

С. Жаринов

О теоретических основах статистического управления запасами

**Если вы занимаетесь наукой, вам требуется понять мир.
Если же вы занимаетесь бизнесом, вам нужно, чтобы мир
не понимали другие.**

Нассим Талеб¹

Консалтинг – такая штука, где бизнес явным образом делается на непонимании мира консультируемыми. И это нормально, ведь понимать всё не под силу никому. Поэтому и обращаются к экспертам, в данном случае, по управлению производством, запасами, цепочками поставок и прочим смежным вопросам. Как человеку с академическим бэкграундом, лично мне всегда удобнее «плясать от печки», то есть от теории. Однако довольно часто приходится сталкиваться с недоумением аудитории:

– Раньше нас учили, что нужно умножать на полтора, а потом ещё скорректировать на «коэффициент паранойи»

– А почему на полтора – не объясняли? И где взять сей экзотический коэффициент?

– Так это же сокровенные знания. Почти как Священное Писание. Поэтому предлагали не заморачиваться и просто принять на веру.

Между тем, указанная предметная область имеет весьма солидный научный фундамент², построенный на принципах и моделях теории вероятностей и математической статистики. Только вот наши менеджеры (и даже консультанты) бывают «не в курсе» теоретических основ.

О ЧЁМ, СОБСТВЕННО, РЕЧЬ?

За последние полтора года мне выпала возможность заниматься разработкой методик управления запасами для двух крупных российских производственных компаний и одной популярной сети аптек. Казалось бы, дело нехитрое. Традиционный подход к повышению доходности в цепочках поставок обычно предполагает сокращение запасов за счёт более частого их пополнения, соответственно, более мелкими партиями. Хотя нередко случается так, что из-за технологических либо организационных особенностей заметное увеличение частоты заказов по сравнению с текущим невозможно либо нежелательно. В подобных обстоятельствах для обеспечения требуемого уровня сервиса приходится применять специальные алгоритмы³.

В кейсе с аптечной сетью, кроме того, возникла необходимость объяснить руководству компании, что иногда целесообразно запасы не сокращать, а повышать. В качестве примера на рис. 1 показаны два набора данных ежедневных продаж и запасов одного и того же медицинского препарата для двух территориально близких аптек сети за первые восемь месяцев (243 дня) 2021 года в штуках. Здесь S^* – средний размер запасов, $\bar{x}$ и s – среднее значение и коэффициент вариации продаж, K – количество поставок, $k = 243/(K+1)$ – среднее время между поставками в днях, $\{0\}$ – число стокаутов на конец операционного дня за весь период. Как видно, первая аптека (диаграмма А рис. 1) при средних 19,5-дневных запасах работала без стокаутов, тогда как вторая (диаграмма Б рис. 1) имела 14 стокаутов, – однако при более низких 15,6-дневных

