{A_s} \\ \left|E_x + \frac{6}{n+1}\right| \geq 2 \cdot \sigma_{E_x}, \end{cases} \qquad (1.33)$$

то данные не являются даже приблизительно нормальными, и их применение в анализе не рекомендуется.

1.5 Модели периодических колебаний.

При рассмотрении квартальных и месячных данных часто обнаруживаются периодические колебания, вызываемые сменой времен года. Их называют сезонными.

Изучение сезонных колебаний имеет самостоятельное значение как исследование особого типа динамики.

***Сезонность** можно понимать как внутригодовую динамику вообще.*

Во многих случаях сезонность приносит ущерб народному хозяйству в связи с неравномерным использованием оборудования и рабочей силы, с неравномерной нагрузкой транспорта, поставкой сырья для других отраслей, связанных с сезонными отраслями.

Можно построить модель сезонной волны и численно определить размах сезонных колебаний, характер их проявления в различных отраслях народного хозяйства.

Моделью периодически изменяющихся уровней служит **ряд Фурье:**

$$\bar{y}_t = a_0 + \sum_{k=0}^{m} \left(a_k \cos kt + b_k \sin kt \right),$$

где k — определяет номер гармоники ряда Фурье и может быть взята с разной степенью точности (чаще от ‘‘1’’ до ‘‘4’’).

Параметры уравнения определяются методом наименьших квадратов, то есть по условию $\sum \left(y - \bar{y}_t \right)^2 \Rightarrow \min$. Решая систему нормальных уравнений, получим:

$$\begin{aligned} a_0 &= \frac{1}{n} \sum y; \\ a_k &= \frac{2}{n} \sum y \cos kt;. \\ b_k &= \frac{2}{n} \sum y \sin kt \end{aligned} \qquad (1.34)$$

Для изучения сезонности берется (n = 12) по числу месяцев в году.

Как правило, при выравнивании по ряду Фурье рассчитывают не более четырех гармоник и затем уже определяют, какая гармоника наилучшим образом отражает периодичность изменения уровней ряда.

Так, при k=1: $\bar{y}_t = a_0 + a_1 \cos t + b_1 \sin t$;

k=2: $\bar{y}_t = a_0 + a_1 \cos t + b_1 \sin t + a_2 \cos 2t + b_2 \sin 2t$. (1.35)

Рассчитав остаточные дисперсии $\sigma^2_{ост} = \dfrac{\sum\left(y_i - \bar{y}_t\right)^2}{n}$ для 2-х случаев, можно сделать вывод, какая гармоника Фурье наиболее близка к фактическим уровням ряда.

1.6 Модели связных временных рядов.

Метод наименьших квадратов, используемый в регрессионном анализе для определения коэффициентов регрессии, основывается на предпосылке независимости друг от друга отдельных наблюдений одной и той же переменной. В динамических же рядах существует еще и автокорреляция. Поэтому величина коэффициентов регрессии, полученных по способу наименьших квадратов, не имеет нужных статистических свойств. Наличие автокорреляции приводит к искажению средних квадратических ошибок коэффициентов регрессии, что в свою очередь затрудняет построение доверительных интервалов по ним и проверку их значимости по соответствующим критериям. Автокорреляция также может привести к сокращению числа наблюдений ввиду невозможности потерять показатели одного и того же объекта за ряд лет, поскольку наблюдение одного объекта за десять лет качественно отличается от наблюдений десяти объектов за одно и то же время. Возникает автокорреляция и в отклонениях от трендов, а также в случайных остатках уравнений регрессии, построенных по многомерным рядам динамики.

Автокорреляция *— это наличие сильной корреляционной зависимости между последовательными уровнями временного ряда.*

Автокорреляция может быть следствием следующих причин:

- Не учтен в модели существенный фактор, при этом его влияние отражается на величине отклонений, которые в этом случае показывают закономерность в изменении, связанную с изменением неучтенного фактора.
- В модели не учитывается несколько факторов, влияние каждого из которых в отдельности не существенно, но при совпадении изменений этих факторов по направлению и по фазе в отклонениях может возникнуть автокорреляция.
- Автокорреляция в отклонениях может появиться в случае, когда неправильно выбрана форма связи между **y** и **x**.
- Неверно выбран порядок авторегрессионой модели.
- Вследствие специфичности внутренней структуры случайного компонента.

Прежде чем делать вывод о тесноте связи между рассматриваемыми рядами динамики, необходимо проверять наличие автокорреляции в них, чтобы оценить степень зависимости между соседними уровнями временного ряда. Наличие автокорреляции устанавливается с помощью **коэффициента автокорреляции**, который определяется на основе формулы коэффициента корреляции для парной (линейной) связи между уровнями исходного ряда и того же ряда, но сдвинутого на τ шагов во времени:

$$r_a = \frac{\overline{y_t \cdot y_{t+1}} - \overline{y}_t \cdot \overline{y}_{t+1}}{\sigma_{y_t} \cdot \sigma_{y_{t+1}}}, \qquad (1.36)$$

где **y_t** — эмпирические значения уровней ряда;

y_{t+1} — эмпирические значения уровней, сдвинутые на один период времени (τ = 1).

Возникает проблема заполнения последнего уровня ряда **y_{t+1}**. В данном случае возможны два варианта:

1. Если значение последнего уровня мало отличается от первого, то чтобы ряд не укорачивался, его можно условно дополнить $y_{t+1} = y_{t_1}$. Тогда

$$\begin{aligned} \overline{y}_t &= \overline{y}_{t+1} \\ \sigma_{y_t} &= \sigma_{y_{t+1}} \end{aligned} \qquad (1.37)$$

И коэффициент автокорреляции будет равен:

$$r_a = \frac{\overline{y_t \cdot y_{t+1}} - \left(\overline{y}_t\right)^2}{\sigma^2_{y_t}} \text{ или} \qquad (1.38)$$

$$r_a = \frac{\sum y_t \cdot y_{t+1} - n\left(\overline{y}_t\right)^2}{\sum y_t^2 - n\left(\overline{y}_t\right)}, \qquad (1.39)$$

где

$$\begin{aligned} \overline{x_t \cdot x_{t+1}} &= \frac{\sum x_t \cdot x_{t+1}}{n} \\ \overline{x}_t &= \frac{\sum x_t}{n} \\ \sigma^2_{x_t} &= \frac{\sum x_t^2}{n} - \left(\overline{x}_t\right)^2 \end{aligned} \quad . \qquad (1.40)$$

Затем аналогично рассчитывается коэффициент автокорреляции для всех временных рядов, входящих в связный.

Если $r_a > r_{a\ кр}$ при заданном уровне значимости α и n, то в исходном временном ряду существует автокорреляция, в противном случае она отсутствует.

Последовательность значений коэффициентов автокорреляции r_τ, вычисленных при τ = 1, 2, ..., l, называют ***автокорреляционной функцией***. Эта функция дает представление о внутренней структуре изучаемого экономического явления.

1.7 Для проверки автокорреляции в уровнях ряда также используется

и **критерий Дарбина-Уотсона**. Гипотеза о наличии автокорреляции проверяется с помощью случайной величины:

$$d = \frac{\sum_{t=1}^{n-1}\left(y_{t+1} - y_t\right)^2}{\sum_{t=1}^{n} y_t^2} \qquad (1.41)$$

$$0 \le d \le 4$$

Если автокорреляции в ряду нет, то значения критерия d колеблются вокруг 2.

Эмпирическое значение d сравнивается с табличным значением.

В таблице есть два значения критерия — d_1 и d_2, v и n,

где d_1 и d_2 — нижняя и верхняя границы теоретических значений;

v — число факторов в модели.

n — число членов временного ряда.

Если 1) $d < d_1$ — в ряду есть автокорреляция;

2) $d > d_2$ — автокорреляции нет;

3) $d_1 \le d \le d_2$ — необходимо дальше исследовать автокорреляцию.

Иногда приходится при анализе рядов динамики исследовать вопрос о наличии или отсутствии автокорреляции не между самими уровнями ряда, а между их отклонениями от среднего уровня или от выравненного уровня.

При значении $r_a \le 0{,}3$ необходимо проверять *наличие автокорреляции в остатках* с помощью следующего **коэффициента Дарбина-Уотсона для остаточных величин:**

$$d = \frac{\sum\left(\xi_t \cdot \xi_{t+1}\right)^2}{\sum \xi_t^2}, \qquad (1.42)$$

где ξ_t — отклонения эмпирических значений уровней от теоретических, полученных по уравнению тренда.

Существует теоретическое распределение значений d_p для положительной автокорреляции с вероятностью 0,95,

где d_1 и d_2 — нижняя и верхняя границы теоретических значений.

ν — число факторов в модели;

n — число членов временного ряда.

При применении критерия Дарбина-Уотсона расчетное значение d_p сравнивается с табличными d_1 и d_2. При этом возникает три исхода:

1) $d < d_1$ => вывод о наличии автокорреляции в отклонениях;
2) $d > d_2$ => вывод об отсутствии автокорреляции;
3) $d_1 \le d \le d_2$ => необходимо дальше исследовать автокорреляцию.

Возможные значения критерия находятся $0 \le d \le 4$. Они различны для положительной и отрицательной автокорреляции. Так как при отрицательной автокорреляции $d \in [2; 4]$, для проверки следует определять величину $(4 - d)$.

Если в рядах динамики или в остаточных величинах имеется автокорреляция, то оценки коэффициентов регрессии, полученные методом наименьших квадратов, будут несмещенными, но неэффективными, так как наличие автокорреляции увеличивает дисперсии коэффициентов регрессии. Это затрудняет построение доверительных интервалов для коэффициентов регрессии и проверку их значимости.

Из этого следует сделать вывод, что прежде чем проводить корреляционно-регрессионый анализ временных рядов, необходимо исключить из исследуемых рядов автокорреляцию.

После того как установлено наличие автокорреляции следует приступить к построению модели.

Основными моделями связных рядов динамики являются модели авторегрессии. Для того, чтобы получить эти модели необходимо исключить автокорреляцию.

В настоящее время разработано **четыре способа исключения автокорреляции:**

1. Основан на использовании, так называемых, **последовательных** или **конечных разностей**.

