*Известия Кабардино-Балкарского научного центра РАН № 3 (89) 2019*

МАТЕМАТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ ЭКОНОМИКИ

*УДК: 330.42(045), 51-77(045)*
*DOI: 10.35330/1991-6639-2019-3-89-48-77*

# ПОГРЕШНОСТЬ ЕМКОСТНОГО МЕТОДА АНАЛИЗА РЕДКИХ СОБЫТИЙ, УДАЛЕННОСТЬ ОТ КОНЕЧНОГО ПОТРЕБИТЕЛЯ*

**Ю.А. КОРАБЛЕВ**

Финансовый университет при Правительстве Российской Федерации
125993, г. Москва, Ленинградский проспект, 49
E-mail: academy@fa.ru

*Работ, посвященных методам исследования редких событий, крайне мало. Наиболее популярным методом анализа редких событий является теория случайных процессов, когда события представляются потоками Пуассона или Пальма. Другие методы обладают еще меньшей точностью и обоснованностью. Тем не менее теория случайных процессов способна определить не момент возникновения события, а только вероятность заданного количества событий за интервал времени фиксированной длины.*

*В работе описана методика исследования редких событий, которая основана на различии источников событий и восстановлении параметров предполагаемого процесса, лежащего в основе возникновения этих событий. После восстановления параметров процесса ищется закономерность любыми другими известными методами, после чего закономерности экстраполируются на будущее. После экстраполяции параметров процессов запускается сам процесс для получения прогноза моментов времени возникновения следующих событий.*

*Наиболее распространёнными в экономике являются процессы потребления, или расхода продукции, или накопления возмущения до определенного уровня. В этом случае источники событий можно моделировать как емкости. Параметром процесса является скорость опустошения этой емкости. Предложен метод восстановления этой скорости, после чего можно прогнозировать будущие события. Такой метод анализа и прогнозирования редких событий автор называет «емкостным» методом.*

*В статье проводится анализ влияния позиции в цепочке распространителей на точность восстановления исходной неизвестной функции скорости потребления продукции с помощью емкостного метода. Другой целью является нахождение величины относительной погрешности восстановления исходной неизвестной функции расхода продукции.*

*С помощью математического анализа рассматривается процесс потребления для цепочки распространителей, строится обратная задача, анализируется погрешность. В результате текущего исследования получены значения величины погрешности восстановления исходной зависимости при реализации продукции через одного посредника, а также при реализации продукции через двух последовательно расположенных посредников. Получены крайние значения интервалов для ошибки восстановления исходной зависимости. На конкретном численном примере подтверждена справедливость полученных формул. Показано, что ошибка не является систематической и что возрастание ошибки из-за удаленности от конечного потребителя при неизменности всех других факторов растет как сумма геометрически убывающей прогрессии. Рассчитаны значения дисперсии и среднего квадратичного отклонения для относительной ошибки, показано, что они растут очень медленно.*

**Ключевые слова:** редкие события, емкостный метод, скорость потребления, точность, ошибка, погрешность, дисперсия, последовательность распространителей, посредники.

## Введение

В экономике методы получения дополнительной информации [1] и построения прогнозов [2] играют чрезвычайно важную роль. Чем точнее окажется прогноз, тем больше может быть выгода лица, принимающего решение. Выгоду получают не только сами отдельные лица, руководители компаний или организаций, но и рядовые служащие этих компаний, которые напрямую не связаны с принятием важных решений, а также их родные и близкие. Принятие

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-010-00154

каждого решения влияет на стабильность, спокойствие, уверенность в завтрашнем дне, общественное настроение, на ситуацию в стране в целом, что во все времена являлось ключевым в исторических свершениях стран, их расцвете или разрушении.

Развитие и совершенствование методов выявления закономерностей и способностей прогнозировать будущие события или некоторые показатели является чрезвычайно важным как с практической, так и с теоретической точки зрения. Среди методов анализа и прогнозирования отдельно выделяют методы для работы с регулярными (частыми) и редкими событиями. Как правило, в математических методах под регулярными событиями понимается набор данных, где каждая порция данных указывает количество событий. Чаще всего это временной ряд. Для работы с регулярными событиями существует большое количество проверенных и зарекомендовавших себя методов: методы фильтрации и сглаживания [3], спектральный анализ [4], фазовый анализ [5, с. 240-247], методы аппроксимации и интерполяции [6], методы исследования временных рядов, факторный анализ [7, 8], адаптивные методы [5] и другие.

Редкие события в математических методах, как правило, представляются не количеством событий за каждый период времени, а длительностью времени между событиями. События возникают дискретно в разные моменты времени. Время между событиями может быть совершенно различным, где-то это могут быть годы или месяцы, где-то секунды или миллисекунды. Какое количество событий за единицу времени является редким, а какое уже нет, является диалектическим вопросом. Важным является само представление данных. Если редкие события представить в виде стандартного временного ряда, то такой временной ряд будет содержать множество нулей. Классические методы обработки и анализа данных к редким событиям не подходят. Анализ редких событий представляет большую сложность для аналитиков. Редкие события характерны своей нерегулярностью, отсутствием четкой закономерности между интервалами возникновения. Сами по себе редкие события сложно анализировать и прогнозировать.

В торговле для анализа редких продаж иногда применяют метод логистической регрессии [9], метод ближайшего соседа [10], метод Кростона [11, 12], метод бутстрэппинга (Виллемейна) [13, 14]. Иногда используются селективные методы [15], которые в зависимости от входных данных переключают модель работы с редкими событиями на одну из нескольких возможных (с выбором самого метода по значению ошибки прогноза на предыдущем шаге). В этих методах осуществляются определенные манипуляции над все теми же временными рядами, но которые содержат уже большое количество нулевых значений. Хорошо, если эти методы позволят хотя бы наполовину приблизиться к истинному значению, но на практике получается, что полученный прогноз очень сильно отличается от наблюдаемых значений. Иногда получается, что там, где прогнозировался спрос, спроса не было, там, где значение спроса должно было расти, упало в 10 раз и т.д. Более того, эти методы не выявляют закономерности, почему редкие события происходят, а лишь ограничиваются статистическим законом распределения.

Для анализа редких событий иногда используют теорию случайных процессов [16], когда анализируется процесс поступления событий, как правило, в виде Пуассоновского потока или обобщенного потока Эрганга, в более сложном варианте это может быть поток Пальма с ограниченным последействием. Иногда используют бета-распределение для моделирования редких событий [17], иногда вводят модифицированные Пуассоновские процессы для «сверхредких» событий [18]. Наиболее распространены классические Пуассоновские процессы. Так, в работе [19, 20, 21] на основе статистических данных редких продаж определяют параметры потока событий, после чего определяют размер собственных запасов. Зная характеристики потоков, можно определить вероятность возникновения события за определенный период времени. Однако сами моменты времени наступления событий не определяются. Вместе с тем не определяется и величина воздействия каждого события. В итоге компании вынуждены разрабатывать собственные эвристические методы для анализа редких продаж, состоящие из статистических математических методов и собственного опыта. До сих пор нет удовлетворительного математического инструментария, способного с приемлемой точностью и достоверностью выявлять соответствующие закономерности и предсказывать будущие редкие события.

Неужели совсем нет никакой информации о том, как возникают эти события? Почему процессы случайны? Откуда взялась эта случайность? Ведь случайность – это лишь мера неопределенности или отсутствия знания. Добавив некоторые знания и представления, можно избавится от большей степени существующей неопределенности.

В предлагаемом подходе автор как раз вносит дополнительную информацию, которая элиминирует большую часть неопределенности. Для этого предлагается следующая схема этапов.

1. Первым шагом мы предлагаем разделять источники событий, а не смешивать их. На самом деле причина огромных ошибок перечисленных ранее методов в том, что анализируются не исходные данные (в торговле – продаж), а агрегированные. Например, когда события (продажи) суммируются по временным интервалам от всех источников (покупателей), таким как дни, недели, месяцы, или когда потоки событий объединяются в один общий поток событий. После такого суммирования данные превращаются во временной ряд или в перемешанный поток событий и в дальнейшем анализируются. В результате такого агрегирования теряется много важной информации. Если каждое событие характеризуется своим источником, моментом времени и воздействием, то при агрегировании во время построения временного ряда получается единственное значение, относящееся к заданному интервалу времени. При получении агрегированного потока событий моменты времени и (иногда) величины воздействия этих событий остаются, а вот информация об источниках теряется. Ситуация с агрегированием данных от всех участников похожа на картину (рис. 1), когда следопыт вместо того, чтобы распознавать и отслеживать каждый оставленный след, начинает суммировать площадь следов на каждом метре и затем разбираться в полученной мешанине данных. В современную эпоху цифровых технологий обмена и хранения данных этот метод представляется крайне расточительным, теряется очень много информации.

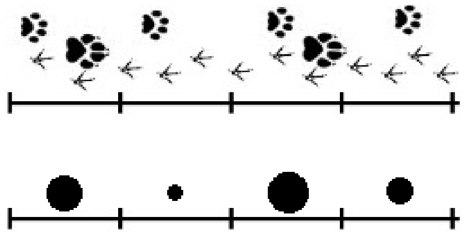

***Рис. 1.*** *Суммирование площади следов*

2. На следующем шаге надо использовать имеющуюся информацию о характере протекающего в источнике процесса, приводящего к возникновению событий. Наверное, это самый главный и ответственный этап. В случае если информации недостаточно, можно использовать предположение о процессе. Чем ближе предположение окажется к действительности, тем точнее будет результат в дальнейшем.

3. После того как процесс установлен, необходимо из имеющихся статистических данных восстановить параметры этого процесса, то есть выполнить регрессию параметров процесса из данных редких событий.

4. Когда значения параметров процесса восстановлены, в дальнейшем следует произвести поиск закономерностей в том, как эти параметры изменялись со временем или в зависимости от других внешних факторов. Цель этого этапа – произвести экстраполяцию параметров процесса на будущее. Для этого можно использовать любые другие известные математические или статистические методы.

5. Завершающим этапом будет запуск самого процесса с экстраполированными значениями параметров, определенными на предыдущем этапе. В результате этого этапа получается прогноз будущих событий, возникающих в конкретном источнике.

Такие действия необходимо выполнить с каждым источником. Затем, если в дальнейшем потребуется получить прогноз будущих событий в виде временного ряда, не составит труда просуммировать эти же события за выбранные интервалы времени. Схема этого метода анализа и прогнозирования редких событий представлена на рисунке ниже (рис. 2).

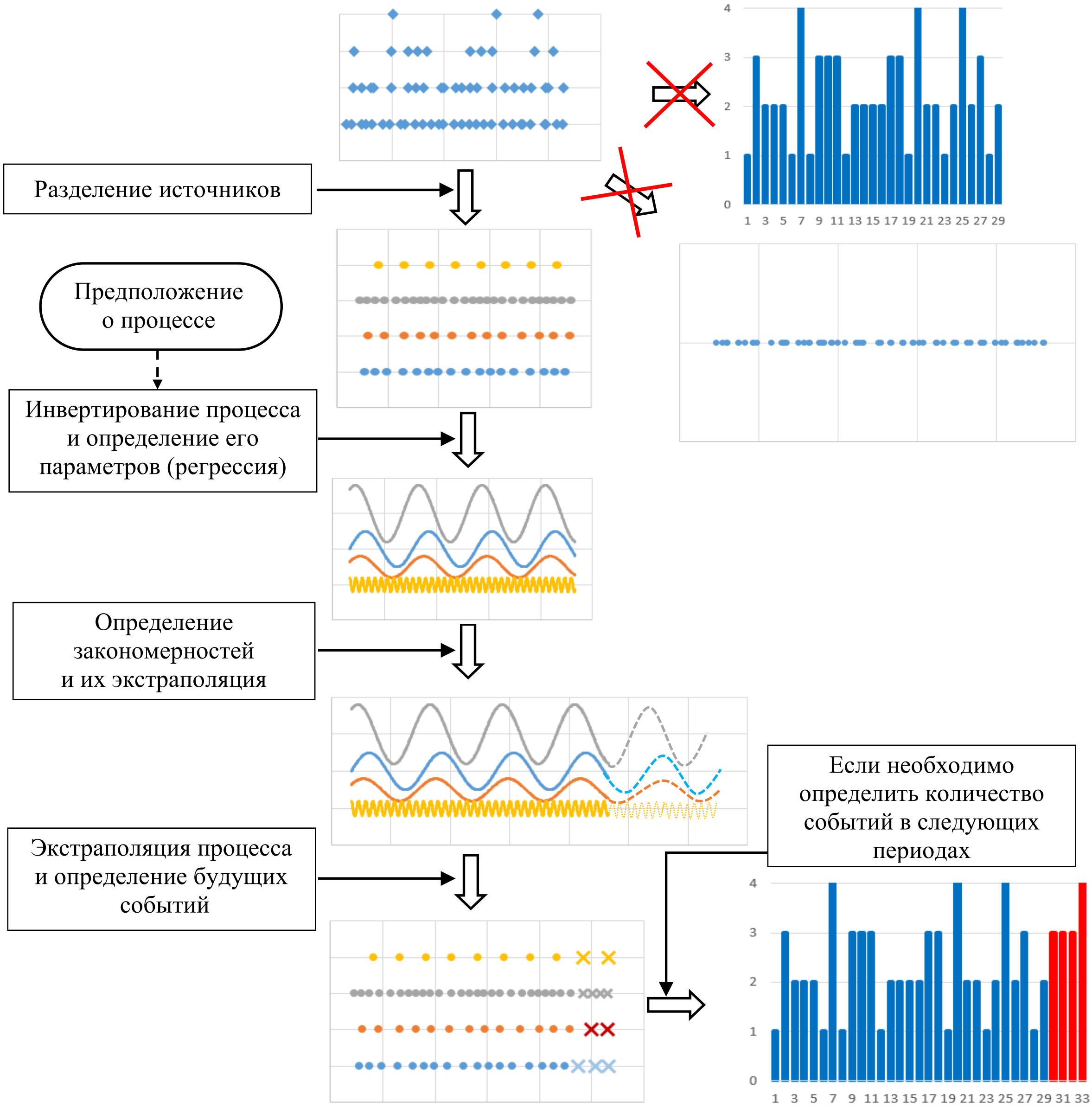


***Рис. 2.*** *Схема анализа и прогнозирования редких событий*

Необходимо отметить, что процесс, протекающий в каждом источнике, может быть полностью уникальным, а не только различаться параметрами. Но если информации об источнике недостаточно, можно по-прежнему предположить, что процесс в этом источнике соответствует некоторому ранее определенному процессу.

