

2.1. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ

В первой главе было показано, что ряды социально-экономической динамики могут быть классифицированы в две больших группы: обратимые и необратимые. К группе стационарных обратимых процессов в полном объёме применим выборочный метод, поскольку со временем статистические характеристики процесса не меняются, а выборочная совокупность даёт представление о генеральной совокупности в целом. При этом увеличение числа элементов выборочной совокупности способствует улучшению математического описания законов изменения генеральной совокупности в целом. Но, к сожалению, чаще всего увеличение объёма выборки, к чему стремится прогнозист, желающий большей точности в прогнозировании стационарных процессов, ведёт и к существенному увеличению затрат на получение этих дополнительных элементов выборки, их обработки и включения в расчётные процедуры. Поэтому прогнозисту приходится решать задачу нахождения приемлемого компромисса между стремлением к повышению точности прогноза и стремлением минимизировать затраты на выполнение прогноза. Поскольку процессы случайны и стационарны, то для их анализа и прогнозирования лучшими методами будут являться методы теории вероятностей и математической статистики.

Первая задача, которую необходимо решить при прогнозировании – собрать необходимую информацию о прогнозируемом объекте и обработать её так, чтобы она была пригодна для дальнейшего использования в моделировании. Несмотря на кажущуюся простоту этой задачи, прогнозисту весьма часто приходится сталкиваться с ситуацией, когда он не может решить её сразу. При постановке задачи прогнозирования, формируется ряд показателей, отражающих основные характеристики объекта прогнозирования. Информацию об этих показателях за некоторый определённый период времени прогнозирование промежуток времени должен собрать прогнозист.

Различают первичные (полевые) и вторичные (кабинетные) методы исследования и сбора необходимой информации¹.

Основными методами получения первичных данных являются опрос, наблюдение, эксперимент и панельные данные. Методы опроса, наблюдения, эксперимента, панельных исследований могут использоваться как отдельно, так и в сочетании с другими методами. Более подробно с этими методами можно познакомиться в дисциплине «Маркетинговые исследования». Первичная информация отличается особой достоверностью, поскольку исследователь сам эту информацию собирает и сам же её обрабатывает. Но следует отметить, что она касается исключительно микроуровня: поведение потребителей; изменение цен в группе магазинов или на локальном рынке (например, ММВБ); передвижение

¹ Светульников С.Г. Методы маркетинговых исследований. – СПб.: Изд-во ДНК, 2003.

товаров по магистралям и т.п. Способ получения первичной информации в ходе эмпирических исследований обладает существенными недостатками.

Во-первых, стоимость информации оказывается весьма высокой, поскольку она получается в результате «полевых» исследований, для которых привлекаются значительные силы; на подготовку, проведение и обработку полученных результатов затрачивается много времени.

Во-вторых, объём полученной информации является незначительным и не всегда удовлетворяет прогнозиста. Он может, например, получить информацию о динамике изменения потребительских предпочтений, но не получить знание о том, почему происходит эта динамика, не выявленными могут оказаться факторы, оказывающие влияние на эту динамику. А это означает, что он не в состоянии будет сделать верный прогноз.

В-третьих, для того, чтобы в ходе проведения полевых исследований получить достоверную информацию, необходимо проводить эти исследования на довольно высоком уровне, то есть, обязательно следует привлекать для этого специалистов соответствующей квалификации, чаще всего – социологов, когда речь идёт о непосредственном контакте с некоторым социумом или его представителями для получения от них первичной информации. Если же эту задачу попытается решить экономист самостоятельно, чаще всего он её решит неудовлетворительно.

Поэтому сбор первичной информации для прогнозирования могут позволить себе только те организации, которые имеют для этого достаточное количество свободных средств и для которых точность прогноза имеет определяющее значение.

Чаще всего в связи с вышесказанным прогнозисты удовлетворяются использованием вторичной информации.

Вторичные исследования, как правило, базируются на уже имеющейся информации и поэтому носят название кабинетных исследований. Вторичные исследования по своему содержанию - это анализ имеющихся источников об изучаемой проблеме.

Различают внешние и внутренние (по отношению к объекту социально-экономического прогнозирования) источники для вторичных исследований. В качестве внутренних источников информации могут быть:

- внутренняя статистика (характеристика товарооборота, объём сбыта, объём распродаж, импорт, экспорт, рекламации),
- данные о производственных затратах,
- прочие данные (о производительности установок, оборудования, прайслисты на сырьё и материалы, характеристика системы складирования, карты потребителей и др.).