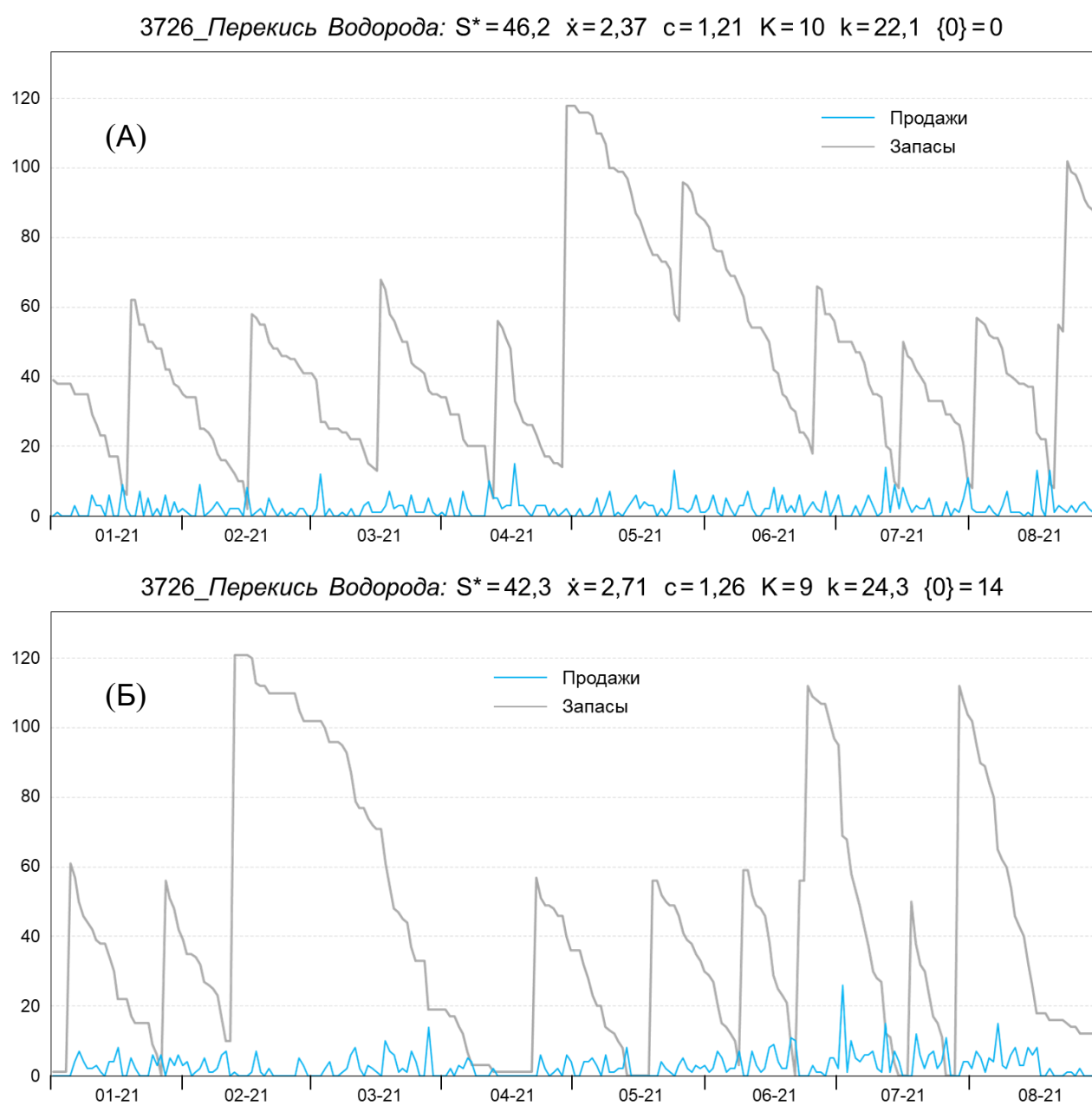


Рис. 1. Продажи и запасы (две аптеки, одна SKU, период 01.01.21 – 31.08.21).

средних запасах. При этом (в предположении, что продажи соответствуют реальному спросу) статистические характеристики потоков потребностей похожи и близки к пуассоновским, но с тяжёлыми хвостами из-за выбросов. В целом, качественно всё логично и понятно: при той же частоте поставок в первом случае запасы, скорее всего, можно уменьшить, а во втором – для работы без стокауты – нужно увеличить. Вопрос: насколько уменьшить и насколько увеличить? Но не за счёт какого-то особого способа или приёма, а в принципе? И как это доказать и наглядно продемонстрировать неспециалистам, особенно если у каждой из аптек по несколько тысяч таких SKU?

Пришлось подготовить специальную презентацию, основное содержание которой приводится ниже и (несмотря на её технический характер), надеюсь, будет полезно в более широком контексте.

ЧТО ТАКОЕ «МНОГО» И ЧТО ТАКОЕ «МАЛО»?