Процессов, протекающих в источниках, может быть большое множество, но в экономике, на наш взгляд, самым распространенным является процесс потребления и расхода продукции, когда запас ведет себя как опустошающаяся емкость, или процесс накопления воздействия до определенного уровня. В случае использования такого процесса, когда источники событий можно моделировать как емкости, предложенный метод анализа и прогнозирования редких событий мы называем «ёмкостным методом» [22-24]. Параметром процесса является нестационарная функция скорости расхода запаса или накопления воздействия $f(t)$, подлежащая определению. Такой функцией могут являться спрос от времени, индивидуальная скорость потребления продукции, интенсивность покупок у выбранного покупателя (источника, не путать со спросом или интенсивностью покупок у нас самих).

После определения вида процесса восстанавливаем его параметры, то есть функцию $f(t)$ скорости изменения запаса или накопления воздействия. Для этого инвертируем процесс потребления продукции, получаем обратную задачу к задаче управления запасами (алгоритм в минус первой степени), когда по имеющимся данным $(t_i, y_i)$ – моментам времени и величинам воздействия события (покупок) – определяется скорость воздействия $f(t)$. Для этого используем основное предположение – величина совершенного события $y_i$ есть интеграл от функции $f(t)$ за время от момента совершения этого события $t_i$ до момента совершения следующего события $t_{i+1}$ (рис. 3).

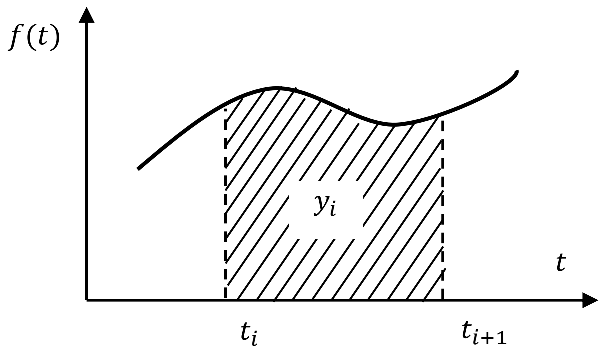


***Рис. 3.*** *Основное предположение*

С использованием этого предположения задача определения (регрессии) $f(t)$ представляет собой оптимизационную задачу восстановления неизвестной функции, для которой известна последовательность интегралов за известные непересекающиеся периоды времени:

$$\sum_{i=1}^{N}\left(y_i - \int_{t_i}^{t_{i+1}} f(t)dt\right)^2 + C\int_{t_1}^{t_n}\left(f''(t)\right)^2 dt \to \min,$$

где: второе слагаемое – штраф на нелинейность; $C$ – произвольная большая константа, влияющая на сглаживание; $n$ – номер последнего события.

Решение этой задачи должно осуществляться сложными математическими методами (данная задача автором пока не решена, но имеются некоторые соображения по этому поводу). Можно предложить упрощенный способ, где достаточно ограничиться рассмотрением лишь средней скорости $y_i/(t_{i+1} - t_i)$, с которой накапливается воздействие на каждом интервале, после чего сгладить полученную ступенчатую функцию любым известным методом для дальнейшего анализа. Последовательность средних скоростей будет представлять собой ступенчатую функцию (рис. 4).

Для получения гладкой функции можно воспользоваться кубическим регрессионным сплайном (со штрафной функцией на нелинейность) [25], который нивелирует все неровности и в то же время при построении учтет погрешности в исходных данных. В результате функция скорости расхода продукции $f(t)$ будет представлена набором полиномов третьей степени с непрерывной первой производной в точках сочленения, что приблизительно эквивалентно решению исходной задачи. Для практического использования этого будет достаточно. Причем по результатам моделирования точность восстановления упрощенным способом оказывается очень высокой (рис. 4), причем чем чаще происходят покупки, тем больше точность.

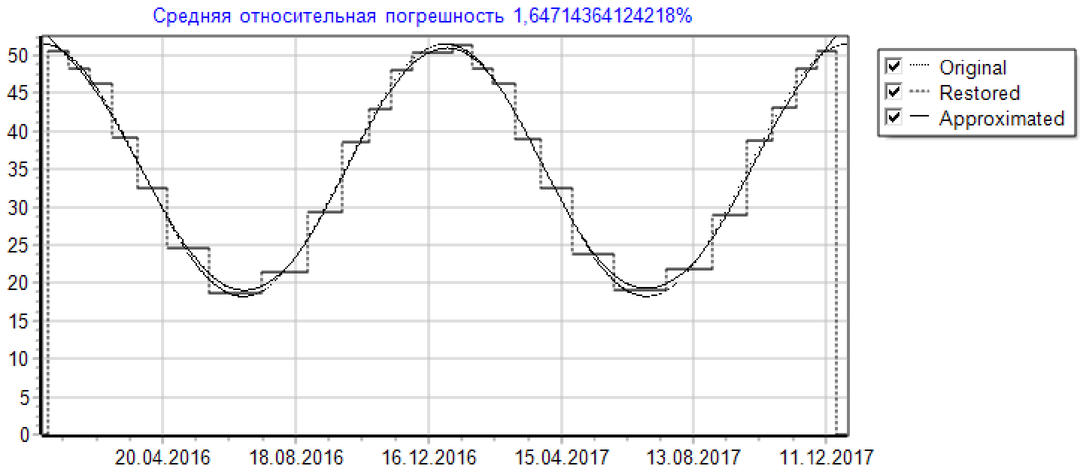


***Рис. 4.*** *Точность восстановления упрощенным способом*

Отдельно несколько слов надо сказать о системах управления запасами [26], в которых на первый взгляд может показаться есть что-то похожее. Во-первых, в системах управления запасами моделируется расход собственных запасов или запасов подконтрольных нам пунктов распространения, а в емкостном методе анализируется запас неподконтрольных нам покупателей. Во-вторых, в задачах управления запасами прогноз спроса уже задан в каждом подконтрольном пункте распространения, когда здесь, наоборот, происходит определение спроса покупателей (скорость расхода продукции, а затем можно определить дискретные моменты возникновения спроса каждого покупателя). В-третьих, используя данный метод, мы не управляем ничьими запасами, не считаем издержки и не определяем точку заказа. В-четвертых, нам не надо знать ни максимальную величину заказа, ни критическую точку запасов, ни величину страховых запасов. Предлагаемый метод является как бы обратным алгоритмом, алгоритмом в минус первой степени, по сравнению с системами управления запасами. Получается некий регрессионный анализ, цель которого – определить скорость расхода продукции, причем индивидуально по каждому нашему покупателю. С другой стороны, этот метод можно отнести к непараметрическим методам прогнозирования. Его можно использовать для любых видов редких событий, для которых выполняется основное предположение, не только для анализа продаж (например, социальных, политических и др. событий).

Подробнее об условии применения этого метода, предположениях о поведении покупателей и др. можно ознакомиться в работах [22-24]. В [23] дано математическое обоснование преимуществ этого метода по сравнению с классическим подходом, когда данные представляются в виде временного ряда и агрегируются за фиксированные интервалы времени.

Отдельный интерес представляет изучение вопросов точности, в каких случаях и на сколько падает точность восстановления исходной зависимости $f(t)$. Говорить будем в контексте анализа редких продаж в торговле. В этом исследовании определим, как изменяется точность в случае, если процесс распространения продукции происходит через цепочку распространителей, причем промежуточные распространители продукцию не потребляют, потребление происходит лишь конечными потребителями. Другими словами, у нас будет иметься серия посредников, которые реализуют товар, и товару необходимо пройти по цепочке от распространителя к распространителю прежде чем попасть к конечному потребителю, где он и будет расходован.

## Постановка задачи

Под определением точности тогда будем понимать нахождение ошибки восстановления исходной суммарной скорости расхода продукции. Причем истинная исходная функция на практике не известна, поэтому определение ошибки возможно только во время моделирования. С помощью имитационного моделирования были получены интересные результаты, для объяснения которых воспользуемся математическим моделированием.

При определении ошибки восстановления исходной функции отдельно надо сказать о сравниваемых величинах. Так как промежуточными распространителями потребление не осуществляется, будем сравнивать восстановленную функцию с функцией расхода продукции конечными потребителями. Однако в иерархической цепочке распространителей у каждого распространителя может быть по несколько покупателей (посредников, дочерних распространителей). У выбранного распространителя может быть несколько посредников, у тех в свою очередь также какое-то количество посредников и так далее, пока не дойдем до конечных потребителей. В результате у выбранного распространителя через серию посредников пополняет продукцию множество конечных потребителей. Выбранный распространитель имеет данные о покупках только своих непосредственных покупателей (посредников). Тогда восстанавливаемую функцию скорости расхода продукции из этих данных будем сравнивать с суммарной исходной функцией абсолютно всех конечных потребителей, которые через цепочку посредников покупают продукцию у распространителя, чьи данные покупок есть у выбранного распространителя. Пример такого моделирования изображен ниже (рис. 5). В этом примере у выбранного распространителя (у нас) есть посредник, у этого посредника (нашего покупателя) есть 3 конечных потребителя, скорость расхода продукции которых задана некоторой функцией, в данном примере это гармоническая функция, отличающаяся амплитудой, частотой и фазой. По данным продаж посреднику выбранный распространитель восстанавливает суммарную скорость расхода продукции всеми покупателями, которые потребляют товар через этого посредника. При увеличении длины цепочки (уровней) посредников увеличивается абсолютное значение суммарной скорости расхода продукции, тогда для оценки зависимости точности от позиции в цепочке распространителей нас будет интересовать значение относительной ошибки, причем усредненной по всему времени моделирования.

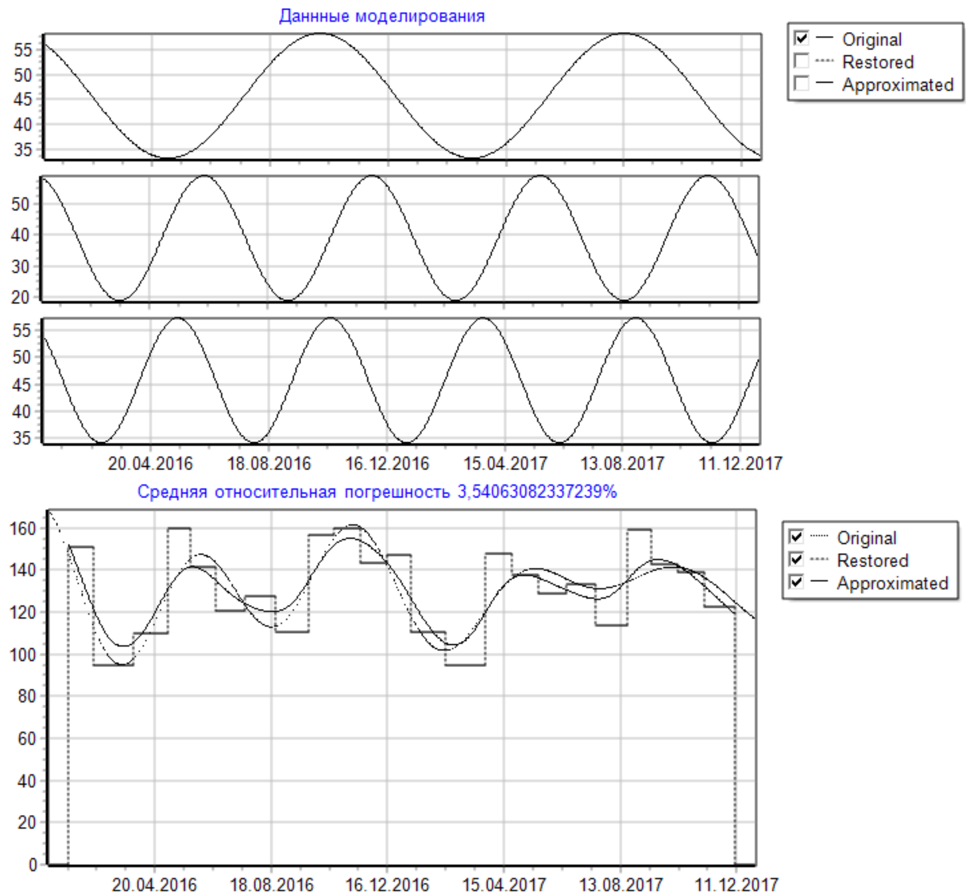


***Рис. 5.*** *Пример моделирования с одним уровнем посредников, сверху скорость расхода продукции 3 конечными потребителями, снизу восстановленная (упрощенным способом) функция из данных о покупках вышестоящего распространителя*

Подробно описывать этапы создания модели и процесса проведения моделирования не будем. Стоит лишь сказать, что при моделировании были приняты меры по обеспечению достоверности результатов моделирования таким образом, чтобы на ошибку влияла непосредственно только позиция в цепочке распространения, а не многие другие факторы. Более подробно об этом исследовании можно прочитать в работе [24]. Приведем лишь результаты исследования, которые показали, что для распространителей, удаленных на несколько позиций от конечного потребителя, точность метода падает незначительно при удалении от конечного потребителя. Результаты исследования были следующими (табл. 1, 2, 3).