В качестве внешних источников выступают:

- публикации национальных и международных официальных организаций;
- публикации государственных органов, министерств, муниципальных комитетов и организаций;
- публикации торгово-промышленных палат и объединений;

- ежегодники статистической информации;
- отчеты и издания отраслевых фирм и совместных предприятий;
- книги, сообщения в журналах и газетах;
- публикации учебных, научно-исследовательских, проектных институтов и общественно-научных организаций, симпозиумов, конгрессов, конференций;
- прайс-листы, каталоги, проспекты и другие фирменные публикации.

Внешнюю информацию подразделяют на общедоступную информацию и специализированную информацию, доступную только ее подписчикам. Последнюю информацию называют "синдикативной" (от слова "синдикат"). Синдикативная информация готовится специализированными центрами по определенной тематике – товары народного потребления, горюче-смазочные материалы, автомобили и т.п. В узкой специализации такой информации заключается ее главное преимущество. К ним же можно отнести относительную дешевизну информации, так как затраты на сбор информации делятся между подписчиками. Как правило, синдикативная информация очень высокого качества.

К недостаткам синдикативной информации следует отнести то, что, во-первых, подписчики не могут оперативно влиять на объем и качество информации, во-вторых, сборники информации выписываются на длительный период времени и при ее ненужности от нее сложно отказаться, в-третьих, указанная информация становится доступной конкурентам.

Значимость для вторичных исследований внутренней или внешней информации в каждом конкретном случае определяется исследователем. Основным достоинством вторичных исследований является то обстоятельство, что затраты на проведение кабинетных исследований существенно меньше, чем на получение такого же объема информации с помощью полевых исследований. Другим существенным преимуществом кабинетных исследований для сбора информации, пригодной для прогнозирования, является то, что времени на сбор и обработку информации здесь по сравнению с полевыми исследованиями необходимо потратить на порядок меньше. Поэтому прогноз можно получить в оперативном порядке, чего не скажешь о прогнозах, получаемых на базе первичной информации.

К сожалению, кабинетные исследования могут привести к получению недостоверной информации, поскольку они опираются на вторичную информацию, которая была собрана другими исследователями. А качество этих исследований может быть низкого уровня.

В тех случаях, когда вторичное исследование не дает нужного результата, проводят первичное (полевое) исследование.

В любом случае полевые исследования всегда дороже кабинетных исследований. Поэтому они применяются в случаях, когда:

- в результате вторичного исследования не достигнут требуемый результат и не возможно прогнозирование динамики;
- высокие затраты на полевые исследования могут быть компенсированы существенным повышением точности за счёт получения в качестве базы достоверной и представительной информации.

На практике поступают следующим образом. Вначале проводят кабинетные исследования, собирают и анализируют максимально возможное число вторичной информации. В том случае, когда этой информации оказалось недостаточно, проводят полевые исследования. Полевое исследование может быть полным или сплошным, если им охвачена вся группа интересующих исследователя респондентов и частичным, если им охвачен определенный процент респондентов. Сплошные исследования обычно используются для исследования относительно небольшого числа респондентов, например, крупные потребители, крупные фирмы.

Если взять объём информации, с которым работают современные прогнозисты, за сто процентов, то с уверенностью можно утверждать, что в 95% случаев прогнозисты работают со вторичной информацией, то есть с информацией, которую уже кто-то собрал и обработал.

И здесь он может столкнуться с проблемой отсутствия достоверности той информации, которая имеется в его распоряжении. С сожалением это можно констатировать применительно к отечественной статистике. Обращаясь к опубликованным источникам, прогнозист, прежде всего, отметит, что информации много, очень много, чрезвычайно много, но той информации, которая ему нужна, в наличии нет. Например, если прогнозист в качестве одного из показателей использует стоимость основных производственных фондов России, то найти этот показатель будет очень не просто. В лучшем случае он найдёт значения этих показателей за прошлые годы, но информация за последние два-три года оказывается ему недоступной. Как в этом случае осуществить прогноз?