Главная идея статистического управления запасами восходит к классической задаче теории вероятностей в следующей самой общей формулировке:

Предполагается, что значение ежедневного потребления x по рассматриваемой SKU является случайной величиной с известным распределением вероятностей. Каков минимальный размер запасов, позволяющий удовлетворить спрос по этой SKU на ближайшие k дней с заданной вероятностью p ?

В простейшей ситуации, когда вариабельность отсутствует (то есть дисперсия случайной величины равна 0) и ежедневно потребляется ровно одна условная единица (у.е.) данной SKU, то, очевидно, достаточно минимального запаса S_0 в размере k у.е.⁴. Естественно, в условиях вариабельности спроса (то есть при ненулевой дисперсии случайной величины) минимальный размер начальных запасов должен быть больше k , причём чем больше дисперсия, тем при тех же значениях k и той же доверительной вероятности p этот размер должен быть выше. Но насколько выше?

Как известно, если случайная величина x имеет распределение f_1 со средним $\bar{x}$ и дисперсией σ^2 , то значение суммы k таких *независимых* случайных величин имеет некоторое распределение f_k со средним $k\bar{x}$ и дисперсией $k\sigma^2$. Поэтому при заданной доверительной вероятности p минимальный размер начальных запасов S_0 определяется квантилем⁵ уровня p распределения f_k . Остаётся только один вопрос: где взять этот квантиль?

Рассмотрим два типичных распределения спроса, – одно непрерывное и одно дискретное, – для которых точное решение задачи существует и может быть легко воспроизведено.

На рис. 2 показана серия f_k для экспоненциального распределения $f_1 = \text{Exp}(1)$ со средним $\dot{x} = 1$ и дисперсией $\sigma^2 = 1$. Известно, что сумма k таких независимых случайных величин имеет гамма-распределение с параметрами k и $\theta = 1$, то есть $f_k = \Gamma(k; 1)$. Следовательно, например, для $k = 15$ дней при $p = 90\%$ достаточно иметь запас наших SKU в размере $\sim 20,1$ у.е., а при $p = 95\%$ – в размере $\sim 21,9$ у.е. (см. рис. 3).