***Таблица 1***

ДАННЫЕ 20 ПРОГОНОВ ДЛЯ РАСПРОСТРАНИТЕЛЯ С 3 ПОКУПАТЕЛЯМИ У КАЖДОГО

| Номер прогона | Среднее относительное отклонение в % | | | |
|---|---|---|---|---|
| | 0 посредников | 1 посредник | 2 посредника | 3 посредника |
| 1 | ~1-3% | 6,0589 | 8,1665 | 7,7042 |
| 2 | | 4,4492 | 7,0122 | 8,5183 |
| 3 | | 3,8206 | 12,1525 | 8,9015 |
| 4 | | 9,6343 | 6,9123 | 10,9366 |
| 5 | | 5,3771 | 11,1905 | 9,6241 |
| 6 | | 4,3904 | 8,2253 | 10,9213 |
| 7 | | 5,7136 | 9,2977 | 8,6745 |
| 8 | | 6,018 | 8,3653 | 8,1735 |
| 9 | | 8,688 | 7,0874 | 7,0018 |
| 10 | | 12,2289 | 8,7743 | 4,1247 |
| 11 | | 9,7904 | 8,2171 | 9,5669 |
| 12 | | 6,846 | 12,9519 | 6,3553 |
| 13 | | 3,3225 | 11,818 | 7,9162 |
| 14 | | 7,6846 | 8,6079 | 6,6991 |
| 15 | | 3,1397 | 6,8984 | 9,7203 |
| 16 | | 4,9831 | 8,3273 | 10,0206 |
| 17 | | 7,616 | 4,0429 | 7,1258 |
| 18 | | 10,428 | 7,8877 | 7,1145 |
| 19 | | 7,3258 | 4,9267 | 10,2031 |
| 20 | | 4,1924 | 8,0084 | 8,8516 |

***Таблица 2***

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ ДЛЯ СХЕМЫ РАСПРОСТРАНИТЕЛЕЙ С 3 ПОКУПАТЕЛЯМИ У КАЖДОГО

| | 1 посредник | 2 посредника | 3 посредника |
|---|---|---|---|
| Среднее значение относительного отклонения по 20 прогонам, в % | **6,5854** | **8,3305** | **8,3606** |
| Корень выборочной дисперсии относительного отклонения по 20 прогонам, в % | **2,5562** | **2,2392** | **1,6748** |

***Таблица 3***

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ ДЛЯ СХЕМЫ РАСПРОСТРАНИТЕЛЕЙ С 5 ПОКУПАТЕЛЯМИ У КАЖДОГО

| | 1 посредник | 2 посредника | 3 посредника |
|---|---|---|---|
| Среднее значение относительного отклонения по 20 прогонам | **4,2889** | **5,6544** | **6,0386** |
| Корень выборочной дисперсии относительного отклонения по 20 прогонам | **1,0284** | **1,3967** | **1,2145** |

По 20 прогонам видим, что точность падает, но падает незначительно. Разброс точности относительно среднего значения достаточно велик, поэтому полной уверенности в том, что погрешность растет именно таким образом, нет.

В этой работе исследуется данный феномен. Требуется определить и объяснить погрешность восстановления исходной функции скорости потребления продукции покупателей в зависимости от позиции в цепочке распространителей.

## Один промежуточный распространитель

Итак, пусть расход продукции некоторого покупателя $j$ задан функцией $f^j(t)$, когда запас продукции подходит к концу, покупатели совершают покупки у распространителя верхнего уровня. В свою очередь у распространителя верхнего уровня заканчивается запас продукции, и он также пополняет запас, совершая покупки у распространителя, находящегося на уровень выше. Пусть есть 3 покупателя (посредника), которые покупают продукцию у одного распространителя (верхнего уровня) (рис. 6).

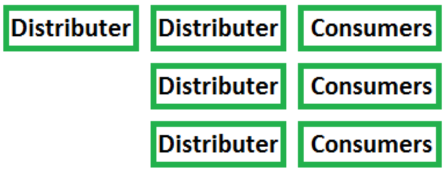


***Рис. 6.*** *Один уровень посредников до конечного потребителя*

Имеются данные наблюдений $y_i^j$, $t_i^j$, которые не связаны друг с другом. Пусть первый раз запас продукции верхнего распространителя заканчивается в момент $t_{k-1}^{\mathrm{II}}$, когда к нему пришел второй покупатель в момент времени $t_{i-1}^2$. Затем покупатели совершали другие покупки, и пусть очередной раз запас закончился в момент времени $t_k^{\mathrm{II}}$, когда пришел первый покупатель в момент времени $t_i^1$. И так далее. Также имеются данные об объеме совершенной покупки $y_k^{\mathrm{II}}$ в моменты времени $t_k^{\mathrm{II}}$. Попробуем определить погрешность определения суммарной скорости потребления продукции $\sum_j f^j(t)$ по отметкам $y_k^{\mathrm{II}}$, $t_k^{\mathrm{II}}$. Изобразим процесс пополнения продукции на рисунке (рис. 7). Нас не интересует уровень запасов посредников и самого распространителя как функция от времени, поэтому привычных пилообразных графиков из моделей управления запасами мы строить не будем, иначе это только усложнит восприятие. На рисунке стрелки показывают покупки распространителей, когда у тех заканчивается продукция, жирными стрелками показаны покупки, которые приводят к тому, что у верхнего распространителя (в самом низу рисунка) заканчивается продукция, и он пополняет ее, делая покупки $y_k^{\mathrm{II}}$, $t_k^{\mathrm{II}}$ уже у своего вышестоящего распространителя. Каждая стрелка обозначает покупку, причем объем покупки по основному предположению соответствует интегралу скорости расхода продукции от времени покупки до времени следующей покупки (см. рис. 3). Такое представление более информативно для объяснения величины ошибки нежели пилообразные картины из систем управления запасами, здесь площадь под графиком функции будет показывать величину расходуемой продукции за определенной период.

Объема сделанной покупки $y_k^{\mathrm{II}}$ хватает на все покупки, совершенные покупателями (посредниками) в период с $t_k^{\mathrm{II}}$ до $t_{k+1}^{\mathrm{II}}$ (рис. 8). На рисунке овалами обведены покупки $y_i^j$, на которые хватает запаса верхнего распространителя $y_k^{\mathrm{II}}$, сделанного в момент времени

$t_k^{\mathrm{II}}$. Причем покупка $y_i^1$ не включается, так как она расходует запасы $y_{k-1}^{\mathrm{II}}$, сделанные в предшествующий период времени $t_{k-1}^{\mathrm{II}}$ (пока не рассматриваем необходимость использования страховых запасов). В то же время покупка $y_{i+5}^2$, сделанная в момент времени $t_{i+5}^2 = t_{k+1}^{\mathrm{II}}$, включается, так как именно она исчерпывает запасы распространителя (пока без использования страховых запасов) и заставляет его делать следующую покупку $y_{k+1}^{\mathrm{II}}$ для восстановления своих запасов.

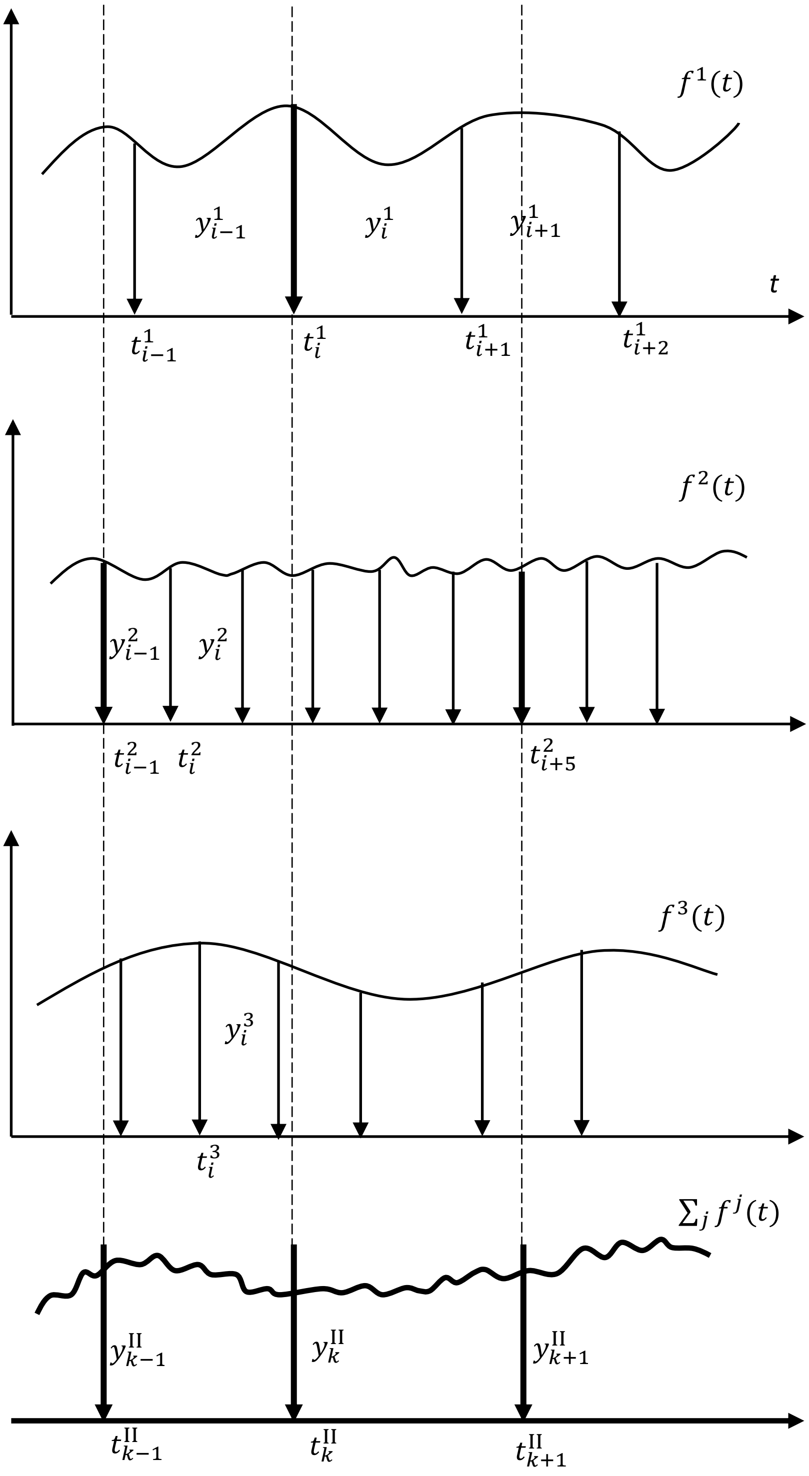


***Рис. 7.*** *Пополнение продукции с одним уровнем промежуточных посредников*

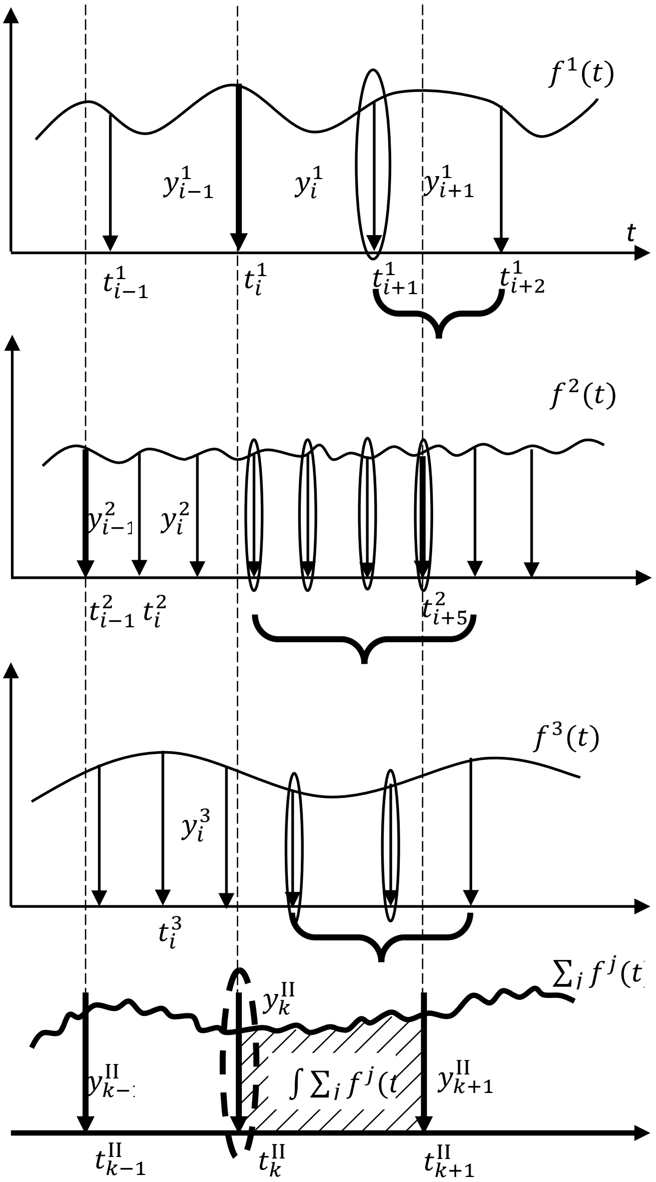


***Рис. 8.*** *Запаса $y_k^{II}$ хватает на обозначенные покупки*

**1.** ***Для начала будем предполагать***, что объема продукции $y_k^{\mathrm{II}}$ хватает на все покупки без необходимости использования страхового запаса, а потом усложним и введем страховой запас. Если объема запаса $y_k^{\mathrm{II}}$ точно хватает на все покупки, включая последнюю (пока без использования страхового запаса), то можно записать:

$$y_k^{\mathrm{II}} = \sum_j \sum_{\substack{i \\ i:\, t_k^{\mathrm{II}} < t_i^j \le t_{k+1}^{\mathrm{II}}}} y_i^j$$

Сравним наблюдаемый объем покупки $y_k^{\mathrm{II}}$ с расходуемым $\int_{t_k^{\mathrm{II}}}^{t_{k+1}^{\mathrm{II}}} \sum_j f^j(t)\,dt$ – объемом потребления за время от $t_k^{\mathrm{II}}$ до $t_{k+1}^{\mathrm{II}}$. Еще раз вспомним основное предположение (рис. 2), объем совершенной покупки $y_i^j$ есть интеграл от функции $f^j(t)$ скорости потребления за время от момента совершения этой покупки $t_i^j$ до момента времени совершения следующей покупки $t_{i+1}^j$ ($y_i^j = \int_{t_i^{\mathrm{j}}}^{t_{i+1}^{\mathrm{j}}} f^j(t)dt$). Так как $\int_a^b f^j(t)dt + \int_b^c f^j(t)dt = \int_a^c f^j(t)dt$, то

$$y_k^{\mathrm{II}} = \sum_j \sum_{\substack{i \\ i:\, t_k^{\mathrm{II}} < t_i^j \le t_{k+1}^{\mathrm{II}}}} \int_{t_i^{\mathrm{j}}}^{t_{i+1}^{\mathrm{j}}} f^j(t)dt = \sum_j \int_{\substack{t_{\min(i)}^{\mathrm{j}} \\ i:\, t_k^{\mathrm{II}} < t_i^j \le t_{k+1}^{\mathrm{II}}}}^{t_{\max(i)+1}^{\mathrm{j}}} f^j(t)dt.$$

Фигурными скобками на рисунке 8 указаны временные промежутки, за которые происходит расходование продукции $y_k^{\mathrm{II}}$.