Модель данным методом имеет вид:

$$\Delta y_{t+1} = a_0 + a_1\Delta x_{1,\,t+1} + a_2\Delta x_{2,\,t+1} + ... + a_k\Delta x_{k,\,t+1}.$$

Сущность метода заключается в последовательном исключении величины предшествующих уровней из последующих:

$\Delta_x = x_i - x_{i-1}$

$\Delta_y = y_i - y_{i-1}$

$\Delta_{y_1} = y_t - y_{t-1}$

. . .

$\Delta_{x_1} = x_t - x_{t-1}$

$\Delta_{x_2} = x_{t-1} - x_{t-2}$

При коррелировании разностей измеряется теснота связи между разностями последовательных величин уровней в каждом динамическом ряду.

Показателем тесноты связей между изучаемыми рядами является **коэффициент корреляции разностей**:

$$r_{\Delta_x \Delta_y} = \frac{\sum \Delta_x \Delta_y}{\sqrt{\sum \Delta_x^2 \cdot \Delta_y^2}}. \qquad (1.43)$$

2. **По отклонениям эмпирических значений от выравненных по тренду.**

Определяется тенденция исходных рядов динамики. Рассчитывается тренд, и его величина исключается из каждого уровня.

Модель в общем виде может быть представлена следующим образом:

$$y - \bar{y}_t = a_0 + a_1\left(x_1 - \bar{x}_{1t}\right) + a_2\left(x_2 - \bar{x}_{2t}\right) + \cdots + a_k\left(x_k - \bar{x}_k\right).$$

При коррелировании отклонений фактических уровней от выравненных необходимо:

1) произвести аналитическое выравнивание сравниваемых рядов по любому рациональному многочлену;
2) определить величину отклонения каждого фактического уровня ряда динамики от соответствующего ему выравненного значения;
3) произвести коррелирование полученных отклонений.

Коэффициент корреляции отклонений определяется по формуле:

$$r_{d_x d_y} = \frac{\sum d_x \cdot d_y}{\sqrt{\sum d_x^2 \cdot d_y^2}}, \qquad (1.44)$$

где $d_y = y_i - \bar{y}_t$

$d_x = x_i - \bar{x}_t$.

Коэффициент корреляции отклонений характеризует степень связи между отклонениями фактических уровней сравниваемых рядов от соответствующих им выравненных уровней коррелируемых рядов динамики.

3. **Метод Фриша-Воу.**

Этот метод заключается в ведении времени как дополнительного факторного признака. Это возможно только в случае, если основные тенденции временных рядов одинаковы. В этом случае парные связи обращаются в связи многофакторные и расчеты коэффициента корреляции и уравнения регрессии проводятся методом многофакторной корреляции.

Коэффициент корреляции рассчитывается как **множественный**:

$$R = \sqrt{1 - \frac{\sigma^2_{ост}}{\sigma^2_y}}, \qquad (1.45)$$

где $\sigma_{ост} = \sqrt{\frac{\left(y_i - \overline{y}_{1,2,t}\right)^2}{n}}$.

$\sigma^2_{ост}$ — остаточная дисперсия;

σ^2_y — общая дисперсия.

При построении многофакторных моделей по динамическим рядам возникает проблема мультиколлинеарности. Под ***мультиколлинеарностью*** в этом случае *понимают наличие сильной корреляционной зависимости между факторами рассматриваемых во взаимосвязи рядов динамики.* Мультиколлинеарность возникает вне зависимости от связи между результативным и факторным признаками. Она часто представляет опасность для правильного определения степени тесноты связи и оценки ее значимости.

Мультиколлинеарность затрудняет проведение анализа, так как усложняется процесс выделения наиболее существенных факторов и искажается смысл коэффициента регрессии.

Мультиколлинеарность возникает в том случае, когда факторными признаками выступают синтетические показатели. Например, в качестве факторов рентабельности могут рассматриваться объем реализации, производительность труда, фондоотдача, которые сильно коррелированы между собой.

На практике считают два фактора сильно коррелированными, если парный коэффициент корреляции между ними по абсолютной величине больше 0,8.

Довольно приблизительным методом обнаружения мультиколлинеарности является следующее правило. Фактор можно отнести к числу мультиколлинеарных, если коэффициент корреляции, характеризующий зависимость результативного признака от этого фактора больше, чем коэффициент множественной корреляции между результативным признаком и множеством остальных факторов.

Меры по устранению мультиколлинеарности в основном сводятся к следующему: построение уравнений регрессии по отклонениям от тренда или по конечным разностям; преобразование множества факторов в несколько ортогональных множеств с использованием методов многомерного анализа (факторного анализа или метода главных компонент); исключение из рассмотрения одного или нескольких линейно связных факторов. Это исключение следует вести с крайней осторожностью, основываясь на тщательном экономическом анализе.

Очистив таким образом уровни ряда динамики от автокорреляции и мультиколлинеарности, остается еще ‘‘подравнять’’ эти уровни по времени. Для этого необходимо рассмотреть вопрос о временном лаге.

Временным лагом *называется запаздывание (или опережение) процесса развития, представленного одним временным рядом, по сравнению с развитием, предоставленным другим рядом.* Временной лаг определяется при помощи перебора парных коэффициентов корреляции между абсолютными уровнями двух рядов динамики. Возможно наличие временного лага и в данных, которые изображают динамику годовых показателей.

Следовательно, приведение данных к сопоставимому виду с точки зрения автокорреляции, коллинеарности и временного лага является предварительным условием построения многофакторной модели динамики.

Построенная с соблюдением этих условий многофакторная регрессионная модель $\overline{y'}_t = f'(x'_1, x'_2, \ldots, x'_m)$, (где знак $'$ показывает номер этапа) будет характеризовать среднее влияние факторных признаков на результативный признак за рассматриваемый интервал времени. Величина этого влияния, выраженная коэффициентами регрессии, частными коэффициентами эластичности и β — коэффициентами будет изменяться от года к году.

При продолжительном времени (свыше 10 лет) это будет означать недоучет влияния НТР, изменение энерговооруженности труда, замещение одного сырья другим и т. д. Эти недостатки отражения связи могут быть устранены несколькими способами. Один из них состоит в разбиении всего периода времени T на пять интервалов. При этом выдвигается гипотеза, что за равные интервалы времени коэффициенты регрессии изменяются несущественно. Исходя из этого, можно построить пять уравнений, аналогичных вышеприведенному. Следовательно, каждое значение коэффициента регрессии a_i будет иметь пять оценок. Итак, получается временной ряд для каждого коэффициента регрессии. По этим рядам динамики можно построить временные модели (тренды) для каждого коэффициента по одному динамическому ряду. Так получается модель по уравнениям регрессии.

Но при построении такой модели возникает ряд проблем. Прежде всего, при расчленении экономических динамических рядов и определяющих их факторов на интервалы, число интервалов должно быть достаточно велико, чтобы ряды динамики, составленные из этих интервалов, правильно отражали тенденцию изменения влияния факторных признаков на результативные. Число лет, входящих в один интервал, должно быть в 3-4 раза больше числа переменных, входящих в регрессионное уравнение. Однако мы часто располагаем более короткими рядами динамики, следовательно, практически применять такие модели крайне затруднительно, а иногда и невозможно.

Поэтому рассмотрим другие методы построения многофакторных моделей. Предположим, что зависимость результативного признака экономического явления от ряда факторных может быть записана уравнением:

$$\overline{y}_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + \ldots + a_k x_k$$

(t = 1, 2,..., k) и коэффициенты регрессии изменяются во времени по линейной функции так, что их можно записать уравнениями:

$$\begin{aligned} a_0 &= b_{00} + b_{01} t \\ a_1 &= b_{10} + b_{11} t \\ &\ldots\ldots\ldots\ldots \\ a_k &= b_k + b_k t \end{aligned}.$$

В этом случае уравнение регрессии имеет другой вид:

$$\overline{y}_t = (b_{00} + b_{01} t) + (b_{10} + b_{11} t) x_{11} + (b_{20} + b_{21} t) x_{21} + \ldots + (b_k + b_k t) x_{k1}.$$

Раскрывая скобки и производя замену переменных произведения **tx** через **z**, так что $tx_{11} = z_{11}$; $tx_{21} = z_{21}$;....., $tx_{k1} = z_{k1}$, а $b_{k0} + b_{k01t} = c_0$, получим новое уравнение:

$$\overline{y}_t = c_0 + \mathrm{b}_{10} x_{1t} + \mathrm{b}_{11} z_{1t} + \mathrm{b}_{20} x_{2t} + \mathrm{b}_{21} z_{2t} + \ldots + \mathrm{b}_{k0} x_{kt} + \mathrm{b}_{k1} z_{kt}.$$

Параметры этого уравнения находятся по способу наименьших квадратов и показывают, как меняется во времени действие отдельных факторов на результативный признак рассматриваемого социально-экономического явления.

Применение приведенного уравнения с большим числом параметров и факторов требует использования рядов в 6-7 раза длиннее числа параметров. Однако в данном случае рассматривались линейные тренды параметров уравнения регрессии, а при криволинейных трендах число параметров самого уровня значительно увеличивается и, следовательно, ряд динамики должен быть еще длиннее. Таким образом, пользоваться только что рассмотренным методом на практике бывает затруднительно. Особенно трудно вести оценку значимости параметров. Обычно имеющиеся в распоряжении исследования временные ряды за 20-25 лет недостаточны. Они должны быть значительно длиннее, чтобы были получены достаточно достоверные выводы.

Раздел 2. Прогнозирование динамики.

2.1 Сущность и классификация статистических прогнозов.

Статистическое прогнозирование, наряду с другими видами прогнозирования социально-экономических явлений и процессов, является инструментом социально-экономического управления и развития.

Прогнозирование – это вид познавательной деятельности человека, направленной на формирование прогнозов развития объектов, на основе анализа тенденций и закономерностей его развития.

Прогнозирование *– это научное, основанное на системе установленных причинно-следственных связей и закономерностей, выявление состояния и вероятностных путей развития явлений и процессов.*

Прогнозирование предопределяет оценку показателей и дает характеристику явлений и процессов в будущем. Прогнозирование распространяется на такие процессы управления, которые в момент выработки прогнозов можно определить в весьма малом диапазоне, либо совсем невозможно, либо возможно, но требует учета действия таких факторов, влияние которых не может быть полностью или однозначно определено.