Но даже если эти данные имеются в распоряжении прогнозиста, то он, осуществив предварительный анализ этих данных, может быть поставлен в тупик. Например, в табл. 2.1 приведены данные по динамике основных фондов и инвестиций по Псковской области¹. Из курса «Экономика предприятия» студент знает, что полная стоимость основных производственных фондов в текущем году определяется так: остаточная стоимость фондов на конец предыдущего периода плюс инвестиции данного года в основные фонды минус выбытие фондов за год.

Из данных таблицы 2.1 легко убедиться в том, что приведённые значения по основным фондам и инвестициям совершенно не вписываются в эту логику. Остаточная стоимость основных фондов, например, в 2006 году (третий столбец), равная 81 641 млн. руб. никак не коррелирует с полной стоимостью основных фондов 2006 года – 164 095 млн. руб. Инвестиции этого года, равные 7 603,3 млн. руб. ещё больше запутывают суть того, что скрывается за этими цифрами, поскольку если к величине остаточной стоимости прибавить величину инвестиций, то будет получено такое число:

$$81\,641 \text{ млн.руб.} + 7\,603,3 \text{ млн.руб.} = 89\,244,3 \text{ млн.руб.}$$

Сравнив это полученное с помощью вычислений число с тем, что приводится во втором столбце, можно убедиться, что числа отличаются друг от друга почти в два раза! Объяснить этот двукратный рост с помощью других вычисле-

¹ Данные представлены О. Коровка и Г. Бомейко.

ний невозможно. Единственное объяснение этому явлению можно найти только в том, что каждый год происходит переоценка основных фондов. Как подтвердили специалисты из областной статистики, фактические значения основных фондов, которые собираются на местах, передаются в Госкомстат России, где и осуществляется их переоценка. Как осуществляется эта переоценка, сообщается в соответствующих специализированных изданиях.

Табл. 2.1.
Некоторые данные по экономике Псковской области

<i>Год</i>	<i>Полная стоимость основных фондов, млн руб</i>	<i>Остаточная стоимость основных фондов в предыдущем периоде, млн руб</i>	<i>Инвестиции в основной капитал в отчетном периоде, млн руб</i>
1	2	3	4
2001	90537	49234	2791
2002	111128	52604	3135
2003	122694	60463	5613,6
2004	129423	67313	5904,9
2005	144880	69984	5546,9
2006	164095	73973	7603,3
2007	188943	81641	11831,2

Эта ситуация с основными фондами, очевидно, характерна для всех областей России, и для некоторых республик бывшего СССР, где также производится переоценка основных фондов по собственным методикам. Поэтому прогнозисту в этом случае необходимо провести дополнительные исследования, чтобы понять, как определить ту величину стоимости основных фондов, которую он собирается применить в своих расчётах. Без такой дополнительной работы использовать ни один из трёх показателей таблицы 2.1 нельзя.

Но если предположить, что подготовленная к анализу информация полностью ясна и отвечает необходимым условиям с позиций её смыслового содержания, то следует помнить, что она может содержать в себе ошибки, которые могут сильно повлиять на точность прогноза. От ошибок необходимо избавиться, но для того, чтобы это сделать, следует понять, какого рода ошибки могут содержаться в статистической информации.

По своим свойствам и характеру влияния на результаты наблюдений ошибки подразделяют на **грубые, систематические и случайные**¹.

Грубые ошибки возникают в случае невнимательности человека, эту информацию записывающего, передающего или получающего. В результате этой невнимательности измеренное число увеличивается или уменьшается на порядок и резко выделяется в общей совокупности данных. В этой связи с такой ошибкой легко бороться – она легко обнаруживается и легко устраняется. Для этого есть несколько возможных процедур:

¹ Большаков В.Д. Теория ошибок наблюдений. – М.: Недра, 1983. - С. 80-81.

1. *Графический анализ.* Статистические данные наносят на некоторый график. Если рассматривается статистическая взаимосвязь между двумя факторами, то изучается эта взаимосвязь. Если изучаются динамические ряды, то строится зависимость изменения показателей во времени. Данные, которые резко выделяются из совокупности наблюдений, отбрасываются. Конечно, надо быть полностью уверенным в том, что это наблюдение не является проявлением некоторых факторов, присущих анализируемому процессу. Ведь возможны ситуации, когда на объект начинают действовать некоторые новые факторы, влияние которых ещё не известно, но в будущем это приведёт к резким отклонениям от общей тенденции. Но резкое отклонение, наблюденное внутри довольно продолжительного отрезка времени, можно смело отбрасывать, а вот в случае, когда отклонение наблюдается в конце ряда наблюдений, необходимо изучить – действительно ли такое отклонение является результатом грубой ошибки, или это предвестник возможных серьёзных изменений?