В итоге видим (рис. 9), что наблюдаемый из данных продаж объем покупки $y_k^{\mathrm{II}}$ занижен на $\sum_j S_k^j$ и превышен на $\sum_j S_{k+1}^j$ по сравнению с ненаблюдаемым расходуемым объемом продукции $\int_{t_k^{\mathrm{II}}}^{t_{k+1}^{\mathrm{II}}} \sum_j f^j(t)\,dt$ на интервале от $t_k^{\mathrm{II}}$ до $t_{k+1}^{\mathrm{II}}$, где величина $S_k^j$ показывает значение объема потребляемой продукции, относящегося к предшествующей покупке:

$$S_k^j = \int_{t_k^{\mathrm{II}}}^{t_{\min(i)}^j,\ t_i^j > t_k^{\mathrm{II}}} f^j(t)dt.$$

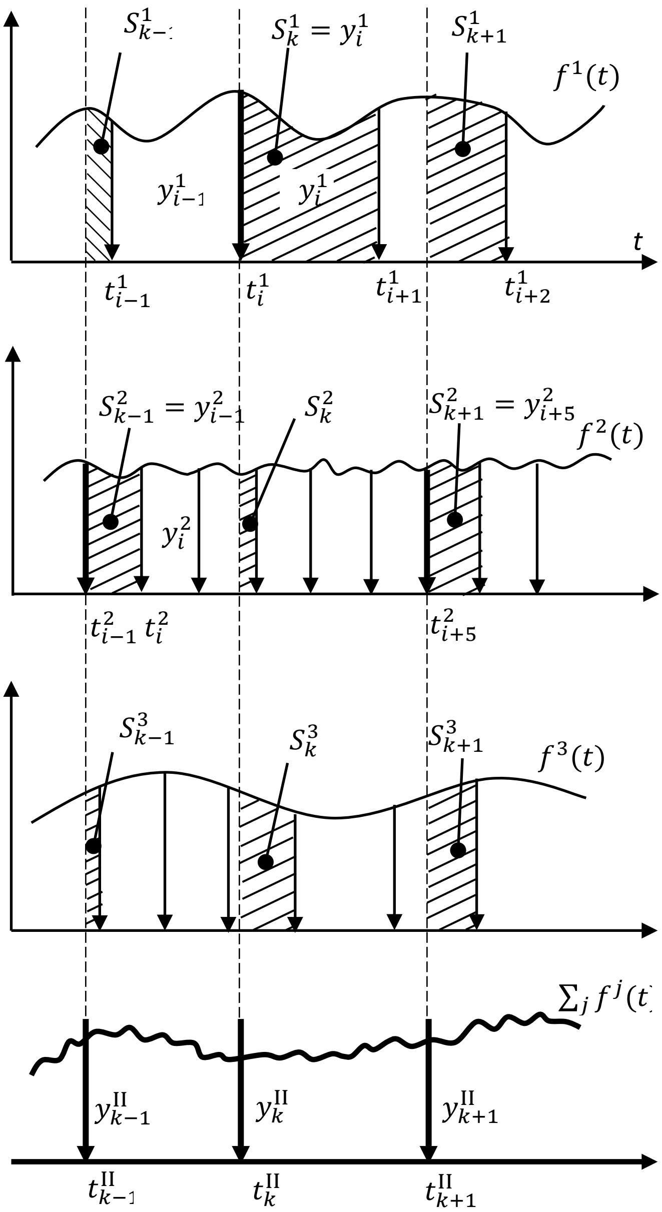


***Рис. 9.*** *Объем $y_k^{II}$ занижен на $\sum_j S_k^j$, и превышен на $\sum_j S_{k+1}^j$*

Величина несоответствия $\Delta Q_S$ между $y_k^{\mathrm{II}}$ и $\int_{t_k^{\mathrm{II}}}^{t_{k+1}^{\mathrm{II}}} \sum_j f^j(t)\,dt$ будет

$$\Delta Q_S = \sum_j S_{k+1}^j - \sum_j S_k^j = \sum_j \left(S_{k+1}^j - S_k^j\right).$$

Оценим величины $S_k^j$. Можно увидеть, что

$$0 < S_k^j \le y^j_{\substack{\min(i)-1 \\ i:\, t_k^{\mathrm{II}} < t_i^j}} = y^j_{\text{предш.}k}.$$

Величина $y^j_{\text{предш.}k}$ показывает объем предшествующей покупки покупателя $j$ относительно момента времени $t_k^{\mathrm{II}}$, причем если запас верхнего распространителя заканчивается как раз из-за покупки в момент $t_i^j = t_k^{\mathrm{II}}$, то $S_k^j = y^j_{\text{предш.}k} = y_i^j$. Есть основания полагать, что величина $S_k^j$ распределена в интервале $(0, y^j_{\text{предш.}k}]$ по равномерному закону. Тогда среднее значение можно грубо посчитать как

$$\overline{S_k^j} \cong \frac{y^j_{\substack{\min(i)-1 \\ i:\, t_k^{\mathrm{II}} < t_i^j}}}{2} = \frac{y^j_{\text{предш.}k}}{2}.$$

Величина несоответствия $\Delta Q_S$ между $y_k^{\mathrm{II}}$ и $\int_{t_k^{\mathrm{II}}}^{t_{k+1}^{\mathrm{II}}} \sum_j f^j(t)\,dt$ лежит в диапазоне

$$\Delta Q_S \in \left(-\sum_j y^j_{\text{предш.}k}, \sum_j y^j_{\text{предш.}k+1}\right).$$

Стоит сказать, что только в самом худшем из худших случаев возможно, что $\Delta Q_S$ будет близко к границам этого интервала, ведь в начале по каждому покупателю $j$ величины $S_{k+1}^j$ и $S_k^j$ могут компенсировать друг друга, а потом, когда складываются по $j$, могут еще раз компенсировать друг друга. Если $S_k^j$ равномерные случайные величины, то $\Delta Q_S = \sum_j (S_{k+1}^j - S_k^j)$ приближается к нормальному распределению (есть основания полагать, что линейная корреляция между $S_{k+1}^j$ и $S_k^j$ отсутствует).

Стоит сказать, что в случае, когда покупатели не меняют свой максимальный запас продукции, т.е. когда величина покупок $y_i^j$ постоянна (пока не считая восполнение страхового запаса), то никакой дополнительной систематической ошибки возникать не будет, так как среднее значение несоответствия будет равно нулю $\overline{\Delta Q_S} = \sum_j (\overline{S_{k+1}^j} - \overline{S_k^j}) = \frac{1}{2}\sum_j (y^j_{\text{предш.}k+1} - y^j_{\text{предш.}k}) = 0$. Но разброс относительно нуля этого несоответствия все равно останется, причем приблизительно по нормальному закону распределения.

Тут стоит сделать несколько замечаний относительно расчета таких характеристик, как $S_k^j = \int_{t_k^{\mathrm{II}}}^{t^j_{\min(i)},\ t_i^j > t_k^{\mathrm{II}}} f^j(t)dt$. Во-первых эту характеристику можно наблюдать только при моделировании, ведь истинных значений функции расхода продукции $f^j(t)$ и запасов продукции покупателей знать невозможно. Во-вторых, при моделировании надо учесть также дискретность исчисления времени, то есть на самом деле запас должен закончится чуть раньше на время $\varepsilon_i^j$ от полной единицы времени, когда происходит проверка и выявляется необходимость пополнения запаса (моделирование проводится по дням, а не миллисекундам). В этом случае каждый покупатель использует собственные страховые запасы $SS_i^j = \int_{t_i^j-\varepsilon_i^j}^{t_i^j} f^j(t)dt$, которые приводят к тому, что величина $S_k^j$ также должна считаться за интервал времени на $\varepsilon_i^j$ короче, так как $SS_i^j$ включаются в наблюдаемую покупку $y_i^j$. Подробнее об этом было рассказано в [27], картина будет похожа на ту, что описана ниже, но относительно торговли посредников и конечных потребителей.

**2.** ***Теперь проанализируем случай использования страховых запасов***, когда запасов $y_k^{\text{II}}$ распространителя верхнего уровня не хватает точь в точь на все покупки и приходится использовать страховые запасы. В моменты $t_k^{\text{II}}$ и $t_{k+1}^{\text{II}}$, когда приходили покупатель 1 и 2, был превышен критический уровень запаса, и запасы брались из страховых запасов (рис. 10).

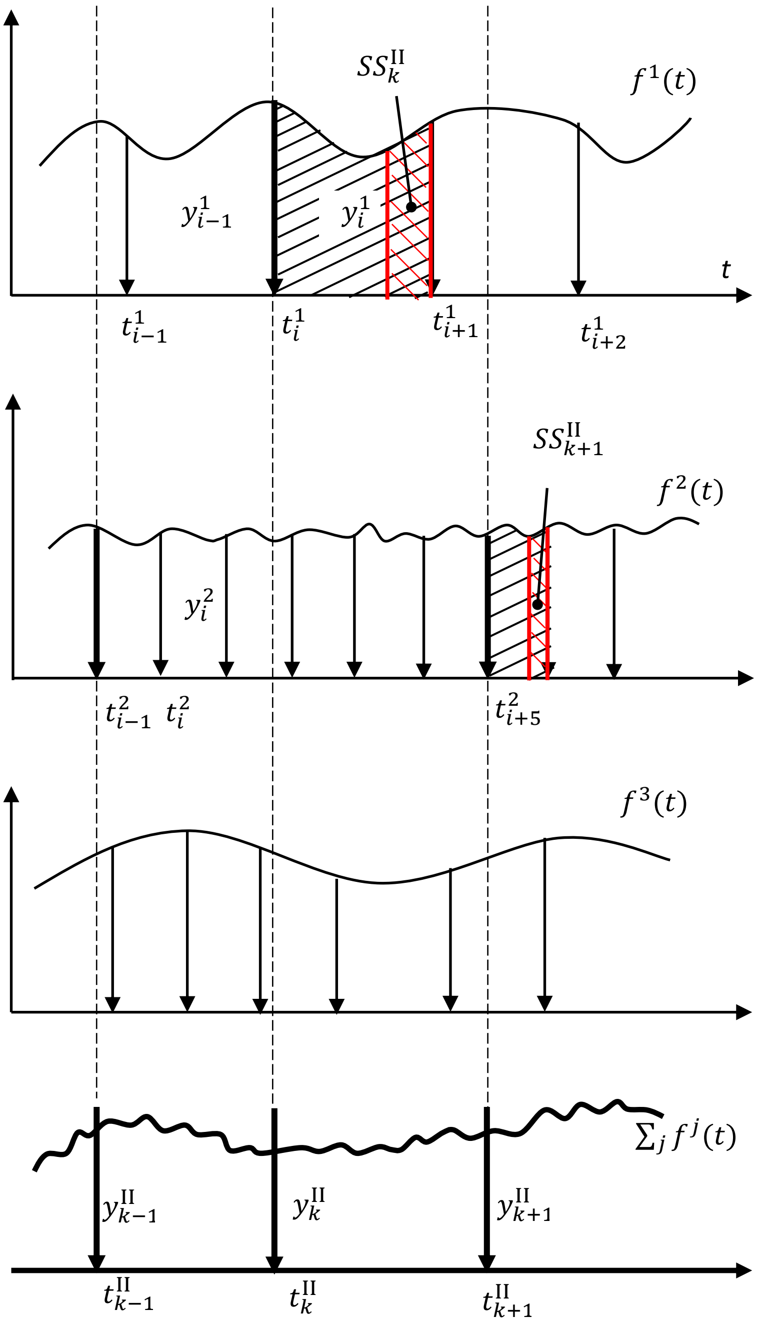


***Рис. 10.*** *Использование страховых запасов*

После того как первый покупатель совершил покупку $y_i^1$ в момент времени $t_i^1 = t_k^{\text{II}}$, наш распространитель выдал часть товара из запаса, затем запас закончился (превышен критический уровень), часть товара $SS_k^{\text{II}}$ он взял из страхового запаса. После этой сделки, так как критический уровень запаса был превышен, распространитель совершил покупку $y_k^{\text{II}} + SS_k^{\text{II}}$ (или сделал заказ) у своего поставщика, из которой часть продукции опять отложилась в страховой запас $SS_k^{\text{II}}$. Но при этом дальше расходовалась только продукция $y_k^{\text{II}}$, то есть наблюдаем превышение на $SS_k^{\text{II}}$.

Когда в момент времени $t_{k+1}^{\text{II}}$приходит второй покупатель и совершает покупку в момент $t_{i+5}^2 = t_{k+1}^{\text{II}}$, когда эта покупка приводит к израсходованию запаса, часть продукции берется из запасов, а часть $SS_{k+1}^{\text{II}}$ берется из страховых запасов. Общий расход продукции за наблюдаемый интервал времени от $t_k^{\text{II}}$ до $t_{k+1}^{\text{II}}$ составит $y_k^{\text{II}} + SS_{k+1}^{\text{II}}$.

Имеем дело с несовпадением наблюдаемого сделанного запаса продукции с расходуемым:

$$y_k^{\text{II}} + SS_k^{\text{II}} \neq y_k^{\text{II}} + SS_{k+1}^{\text{II}}.$$

Можно показать, что

$$0 < SS_k^{\text{II}} \leq \max_j y_{\text{предш}.k}^j.$$

Есть основания предполагать, что величина $SS_k^{II}$ распределена приблизительно равномерно по интервалу $(0; \max_j y_{\text{предш}.k}^j]$.

Среднее значение можно приблизительно посчитать как

$$\overline{SS_k^{\text{II}}} \cong \frac{\sum_{j=1}^n y_{\text{предш}.k}^j}{2n},$$

где $n$ – количество покупателей (в рассматриваемом примере 3).