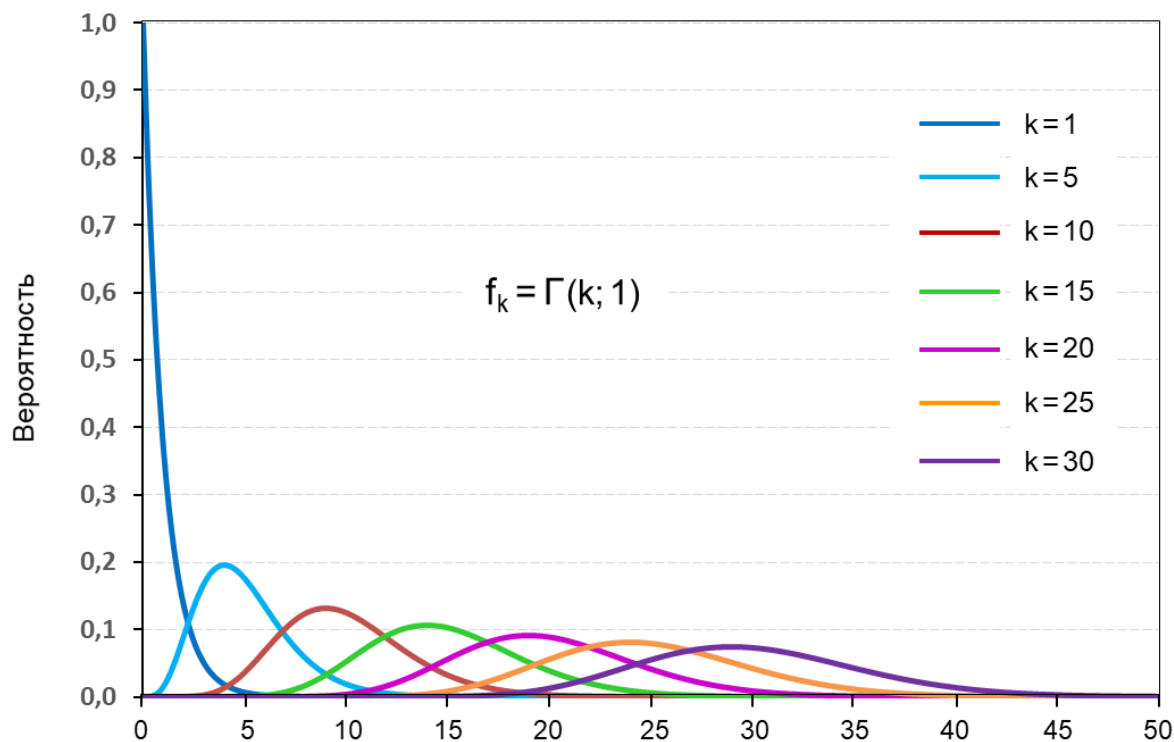


Рис. 2. Пример распределения вероятностей суммы k независимых случайных величин с экспоненциальным распределением.

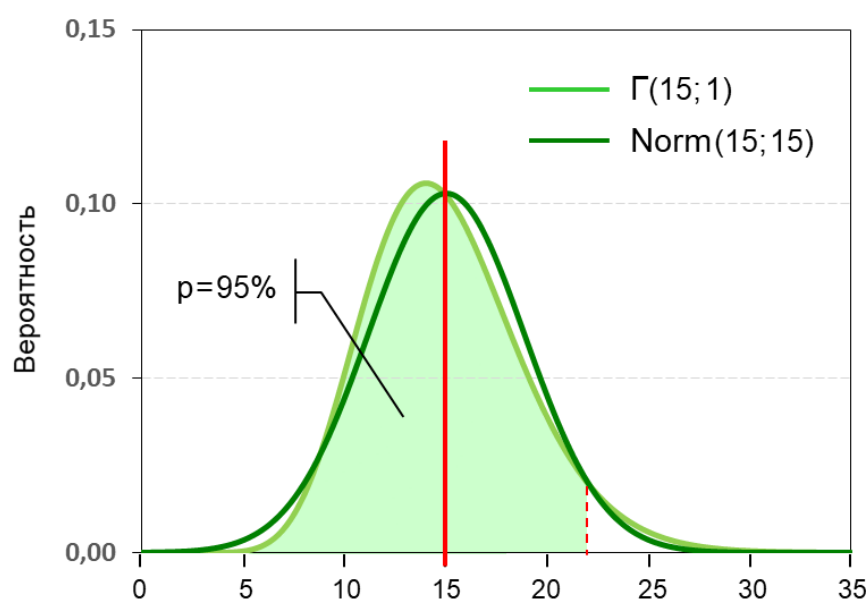


Рис. 3. Гамма-распределение и аппроксимирующее нормальное распределение для $k = 15$ дней.

На рис. 4 показана серия f_k для распределения Пуассона с параметром $\lambda = 0,25$ (то есть ситуация, когда наша SKU потребляется примерно один раз за четыре дня), со средним значением $\bar{x} = 0,25$ и дисперсией $\sigma^2 = 0,25$. Известно, что сумма k таких независимых случайных величин имеет распределение Пуассона с параметром $k\lambda$, то есть $f_k = P(0,25k)$. Следовательно, например, для $k = 15$ дней минимального запаса 6 у.е. будет достаточно при $p = 91\%$, а запаса 7 у.е. при $p = 96\%$ (см. рис. 5).