В зависимости от степени конкретности и характера воздействия на ход исследуемых процессов и явлений можно выделить три основные понятия прогнозирования:

- гипотеза;
- предсказание;
- прогноз.

Данные понятия тесно взаимосвязаны в своих проявлениях друг с другом и с исследуемым объектом и представляют собой последовательные ступени познания поведения явления и объекта в будущем.

Гипотеза *– это научно обоснованное предположение либо о непосредственно ненаблюдаемом факте, либо о закономерном порядке, объясняющем известную совокупность явлений.* На уровне гипотезы дается качественная характеристика объекта, выражающая общие закономерности его поведения. Гипотезой является не всякая догадка, а лишь предположение, которое носит вероятный характер. Установив, что группа явлений, закономерная связь которых неизвестна, имеет ряд тождественных черт с другой группой явлений, закономерная связь которых уже установлена, делается вывод о вероятности частичного сходства искомой закономерной связи с уже определенной.

Развиваясь, гипотеза одновременно подвергается проверке, необходимость которой вытекает из самой сущности гипотезы как предположения. Проверка гипотезы состоит в том, что все следствия, полученные посредством теоретического анализа основного допущения гипотезы сопоставляются с эмпирическими данными. Если по одной и той же задаче, проблеме и так далее возникает одновременно несколько гипотез и известно, какие гипотезы здесь вообще возможны, а какие – нет, то доказательством истинности одной из рассматриваемых гипотез является установление ложности всех остальных.

Степень вероятности гипотезы тем выше, чем разнообразнее и многочисленнее ее следствия, подтвержденные эмпирическим путем.

Достаточность условий реализации гипотез, их вероятность теоретически и практически граничит с высокой степенью достоверности. Гипотеза оказывает воздействие на процесс через прогноз, являясь важным источником информации для его составления.

Предсказание *– это предвидение таких событий, количественная характеристика которых невозможна или затруднена*.

Прогноз *– это количественное, вероятностное утверждение в будущем о состоянии объекта или явления с относительно высокой степенью достоверности, на основе анализа тенденций и закономерностей прошлого и настоящего*.

Прогноз в сравнении с гипотезой имеет большую определенность и достоверность, так как основывается как на качественных так и на количественных характеристиках. В отдельных случаях прогноз может носить качественный характер, но в его основе всегда лежат количественные явления.

Для осуществления прогноза, то есть определения понятий, как будет осуществляться и развиваться прогнозируемые явления в будущем, необходимо знать тенденции и закономерности прошлого и настоящего. При этом, следует помнить, что будущее зависит от многих случайных факторов, сложное переплетение и сочетание которых учесть практически невозможно. Следовательно, все прогнозы носят вероятностный характер.

Прогнозы можно подразделить в зависимости от целей, задач, объектов, времени упреждения, источников информации и так далее.

1. *В зависимости от целей исследования прогнозы делятся на поисковые и нормативные*.

Нормативный прогноз *– это прогноз, который предназначен для указания возможных путей и сроков достижения заданного, желаемого конечного состояния прогнозируемого объекта,* то есть нормативный прогноз разрабатывается на базе заранее определенных целей и задач.

Поисковый прогноз не ориентируется на заданную цель, а *рассматривает возможные направления будущего развития прогнозируемого объекта, то есть выявление того, как будет* *развиваться объект в будущем полностью зависит от сохранения существующих тенденций.*

Таким образом, поисковый прогноз отталкивается при определении будущего состояния объекта от его прошлого и настоящего, а нормативный прогноз осуществляется в обратной последовательности: от заданного состояния в будущем к существенным тенденциям и закономерностям в соответствии с поставленной задачей.

2. *В зависимости от специфики области применения прогноза и от объекта прогнозирования* прогнозы подразделяются на:

- естественноведческие – это прогнозы в области биологии, медицины и так далее;
- научно-технические – это, например, инженерное прогнозирование технических характеристик узлов, деталей и так далее;

3. *В зависимости от масштабности объекта*, прогнозы бывают:

- глобальные – рассматривают наиболее общие тенденции и закономерности в мировом масштабе;
- макроэкономические – анализируют наиболее общие тенденции явлений и процессов в масштабе экономики страны в целом;
- структурные (межотраслевые и межрегиональные) – предсказывают развитие экономики в разрезе отраслей;
- региональные – предсказывают развитие отдельных регионов;
- отраслевые – прогнозируют развитие отраслей;
- микроэкономические – предсказывают развитие отдельных предприятий, производств и так далее.

4. *По сложности* прогнозы различают:

- сверхпростые – прогноз на основе одномерных временных рядов, когда отсутствуют связи между признаками;
- простые – прогнозы, предполагающие учет оценки связей между факторными признаками;
- сложные – прогнозы, оценка связей между признаками в которых определяется на основе системы уравнений или многофакторного динамического прогнозирования.

5. *По времени упреждения* выделяются следующие прогнозы социально-экономических явлений и процессов:

- текущие – до 1 года;
- краткосрочные – 1 – 3 года;
- среднесрочные – 3 – 5 лет;
- долгосрочные – 5 – 10 лет;
- дальнесрочные – 10 и более лет.

Период упреждения прогноза *– это отрезок времени от момента, для которого имеются последние статистические данные об изучаемом объекте, до момента, к которому относится прогноз.*

Период упреждения прогноза зависит от специфики и особенностей изучаемого объекта исследования, от интенсивности изменения показателей, от продолжительности действия выявленных тенденций и закономерностей, от длины временного ряда и от многих других факторов.

Перечисленные виды прогнозов по времени упреждения отличаются друг от друга по своему содержанию и характеру оценок исследуемых процессов. Текущий прогноз основан на предположении о том, что в прогнозируемом периоде не произойдет существенных изменений в исследуемом объекте, а если и произойдут, то количественно несущественные. Краткосрочный и среднесрочные прогнозы предполагают, что произойдут существенные изменения с изучаемым объектом как в количественных, так и в качественных характеристиках. При этом в краткосрочном и среднесрочном прогнозах оценка явлений и процессов дается в разрезе количественно-качественном, а в долгосрочном и дальнесрочном прогнозах – качественно-количественном.

Выбор методов прогнозирования осуществляется в соответствии с характером объекта, требований, предъявляемых к информационному обеспечению, а также на основе сравнения эффективности и оптимальности решения аналогичных задач.

Отличительной чертой социально-экономических явлений и процессов является инерционность, проявляющаяся, с одной стороны в сохранении взаимосвязей прогнозируемого явления с другими явлениями, а с другой – в сохранении тенденции во времени. Для обеспечения научной обоснованности и достоверности социально-экономических прогнозов необходимо, чтобы в ходе их составления раскрывались и познавались причинно-следственные связи и факторы, характеризующие развитие процессов и явлений, изучались их внутренние структурные связи, а также внешняя среда, в которой они проявляются.

Основными этапами разработки статистических прогнозов являются:

1. Анализ объекта прогнозирования. На этом этапе рассматривается состояние, основные элементы, взаимосвязи и факторы, формирующие и оказывающие влияние на исследуемых объект; выдвигается основная рабочая гипотеза; выявляются причинно-следственные связи как внутри явления, так и вне его и определяется их статистическое выражение.

2. Характеристика информационный базы исследования. На данном этапе выдвигаются основные требования, предъявляемые к информационной базе. При этом различают количественную информацию, обработку которой осуществляют статистическими методами, и качественную информацию, сбор и обработка которой

производится преимущественно эвристическими и непараметрическими статистическими методами анализа.

3. Выбор метода прогнозирования. Процесс выбора метода прогнозирования обусловлен объективизацией прогноза, которая обеспечивает реализацию наиболее точного и достоверного прогноза. С этой целью целесообразно использовать различную исходную информацию и несколько методов прогнозирования.

4. Построение исходной модели прогноза и ее реализация. Данный этап предполагает, что основой построения прогноза является разработка достаточно адекватной исходной модели, обладающей прогностическими свойствами.

5. Проверка достоверности, точности и обоснованности прогноза. На данном этапе дается достоверная оценка процесса прогнозирования на основе расчета и анализа абсолютных, относительных и средних показателей точности прогноза. Надежность прогноза определяется, как правило, величиной доверительных интервалов.

6. Принятие решений на основе прогнозной модели и выработка рекомендаций о возможностях ее использования для получения прогнозных оценок.

Построение достаточно точных и надежных прогнозов позволяет на практике наиболее четко сформулировать резервы и пути развития изучаемых социально-экономических явлений и процессов.

Одним из наиболее распространенных методов прогнозирования социально-экономических явлений и процессов является ***экстраполяция***, *то есть продление тенденции и закономерностей, связей и соотношений прошлого и настоящего на будущее.*

Типичным и наиболее применимым примером экстраполяции является прогнозы по одномерному временному ряду, который заключается в продлении на будущий период сложившейся тенденции изучаемого явления. Основная цель данного прогноза заключается в том, чтобы показать, к каким результатам можно прийти в будущем, если развитие явления будет происходить со скоростью, ускорением и так далее, аналогичным прошлого периода. Если прогнозная оценка окажется неудовлетворительной, то сложившаяся в прошлом тенденция должна быть изменена с учётом тех факторов, под влиянием которых она складывается.

Широкое практическое применение методов экстраполяции трендов объясняется простотой метода, сравнительно небольшим объемом информации и четкостью механизма реализации, лежащих в его основе предпосылок.

Теоретической основой распространения тенденции на будущее является свойство социально-экономических явлений, называемое инерционностью.

Инерционность *– это сохранение тенденций, закономерностей, скорости и характера развития явлений и процессов в будущем, измеренных по данным прошлого периода.*

Статистическое прогнозирование предполагает не только качественное предсказание, но и достаточно точное количественное измерение вероятных возможностей, ожидаемых значений признака. Для данной цели важно, чтобы прогностическая модель имела достаточную точность или допустимо малую ошибку прогноза.

Ошибка статистического прогноза будет тем меньше, чем меньше срок упреждения и чем длиннее информационная база прогноза. Оба этих фактора ошибки прогноза имеют следующие условия: состояние и параметры процесса в ближайшем будущем более сходны с фактическими данными и поэтому их предвидеть можно точнее, чем параметры того же процесса в далеком будущем.

Если тенденция динамики сохранялась неизменной 30 лет, есть гораздо большая вероятность ее сохранения и в последующие пять лет, чем если существующая тенденция возникла всего десять лет назад.