2. *Табличный анализ.* Данные, которые невозможно, или нежелательно изображать графически, анализируют непосредственно в таблицах. Например, если данные содержат в себе несколько тысяч наблюдений, их графическое изображение будет представлять собой хаотическое нагромождение точек из-за того, что их общее количество велико, а размеры графика ограничены. Возможен и другой вариант, когда изучается многофакторная зависимость, графическое изображение которой осуществить довольно сложно. Тогда имеющиеся данные располагают в таблице в некотором порядке, присущем анализируемому процессу. Например, в порядке поступления информации (временные ряды) или в порядке возрастания какого-либо признака. После того, как произведено такое упорядочивание, проводится анализ полученной последовательности. Те наблюдения, которые выбиваются из этого порядка, отбрасываются. Делают это визуально. При этом необходимо иметь в виду, что такой анализ субъективен и очень высока вероятность того, что исследователь может не заметить в общей совокупности информационных массивов наблюдение, содержащее грубую ошибку.

3. *Статистический анализ.* В том случае, когда визуальный анализ затруднён из-за большого количества данных, следует использовать статистические процедуры. Полученные данные проверяются на одну из статистических гипотез относительно характера распределения вероятностей появления анализируемой совокупности данных. После этого данные, противоречащие этому закону, отбрасываются. Чаще всего предполагают наличие нормального закона распределения. В самом простом случае, когда измеренные величины колеблются около некоторого значения, рассчитывается средняя арифметическая этого значения, дисперсия, доверительные интервалы. Данные, выходящие за эти интервалы, отбрасываются. Естественно, что доверительные границы на этом этапе предварительного анализа данных должны быть наиболее широкими с очень высоким уровнем доверительной вероятности – равным, например, 99%.

Источниками **систематической ошибки** могут являться как *инструмент сбора и обработки информации*, так и *человеческий фактор* (желание приукра-

ситуацию или скрыть часть неблагоприятной информации). К сожалению, при работе со вторичными данными социально-экономической динамики (официальной и неофициальной статистикой) очень часто приходится иметь дело с ошибками, вызванными влиянием *человеческого фактора*. Особенно это касается данных отечественной статистики и статистики стран – республик бывшего СССР. Дело в том, что большая часть экономических показателей отражает эффективность деятельности того или иного подразделения, той или иной системы, того или иного региона. Классическим примером ошибки информации такого рода являются записываемые год от года в статистические сборники, данные о количественных показателях экономического развития бывшего СССР, которые отражали не столько реальные процессы, сколько желаемые результаты. Известно, что, например, в Узбекской ССР долгие годы шли приписки о сборе невыращенного хлопка, которые попадали в статистические сборники. По отчетам о выполнении плановых заданий по сбору хлопка составлялись планы работы текстильной промышленности, которая из несобранного хлопка не могла, естественно, выпустить несуществующую ткань. Но в статистические отчеты часть этой невыпущенной ткани попадала, а другая списывалась на разного рода потери и брак. Эта цепочка пронизывала весомую часть статистической отчетности, в результате чего ошибочно определялись и производные экономические инструменты - расхода хлопка на единицу ткани, нормы электропотребления, производительность труда и т.п. Но, помимо приписок хлопка в экономике бывшего СССР осуществлялись и другие приписки. Таким образом, практически все обобщающие данные экономического развития (валовой продукт, национальный доход и т.п.) отдельных регионов, да и страны в целом оказывались год от года засоренными ошибками такого рода.

Сегодня причиной возникновения подобной ошибки может быть, например, желание уменьшить выплаты по платежам в бюджеты и внебюджетные фонды, искажение данных в ходе «информационной войны» с конкурентами и т.п. Многие предприятия в результате этого представляют в статистические органы информацию, засоренную именно такой ошибкой и выявить в ней эту ошибку практически невозможно.

Систематические ошибки же, вызванные применением неисправного инструмента измерения, могут быть выявлены и исключены, так как имеют примерно одну и ту же величину, один и тот же знак. Поэтому исходные данные, содержащие этот тип ошибки, всегда несколько завышены или занижены. Объективным источником этой ошибки служат, в основном, измерительные устройства, приёмы или приборы, вносящие одну и ту же погрешность при измерениях.