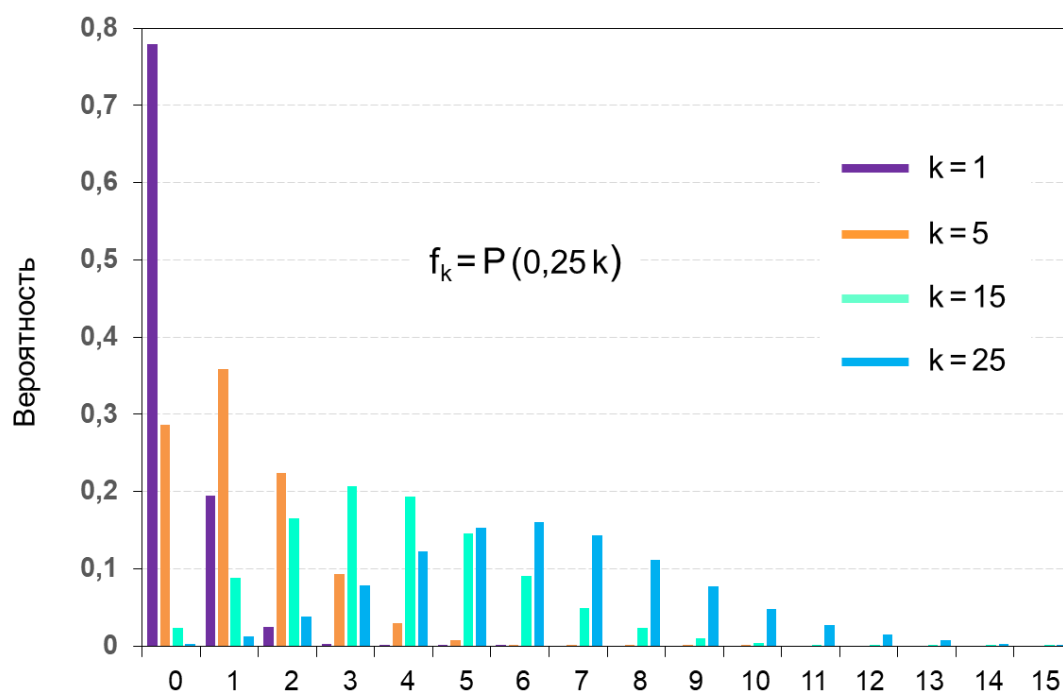


Рис. 4. Пример распределения вероятностей суммы k независимых случайных величин с распределением Пуассона при $\lambda = 0,25$.

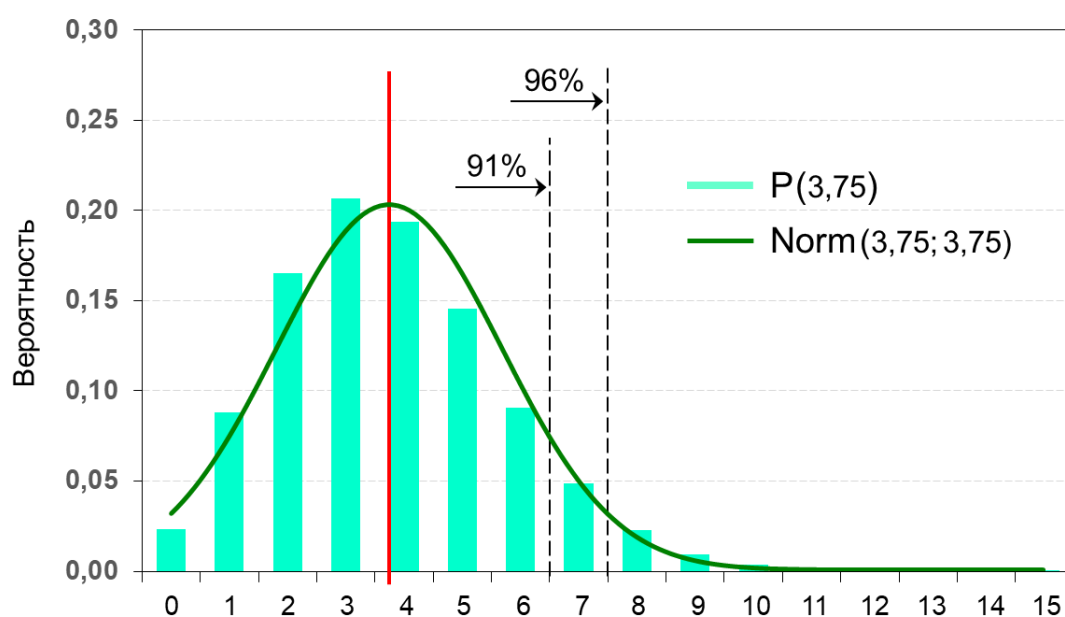


Рис. 5. Распределение Пуассона и аппроксимирующее нормальное распределение для $k = 15$ дней.