Однако из этих условий нельзя однозначно вывести какой-либо универсальный алгоритм определения допустимого срока упреждения при заданной точности прогноза либо наоборот. Приходится на данном этапе ограничиться чисто эмпирическим правилом: в большинстве случаев срок упреждения не должен превышать третей части длины базы прогноза. Иначе говоря, для прогноза на 5 уровней желательно иметь временной ряд для прогноза по длине не менее чем 15 уровней.

В каждом конкретном исследовании соотношение длины базы прогноза и срока упреждения необходимо обосновать, кроме учета вышеперечисленных общих правил, используя еще и всю возможную информацию об особенностях изучаемого объекта.

Прогнозы на основе экстраполяции временных рядов, как и любые статистические прогнозы, могут быть либо точечными либо интервальными.

Экстраполяцию в общем виде можно представить формулой вида:

$$\hat{y}^{*}_{t+L} = f(y_i, L, a_i) \qquad (2.1)$$

где -прогнозируемый уровень;

y_i - текущий уровень исходного временного ряда;
L - период упреждения;
a_i - $\hat{y}^{*}_{t+L}$ параметр уравнения тренда

В зависимости от того, какие принципы и исходные данные положены в основу прогноза, можно выделить следующие группы методов прогнозирования социально-экономических явлений:

1. прогнозирование на основе простейших методов;
2. прогнозирование на основе экстраполяции трендов;
3. прогнозирование на основе дисконтирования информации;
4. прогнозирование на основе кривых роста.

Данные группы методов прогнозирования наиболее подробно будут рассмотрены в следующих параграфах.

2.2 Простейшие методы прогнозной экстраполяции

Наиболее простыми методами прогнозирования по одномерным временным рядам, являются:

- прогнозирование в предложении абсолютной неизменности значений предшествующих уровней в будущем;
- метод среднего уровня ряда;
- метод среднего абсолютного прироста;
- метод среднего темпа роста.

Рассмотрим каждый из названных методов.

Прогнозирование в предположении абсолютной неизменности значений предшествующих уровней исходит из утверждения, что каждое следующее прогнозное значение будет равно предыдущему значению признака, то есть:

$$\hat{y}^*_{t+L} = \hat{y}^*_{t+L-1} \qquad \begin{aligned} \hat{y}^*_{t+1} &= y_t \\ \hat{y}^*_{t+2} &= \hat{y}^*_{t+1} \\ &\dots \end{aligned} \qquad (2.2)$$

где $\hat{y}^*_{t+L}$ - прогнозное значение на период упреждения L.

$\hat{y}^*_{t+L-1}$ - прогнозное значение предшествующее периоду упреждения L.

Данный случай прогнозирования является частным и в практике статистического прогнозирования социально-экономических явлений встречается крайне редко.

Другим простейшим методом прогнозирования социально-экономических явлений является **метод прогнозирования на основе среднего уровня ряда.** Данный метод прогнозирования используется для случаев, когда изменение значений уровней временных рядов носит стационарный характер. При построении прогноза данным методом используется принцип, согласно которому значения всех последующих прогнозируемых уровней принимаются равными среднему значению уровней ряда в прошлом, то есть:

$$\hat{y}^{*}_{t+L} = \bar{y} \qquad (2.3)$$

Таким образом получают точечный прогноз. Однако, рассматривая временный ряд как выборку из некоторой генеральной совокупности, сложно предположить, что прогнозная точечная оценка полностью совпадает с эмпирическими значениями признака. В этом случае целесообразно определить доверительный интервал прогноза путем построения интервального прогноза данным методом по выражению вида:

$$\hat{y}^{*}_{t+L} = \bar{y} \pm t_{\alpha}\sigma_{\bar{y}}\,, \qquad (2.4)$$

где t_{α} - табличное значение t – критерия Стьюдента с (n-1) числом степеней свободы и уровнем значимости α;

$\sigma_{\bar{y}}$ - средняя квадратическая ошибка средней, которая определяется по формуле:

$$\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}}\,, \qquad (2.5)$$

где σ_y - среднее квадратическое отклонение, которое определяется как:

$$\sigma_y = \sqrt{\frac{\sum\left(y_i - \bar{y}\right)^2}{n-1}}\,, \qquad (2.6)$$

где y_i - эмпирические значения уровней временного ряда;

$\bar{y}$ - средний уровень исходного временного ряда;

n – число уровней ряда

Полученный таким образом (2.4) доверительный интервал учитывает колеблемость выборочных средних и предполагает, что каждая следующая прогнозная оценка будет равна среднему уровню ряда динамики. При этом упускается из вида возможность колеблемости эмпирических значений признака вокруг средней, то есть в определении доверительного интервала, в расчете дисперсии необходимо учесть как колеблемость выборочных средних, так и степень варьирования индивидуальных эмпирических значений признака вокруг средней.

В этом случае доверительный интервал прогнозной оценки можно определить по выражению вида:

$$\hat{y}^*_{t+L} = \bar{y} \pm t_\alpha \cdot \sigma_y \sqrt{1+\frac{1}{n}} \qquad (2.7)$$

Как видно, общая вариация прогнозируемого социально-экономического явления, то есть его ошибка, определяется суммой двух дисперсий: общая дисперсия и дисперсия выборочной средней при условии рассмотрения исходного временного ряда как выборки из некоторой генеральной совокупности.

Прогнозирование методом среднего абсолютного прироста предполагает, что общая тенденция развития изучаемого социально-экономического явления наилучшим образом аппроксимируется линейной формой аналитического выражения.

Применение данного метода прогнозирования возможно при предварительной *проверке следующих предпосылок:*

1. Абсолютные цепные приросты ($\Delta_{iy} = y_i - y_{i-1}$, где y_i - значение уровня i-го периода; y_{i-1} - значение уровня предшествующего i-му периоду времени) должны быть приблизительно одинаковыми;
2. Должно выполняться неравенство вида:

$$\sigma^2_{ост} \le \rho^2,$$

где $\sigma^2_{ост}$ - остаточная дисперсия, определяемая по формуле:

$$\sigma^2_{ост} = \frac{\sum_{i=1}^{n}\left(y_i - \bar{y}_{\bar{\Delta}}\right)^2}{n}, \qquad (2.8)$$

где y_i - эмпирические значения уровней ряда динамики;

$\bar{y}_{\bar{\Delta}}$ - теоретические значения уровней ряда, выравненные методом среднего абсолютного прироста.

n – число уровней исходного ряда динамики.

$$\rho^2 = \frac{1}{2} \cdot \frac{\sum_{i=1}^{n}\Delta_i^2}{n}, \qquad (2.9)$$

где Δ_i - цепные абсолютные приросты уровней исходного временного ряда.

После проверки и подтверждения выполнения данной предпосылки можно приступать к прогнозированию методом среднего абсолютного прироста, общая модель прогноза которого имеет вид:

$$\hat{y}^*_{t+L} = y_t + \overline{\Delta} \cdot L, \qquad (2.10)$$

где y_t - последний уровень исходного ряда динамики (для перспективного прогноза) или уровень принятый за базу экстраполяции;

L - период упреждения прогноза;

$\overline{\Delta}$ - средний абсолютный прирост, который определяется по формулам вида:

$$\overline{\Delta} = \frac{y_n - y_1}{n-1} \quad \text{или} \quad \overline{\Delta} = \frac{\sum \Delta_i}{n-1}, \qquad (2.11)$$

где y_n - последний уровень исходного ряда динамики;

y_1 - первый уровень исходного ряда динамики.

Как видно из приведенных преобразований, прогнозирование методом среднего абсолютного прироста заключается в непрерывном увеличении последнего уровня исходного ряда динамики на величину среднего абсолютного прироста на всем периоде упреждения.

Прогнозирование методом среднего темпа роста осуществляется в случае если темпы роста цепные, рассчитанные по данным исходного ряда динамики за исследуемый период времени, имеют приблизительно одинаковое цифровое значение, а тенденция развития явления подчиняется геометрической прогрессии и может быть описана показательной (экспоненциальной) кривой.

Модель прогноза методом среднего темпа роста имеет вид:

$$\hat{y}^*_{t+L} = y_t + \overline{T}_p^{\,L}, \qquad (2.12)$$

где y_t - последний уровень исходного ряда динамики (для перспективного прогноза) или уровень принятый за базу экстраполяции (во всех остальных случаях);

$\overline{T}_p$ - средний темп роста, который определяется по формулам вида:

$$\overline{T}_p = \sqrt[n-1]{\frac{y_n}{y_1}} \text{ или } \overline{T}_p = \sqrt{T_{p_{ц_1}} \times T_{p_{ц_2}} \times ... \times T_{p_{ц_{n-1}}}} = \sqrt[n-1]{ПT_{p_ц}}, \quad (2.13)$$

где y_n - последний уровень исходного ряда динамики;

y_1 - первый уровень исходного ряда динамики;

$T_{p_{ц_{1,2,...,n-1}}}$ - цепные темпы роста;

$ПT_{p_ц}$ - произведение цепных темпов роста

Сумма теоретических значений $\sum y_{\overline{T}p}$, полученных в результате выравнивания по среднему темпу роста, должна совпадать с суммой эмпирических значений исходного временного ряда $\left(\sum y_i\right)$:

$$\sum_{\overline{T}p}^{n-1} \overline{y}_{\overline{T}p} = \sum_{i=1}^{n} y_i \quad (2.14)$$

Несовпадение данных сумм может быть вызвано следующими причинами:

1. исходному временному ряду свойственна другая закономерность, а не экспоненциальная;
2. существенное и значимое влияние на изучаемое социально-экономическое явление оказывают случайные факторы.

Рассмотренные методы прогнозирования являются простейшими и поэтому, прогнозы полученные на их основе являются приближенными и не всегда надежны при увеличении периода упреждения. Как правило, эти методы используются только при краткосрочном прогнозировании.

Применение этих методов в среднесрочном и долгосрочном прогнозировании нецелесообразно, так как они не только не учитывают вариацию, скачки внутри временного ряда, но и в основе построения их моделей прогноза и получения прогнозных оценок на всем периоде упреждения лежит принцип равномерного увеличения или уменьшения (в зависимости от знака абсолютного прироста или допустимых границ темпа роста) исследуемого явления, в частности его последнего уровня в исходном временном ряду, от одного периода упреждения к другому на постоянную величину, количественно выраженную значением среднего абсолютного прироста или среднего темпа роста.