Случайные ошибки неизбежны. Причины их появления многообразны. Они вызваны действием множества случайных неконтролируемых факторов и поэтому не поддаются анализу. В результате этого практически любое измерение содержит случайные ошибки, но так как источников возникновения слу-

чайных ошибок достаточно много, они, как правило, обладают следующими свойствами¹.

Первое свойство. Для ряда результатов наблюдений с известным параметром распределения абсолютные величины случайных ошибок с заданной вероятностью P не превосходят определенного предела. Это значит, что влияние случайных ошибок на результат все-таки незначительно.

Второе свойство. Положительные и отрицательные случайные ошибки равновозможны, т.е. они одинаково часто встречаются при наблюдениях. Из этого вытекает и следующее свойство.

Третье свойство. Математическое ожидание случайной ошибки равно нулю.

Четвертое свойство. Малые по абсолютной величине случайные ошибки встречаются при наблюдениях чаще, чем большие.

Следовательно, можно предполагать в большинстве случаев, что случайные ошибки подчиняются закону нормального распределения вероятностей и их математическое ожидание равно нулю. С учетом перечисленных свойств, создается ситуация, когда проявляются условия действия центральной предельной теоремы теории вероятностей (закон больших чисел), в соответствии с которой «совокупное действие большого числа случайных факторов приводит при некоторых, достаточно широких условиях к результату, почти не зависящему от случая»². Таким образом, избежать влияния случайных ошибок можно, если увеличить объем выборки.

После того, как будут устранены или учтены ошибки информации, перед исследователем возникает проблема ее систематизации и обработки. В достаточно редких случаях необходимая исследователю информация представлена в систематизированном виде и в виде, пригодном для последующего анализа и обработки. Чаще всего информация представлена в виде некоторой неупорядоченной совокупности. Для того чтобы ее обработать и сделать соответствующие выводы, она нуждается в упорядочении и систематизации.

Систематизация информации заключается в ее представлении в виде таблиц, графиков, диаграмм и других формах, удобных для исследователя и показывающих некоторые наиболее очевидные закономерности. В большинстве случаев исследователи предпочитают сведение информации в статистические таблицы – при этом возможен их последующий формализованный анализ с помощью математических методов.

Для того чтобы неупорядоченную совокупность данных можно было свести в таблицу, необходимо определить признак упорядочивания данных. Такими признаками могут являться:

- номер наблюдения;
- время наблюдения;
- номера экспертов, дававших оценку объекту исследования;

¹ Большаков В.Д. Теория ошибок наблюдений. - М.: Недра, 1983. - 223 с.

² Дунин-Барковский И.В., Смирнов Н.В. Теория вероятностей и математическая статистика в технике. – М.: Гос. изд-во технико-теоретической литературы, 1955. - С.104.

- ранги, полученные для свойств товара по шкале отношений или интервалов;
- товарный ряд и т.п.

В первых двух случаях осуществить систематизацию достаточно просто. Информация рассматривается в качестве зависимого от номера наблюдения или времени фактора. Тогда в первой колонке таблицы указывается номер наблюдения или время, а в последующих колонках – соответствующая этим наблюдениям количественная информация об объекте исследования. Точно также систематизируются данные по и другим указанным критериям.

Довольно часто на практике встречаются случаи, когда сведенная в таблицы информация оказывается неполной – часть данных отсутствует. Это может возникнуть, например, в случае, когда при опросе один из респондентов ответил не на все вопросы или в статистическом сборнике не приводятся данные за какой-либо год. В подобном случае возникает необходимость восстановления утерянной в процессе сбора и обработки наблюдений информации. Определить неизвестную величину внутри статистического ряда можно с помощью одного из методов интерполяции.

Теория интерполяции является одним из старейших разделов математики и начиналась она работами И.Ньютона, Ж.Лагранжа, Н.Абея, Ш.Эрмита и др. По определению интерполирование – это способ нахождения какой-либо величины по известным отдельным значениям этой же или других величин, связанных с ней¹.

Теория интерполяции является одним из наиболее разработанных разделов численных методов, и поэтому поставленная задача может быть решена с той или иной степенью точности.

Проще всего воспользоваться методом разностей, хотя это и не самый точный метод экстраполяции. Его суть заключается в следующем.