Несовпадение наблюдаемого и расходуемого объема продукции на интервале времени от $t_k^{\text{II}}$ до $t_{k+1}^{\text{II}}$ из-за необходимости использовать страховые запасы будет:

$$\Delta Q^{SS} = SS_{k+1}^{\text{II}} - SS_k^{\text{II}} \in \left(-\max_j y_{\text{предш}.k}^j, \max_j y_{\text{предш}.k+1}^j\right).$$

Аналогично предыдущему случаю можно сказать, что если покупатель не склонен изменять объем заказа продукции, то систематической ошибки возникать не будет, разброс будет относительно нуля $\overline{\Delta Q^{SS}} = 0$. Причем если величина $SS_k^{\text{II}}$ распределена равномерно, то так как в образовании несовпадения $\Delta Q^{SS}$ участвуют только две величины, распределение погрешности из-за использования страховых запасов будет подчинено треугольному распределению.

В итоге имеем два вида погрешности:

1) $\Delta Q^S$ – из-за того, что часть покупок идет в зачет другого периода времени;
2) $\Delta Q^{SS}$ – из-за того, что приходится использовать страховые запасы.

Причем эти два вида погрешности обязательно вычитаются друг из друга, так как объем покупки $y_k^{\text{II}}$ занижен на $S_k^j$ и завышен на $SS_k^{\text{II}}$. Это приводит к тому, что общая погрешность $\Delta Q$ будет:

$$\Delta Q = \Delta Q^S - \Delta Q^{SS}$$

## Численный пример

Рассмотрим конкретный числовой пример, где один распространитель продает продукцию трем другим более мелким распространителям, которые работают напрямую с конечными покупателями. На рисунке 11 сплошной гладкой линией изображена суммарная исходная функция расхода продукции, ступенчатой функцией изображена восстановленная функция (упрощенным способом), пунктирными горизонтальными линиями указана средняя истинная скорость расхода продукции.


***Рис. 11.*** *Пример расчета погрешности с одним уровнем промежуточных распространителей*

У распространителя имеется информация (табл. 4) только о времени и объеме покупок, столбцы 2 и 3. Общее исходное потребление за время между покупками представлено в столбце 4. Во время моделирования также можно узнать величину используемых страховых запасов, столбец 5, а также величину оставшейся части покупок, которая относится к предыдущему периоду.

***Таблица 4***

Результаты численного примера

| Индекс | Дата $t_k^{\mathrm{II}}$ | Покупка $y_k^{\mathrm{II}}$ | $\int_{t_k^{\mathrm{II}}}^{t_{k+1}^{\mathrm{II}}} \sum_j f^j(t)\,dt$ | $SS_k^{\mathrm{II}}$ | $\sum_j S_k^j$ |
|---|---|---|---|---|---|
| $k-1$ | 13.10.2016 | 4336,58 | 4811,05 | 736,59 | 1784,22 |
| $k$ | 26.11.2016 | 4168,35 | 3506,35 | 568,35 | 1141,51 |
| $k+1$ | 24.12.2016 | 3701,08 | 3696,03 | 101,08 | 1336,24 |
| $k+2$ | 22.01.2017 | 4123,84 | 4470,85 | 523,84 | 1764,05 |
| $k+3$ | 01.03.2017 | | | 762,38 | 1655,59 |

При моделировании наблюдается определенная сложность со сбором данных $S_k^j$, так как посредники тоже используют свои страховые запасы. Для этого надо с момента времени последней покупки $t_k^{\mathrm{II}}$ суммировать $f^j(t)$ до момента времени $t_{\min(i)}^j$, $t_i^j > t_k^{\mathrm{II}}$ (следующей покупки покупателя $j$). После этого от данной суммы надо отнять величину страховых запасов, которые использовал покупатель $j$ перед тем как совершил покупку

$y^j_{\min(i)}$, $t^j_i > t^{\mathrm{II}}_k$ для обслуживания уже своих покупателей (в нашем случае конечных потребителей, потребление которых задается функцией $f^j(t)$).

**Для интервала $k-1$** имеем несоответствие $\Delta Q$ в объемах купленной и израсходованной продукции, наблюдаем средний расход $\bar{f}(t)$ оригинальной и $\tilde{f}(t)$ восстановленной функции (упрощенным способом):

$$\Delta Q = y^{\mathrm{II}}_{k-1} - \int_{t^{\mathrm{II}}_{k-1}}^{t^{\mathrm{II}}_k} \sum_j f^j(t)\,dt = 4336.58 - 4811.05 = -474.47$$

$$\bar{f}(t) = \frac{\int_{t^{\mathrm{II}}_{k-1}}^{t^{\mathrm{II}}_k} \sum_j f^j(t)\,dt}{t^{\mathrm{II}}_k - t^{\mathrm{II}}_{k-1}} = \frac{4811.05}{44} = 109.34$$

$$\tilde{f}(t) = \frac{y^{\mathrm{II}}_{k-1}}{t^{\mathrm{II}}_k - t^{\mathrm{II}}_{k-1}} = \frac{4336.58}{44} = 98.56.$$

Наблюдаем занижение на 10.78, на графике также видим соответствующую разницу (ступенчатая функция лежит ниже пунктирной, которая усредняет оригинальную функцию).

Воспользуемся формулой:

$$\Delta Q^S = \sum_j \left(S^j_k - S^j_{k-1}\right) = 1141.51 - 1784.22 = -\,642.71$$

$$\Delta Q^{SS} = SS^{\mathrm{II}}_k - SS^{\mathrm{II}}_{k-1} = 568.35 - 736.59 = -168.24$$

$$\Delta Q = \Delta Q^S - \Delta Q^{SS} = -642.71 - (-168.24) = -474.47.$$

Что полностью совпадает с разницей наблюдаемой покупки и интегралом от суммарного потребления на временном участке $k-1$.

Проделаем аналогичные расчеты **для интервала $k$**:

$$\Delta Q = y^{\mathrm{II}}_k - \int_{t^{\mathrm{II}}_k}^{t^{\mathrm{II}}_{k+1}} \sum_j f^j(t)\,dt = 4168.35 - 3506.35 = 662$$

$$\bar{f}(t) = \frac{\int_{t^{\mathrm{II}}_k}^{t^{\mathrm{II}}_{k+1}} \sum_j f^j(t)\,dt}{t^{\mathrm{II}}_{k+1} - t^{\mathrm{II}}_k} = \frac{3506.35}{28} = 125.23$$

$$\tilde{f}(t) = \frac{y^{\mathrm{II}}_k}{t^{\mathrm{II}}_{k+1} - t^{\mathrm{II}}_k} = \frac{4168.35}{28} = 148.87$$

Наблюдаем превышение на 23.64, на графике также видим соответствующую разницу (ступенчатая функция лежит выше пунктирной, которая усредняет оригинальную функцию).

Воспользуемся формулой

$$\Delta Q^S = \sum_j \left(S^j_{k+1} - S^j_k\right) = 1336.24 - 1141.51 = 194.73$$

$$\Delta Q^{SS} = SS^{\mathrm{II}}_{k+1} - SS^{\mathrm{II}}_k = 101.08 - 568.35 = -467.27$$

$$\Delta Q = \Delta Q^{S} - \Delta Q^{SS} = 194.73 - (-467.27) = 662.$$

Что также полностью совпадает с разницей наблюдаемой покупки и интегралом от суммарного потребления на временном участке $k$.

Для остальных интервалов $k+1,\ k+2$ значения расчетов $\Delta Q$, полученных по разнице между покупкой и интегралом и по формулам, также будут полностью совпадать, можете в этом убедиться самостоятельно.

Этот пример показывает справедливость полученных нами формул, поясняет, откуда берется разница между восстановленной и оригинальной суммарной функцией скорости потребления. Однако мы еще не ответили на главный вопрос, почему и как с удалением распространителя от конечного потребителя растет относительная погрешность.

### Два промежуточных распространителя

Проведем похожие рассуждения относительно распространителя, у которого два уровня посредников до конечного потребителя. Сравним наблюдаемый объем покупки $y_L^{\mathrm{III}}$ с $\int_{t_L^{\mathrm{III}}}^{t_{L+1}^{\mathrm{III}}} \sum_r \sum_j f^{r,j}(t)\,dt$ – суммарным объемом потребления за время от $t_L^{\mathrm{III}}$ до $t_{L+1}^{\mathrm{III}}$ ($r$ – количество покупателей у распространителя).

На рисунке 12 в самом низу изображены покупки распространителя $y_L^{\mathrm{III}}$ и $y_{L+1}^{\mathrm{III}}$ в моменты времени $t_L^{\mathrm{III}}$ и $t_{L+1}^{\mathrm{III}}$. Причем объема покупки хватает на все продажи своим клиентам вплоть до момента времени $t_{L+1}^{\mathrm{III}}$. Однако каждый клиент также реализует свою продукцию через других посредников. Посредников наших посредников изобразим выше и в стороне. Для удобства восприятия рассмотрим случай, когда только продажи посреднику с номером $r=1$ приводят к завершению запаса нашего распространителя.

Если объема запаса точно хватает на все покупки, включая последнюю (пока без использования страхового запаса), то можно записать:

$$y_L^{\mathrm{III}} = \sum_r \sum_{\substack{k \\ k:\, t_L^{\mathrm{III}} < t_k^{r,\mathrm{II}} \le t_{L+1}^{\mathrm{III}}}} y_k^{r,\mathrm{II}} = \sum_r \sum_{\substack{k \\ k:\, t_L^{\mathrm{III}} < t_k^{r,\mathrm{II}} \le t_{L+1}^{\mathrm{III}}}} \sum_j \sum_{\substack{i \\ i:\, t_k^{r,\mathrm{II}} < t_i^{r,j} \le t_{k+1}^{r,\mathrm{II}}}} y_i^{r,j}.$$

На рисунке 12 обведены овалами покупки $y_k^{r,\mathrm{II}}$ и $y_i^{r,j}$, которые содержатся в большой покупке $y_L^{\mathrm{III}}$ (на которые хватает запаса $y_L^{\mathrm{III}}$). Фигурными стрелками обозначены периоды времени, за которые расходуется этот запас продукции.

$$y_L^{\mathrm{III}} = \sum_r \sum_j \int_{\substack{t_{\min(i)}^{r,j} \\ i:\, t_k^{r,\mathrm{II}} < t_i^{r,j} \le t_{k+1}^{r,\mathrm{II}} \\ k:\, t_L^{\mathrm{III}} < t_k^{r,\mathrm{II}} \le t_{L+1}^{\mathrm{III}}}}^{t_{\max(i)+1}^{r,j}} f^{r,j}(t)dt.$$