2.3 Прогнозирование на основе экстраполяции тренда

Наиболее распространенным методом прогнозирования выступает аналитическое выражение тренда. При этом, для выхода за границы исследуемого периода достаточно продолжить значения независимой переменной времени.

При таком подходе к прогнозированию предполагается, что размер уровня, характеризующего явление, формируется под воздействием множества факторов, причем не представляется возможным выделить порознь их влияние. В связи с этим ход развития связывается не с какими-либо конкретными факторами, а с течением времени, то есть:

$$\overline{y}_t = f(t). \qquad (2.15)$$

Экстраполяция дает возможность получить точечное значение прогноза. Точечный прогноз есть оценка прогнозируемого показателя в точке (в конкретном году, месяце, дне) по уравнению, описывающему тенденцию показателя. Точечная оценка рассчитывается путем подстановки номера года t, на который рассчитывается прогноз, в уравнении тренда. Она является средней оценкой для прогнозируемого интервала времени. Совпадение фактических данных и прогностических оценок – явление маловероятное, поэтому целесообразно определить доверительные интервалы прогноза.

Величина доверительного интервала определяется следующим образом:

$$\widehat{y}_t \pm t_\alpha \cdot \sigma_{\overline{y}_t}, \qquad (2.16)$$

где $\sigma_{\overline{y}_t}$ - средняя квадратическая ошибка тренда;

$\widehat{y}_t$ - расчетное значение уровня;

t_α - доверительное значение критерия Стьюдента.

Метод прогнозирования на основе экстраполяции тренда базируется на следующих предпосылках:

1. исходный временной ряд должен описываться плавной кривой;
2. общие условия, определяющие тенденцию развития изучаемого явления в прошлом и настоящем не должны претерпевать значительных изменений в будущем;
3. исходный ряд динамики должен иметь достаточное число уровней, с тем, чтобы отчетливо проявилась тенденция.

Трендовые модели выражаются различными функциями $\bar{y}_t = f(t)$, на основе которых строятся модели прогноза и осуществляется их оценка.

На практике наибольшее распространение получили следующие виды трендовых моделей:

1. линейная $\hat{y}^*_{t+L} = a_0 + a_1 t$;
2. параболы различных степеней:
   - 2-го порядка $\hat{y}^*_{t+L} = a_0 + a_1 t + a_2 t^2$;
   - 3-го порядка (кубическая) $\hat{y}^*_{t+L} = a_0 + a_1 t + a_2 t^2 + a_3 t^3$ и т.д.
3. степенная: $\hat{y}^*_{t+L} = a_0 + a_1 t^{a_2}$

   $\hat{y}^*_{t+L} = a_0 + a_1 t^{-a_2}$
4. показательная: $\hat{y}^*_{t+L} = a_0 \cdot a_1^{\,t}$ (2.17)

   $\hat{y}^*_{t+L} = a_0 \cdot e^t$

   $\hat{y}^*_{t+L} = a_0 \cdot e^{a_1 t}$
5. логарифмическая: $\hat{y}^*_{t+L} = a_0 + a_1 \lg t$.

При этом наиболее существенным вопросом прогнозирования по трендовым моделям является проблема точного прогноза.

Точная оценка прогноза весьма условна в силу следующих причин:

1. Выбранная для прогнозирования функция дает лишь приближенную оценку тенденции, так как она не является единственно возможной.
2. Статистическое прогнозирование осуществляется на основе ограниченного объема информации, что, в свою очередь, сказывается на величине доверительных интервалов прогноза.
3. Наличие в исходном временном ряду случайного компонента приводит к тому, что любой прогноз осуществляется лишь с определенной долей вероятности.

Рассматривая получение интервальных или точечных оценок прогноза следует учитывать, что в отдельных случаях получение более точных оценок не гарантирует надежности прогноза.

Применение трендовых моделей прогнозирования социально-экономических явлений имеют большую значимость и, несмотря на определенную простоту их реализации, часто применяются для прогнозирования сложных социально-экономических явлений.

Если выбранная модель тренда достаточно правильно отражает тенденцию развития, то полученные на ее основе прогнозы практически всегда надежны.

Прогнозирование методом экстраполяции тренда основывается на анализе тенденций развития одномерных временных рядов социально-экономических явлений и процессов.

Однако прогноз по аналитическому выражению тренда имеет один существенный недостаток, который иногда приводит к большим ошибкам при прогнозировании явления. Дело заключается в том, что в данном случае прогнозируется только детерминированная составляющая ряда динамики и не учитывается случайный компонент. Чтобы избежать этой ошибки и сделать прогноз более точным, надо отыскать закономерность изменения во времени случайного компонента. Для этого принято вначале находить отклонения от тренда и определять закономерность их изменения во времени, а затем делать прогноз случайной составляющей динамического ряда. Результаты обоих прогнозов объединяются. Рассматриваемый метод тогда дает удовлетворительные результаты, когда в эмпирическом ряду случайные колебания будут небольшими и между ними отсутствует автокорреляция.

2.4 Прогнозирование с учетом дисконтирования информации

Рассмотренные выше методы прогнозирования на основе временных рядов были основаны на равнозначной оценке исходной информации, независимо от того отражала эта информация последние или прошлые тенденции развития социально-экономических явлений и процессов.

Для получения достоверных прогнозов существенно: какая, по времени отражения прогнозируемых явлений, информация используется для получения прогноза. Практика показывает, что для точных и надежных прогнозных оценок наиболее ценной является информация последних уровней. Следовательно и оценивать исходную информацию необходимо по-разному: наиболее позднюю (последнюю) информацию необходимо оценивать выше, чем информацию, характеризующую тенденцию явления в прошлом. Такая оценка информации может быть произведена путем взвешивания или дисконтирования.

Принцип дисконтирования *предполагает, что для построения точных и надежных прогнозов более поздняя информация имеет больший удельный вес по степени информативности, чем более ранняя информация.*

На принципе дисконтирования информации разработаны следующие методы статистического прогнозирования:

1. метод простого экспоненциального сглаживания;
2. метод гармонических весов.

Данные методы могут быть использованы при прогнозировании социально-экономических явлений и процессов только при условии выполнения следующих предпосылок их реализации:

- исходные ряды динамики должны быть достаточно длинными с тем, чтобы более четко проявилась тенденция изменения социально-экономических явлений;
- в уровнях исходных временных рядов должны отсутствовать скачки в развитии явления;
- должен соблюдаться принцип инерционности, то есть тенденция и закономерности прошлого и настоящего могут продлеваться на будущее и для получения значительных изменений в основных характеристиках социально-экономических явлений необходимо, чтобы существовал значительный период упреждения.

Метод простого экспоненциального сглаживания заключается в том, что уровни исходного временного ряда взвешиваются с помощью скользящей средней, веса которой подчиняются экспоненциальному закону распределения. Данная скользящая средняя получила название экспоненциальной средней ($S_t(y)$) и позволяет проследить закономерности изменения явления в динамике по наиболее существенным, последним уровням. Особенность метода заключается в том, что при расчете теоретических значений полученных по модели тренда, учитываются только значения предыдущих уровней временного ряда, взятых с определенным весом.

Общая формула расчета экспоненциальной средней имеет вид:

$$S_t(y) = \alpha \cdot y_t + (1-\alpha) \cdot S_{t-1}(y), \qquad (2.18)$$

где $S_t(y)$ - значение экспоненциальной средней временного ряда для момента t;

$S_{t-1}(y)$ - значение экспоненциальной средней для момента (t-1);

y_t- значение последнего уровня исходного ряда динамики (для перспективного прогнозирования) или значение уровня временного ряда социально-экономического явления в момент t;

α- параметр сглаживания (вес t-го значения уровня временного ряда).

Из формулы (2.18) видно, что при вычислении экспоненциальной средней $S_t(y)$ используется значение только предыдущей экспоненциальной средней $S_{t-1}(y)$ и значение последнего уровня временного ряда, а все предыдущие уровни ряда опускаются.

Одной из проблем практической реализации метода простого экспоненциального сглаживания является определение значения параметра сглаживания α. От значения параметра α зависят веса предшествующих уровней временного ряда и в соответствии с этим степень их влияния на сглаживаемый уровень, а следовательно и значения прогнозных оценок. Чем больше значение параметра сглаживания α, тем меньше влияние на прогнозные оценки предшествующих уровней и тем следовательно меньше сглаживающее влияние экспоненциальной средней.

Если α стремится к 1 – это означает, что при прогнозе в основном учитывается влияние только последних уровней временного ряда.

Если α стремится к 0 – это означает, что при прогнозе учитываются прошлые уровни временного ряда.

Автор метода простого экспоненциального сглаживания Р.Г. Браун предложил следующую формулу расчета α:

$$\alpha = \frac{2}{n+1}, \qquad (2.19)$$

где n – число уровней временного ряда, вошедших в интервал сглаживания.

Пределы изменения α установлены эмпирическим путем и изменяются в пределах: $0{,}1 \le \alpha \le 0{,}3$.

Однако, следует учитывать, что в этом случае параметр α полностью зависит от числа наблюдений n. Часто на практике при решении конкретных задач параметр α применяется равным: α = 0,1; 0,15; 0,2; 0,25; 0,3.

Параметр сглаживания α может быть также определен на основе метода перебора различных его значений. При этом, в качестве оптимального значения α выбирается то значение α, при котором получена наименьшая средняя квадратическая ошибка прогноза, рассчитанная по данным всего сглаживаемого временного ряда или по данным части временного ряда, специально оставленной для проверки качества прогнозной модели. То есть путем построения ретроспективного прогноза, сущность которого заключается в том, что весь исходный ряд динамики разбивается на две части в соотношении 2/3 к 1/3. Для различных значений α строится модель прогноза по первой части ряда (2/3) и по ней осуществляется прогноз на вторую (1/3 от исходной) часть ряда, по которой определяются отклонения прогнозных значений ($\widehat{y}^{*}_{t}$) временного ряда от эмпирических значений уровней (y_i) и определяется средняя квадратическая ошибка этих отклонений по формуле:

$$\sigma_{ост} = \sqrt{\frac{\Sigma\left(\widehat{y}^{*}_{t} - y_i\right)^2}{n-k-1}}\,. \qquad (2.20)$$

Наиболее оптимальным считается тот параметр сглаживания α, которому соответствует наименьшее значение средней квадратической ошибки.