Более того, на основании центральной предельной теоремы теории вероятностей сумма k независимых одинаково распределённых случайных величин имеет распределение, близкое к нормальному, то есть $f_k \approx \text{Norm}(k\bar{x}; k\sigma^2)$. Для двух рассмотренных выше ситуаций это утверждение проиллюстрировано на рис. 3 и рис. 5. В случае непрерывного экспоненциального распределения $f_1 = \text{Exp}(1)$ квантиль уровня 0,9 соответствующего нормального распределения $\text{Norm}(15; 15)$ даёт для минимального размера запасов S_0 величину $\sim 20,0$ у.е., а для уровня 0,95 $\sim 21,4$ у.е. (напомню, при точных значениях 20,1 и 21,9). В случае дискретного распределения Пуассона с параметром 0,25 процентные точки аппроксимирующего нормального распределения $\text{Norm}(3,75; 3,75)$ для $p = 90\%$ составляют примерно 6,25 у.е., а для $p = 95\%$ – 6,95 (после округления до целого это будет 7) у.е., что совпадает с точным решением. При этом, несмотря на заметные отклонения при малых k самих распределений f_k от нормальных, значения их квантилей высокого уровня могут отличаться не так уж сильно.

На рис. 6 приведена диаграмма, показывающая отклонение квантиля уровня p от среднего значения (очевидно, одинакового) для рассмотренного выше гамма-распределения $f_k = \Gamma(k; 1)$ и аппроксимирующего его нормального $\text{Norm}(k; k)$ при k от 1 до 30 и четырёх вариантах доверительной вероятности p (80%,