Первая производная функции, как известно, находится по формуле:

$$Y'_t = \lim_{x \rightarrow 0} \frac{\Delta Y_t}{\Delta X_t}. \quad (2.1.1)$$

Эта производная остаётся постоянной и не равной нулю, если между двумя переменными существует линейная функциональная зависимость. Если между двумя переменными существует линейная регрессионная зависимость, то в каждой конкретной точке наблюдения за переменными в момент t первая производная будет иметь значения, в общем случае отличающиеся от значений первой производной в другие моменты времени. Эти отклонения вызваны действием множества случайных факторов и поэтому значения первой производной в разных точках будут колебаться вокруг своего математического ожидания, лучшей оценкой которой в данном случае является средняя арифметическая.

Как вычислить первую производную регрессионной зависимости, если не известны коэффициенты линейной функции, которая описывает эту зависи-

¹ Янович Л.А. Интерполирование // Математическая энциклопедия, т.2. – М.: Советская Энциклопедия, 1979. – С.622.

мость? У исследователя имеются в распоряжении только эмпирические значения X_t и Y_t . Для решения поставленной задачи первую производную заменяют отношением конечных разностей:

$$Y'_t \approx \frac{\Delta Y_t}{\Delta X_t}. \quad (2.1.2)$$

Такая замена возможна только в том случае, когда приращения ΔX_t (конечные разности первого порядка) достаточно малы. Обычно с такой ситуацией и приходится иметь дело на практике. Поэтому легко найти первые разности ΔX_t и ΔY_t :

$$\Delta X_t = X_t - X_{t-1}, \quad \Delta Y_t = Y_t - Y_{t-1}, \quad (2.1.3)$$

а затем найти их отношение

$$\Delta Y_t / \Delta X_t = (Y_t - Y_{t-1}) / (X_t - X_{t-1}). \quad (2.1.4)$$

Если это отношение при разных значениях t действительно колеблется около некоторого значения, можно найти одну из оценок этого значения, а именно, среднюю арифметическую:

$$\bar{\Delta} = \frac{1}{n-1} \sum_{t=1}^n \frac{\Delta X_t}{\Delta Y_t}, \quad (2.1.5)$$

которая будет в среднем характеризовать первую производную зависимости, а значит, будет являться одной из оценок коэффициента линейной регрессии Y на X . С помощью этого коэффициента можно решать задачи интерполяции подобных линейных зависимостей. Однако точность такой интерполяции не очень высока, к тому же линейные зависимости встречаются крайне редко в практике прогнозирования социально-экономических процессов. Поэтому метод вычисления конечных разностей для целей интерполяции в настоящее время применяется на практике в простых случаях, но именно конечные разности легли в основу интерполяции методом полинома Ньютона.

Интерполяционная формула Ньютона применяется в том случае, когда упорядоченные значения X_t находятся на равном расстоянии друг от друга, то есть, когда $\Delta X_t = X_{t+1} - X_t = h = \text{const}$ для всех t . Константа h получила название шага наблюдений (шаг таблицы наблюдений). С учётом этого свойства значения функции двух переменных X_t и Y_t характеризуются только изменением переменной Y_t . Эти изменения можно определить, вычислив значения конечных разностей. Сами разности можно осуществлять с шагом назад, как это было сделано в (2.1.3), а можно делать и с шагом вперёд, как это предусмотрено методом Ньютона. При этом формулы для расчёта первых разностей будут иметь вид:

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

вторых разностей:

$$\Delta^2 Y_t = \Delta Y_{t+1} - \Delta Y_t;$$

третьих разностей:

$$\Delta^3 Y_t = \Delta^2 Y_{t+1} - \Delta^2 Y_t$$

и так далее для других конечных разностей. Этот процесс показан в таблице 2.2.

Таблица 2.2.
Конечные разности различных порядков

X_t	Y_t	ΔY_t	$\Delta^2 Y_t$	$\Delta^3 Y_t$	$\Delta^4 Y_t$	$\Delta^5 Y_t$
X_1	Y_1	ΔY_1	$\Delta^2 Y_1$	$\Delta^3 Y_1$	$\Delta^4 Y_1$	$\Delta^5 Y_1$
X_2	Y_2	ΔY_2	$\Delta^2 Y_2$	$\Delta^3 Y_2$	$\Delta^4 Y_2$	
X_3	Y_3	ΔY_3	$\Delta^2 Y_3$	$\Delta^3 Y_3$		
X_4	Y_4	ΔY_4	$\Delta^2 Y_4$			
X_5	Y_5	ΔY_5				
X_6	Y_6					

Легко убедиться в том, что если будет не шесть наблюдений, как это показано в табл. 2.2, а, например, восемь, то можно будет вычислить ещё и разности шестого и седьмого порядков; если будет n наблюдений, то можно вычислить разности $(n - 1)$ -го порядка. Разность каждого порядка некоторым образом характеризует производную степени, соответствующей данному порядку.