Прежде чем приступать к определению экспоненциальных средних, необходимо, кроме параметра α определить $S_{t-1}(y)$, то есть возникает проблема определения начальных условий. Таким образом прогнозирование методом простого экспоненциального сглаживания может быть реализован в двух возможных вариантах:

- начальные условия (y_0) известны.
- начальные условия не известны.

В случае если начальные условия известны также возможны два случая реализации этого варианта:

- В качестве начального условия y_0 может быть использована

средняя арифметическая, определенная по всем значениям уровней исходного временного ряда по формуле вида:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}. \qquad (2.21)$$

Использование средней арифметической в качестве начального условия возможно только в том случае, когда известны данные о развитии изучаемого социально-экономического явления в прошлом.

- В качестве начального условия y_0 возможно использование

значения первого уровня исходного временного ряда – y_1. При этом вес данного уровня будет уменьшаться по мере скольжения по уровням исходного временного ряда от уровня к уровню, а следовательно будет уменьшаться влияние каждого следующего уровня на величину экспоненциальной средней.

В случае если начальные условия не известны, то они могут быть определены по формулам, разработанным Р.Г. Брауном. При этом возможны различные модификации их расчета в зависимости от того, какая модель тренда наилучшим образом описывает реально существующую тенденцию развития изучаемого социально-экономического явления.

Так, если **тенденция** исходного временного ряда **описывается уравнением линейного тренда** вида:

$$\bar{y}_t = a_0 + a_1 t,$$

то *прогнозирование методом простого экспоненциального сглаживания осуществляется в следующей последовательности:*

1. Определяются параметры линейного тренда a_0 и a_1, описывающего тенденцию исходного временного ряда:

$$\bar{y}_t = a_0 + a_1 t\,.$$

Параметры a_0 и a_1 определяются путем решения следующей системы нормальных уравнений методом наименьших квадратов:

$$\begin{cases} n\,a_0 + a_1 \sum t = \sum y \\ a_0 \sum t + a_1 \sum t^2 = \sum yt \end{cases} \qquad (2.22)$$

2. Определяются начальные условия первого и второго порядков (порядок начальных условий определяется числом параметров уравнения тренда – линейного тренда - a_0 и a_1) по формулам вида:

- начальное условие первого порядка:

$$S_0^{[1]}(y) = a_0 - \frac{1-\alpha}{\alpha} a_1 \qquad (2.23)$$

- начальное условие второго порядка:

$$S_0^{[2]}(y) = a_0 - \frac{2(1-\alpha)}{\alpha} a_1 \qquad (2.24)$$

где a_0 и a_1 – параметры уравнения тренда (2.22), полученные методом наименьших квадратов.

3. Рассчитываются экспоненциальные средние первого и второго порядка:

- экспоненциальная средняя первого порядка:

$$S_t^{[1]}(y) = \alpha y_t + (1-\alpha)\cdot S_0^{[1]}(y); \qquad (2.25)$$

где y_t – значение последнего фактического уровня исходного временного ряда

- экспоненциальная средняя второго порядка:

$$S_t^{[2]}(y) = \alpha \cdot S_t^{[1]}(y) + (1-\alpha)\cdot S_0^{[2]}(y) \qquad (2.26)$$

4. Прогноз строится по модели вида:

$$\hat{y}^{*}t + L = \hat{a}_0 + \hat{a}_1 t\,,$$

где

$$\hat{a}_0 = 2S_t^{[1]}(y) - S_t^{[2]}(y);$$
$$\hat{a}_1 = \frac{\alpha}{1-\alpha}\left[S_t^{[1]}(y) - S_t^{[2]}(y)\right] \qquad (2.27)$$

5. Ошибка прогноза определяется по следующей формуле:

$$\sigma_{\hat{y}^*_{t+L}} = \sigma_y \cdot \sqrt{\frac{\alpha}{(2-\alpha)^5} \times \left[1 - 4(1-\alpha) + 5(1-\alpha)^2 + 2\alpha(4-3\alpha) \times t + 2\alpha^2 L^2\right]}, \qquad (2.28)$$

где σ_y - средняя квадратическая ошибка, рассчитанная по отклонениям эмпирических значений признака от теоретических, полученных по уравнению линейного тренда, то есть по следующей формуле:

$$\sigma_y = \sqrt{\frac{\Sigma\left(y_i - \hat{y}^* t + L\right)^2}{n-K}}, \qquad (2.29)$$

где К – число степеней свободы, определяемое в зависимости от длины исходного временного ряда (n) и числа параметров уравнения тренда.

Метод гармонических весов был разработан польским статистиком З. Хелвингом, близок к методу простого экспоненциального сглаживания и использует тот же принцип. В его основе лежит взвешивание скользящего показателя, но вместо скользящей средней используется идея скользящего тренда. Экстраполяция проводится по скользящему тренду, отдельные точки ломаной линии взвешиваются с помощью гармонических весов, что позволяет более поздним наблюдениям придавать большой вес.

Метод гармонических весов базируется на следующих предпосылках:

1. Период времени, за который изучается экономический процесс, должен быть достаточно длительным, чтобы можно было определить его закономерности.
2. Исходный ряд динамики не должен иметь скачкообразных изменений.
3. Прогнозируемое социально-экономическое явление должно обладать инерционностью, то есть для наступления большого изменения

в характеристиках процесса необходимо, чтобы прошло значительное время.

4. Отклонения от скользящего тренда (ε_t) имеют случайный характер.

5. Автокорреляционная функция, рассчитанная на основе последовательных разностей, должна уменьшаться с увеличением уровней временного ряда, то есть влияние более поздней информации должно отражаться на прогнозируемой величине сильнее, чем ранней информации.

Для получения точного прогноза по методу гармонических весов необходимо выполнение всех вышеуказанных предпосылок для исходного ряда динамики.

Для осуществления прогноза данным методом исходный временной ряд разбивается на фазы (к). Число фаз должно быть меньше числа членов ряда (n), то есть к < n. Обычно фаза равна 3-5 уровням. Для каждой фазы рассчитывается линейный тренд, то есть

$$y_i(t) = a_i + b_i t, \qquad (i = 1, 2, \ldots, n - k + 1); \qquad (2.30)$$

при этом для $i = 1$, $t = 1, 2, 3, \ldots, k$;

для $i = 2$, $t = 2, 3, \ldots, k + 1$;

для $i = n-k+1$ $t = n - k + 1, n - k + 2, \ldots, n$.

Для оценки параметров используется способ наименьших квадратов.

С помощью полученных (n – k + 1) уравнений определяются значения скользящего тренда. С этой целью выделяются те значения $y_i(t)$, для которых $t = i$, их обозначают $y_j(t)$. Пусть их будет n_j.

Затем находится среднее значение $\bar{y}_j(t)$ по формуле:

$$\bar{y}_j(t) = \frac{1}{n_j} \sum_{j=1}^{n_j} y_j(t), \qquad (2.31)$$

где $j = 1, 2, \ldots, n_j$.

После этого необходимо проверить гипотезу о том, что отклонения от скользящего тренда представляют собой стационарный процесс. С этой целью рассчитывается автокорреляционная функция. Если значения автокорреляционной функции уменьшаются от периода к периоду, то пятая предпосылка данного метода выполняется.

Далее рассчитываем приросты по формуле

$$w_{t+1} = f_{(t+1)} - f_{(t)} = \bar{y}_{t+1} - \bar{y}_t \qquad (2.32)$$

Средняя приростов вычисляется по формуле

$$\overline{w} = \sum_{t=1}^{n-1} C^{n}{}_{t+1} \cdot w_{t+1}, \qquad (2.33)$$

где $C^{n}{}_{t+1}$ - гармонические коэффициенты, удовлетворяющие следующим условиям:

$$C^{n}{}_{t+1} > 0; \quad (t = 1, 2, \dots, n - 1); \qquad (2.34)$$

$$\sum_{t=1}^{n-1} C^{n}{}_{t+1} = 1.$$

Данное выражение позволяет более поздней информации придавать большие веса, так как приросты весов обратно пропорциональны времени, которое отделяет раннюю информацию от поздней для момента t = n.

Если самая ранняя информация имеет вес $m_2 = \dfrac{1}{n-1}$, то вес информации, относящейся к следующему моменту времени, равен:

$$m_3 = m_2 + \frac{1}{n-2} = \frac{1}{n-1} + \frac{1}{n-2}. \qquad (2.35)$$

В общем виде ряд гармонических весов определяют по формуле:

$$m_{t+1} = m_t + \frac{1}{n-t}, \quad (t = 2, 3, \dots, n - 1); \qquad (2.36)$$

или $m_{t+1} = \sum_{i=1}^{t} \dfrac{1}{n-i}$.

Отсюда $$\sum_{t=1}^{n-1} m_{t+1} = n - 1.$$

Для того чтобы получить гармонические коэффициенты $C^{n}{}_{t+1}$ нужно гармонические веса m_{t+1} разделить на (n - 1), то есть:

$$C^{n}{}_{t+1} = \frac{m_{t+1}}{n-1}. \qquad (2.37)$$

Далее прогнозирование сводится, так же как и при простых методах прогноза, путем прибавления к последнему значению ряда динамики среднего прироста, то есть:

$$\hat{y}_{t+1} = \bar{y}^{*}t + \bar{w} \qquad (2.38)$$

При начальном условии $\bar{y}^{*}t = \bar{y}_{j(t)}$.

2.5 Прогнозирование на основе кривых роста

Прогнозирование социально-экономических явлений на основе кривых роста (кривых насыщения) стало применяться сравнительно недавно. Впервые эти методы были использованы в начале XX века для прогнозирования роста биологических популяций. Однако кривые роста хорошо себя зарекомендовали и при прогнозировании социально-экономических явлений. Однако их применение в этом случае требует соблюдения определенных условий.

1. Исходный временной ряд должен быть очень длинным (30-40 лет).
2. Исходный временной ряд не должен иметь скачков, и тенденция такого ряда должна описываться достаточно плавной кривой.
3. Использование кривых роста в прогнозировании социально-экономических явлений может давать достаточно хорошие результаты, если предел насыщения будет определен сравнительно точно.