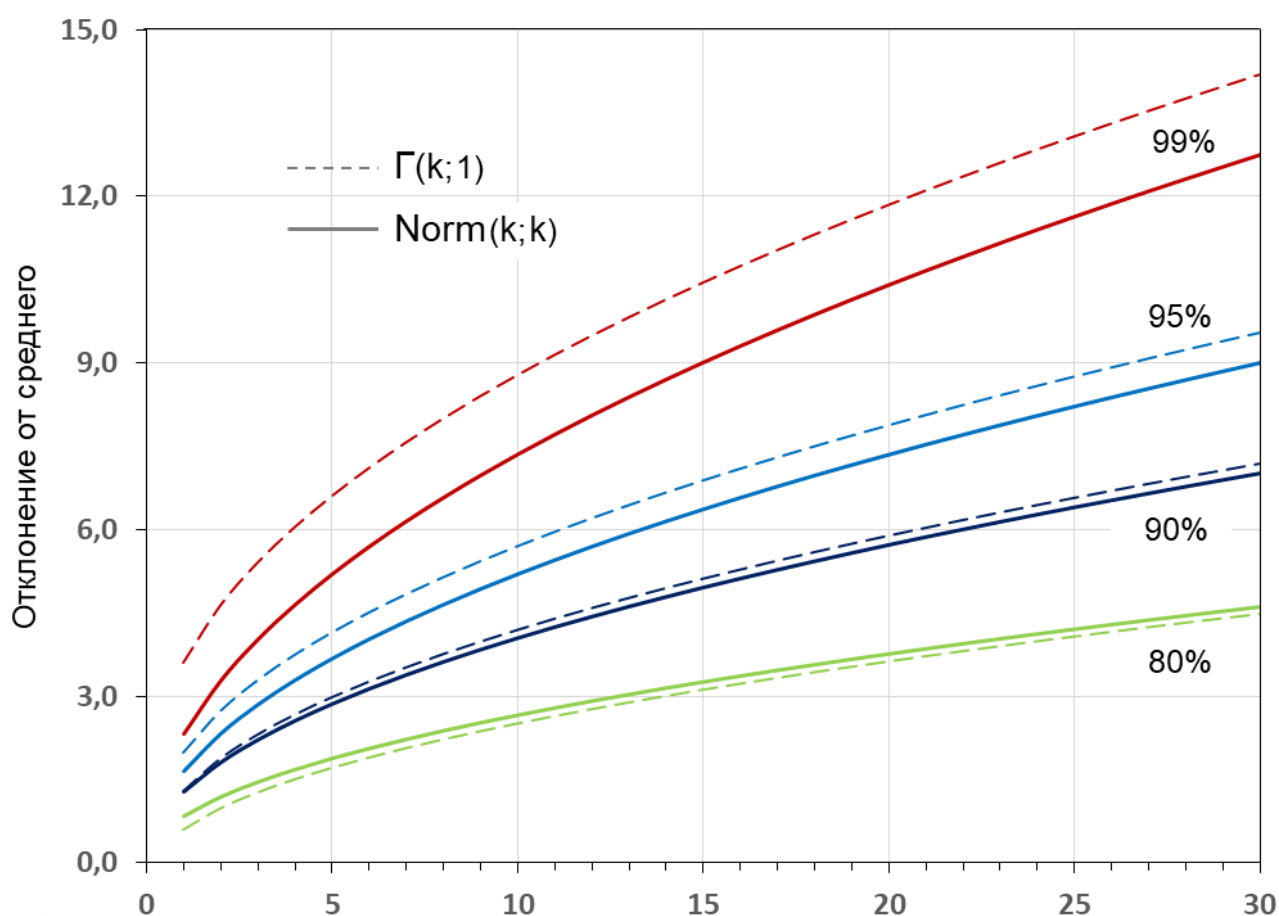


Рис. 6. Отклонение p -квантиля от среднего для гамма-распределения и аппроксимирующего нормального распределения в зависимости от k .

90%, 95% и 99%). Как видно, для p порядка 80–90% расхождение оказывается не существенным, а для $p = 95\%$ небольшим. При $p \sim 99\%$ и выше расхождение растёт, однако работа на самых краях распределений при их слишком малом «весе» лишена практического смысла.

На рис. 7 приведены аналогичные сведения по распределению Пуассона $f_k = P(0,25k)$ и аппроксимирующему его нормальному $\text{Norm}(0,25k; 0,25k)$, но в абсолютных величинах. Видно, что в целом картина такая же, а именно, достаточно хорошее качество приближения, – несмотря на ломаный характер линий для точного распределения из-за дискретности описываемой им случайной величины.

Реальные и обычно неизвестные распределения вероятностей спроса часто бывают асимметричными (сильно «скошенными» вправо или влево) и, кроме того, – в отличие от экспоненциального, – характеризуются коэффициентами вариации (отношением среднеквадратичного отклонения к среднему значению) c больше 1. Так, для рассмотренного выше распределения Пуассона с параметром 0,25 $c = 2$. В подобных ситуациях известны как минимум три способа оценки нужного p -квантиля по первым двум моментам распределения:

- (I) На основе неравенства Чебышёва⁶, справедливого для любого (непрерывного либо дискретного) распределения и постулирующего, что веро-