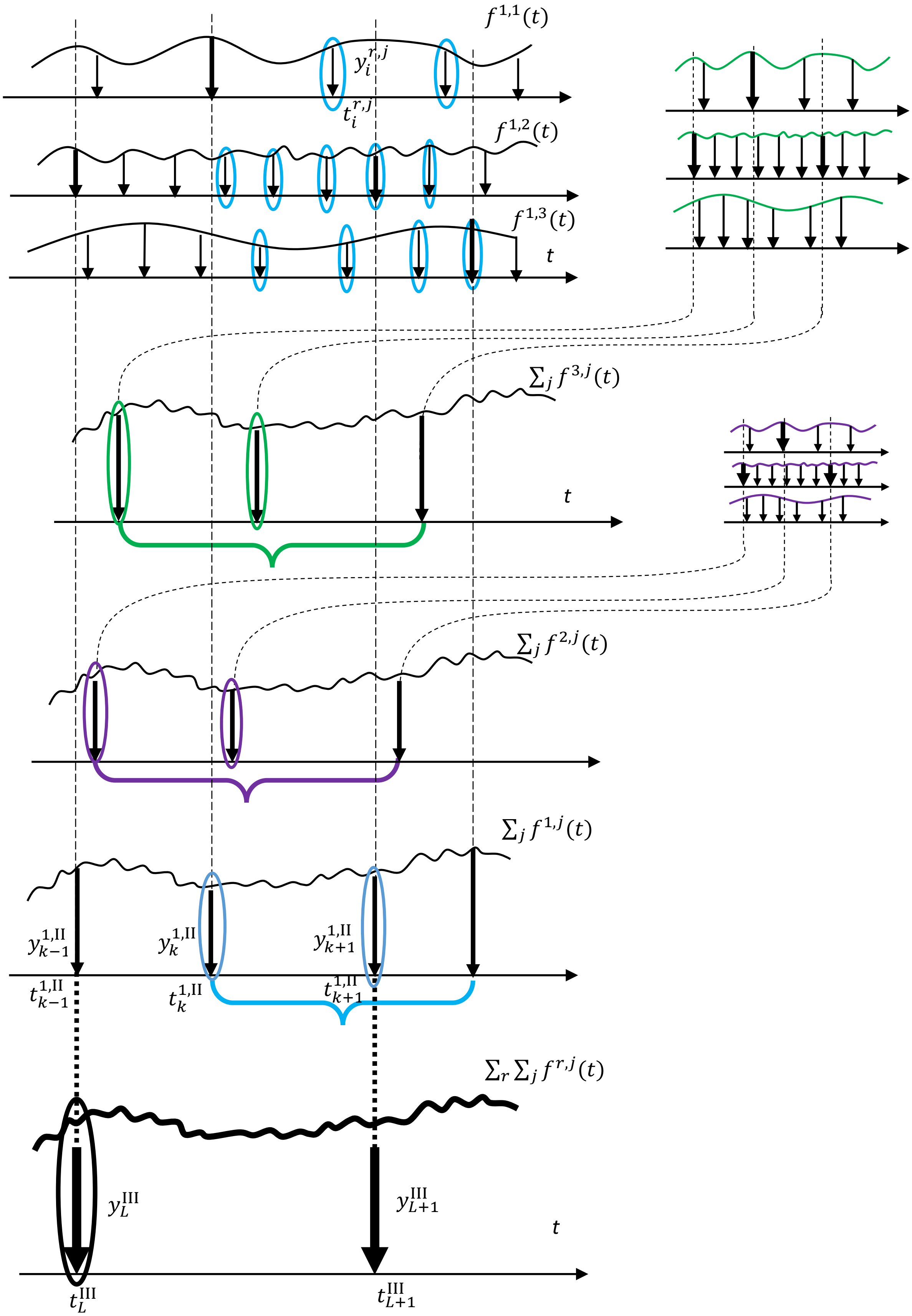


***Рис. 12.*** *Два уровня посредников до конечного потребителя*

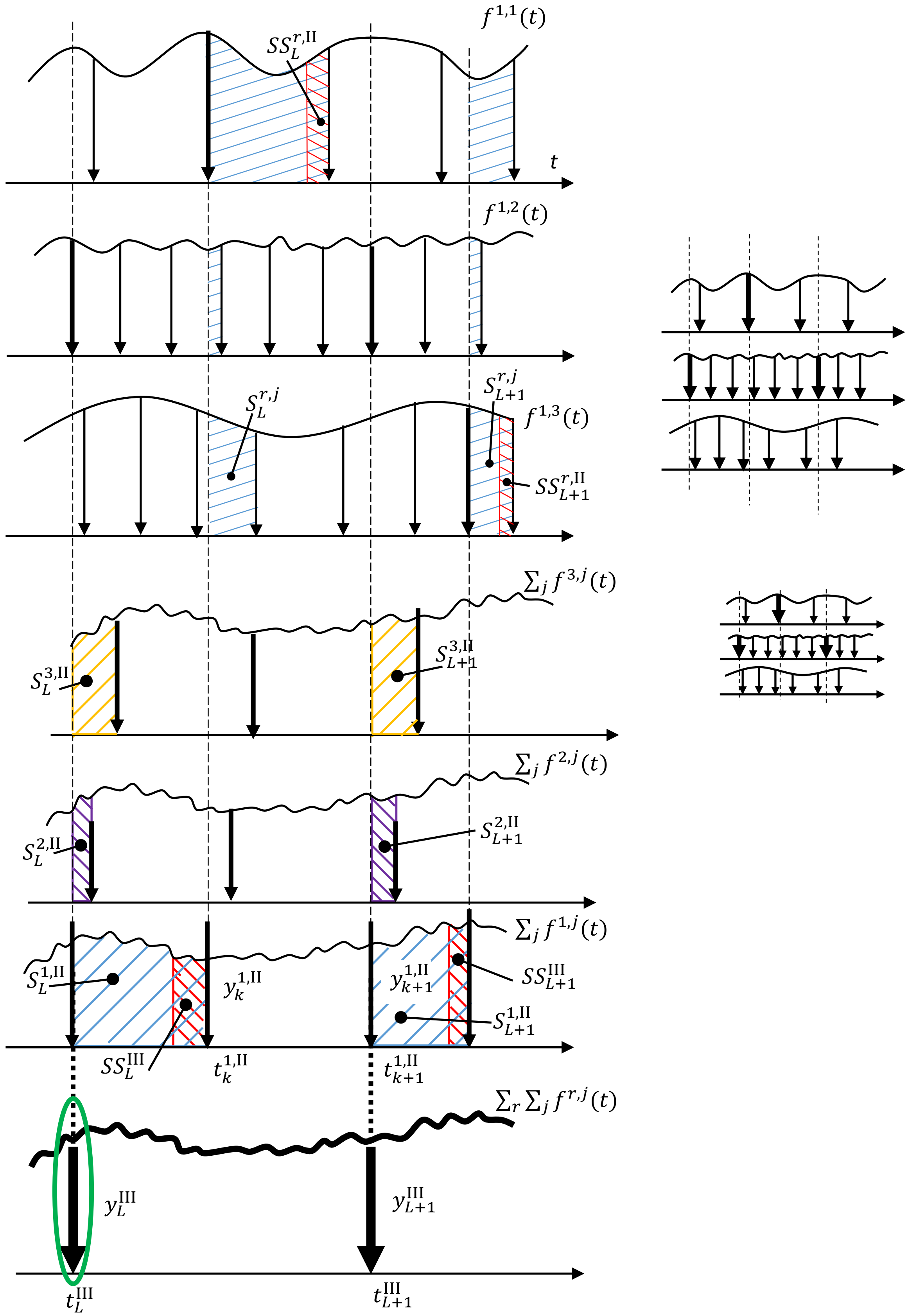


***Рис. 13.*** *Несоответствие наблюдаемого и расходуемого объема продукции*

Опять видим, что объем покупки $y_L^{\mathrm{III}}$ занижен на $\sum_r S_L^{r,\mathrm{II}}$ и превышен на $\sum_r S_{L+1}^{r,\mathrm{II}}$, где

$$S_L^{r,\mathrm{II}} = \int_{t_L^{\mathrm{III}}}^{t_{\min(k)}^{r,\mathrm{II}},\ t_k^{r,\mathrm{II}} > t_L^{\mathrm{III}}} \sum_j f^{r,j}(t).$$

Но на предыдущем уровне также наблюдается занижение на $\sum_j S_L^{r,j}$ и превышение на $\sum_j S_{L+1}^{r,j}$ аналогично рассмотренному примеру с одним промежуточным распространителем.

Величина несоответствия из-за того, что часть покупок идет в зачет других периодов, будет:

$$\Delta Q_S = \sum_r S_{L+1}^{r,\mathrm{II}} - \sum_r S_L^{r,\mathrm{II}} + \sum_r \sum_j \left(S_{L+1}^{r,j} - S_L^{r,j}\right).$$

Помимо этого, возникает та же ситуация с использованием страховых запасов, наблюдается завышение на $SS_L^{\mathrm{III}}$ и занижение на $SS_{L+1}^{\mathrm{III}}$. Но на предыдущем уровне также возникает потребность использовать страховые запасы, поэтому наблюдается завышение на $SS_L^{r,\mathrm{II}}$ и занижение на $SS_{L+1}^{r,\mathrm{II}}$.

Величина несоответствия в связи с использованием страховых запасов:

$$\Delta Q^{SS} = SS_{L+1}^{\mathrm{III}} - SS_L^{\mathrm{III}} + \sum_r \left(SS_{L+1}^{r,\mathrm{II}} - SS_L^{r,\mathrm{II}}\right).$$

Эти два вида погрешности вычитаются друг из друга:

$$\Delta Q = \Delta Q^S - \Delta Q^{SS}.$$

Величины $S_L^{r,\mathrm{II}}$ и $SS_L^{\mathrm{III}}$ имеют такой же интервал значений относительно предшествующих покупок, как и в случае с одним промежуточным распространителем:

$$0 < S_L^{r,\mathrm{II}} \le y_{\text{предш}.L}^{r,\mathrm{II}}.$$

$$0 < SS_L^{\mathrm{III}} \le \max_r y_{\text{предш}.L}^{r,\mathrm{II}}$$

А величины $S_L^{r,j}$ и $SS_L^{r,\mathrm{II}}$ привязаны к покупкам на предыдущем (ближе к потребителям) уровне и имеют то же самое значение, как и в случае с одним промежуточным распространителем.

Справедливость полученных формул можно подтвердить численным примером, однако он не нагляден, его трудно изобразить одним рисунком, и он занимает много объема. Если есть желание, можете самостоятельно представить ход моделирования с помощью таблиц Excel, это займет более 400 строк и 30 столбцов, или какого-нибудь языка программирования. При расчете величин $S_L^{r,\mathrm{II}}$ во время моделирования как сумму значений функции $f^{r,j}(t)$, надо делать корректировку на величину используемых страховых запасов (величина покупки минус максимальный запас), так как они расходуются в текущем периоде, а не в предыдущем. С помощью численного примера убеждаемся, что полученные формулы для величины несоответствия наблюдаемой и расходуемой продукции верны.

Теперь нам достаточно абсолютно всего, чтобы можно было объяснить рост погрешности при удалении от конечного потребителя.

## Определение скорости роста погрешности с удалением от конечного потребителя

Как было показано здесь, в текущем исследовании, абсолютное значение разброса погрешности определения исходной зависимости зависит от объема совершаемых покупок.

При отдалении от конечного потребителя объемы покупок растут. Причем, если бы в погрешности не участвовали покупки от более низких уровней, то относительная погрешность оставалась бы одинаковой. Но у нас в итоговую погрешность включаются погрешности с предшествующих уровней в цепочке распространения, поэтому погрешность растет.

Стоит заметить, что влияние значения объема покупок предшествующих распространителей на относительную погрешность падает с экспоненциальной скоростью при удалении от конечного потребителя. Это происходит из-за того, что объем покупок распространителей верхнего уровня всегда в несколько раз больше, чем предшествующих распространителей. Вклад предшествующих распространителей в погрешность при расчете относительной погрешности надо будет делить на значение, зависящее от объемов покупок на текущем уровне.

Рассмотрим еще раз величину несоответствия из-за того, что часть покупок относилась к другому временному интервалу:

1 уровень посредников – $\Delta Q_S = \sum_j \left(S_{k+1}^j - S_k^j\right)$.

2 уровня посредников – $\Delta Q_S = \sum_r \left(S_{L+1}^{r,\mathrm{II}} - S_L^{r,\mathrm{II}}\right) + \sum_r \sum_j (S_{L+1}^{r,j} - S_L^{r,j})$.

Для сокращения записи обозначим разницу в скобках как $dS^j, dS^{r,\mathrm{II}}, dS^{r,j}$. Величина несоответствия будет:

1 уровень посредников – $\Delta Q_S = \sum_j dS^j$.

2 уровня посредников – $\Delta Q_S = \sum_r dS^{r,\mathrm{II}} + \sum_r \sum_j dS^{r,j}$.

Если повторить все рассуждения для 3 уровней посредников, можно показать, что величина несоответствия будет:

3 уровня посредников – $\Delta Q_S = \sum_m dS^{m,\mathrm{III}} + \sum_m \sum_r dS^{m,r,\mathrm{II}} + \sum_m \sum_r \sum_j dS^{m,r,j}$.

Теперь рассмотрим величину несоответствия из-за использования страховых запасов:

1 уровень посредников – $\Delta Q_{SS} = \left(SS_{k+1}^{\mathrm{II}} - SS_k^{\mathrm{II}}\right)$.

2 уровня посредников – $\Delta Q_{SS} = (SS_{L+1}^{\mathrm{III}} - SS_L^{\mathrm{III}}) + \sum_r (SS_{L+1}^{r,\mathrm{II}} - SS_L^{r,\mathrm{II}})$.

Также для сокращения записи сделаем замену:

1 уровень посредников – $\Delta Q_{SS} = dSS^{\mathrm{II}}$.

2 уровня посредников – $\Delta Q_{SS} = dSS^{\mathrm{III}} + \sum_r dSS^{r,\mathrm{II}}$.

Повторяя все рассуждения для трех уровней посредников, можно показать, что величина несоответствия будет:

3 уровня посредников – $\Delta Q_{SS} = dSS^{\mathrm{IV}} + \sum_m dSS^{m,\mathrm{III}} + \sum_m \sum_r dSS^{m,r,\mathrm{II}}$.

Величины превышений/занижения наблюдаемой и реально расходуемой продукции зависят от величин предшествующих покупок: $S_k^j$ распределена в интервале $(0, y_{\text{предш}.k}^j]$; $SS_k^{\mathrm{II}}$ распределена в интервале $(0, \max_j y_{\text{предш}.k}^j]$. Значит, разницы $dS^j$, $dSS^{\mathrm{II}}$ также пропорциональны предыдущим покупкам.

Зафиксируем все условия, чтобы определить влияние только позиции в цепочке распространения на величину ошибки, а все остальные факторы не вносили своего влияния. Для этого предположим, что на каждом уровне посредников, у каждого распространителя одинаковое количество покупателей $\max m = \max r = \max j = n$. Также предположим, что запасов распространителей хватает на одинаковое количество покупок $P$ более мелких распространителей. То есть величина покупки распространителя $y_{\text{предш}.L}^{r,\mathrm{II}}$ всегда относится к величине покупки $y_{\text{предш}.k}^j$ (дочерних распространителей) как некоторая постоянная величина $P$. Тогда величина $dS^{r,\mathrm{II}}$ относится к $dS^j$ так же, как $P$, и $dS^{m,\mathrm{III}}$ относится к $dS^{r,\mathrm{II}}$ как $P$. Аналогично остальные величины $dSS^{\mathrm{IV}}$, $dSS^{\mathrm{III}}$, $dSS^{\mathrm{II}}$ также относятся последовательно друг к другу как $P$.

Величина несоответствия $\Delta Q_S$ в зависимости от количества уровней посредников будет выглядеть следующим образом:

1 уровень – $\Delta Q_S = \sum_j dS^j = n \cdot dS$.

2 уровня – $\Delta Q_S = \sum_r dS^{r,\mathrm{II}} + \sum_r \sum_j dS^{r,j} = n \cdot P \cdot dS + n^2 dS$.

3 уровня – $\Delta Q_S = \sum_m dS^{m,\mathrm{III}} + \sum_m \sum_r dS^{m,r,\mathrm{II}} + \sum_m \sum_r \sum_j dS^{m,r,j} =$

$$= n \cdot P^2 \cdot dS + n^2 \cdot P \cdot dS + n^3 \cdot dS.$$

Для величины несоответствия $\Delta Q_{SS}$ в связи с использованием страхового запаса аналогично получаем следующее:

1 уровень посредников – $\Delta Q_{SS} = dSS^{\mathrm{II}} = dSS$.

2 уровня посредников – $\Delta Q_{SS} = dSS^{\mathrm{III}} + \sum_r dSS^{r,\mathrm{II}} = P \cdot dSS + n \cdot dSS$.

3 уровня посредников – $\Delta Q_{SS} = dSS^{\mathrm{IV}} + \sum_m dSS^{m,\mathrm{III}} + \sum_m \sum_r dSS^{m,r,\mathrm{II}} =$

$$= P^2 \cdot dSS + n \cdot P \cdot dSS + n^2 \cdot dSS.$$

Общее несоответствие $\Delta Q = \Delta Q_S - \Delta Q_{SS}$ тогда будет:

1 уровень – $\Delta Q = n \cdot dS - dSS$.