Следует отметить, что кривые роста отражают кумулятивные возрастания к определенному заранее максимальному пределу. Особенностью кривых роста является то, что абсолютные приращения уменьшаются по мере приближения к пределу. Однако процесс роста идет до конца. Значение кривых роста как методов статистического прогнозирования социально-экономических явлений состоит в том, что они способствуют эмпирически правильному воспроизводству тенденции развития исследуемого явления.

Наиболее распространенными кривыми роста, используемыми в статистической практике прогнозирования, являются кривая Гомперца и кривая Перля-Рида.

Обе кривые, в общем, похожи одна на другую и графически изображаются S-образной кривой. Особенностью уравнений этих кривых является то, что их параметры могут быть определены методом наименьших квадратов лишь приближенно. Поэтому для расчета этих кривых используется ряд искусственных методов, основанных на разбиении исходного ряда динамики на отдельные группы.

Например, *для того чтобы осуществить прогноз на основе кривой Гомперца* (она названа так в честь английского статистика и математика, впервые применившего эту кривую для прогнозирования в страховании), *необходимо выполнить следующее:*

1. кривая описывается уравнением

$$y = a \cdot b^{cx}; \qquad (2.39)$$

2. прологарифмировав уравнение, получаем

$$lg\, y = lg\, a + (lg\, b) \cdot c^{x},$$

где *lg a* – логарифм максимального значения, к которому приближается прогнозный уровень явления;
lg b – расстояние, которое отделяет в каждый данный момент значение уровня от его максимального значения;
c – имеет значение от нуля до единицы;
x – начало на шкале x, то есть время, год, к которому относится первое значение уровня (t = 0, 1, 2, … , n);

3. затем весь ряд динамики разбивается на три части:

$$n_i = \frac{1}{3} \text{ длины ряда}; \qquad (2.40)$$

4. для каждой выделенной группы рассчитываются суммы S_1, S_2, S_3;
5. затем рассчитываются первые разности по этим суммам:

$$d_1 = S_2 - S_1; \qquad d_2 = S_3 - S_2; \qquad (2.41)$$

6. на основании этих расчетов получим параметры уравнения c, *lga, lg b,* которые рассчитываются следующим образом:

$$c^n = \frac{d_2}{d_1},$$

где n – число уровней ряда в каждой части;

Отсюда

$$c = \frac{d_1}{\sqrt[n]{c^n}}$$

$$\lg b = \frac{d_1 (c-1)}{(c^n - 1)^2}, \qquad (2.42)$$

$$\lg a = \frac{1}{n}\left(S_1 - \frac{d_1}{c^n - 1}\right)$$

Чтобы использовать данную кривую для экстраполяции за пределы исходного ряда динамики, достаточно подставить соответствующее значение x_t в уравнение кривой.

Наряду с кривой Гомперца достаточно широкое распространение получила также кривая Перля-Рида, которая в социально-экономической статистике впервые была использована для демографических расчетов американским учеными – биологом Р. Перлем и математиком Л. Ридом. Эта кривая выражает модифицированную геометрическую прогрессию, в которой возрастание затухает по мере приближения к некоторому определенному пределу. Максимальный предел устанавливается, прежде всего, на основании конкретного изучения исследуемого социально-экономического явления.

Так же как и кривая Гомперца, кривая Перля-Рида использует тот же искусственный прием для определения параметров кривой. Однако следует отметить, что по сравнению с кривой Гомперца прогнозные данные, полученные по этой кривой, имеют некоторую неопределенность.

Кривая Перля-Рида описывается уравнением:

$$\frac{1}{y} = a + bc^x \qquad (2.43)$$

Параметры уравнения находятся следующим образом:

$$c^n = \frac{d_2}{d_1}; \qquad b = \frac{d_1(c-1)}{\left(c^n - 1\right)^2}; \qquad a = \frac{1}{n}\left(S_1 - \frac{d_1}{c^n - 1}\right) \qquad (2.44)$$

Из приведенных расчетов видно, что параметры уравнения кривой Перля-Рида определяются так же, как и параметры кривой Гомперца, за исключением того, что в последнем случае не используется прием логарифмирования. Кроме того, нужно иметь в виду, что в зависимости от масштаба данных величина $\frac{1}{y}$ умножается на 10000, 100000 или 1000000.

2.6 Прогнозирование рядов динамики, не имеющих тенденции

При решении конкретных прикладных задач анализа социально-экономических явлений исследователь сталкивается с временными рядами социально-экономических показателей, в которых отсутствует тенденция развития, то есть изменение значений уровней исходного ряда динамики носит стационарный характер.

Однако временные ряды не имеющие тенденции, на практике, встречаются крайне редко.

В этой связи, прежде чем приступать к прогнозированию, необходимо всеми известными методами убедиться в том, что тенденция в исследуемом временном ряду действительно отсутствует. Только после того, как установлено отсутствие тенденции и гипотезы о наличии тенденции отвергнуты всеми методами, следует использовать те методы прогнозирования, которые дают возможность установить развитие явления при отсутствии тенденции.

Особенность прогнозирования данных временных рядов заключается в том, что использование методов статистического прогнозирования, основанных на получении точечной или интервальной количественной вероятностной характеристики изучаемого явления в будущем с относительно высокой степенью достоверности, невозможно.

В этом случае для прогнозирования таких рядов применяются вероятностные статистические методы прогнозного оценивания.

Вероятностные методы оценивания не позволяют дать точечную количественную характеристику прогнозируемого явления. Они дают возможность лишь оценить вероятность того, что значение прогнозируемого явления на каждый следующий (с отдалением) период упреждения будет больше или меньше значения последнего уровня исходного временного ряда. Вероятностные методы прогнозирования дают менее точные прогнозные оценки и обладают большей степенью неопределенности.

На практике, в анализе временных рядов социально-экономических явлений, не имеющих тенденции, наибольшее распространение среди вероятностных методов прогнозирования, получил метод, в основе которого лежит использование закона распределения Пуассона (распределение редких явлений) с плотностью

$$\rho = e^{-x}. \qquad (2.42)$$

Особенность метода заключается в том, что всегда прогнозируется благоприятная тенденция.

Этапы реализации данного метода следующие:

1. Осуществляется последовательное сравнение каждого следующего значения уровня исходного временного ряда со значением

предыдущего уровня. При этом знаком «+» отмечается возрастание значения уровня, а «-» - убывание. Если последующий уровень больше предыдущего, то ставится знак «+», меньше предыдущего – «-». Причем первый уровень всегда отмечается знаком «-». Знак «+» показывает, сколько периодов времени исследуемое явление возрастает и этот временный период принято считать благоприятной тенденцией.

2. Строится специальная таблица, характеризующая виды тенденции, длину благоприятной тенденции (τ) и частоту повторения благоприятной тенденции (f):

Виды тенденций	**Длина благоприятной тенденции, τ**	**Частота, f**
- - - + - - + + - - + + + - …	0 1 2 3 …	

При этом две первые графы таблицы: вид тенденции и длина благоприятной тенденции существуют априори и исследователь только частотой определяет наличие того или иного вида тенденции в исследуемом временном ряду.

Длина же благоприятной тенденции (τ) определяется числом плюсов между двумя минусами в ряду динамики «+» и «-».

3. На основе данных таблицы определяется средняя длина благоприятной тенденции по формуле вида:

$$\bar{\tau} = \frac{\sum \tau f}{\sum f}, \qquad (2.45)$$

где τ - длина благоприятной тенденции;

f - частота повторения благоприятной тенденции.

Средняя длина благоприятной тенденции показывает, сколько в среднем в рассматриваемом временном ряду, наблюдалось совершение благоприятной тенденции.

На основе полученной средней длины благоприятной тенденции $\bar{\tau}$ определяется показатель, характеризующий интенсивность прерываний этой благоприятной тенденции (λ), который определяется по формуле:

$$\lambda = \frac{1}{\bar{\tau}} \qquad (2.46)$$

Данный показатель характеризует сколько в среднем раз за рассматриваемый период времени совершалось прерывание благоприятной тенденции.

4. Вероятность благоприятной тенденции определяется на основе следующей модификации закона распределения Пуассона:

$$p = e^{-\lambda L}, \qquad (2.47)$$

где p - вероятность совершения благоприятной тенденции;
λ - интенсивность прерываний благоприятной тенденции;
L - период упреждения (число лет сохранения благоприятной тенденции).

2.7 Оценка точности и надежности прогнозов.

Важным этапом прогнозирования социально-экономических явлений является оценка точности и надежности прогнозов.

Эмпирической мерой точности прогноза, служит величина его ошибки, которая определяется как разность между прогнозными ($\hat{y}^{*}_{t}$) и фактическими (y_t) значениями исследуемого показателя. Данный подход возможен только в двух случаях:

а) период упреждения известен, уже закончился и исследователь располагает необходимыми фактическими значениями прогнозируемого показателя;

б) строится ретроспективный прогноз, то есть рассчитываются прогнозные значения показателя для периода времени за который уже имеются фактические значения. Это делается с целью проверки разработанной методики прогнозирования.

В данном случае вся имеющаяся информация делится на две части в соотношении 2/3 к 1/3. Одна часть информации (первые 2/3 от исходного временного ряда) служит для оценивания параметров модели прогноза. Вторая часть информации (последняя 1/3 части исходного ряда) служит для реализации оценок прогноза. Полученные, таким образом, ретроспективно ошибки прогноза в некоторой степени характеризуют точность предлагаемой и реализуемой методики прогнозирования. Однако величина ошибки ретроспективного прогноза не может в полной мере и окончательно характеризовать используемый метод прогнозирования, так как она рассчитана только для 2/3 имеющихся данных, а не по всему временному ряду.

В случае если, ретроспективное прогнозирование осуществлять по связным и многомерным динамическим рядам, то точность прогноза, соответственно, будет зависеть от точности определения значений факторных признаков, включенных в многофакторную динамическую модель, на всем периоде упреждения. При этом, возможны следующие подходы к прогнозированию по связным временным рядам: можно использовать как фактические, так и прогнозные значения признаков.

Все показатели оценки точности статистических прогнозов условно можно разделить на три группы:

- аналитические;
- сравнительные;
- качественные.