Так как значения X_t в рассматриваемом случае представляют собой арифметическую прогрессию, то, введя обозначение:

$$q = \frac{(X - X_1)}{h}, \quad (2.1.6)$$

получим интерполяционную формулу Ньютона на основе вычисленных значений конечных разностей:

$$P_n(X) = Y_1 + q\Delta Y_1 + \frac{q(q-1)}{2!} \Delta^2 Y_1 + \dots + \frac{q(q-1)\dots(q-n+1)}{n!} \Delta^n Y_1. \quad (2.1.7)$$

Подставляя в (2.1.6) известное значение X_k , легко найти q , подставляя которое в (2.1.7), вычисляется интерполируемое значение Y_k .

Формула (2.1.7) называется интерполяционной формулой Ньютона для интерполирования вперёд¹. Существует также формула Ньютона для интерполирования назад, которая использует разности, вычисленные по принципу (2.1.3).

В случаях, когда необходимо получить более точные результаты интерполяции, рекомендуется использовать и более сложные нелинейные интерполяционные формулы, в первую очередь, интерполяционные формулы Лагранжа и Ньютона. Методика интерполяций этими методами исходит из необходимости построения интерполирующей функции, проходящей через все точки X_t и Y_t . Подобной функцией является многочлен $(n-1)$ -ой степени, который, очевидно, пройдет через все n точек:

$$L(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_{n-1}x^{n-1}. \quad (2.1.8)$$

Однако достаточно часто построение подобных функций оказывается излишним, поскольку с подобной задачей успешно могут справиться и функции с более низкими степенями. Именно эту задачу решает методы вычисления ин-

¹ Самарин М.К. Ньютона интерполяционная формула // Математическая энциклопедия, т.3. – М.: «Советская энциклопедия», 1982. – С.1092.

терполяционного многочлена (полинома) Лагранжа, который рассчитывается по формуле¹:

$$L_n(x) = \sum_{i=1}^n \frac{(x-x_1)(x-x_2)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_1)(x_i-x_2)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} y_i. \quad (2.1.9)$$

Этот интерполяционный многочлен вычисляется для имеющихся пар значений, а затем, по известному значению X_k ($k \in t$) интерполируют значение Y_k , подставляя это значение X_k в формулу (2.1.9).

Каждая из интерполяционных формул (2.1.7) и (2.1.9) даёт при вычислении ошибки, которые при необходимости можно рассчитать и учесть в вычислениях². Однако, величина этих ошибок достаточно мала, поэтому их влиянием в практике прогнозирования пренебрегают, тем более, что эмпирические данные загрязнены другими ошибками, о которых писалось ранее.

К числу методов интерполирования относят также и интерполирование методами моделирования с использованием математической статистики – чаще всего с помощью МНК. Суть этого метода заключается в следующем. По имеющимся статистическим данным двух переменных X_t и Y_t строится регрессионная модель зависимости Y_t от X_t или наоборот. Затем, по известным данным X_k с помощью модели рассчитывается интерполируемая величина Y_k или наоборот.

С учётом того, что сегодня в распоряжении прогнозиста имеются разнообразные пакеты прикладных программ, реализованных применительно к персональным компьютерам, в которых вычисление оценок МНК представляет собой элементарную процедуру (например, Microsoft Office Excel), этот способ используют в практике прогнозирования довольно часто.

После того, как информация об объекте прогнозирования собрана, очищена по возможности от ошибок, систематизирована и, при необходимости, восстановлена, прогнозист становится обладателем данных, которые можно использовать для последующих вычислений.

¹ Кудрявцев Л.Д., Самарин М.К. Лагранжа интерполяционная формула // Математическая энциклопедия, т.3. – М.: «Советская энциклопедия», 1982. – С. 170-171.

² Хемминг Р.В. Численные методы для научных работников и инженеров. – М.: Наука, 1968. – С. 110–112.