2 уровня – $\Delta Q = n \cdot P \cdot dS + n^2 dS - P \cdot dSS - n \cdot dSS =$

$$= P \cdot (n \cdot dS - dSS) + n \cdot (n \cdot dS - dSS) = (n \cdot dS - dSS)(P + n).$$

3 уровня –

$$\Delta Q = n \cdot P^2 \cdot dS + n^2 \cdot P \cdot dS + n^3 \cdot dS - P^2 \cdot dSS + n \cdot P \cdot dSS + n^2 \cdot dSS =$$
$$= P^2(n \cdot dS - dSS) + nP(n \cdot dS - dSS) + n^2(n \cdot dS - dSS) =$$
$$= (n \cdot dS - dSS)(P^2 + nP + n^2).$$

Так как нас интересует относительная погрешность, то надо делить величину несоответствия на объем покупки рассматриваемого распространителя. То есть для 1 уровня посредников величину $\Delta Q$, которая пропорциональна объему покупок $y_i^j$ на предшествующем уровне, надо делить на объем покупки $y_k^{\mathrm{II}}$, который, как было условлено, будет в $P$ раз больше.

1 уровень посредников

$$\frac{\Delta Q}{y_k^{\mathrm{II}}} = \frac{(n \cdot dS - dSS)}{y_k^{\mathrm{II}}} = \frac{(n \cdot dS - dSS)}{P \cdot y_i^j}.$$

2 уровня посредников

$$\frac{\Delta Q}{y_L^{\mathrm{III}}} = \frac{(n \cdot dS - dSS)(P + n)}{y_L^{\mathrm{III}}} = \frac{(n \cdot dS - dSS)(P + n)}{P \cdot y_k^{\mathrm{II}}} = \frac{(n \cdot dS - dSS)(P + n)}{P^2 \cdot y_i^j}$$
$$= \frac{(n \cdot dS - dSS)}{P \cdot y_i^j}\left(1 + \frac{n}{P}\right).$$

3 уровня посредников

$$\frac{\Delta Q}{y^{\mathrm{IV}}} = \frac{(n \cdot dS - dSS)(P^2 + nP + n^2)}{y_L^{\mathrm{IV}}} =$$

$$= \frac{(n \cdot dS - dSS)(P^2 + nP + n^2)}{P^3 \cdot y_i^j} = \frac{(n \cdot dS - dSS)}{P \cdot y_i^j}\left(1 + \frac{n}{P} + \left(\frac{n}{P}\right)^2\right).$$

Если продолжить эту последовательность дальше, то получим следующее:

$N$ уровней посредников

$$\frac{\Delta Q}{y^{[N+1]}} = \frac{(n \cdot dS - dSS)}{P \cdot y_i^j}\left(1 + \frac{n}{P} + \left(\frac{n}{P}\right)^2 + \cdots + \left(\frac{n}{P}\right)^{N-1}\right).$$

Выражение в скобках является суммой геометрически убывающей прогрессии при условии, что величина $\left(\frac{n}{P}\right)$ меньше 1. Эта величина будет меньше единицы, так как $P$ показывает, на сколько покупок хватает запаса продукции распространителя, а $n$ показывает количество покупателей у распространителя. Только в случае, если $n \geq P$ , выражение в скобках не будет убывающей прогрессией, это возможно, если объема покупок распространителя не хватает, чтобы удовлетворить лишь по одной покупке своих покупателей, то есть объем покупок покупателей превосходит доли $1/n$ от объема покупки распространителя.

## Расчет дисперсии

При увеличении количества уровней посредников увеличивается общее количество задействованных покупателей и увеличивается общее количество случайных (неопределенны) величин. Если дисперсия будет расти как убывающая геометрическая прогрессия, то относительно среднее квадратичное отклонение должно расти как корень из этой дисперсии, то есть еще медленнее (ошибки могут компенсировать друг друга). Однако большая часть вклада в дисперсию осуществляется первыми $2(n+1)$ случайными величинами (по 2 величины на значения $dS$ и $dSS$), вклад следующих $2(n^2+n)$ случайных величин падает в $P$ раз, а очередных $2(n^3+n^2)$ падает уже в $P^2$ раз и так далее. Закон распределения относительной погрешности не будет соответствовать нормальному распределению, так как в нормальном распределении предполагается, что вклад всех случайных величин одинаков. Закон распределения относительной погрешности будет характеризоваться в основном суммой первых $2(n+1)$ равномерных случайных величин (распределенных от 0 до $\frac{1}{P}$, так как каждая величина лежит в диапазоне от 0 до $y^{j}_{\text{предш.}k}$, а затем она делится на $y^{\text{II}}_{k}$, которое в $P$ раз больше).

Попробуем определить дисперсию относительного отклонения $\frac{\Delta Q}{y^{[N+1]}}$. Дисперсия базовых равномерных величин равна 1/12. Тогда суммарная дисперсия первых $2(n+1)$ величин будет $\frac{2(n+1)}{12P^2}$ (дисперсия суммы/разницы независимых случайных величин равна сумме дисперсий, умножение случайной величины на константу приводит к увеличению дисперсии в константу в квадрате раз). Следующие $2(n^2+n)$ случайных величин увеличивают дисперсию на $\frac{2(n^2+n)}{12P^4}$, а очередные $2(n^3+n^2)$ случайные величины на $\frac{2(n^3+n^2)}{12P^6}$ и так далее.

В результате дисперсия относительного отклонения в связи с наличием $N$ уровней посредников будет:

$$D\left[\frac{\Delta Q}{y^{[N+1]}}\right] = \frac{2(n+1)}{12P^2} + \frac{2(n^2+n)}{12P^4} + \frac{2(n^3+n^2)}{12P^6} + \cdots + \frac{2(n+1)n^{N-1}}{12P^{2N}} =$$

$$= \frac{n+1}{6P^2}\left(1 + \frac{n^1}{P^2} + \frac{n^2}{(P^2)^2} + \cdots + \frac{n^{N-1}}{(P^2)^{N-1}}\right).$$

Выражение в скобках опять является суммой геометрически убывающей прогрессии. Сумма первых $N$ слагаемых будет $\frac{1-q^N}{1-q}$, где $q = \frac{n}{P^2}$ – знаменатель прогрессии. В результате можем получить значение дисперсии для относительного отклонения:

$$D\left[\frac{\Delta Q}{y^{[N+1]}}\right] = \frac{(n+1)\left(1-\left(\frac{n}{P^2}\right)^N\right)}{6P^2\left(1-\frac{n}{P^2}\right)}.$$

Среднее квадратичное отклонение получаем как корень из дисперсии:

$$\sigma\left[\frac{\Delta Q}{y^{[N+1]}}\right] = \sqrt{\frac{(n+1)\left(1-\left(\frac{n}{P^2}\right)^N\right)}{6P^2\left(1-\frac{n}{P^2}\right)}}.$$

Для рассмотренного абстрактного примера, когда у распространителя по 3 покупателя ($n = 3$ более мелких распространителей), если объем покупок распространителей при удалении от конечного потребителя растет как $P = 10$, то можем посчитать дисперсию и среднее квадратичное отклонение для относительного отклонения, столбец 1 (табл. 5). График зависимости относительной погрешности от удаленности от конечных потребителей представлен на рисунке ниже (рис. 14). Как видно из таблицы и графика, относительная погрешность растет незначительно, уже при удаленности на 3 позиции погрешность изменяется лишь в четвертом знаке.

***Таблица 5***

РАСЧЕТ ДИСПЕРСИИ И СРЕДНЕГО КВАДРАТИЧНОГО ОТКЛОНЕНИЯ ОТНОСИТЕЛЬНОЙ ОШИБКИ ПРИ УДАЛЕНИИ ОТ КОНЕЧНОГО ПОТРЕБИТЕЛЯ

| $N$ | $n = 3, P = 10$ | | $n = 3, P = 4$ | | $n = 20, P = 20$ | |
|---|---|---|---|---|---|---|
| | $D$ | $\sigma$ в % | $D$ | $\sigma$ в % | $D$ | $\sigma$ в % |
| 1 | 0,006666667 | 8.164966 | 0,041666667 | 20.412415 | 0,00875 | 9.3541435 |
| 2 | 0,006866667 | 8.286535 | 0,049479167 | 22.243910 | 0,0091875 | 9.5851448 |
| 3 | 0,006872667 | 8.290155 | 0,05094401 | 22.570780 | 0,009209375 | 9.5965489 |
| 4 | 0,006872847 | 8.290263 | 0,051218669 | 22.631542 | 0,009210469 | 9.5971187 |
| 5 | 0,006872852 | 8.290267 | 0,051270167 | 22.642917 | 0,009210523 | 9.5971472 |

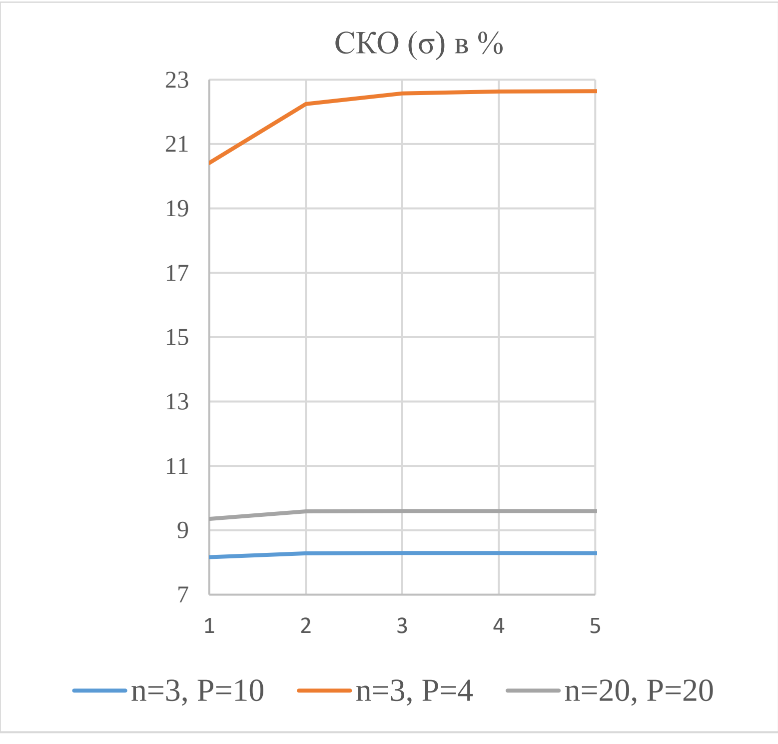


***Рис. 14.*** *Относительная погрешность в % в зависимости от удаленности от конечных потребителей (длины цепочки посредников)*

## Заключение

Подведем итоги. Получены формулы для расчета ошибки восстановления исходной зависимости в зависимости от позиции в цепочке распространителей. Получены значения интервалов, в которых лежит ошибка. Ошибка имеет нулевое математическое ожидание при неизменности исходных факторов. На численном примере показана справедливость формул. Определено, что относительная погрешность при удалении от конечного потребителя (без влияния других факторов) будет расти приблизительно так же, как растет сумма геометрически убывающей прогрессии. Однако так как элементы суммы являются случайными величинами, дисперсия и среднее квадратичное отклонение растут еще медленнее. Получены формулы для их расчета в идеальных условиях, когда никакие другие факторы не оказывают никакого воздействия. Показано что при удалении от конечного потребителя дисперсия и среднее квадратичное отклонение для относительной ошибки растут очень медленно. Однако даже в худшем случае, когда все случайные величины накладываются и ошибка увеличивается, а не компенсируется, погрешность все равно не превысит суммы геометрической убывающей прогрессии.

Тем не менее в реальных условиях при изменении позиции в цепочке распространителей могут происходить разные вещи, например, может изменяться количество дочерних распространителей, меняться схема цепи распространителей, объем покупок текущего распространителя может непропорционально меняться относительно объемов покупок дочерних распространителей (или исходной суммарной функции потребления), может меняться максимальный запас продукции в зависимости от спроса и т.д. Все это приведет к увеличению погрешности, однако это уже другая история, и к текущему исследованию это отношения не имеет.

Что касается погрешности за множество интервалов времени между покупками, например, за время моделирования 2 года, когда было совершено много покупок, в разных покупках ошибка могла быть либо отрицательной, либо положительной, причем маловероятно, чтобы в соседних покупках ошибка была с одинаковым знаком. Тогда при сглаживании восстановленная функция будет еще ближе к исходной функции на всем периоде наблюдений.

Результаты исследования с уверенностью позволяют сказать, что при анализе данных из истории продаж по возможности следует всегда переходить на емкостный метод анализа редких событий, вместо агрегирования данных по временным интервалам и дальнейшего анализа этих временных интервалов. Для этого необходимо от всех покупателей раздельно собирать такие данные, как дата, объем покупки и идентификатор покупающего, все это, наверное, уже давно реализовано в условиях цифровой экономики, когда большой брат следит за всеми твоими действиями, осталось это только грамотно использовать. Не стоит лишний раз агрегировать в кучу данные продаж по временным интервалам, так как при таком агрегировании теряется важная индивидуальная информация.

Следующие исследования будут направлены на изучение вопросов точности в зависимости от влияния конкуренции, когда покупатель при пополнении своего запаса переключается с одного распространителя на другого, а также на определение точности при потере части данных, когда часть данных теряется бесследно. Еще интересным кажется определить, как изменяется точность, если покупатели нестрого следят за своими запасами, а проверка запасов осуществляется с разбросом критического уровня.

## ЛИТЕРАТУРА

1. *Dudewicz Edward J., Karian Zaven A.* The role of statistics in is/it: practical gains from mined data // Information systems frontiers. 1999. № 1(3). P. 259-266.

2. *Makridakis S.* Forecasting: its role and value for planning and strategy // International journal of forecasting. 1996. № 12(4). P. 513-537.

3. *Saerkkae S.* Bayesian Filtering and Smoothing. Cambridge University Press. 2013. URL: http://www.cambridge.org/9781107030657 (дата обращения: 03.03.2019).

4. *Lambert Koopmans*. The Spectral Analysis of Time Series. 1st Edition. University of New Mexico. Academic Press. 1995. URL: https://www.elsevier.com/books/the-spectral-analysis-of-time-series/koopmans/978-0-12-419251-5 (дата обращения: 03.03.2019).

5. *Лукашин Ю.П.* Адаптивные методы краткосрочного прогнозирования временных рядов. М: Финансы и статистика. 2003.

6. *Голубинский А.Н.* Методы аппроксимации экспериментальных данных и построения моделей // Вестник Воронежского института МВД России. 2007. № 2. С. 138-143.