Аналитические показатели точности прогноза позволяют количественно определить величину ошибки прогноза. К ним относятся следующие показатели точности прогноза:

Абсолютная ошибка прогноза (Δ^*) определяется как разность между эмпирическим и прогнозным значениями признака и вычисляется по формуле:

$$\Delta^* = y_t - \widehat{y}^*_t \ , \qquad (2.48)$$

где y_t – фактическое значение признака;
$\widehat{y}^*_t$ - прогнозное значение признака.

Относительная ошибка прогноза ($d^*_{отн}$) может быть определена как отношение абсолютной ошибки прогноза (Δ^*):

- к фактическому значению признака (y_t):

$$d^*_{отн} = \frac{\Delta^*}{y_t} = \frac{\left| y_t - \widehat{y}^* t \right|}{y_t} \times 100\% \qquad (2.49)$$

- к прогнозному значению признака ($\widehat{y}^*_t$)

$$d^*_{отн} = \frac{\Delta^*}{y_t} = \frac{\left| y_t - \widehat{y}^* t \right|}{\widehat{y}^* t} \times 100\% \qquad (2.50)$$

Абсолютная и относительная ошибки прогноза являются оценкой проверки точности единичного прогноза, что снижает их значимость в оценке точности всей прогнозной модели, так как на изучаемое социально-экономическое явление подвержено влиянию различных факторов внешнего и внутреннего свойства. Единично удовлетворительный прогноз может быть получен и на базе реализации слабо обусловленной и недостаточно адекватной прогнозной модели и наоборот – можно получить большую ошибку прогноза по достаточно хорошо аппроксимирующей модели.

Поэтому на практике иногда определяют не ошибку прогноза, а некоторый коэффициент качества прогноза ($К_к$), который показывает соотношение между числом совпавших (с) и общим числом совпавших (с) и несовпавших (н) прогнозов и определяется по формуле:

$$К_к = \frac{с}{с + н} \qquad (2.51)$$

Значение $К_к$ = 1 означает, что имеет место полное совпадение значений прогнозных и фактических значений и модель на 100% описывает изучаемое явление. Данный показатель оценивает удовлетворительный вес совпавших прогнозных значений в целом по временному ряду и изменяющегося в пределах от 0 до 1.

Следовательно, оценку точности получаемых прогнозных моделей целесообразно проводить по совокупности сопоставлений прогнозных и фактических значений изучаемых признаков.

Средним показателем точности прогноза является средняя абсолютная ошибка прогноза ($\overline{\Delta^*}$), которая определяется как средняя арифметическая простая из абсолютных ошибок прогноза по формуле вида:

$$\overline{\Delta^*} = \frac{\sum\limits_{t=1}^{n} \Delta^*}{n} = \frac{\sum\limits_{t=1}^{n} \left| y_t - \hat{y}_t^* \right|}{n}, \qquad (2.52)$$

где n – длина временного ряда.

Средняя абсолютная ошибка прогноза показывает обобщенную характеристику степени отклонения фактических и прогнозных значений признака и имеет ту же размерность, что и размерность изучаемого признака.

Для оценки точности прогноза используется **средняя квадратическая ошибка прогноза**, определяемая по формуле:

$$\sigma_{ош} = \sqrt{\frac{\sum\limits_{t=1}^{n} \left(y_t - \hat{y}_t^* \right)^2}{n}} \qquad (2.53)$$

Размерность средней квадратической ошибки прогноза также соответствует размерности изучаемого признака. Между средней абсолютной и средней квадратической ошибками прогноза существует следующее примерное соотношение:

$$\sigma_{ош} = 1{,}25\overline{\Delta}^{*} \qquad (2.54)$$

Недостатками средней абсолютной и средней квадратической ошибками прогноза является их существенная зависимость от масштаба измерения уровней изучаемых социально-экономических явлений.

Поэтому на практике в качестве характеристики точности прогноза определяют **среднюю ошибку аппроксимации**, которая выражается в процентах относительно фактических значений признака, и определяется по формуле вида:

$$\bar{\varepsilon}^{*} = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|y_t - \hat{y}_t^{*}\right|}{y_t} \times 100\% \qquad (2.55)$$

Данный показатель является относительным показателем точности прогноза и не отражает размерность изучаемых признаков, выражается в процентах и на практике используется для сравнения точности прогнозов полученных как по различным моделям, так и по различным объектам. Интерпретация оценки точности прогноза на основе данного показателя представлена в следующей таблице:

$\bar{\varepsilon}$,%	Интерпретация точности
< 10	Высокая
10 – 20	Хорошая
20 – 50	Удовлетворительная
> 50	Не удовлетворительная

В качестве **сравнительного показателя точности прогноза** используется **коэффициент корреляции** между прогнозными и фактическими значениями признака, который определяется по формуле:

$$R = \frac{\frac{1}{n}\sum_{t=1}^{n}\left(\hat{y}_t^{*} - \bar{\hat{y}}_t^{*}\right)\cdot\left(y_t - \bar{y}\right)}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\hat{y}_t^{*} - \bar{\hat{y}}_t^{*}\right)^2 \cdot \frac{1}{n}\sum_{t=1}^{n}\left(y_t - \bar{y}\right)^2}}, \qquad (2.56)$$

где $\bar{\bar{y}}^{*}_{t}$ - средний уровень ряда динамики прогнозных оценок.

Используя данный коэффициент в оценке точности прогноза следует помнить, что коэффициент парной корреляции в силу своей сущности отражает линейное соотношение коррелируемых величин и характеризует лишь взаимосвязь между временным рядом фактических значений и рядом прогнозных значений признаков. И даже если коэффициент корреляции R = 1, то это еще не предполагает полного совпадения фактических и прогнозных оценок, а свидетельствует лишь о наличии линейной зависимости между временными рядами прогнозных и фактических значений признака.

Одним из показателей оценки точности статистических прогнозов является **коэффициент несоответствия (КН)**, который был предложен Г. Тейлом и может рассчитываться в различных модификациях:

1. **Коэффициент несоответствия ($КН_1$),** определяемый как отношение средней квадратической ошибки к квадрату фактических значений признака:

$$KH_1 = \sqrt{\frac{\sum_{t=1}^{n}\left(\hat{y}_t^{*} - y_t\right)^2}{\sum_{t=1}^{n} y_t^2}} \qquad (2.57)$$

КН = о, если $\hat{y}^{*}_{t} = y_t$, то есть полное совпадение фактических и прогнозных значений признака.

КН = 1, если при прогнозировании получают среднюю квадратическую ошибку адекватную по величине ошибке, полученной одним из простейших методов экстраполяции неизменности абсолютных цепных приростов.

КН > 1, когда прогноз дает худшие результаты, чем предположение о неизменности исследуемого явления. Верхней границы коэффициент несоответствия не имеет.

2.Коэффициент несоответствия $КН_2$ определяется как отношение средней квадратической ошибки прогноза к сумме квадратов отклонений фактических значений признака от среднего уровня исходного временного ряда за весь рассматриваемый период:

$$KH_2 = \sqrt{\frac{\sum_{t=1}^{n}\left(\hat{y}_t^{*} - y_t\right)^2}{\sum_{t=1}^{n}\left(\bar{y} - y_t\right)^2}}\ , \qquad (2.58)$$

где $\bar{y}$ - средний уровень исходного ряда динамики.

Если КН > 1, то прогноз на уровне среднего значения признака дал бы лучший результат, чем имеющийся прогноз.

3.Коэффициент несоответствия (КН$_3$), определяемый как отношение средней квадратической ошибке прогноза к сумме квадратов отклонений фактических значений признака от теоретических, выравненных по уравнению тренда:

$$KH_3 = \sqrt{\frac{\sum\limits_{t=1}^{n}\left(\hat{y}_t^{\,*} - y_t\right)^2}{\sum\limits_{t=1}^{n}\left(y_t - \bar{y}_t\right)^2}}\,, \qquad (2.59)$$

где $\bar{y}_t$ - теоретические уровни временного ряда, полученные по модели тренда.

Если КН > 1, то прогноз методом экстраполяции тренда дает хороший результат.

Заключение

В предложенном учебном пособии рассмотрена методология комплексной оценки и анализа реальных социально-экономических явлений и процессов, представленных одномерными и многомерными временными рядами, в разрезе выявления, характеристики и моделирования тенденции и методов ее прогнозирования с учетом особенностей и специфики применяемых методов и изучаемого объекта исследования.

Статистическая методология анализа временных рядов и прогнозирования находит широкое применение во многих областях знаний как на макро- так и на микроуровнях экономического развития, оценке эффективности, финансового состояния и финансовой устойчивости, деловой активности сегментов различных рынков и организационно-правовых структур.

Наиболее эффективным и целесообразным является широкое использование прикладного программного обеспечения при решении задач исследования конкретных социально-экономических явлений и процессов, что существенно ускоряет проведение расчетов. В этой связи распространены следующие программные продукты, такие как стандартные пакеты прикладных программ STATISTIKA, ОЛИМП, SPSS, STATGRAPHICS и другие.

Список литературы

1. Андерсен Т. Статистический анализ временных рядов. М., Мир, 1976.

2. Вайну Я. Я. Корреляция рядов динамики. М., Статистика, 1977.

3. Венсель В.В. Интегральная регрессия и корреляция: статистическое моделирование рядов динамики. М., Финансы и статистика, 1983.

4. Гранберг Д. Статистическое моделирование и прогнозирование. М., Финансы и статистика,1990.

5. Иващенко Т.А., Кильдишев Г.С., Шмойлова Р.А .Статистическое изучение основной тенденции и взаимосвязи в рядах динамики. Томск, издательство Томского университета, 1985.

6. Лизер Р. Эконометрические методы краткосрочного прогнозирования. М., Статистика, 1979.

7. Льюис Х.Д. Методы прогнозирования экономических показателей. М., Статистика,1986.

8. Королев Ю.Г., Рабинович П.М., Шмойлова Р.А. Статистическое моделирование и прогнозирование. Учебное пособие. МЭСИ, 1985.

9. Маленво Э. Статистические методы эконометрии. Вып. 1, 2. М. Статистика, 1976.

10. Рябушкин Т.В. Методологические методы анализа и прогноза краткосрочных процессов. М., Статистика, 1979.

11. Манелля А.В., Юзбашев М.М. Статистический анализ тенденций колеблемости. М., Финансы и статистика, 1983.

12. Четыркин Е.М. Статистические методы прогнозирования. М., Статистика, 1977.