7. *Jackson J.E.* Principal components and factor analysis: Part I – principal components // Journal of Quality Technology. 1980. № 12. P. 201-213.

8. *Jackson J.E.* Principal components and factor analysis: Part II – additional topics related to principal components // *Journal of Quality Technology*. 1981. № 13. C. 46-58.

9. Корректировка числа редких событий в логистической регрессии. URL: http://www.statmethods.ru/stati/178-korrektirovka-chisla-redkikh-sobytij-v-modeli-logisticheskoj-regressii.html (дата обращения: 03.03.2019).

10. *Altman N.S.* An introduction to kernel and nearest-neighbor nonparametric regression // The American Statistician. 1992. № 46(3). C. 175-185. doi:10.1080/00031305.1992.10475879.

11. *Croston J.D.* Forecasting and stock control for intermittent demands // Operational Research Quarterly (1970-1977). 1972. № 23(3). C. 289-303.

12. *Johnston F.R., Boylan J.E.* Forecasting intermittent demand: a comparative evaluation of Croston's method. Comment // International journal of forecasting. 1996. № 12(2). P. 297-298.

13. *Efron B. and Tibshirani R.J.* An introduction of the Bootstrap. New York: Chapman & Hall, 1993.

14. *Willemain T.R., Park D.S., Kim Y.B., Shin K.I.* Simulation output analysis using the threshold bootstrap. 2001. № 134(1). P. 17-28.

15. *Иванько Р.С.* Краткосрочное прогнозирование нестационарного спроса в оптовой торговле: дис. … канд. эконом. наук. М., 2005.

16. *Вентцель Е.С., Овчаров Л.А.* Теория случайных процессов и ее инженерные приложения: Учеб. пособие для втузов. 2-е изд., стер. М.: Высш. шк., 2000. 383 с.

17. *Жуков М.М., Кудряш В.И.* Математическая модель рисков распределенных технических систем при бета-распределении плотности вероятности наступления ущерба // Общественная безопасность, законность и правопорядок в III тысячелетии. Воронеж: Воронежский институт МВД РФ, 2017. № 3-3. С. 278-281.

18. *Дзанагова И.Т., Хугаева Л.Т.* Информационно-статистические методы построения экстремальных моделей редких событий // Фундаментальные исследования. Пенза: Академия естествознания. 2015. № 11-6. С. 1081-1084.

19. *Лукинский В., Замалетдинова Д.* Методы управления запасами: расчет показателей запаса для товарных групп, относящихся к редким событиям (часть I). Логистика. 2015. № 1 (98). С. 28-33.

20. *Лукинский В., Замалетдинова Д.* Методы управления запасами: расчет показателей запаса для товарных групп, относящихся к редким событиям (часть II). Логистика. 2015. № 2 (99). С. 24-27.

21. *Вожжов А.П., Луняков О.В., Вожжов С.П.* Формирование страховых запасов предприятия при пуассоновском характере поступающих и выдаваемых потоков // Экономика и управление: теория и практика. 2015. № 1(1). С. 30-35.

22. *Кораблев Ю.А.* Емкостный метод определения функции скорости потребления // Экономика и менеджмент систем управления. Воронеж: Изд-во «Научная книга». 2015. № 15(1.1). С. 140-150.

23. *Кораблев Ю.А.* Обоснование емкостного метода определения спроса // Экономика, статистика и информатика. Вестник УМО. М.: РЭУ им. Плеханова. 2015. № (5). С. 96-101.

24. *Кораблев Ю.А.* Исследование точности емкостного метода от позиции в цепочке распространителей // Экономика и управление: проблемы, решения. М.: Научная библиотека. 2018. №7(5). С. 106-121.

25. Аппроксимация линейным или нелинейным МНК. Кросс-платформенная библиотека численного анализа ALGLIB. URL: http://alglib.sources.ru/interpolation/leastsquares.php#header0 (дата обращения: 03.03.2019).

26. *Бауэрсокс Дональд Дж., Клосс Девид Дж.* Логистика: интегрированная цепь поставок. 2-е изд. / Пер. с англ. Н.Н. Барышниковой, Б.С. Пинскера. М.: ЗАО «Олимп-Бизнес», 2008. 640 с.

27. *Кораблев Ю.А.* Разбор причин и оценка погрешности аномальных картин в емкостном методе анализа редких событий // Экономика и управление: проблемы, решения. М.: Научная библиотека. 2017. № 8(6). С. 8-12.

## REFERENCES

1. Dudewicz Edward J., Karian Zaven A. The role of statistics in is/it: practical gains from mined data // Information systems frontiers. 1999. № 1(3). P. 259-266.

2. Makridakis S. Forecasting: its role and value for planning and strategy // International journal of forecasting. 1996. №12(4). P. 513-537.

3. Saerkkae S. Bayesian Filtering and Smoothing. Cambridge University Press. 2013. URL: http://www.cambridge.org/9781107030657 (дата обращения: 03.03.2019).

4. Lambert Koopmans. The Spectral Analysis of Time Series. 1st Edition // University of New Mexico. Academic Press. 1995. URL: https://www.elsevier.com/books/the-spectral-analysis-of-time-series/koopmans/978-0-12-419251-5 (дата обращения: 03.03.2019).

5. Lukashin Yu.P. *Adaptivnyye metody kratkosrochnogo prognozirovaniya vremennykh ryadov* [Adaptive methods of short-term forecasting of time series]. M: Finance and statistics. 2003.

6. Golubinsky A.N. *Metody approksimatsii eksperimental'nykh dannykh i postroyeniya modeley* [Methods of approximation of experimental data and model building] // *Vestnik Voronezhskogo instituta MVD Rossii* [Bulletin of the Voronezh Institute of the Ministry of Internal Affairs of Russia]. 2007. No. 2. P. 138-143.

7. Jackson J.E. Principal components and factor analysis: Part I – principal components // Journal of Quality Technology. 1980. № 12. P. 201-213.

8. Jackson J. E. Principal components and factor analysis: Part II – additional topics related to principal components // Journal of Quality Technology. 1981. № 13. P. 46-58.

9. *Korrektirovka chisla redkikh sobytiy v logisticheskoy regressii* [Adjustment of the number of rare events in the logistic regression. URL: http://www.statmethods.ru/stati/178-korrektirovka-chisla-redkikh-sobytij-v-modeli-logisticheskoj- regressii.html (appeal date: 03/03/2019)].

10. Altman N.S. An introduction to kernel and nearest-neighbor nonparametric regression // The American Statistician. 1992. № 46(3). P. 175-185. doi:10.1080/00031305.1992.10475879.

11. Croston J.D. Forecasting and stock control for intermittent demands // Operational Research Quarterly (1970-1977). 1972. № 23(3). P. 289-303.

12. Johnston F.R., Boylan J.E. Forecasting intermittent demand: a comparative evaluation of Croston's method. Comment // International journal of forecasting. 1996. № 12(2). P. 297-298.

13. Efron B. and Tibshirani R.J. An introduction of the Bootstrap. New York: Chapman & Hall, 1993.

14. Willemain T.R., Park D.S., Kim Y.B., Shin K.I. Simulation output analysis using the threshold bootstrap. 2001. № 134(1). P. 17-28.

15. Ivanko R.S. *Kratkosrochnoye prognozirovaniye nestatsionarnogo sprosa v optovoy torgovle: dis. ... kand. ekonom. nauk* [Short-term forecasting of non-stationary demand for wholesale trade: thesis for Cand. of economics]. Moscow, 2005.

16. Wentzel E.S., Ovcharov L.A. *Teoriya sluchaynykh protsessov i yeye inzhenernyye prilozheniya: Ucheb. posobiye dlya vtuzov* [The theory of random processes and its engineering applications. Training manual for technical colleges]. 2nd ed., Sr. M.: Higher. school, 2000. 383 P.

17. Zhukov M.M., Kudryash V.I. *Matematicheskaya model' riskov raspredelennykh tekhnicheskikh sistem pri beta-raspredelenii plotnosti veroyatnosti nastupleniya ushcherba* [Mathematical model of risks of distributed technical systems for beta-distribution of density of probability of suffering of damage] // *Obshchestvennaya bezopasnost', zakonnost' i pravoporyadok v III tysyacheletii*. Voronezh: Voronezhskii institut MVD RF, 2017. № 3-3. Pp. 278-281.

18. Dzanagova I.T., Khugaeva L.T. *Informatsionno-statisticheskiye metody postroyeniya ekstremal'nykh modeley redkikh sobytiy* [Method of operator series for constructing extremal models of rare events] // *Fundamental'nye issledovaniya. Penza: Akademiya Estestvoznaniya.* 2015. № 11-6. Pp. 1081-1084.

19. Lukinsky V., Zamaletdinova D. *Metody upravleniya zapasami: raschet pokazateley zapasa dlya tovarnykh grupp, otnosyashchikhsya k redkim sobytiyam (chast' I)* [Methods of inventory management: the calculation of inventory indicators for product groups related to rare events (Part I)] // *Logistika.* 2015. № 1 (98). Pp. 28-33.

20. Lukinsky V., Zamaletdinova D. *Metody upravleniya zapasami: raschet pokazateley zapasa dlya tovarnykh grupp, otnosyashchikhsya k redkim sobytiyam (chast' II)* [Methods of inventory management: the calculation of inventory indicators for product groups related to rare events (Part II)] // *Logistika.* 2015. № 2 (99). Pp. 24-27.

21. Vozhzhov A.P., Lunyakov O.V., Vozhzhov S.P. *Formirovaniye strakhovykh zapasov predpriyatiya pri puassonovskom kharaktere postupayushchikh i vydavayemykh potokov* [Formation of safety stock with application of the Poisson processes to incoming and outgoing flows] // *Ekonomika i upravlenie: teorija i praktika* [Economics and Management: Theory and Practice]. 2015. № 1(1). S. 30-35.

22. Korablev Yu.A. *Yemkostnyy metod opredeleniya funktsii skorosti potrebleniya* [Capacity method of determination consumption rate function] // *Ekonomika i menedzhment sistem upravleniya* [Economics and management systems management] Voronezh: Izd-vo «Nauchnaya kniga». 2015. № 15(1.1). S. 140-150.

23. Korablev Yu.A. *Obosnovaniye yemkostnogo metoda opredeleniya sprosa* [Argumentation of capacity method demand determination] // *Ekonomika, statistika i informatika. Vestnik UMO* [Economy, statistics and computer science. Herald UMO]. M.: REU Im. Plehanova. 2015. № (5). S. 96-101.

24. Korablev Yu.A. *Issledovaniye tochnosti yemkostnogo metoda ot pozitsii v tsepochke rasprostraniteley* [The study of the capacitive method accuracy from the position in the chain of distributors] // *Ekonomika i upravlenie: problemy, resheniya* [Economics and Management: problems, solutions]. M.: Nauchnaya biblioteka. 2018. № 7(5). S. 106-121.

25. *Approksimatsiya lineynym ili nelineynym MNK. Kross-platformennaya biblioteka chislennogo analiza ALGLIB* [Approximation of linear or nonlinear OLS. Cross-platform biblio-tech numerical analysis ALGLIB]. URL: http://alglib.sources.ru/interpolation/leastsquares.php # header0 (appeal date: 03/03/2019).

26. Bowersox, Donald J., Kloss David J. *Logistika: integrirovannaya tsep' postavok. 2-ye izd* [Logistics: an integrated supply chain. 2nd ed.] / Per. from English N.N. Baryshnikova, B.S. Pinsker. M.: ZAO Olimp-Business, 2008. 640 p.

27. Korablev Yu.A. *Razbor prichin i otsenka pogreshnosti anomal'nykh kartin v yemkostnom metode analiza redkikh sobytiy* [The causes analysis and error estimation of the anomalous pictures in the capacity method for the analysis of rare events] // *Ekonomika i upravlenie: problemy, resheniya* [Economics and Management: problems, solutions]. M.: Nauchnaya biblioteka. 2017. № 8(6). S. 8-12.

## ERROR OF THE CAPACITY METHOD OF RARE EVENTS ANALYSIS, REMOTENESS FROM THE END USER

**Yu.A. KORABLEV**

Financial University under the Government of the Russian Federation
125993, Moscow, Russia, Leningradsky Prospect, 49
Email: academy@fa.ru

*The works devoted to the study of rare events are very few. The most popular method for analyzing rare events is the theory of random processes, when events are represented by Poisson or Palm flows. Other methods have even less accuracy and validity. Nevertheless, the theory of random processes is not able to determine the event occurrence moment but only the probability of a given number of events for a fixed length time interval.*

*The paper describes the methodology for the rare events study, which is based on the difference in the events sources and the restoration of the proposed process parameters underlying the occurrence of these events. After the restoration of the process parameters, a pattern is sought by any other known methods, after which the patterns are extrapolated for the future. After extrapolating the process parameters, the process starts to obtain a forecast of the time moments of the following events occurrence.*

*The most common process in the economy is the process of consuming, or expenditure of products, or accumulating disturbances to a certain level. In this case, event sources can be modeled as capacities. The process parameter is the emptying speed of this capacity. A method for restoring this speed is proposed, after which future events can be predicted. I call this method of analyzing and predicting rare events the "capacity" method.*

*The article analyzes the influence of the position in the chain of distributors on the accuracy of restoring the original unknown function of the products consumption rate using the capacity method. Another goal is to find the magnitude of the relative error in the restoration of the original unknown function of product consumption.*

*With the help of mathematical analysis, the consumption process for the chain of distributors is considered, the inverse problem is constructed, the error is analyzed. As a result of the current study, the values of the error in restoring the original dependence in the sale of products through one intermediary, as well as in the sale of products through two successive intermediaries, are obtained. The extreme values of the intervals for the error of restoring the original dependence were obtained. On a specific numerical example, the validity of the formulas obtained is confirmed. It is shown that the error is not systematic. It is shown that the increase in error from the distance from the end user, with all other factors remaining unchanged, grows as a sum of a geometrically decreasing progression. Values of the variance and standard deviation for the relative error were calculated; it is shown that they grow very slowly.*

**Keywords:** rare events; capacity method; consumption rate; accuracy; error; dispersion; sequence of distributors; intermediaries.

*Работа поступила 07.06.2019 г.*