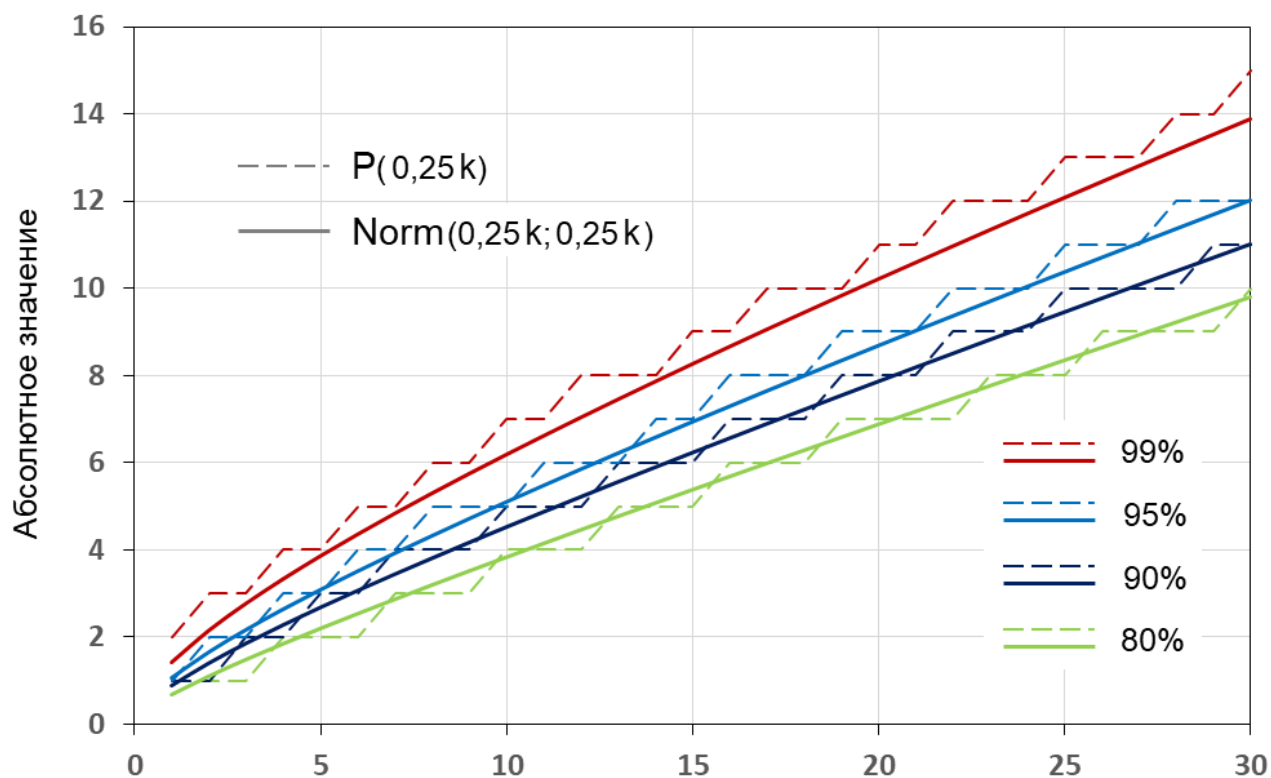


Рис. 7. Значение p -квантиля для распределения Пуассона с параметром 0,25 и аппроксимирующего нормального распределения в зависимости от k .

ятность отклонения случайной величины от своего среднего значения на величину больше $\delta\sigma$ не превышает $(1/\delta^2)$. Например, вероятность отклонения больше, чем на 3σ , будет не выше $1/9 \approx 11\%$.

- (II) На основе неравенства Высочанского–Петунина⁷, усиливающего результат Чебышёва для непрерывных унимодальных распределений и при $\delta > \sqrt{8/3} \approx 1,63$ на коэффициент $4/9$. В таком случае вероятность отклонения больше, чем на 3σ , будет уже не выше $(4/9) \cdot (1/9) \approx 5\%$.

Следует заметить, что несмотря на формальное условие унимодальности, 95-процентные границы 3-х сигмовых интервалов практически действуют и для более широкого класса распределений реального спроса. В качестве иллюстрации на рис. 8 приведен пример смеси двух нормальных распределений с единичной дисперсией, средними значениями $-0,8$ и $3,2$ и соотношением весовых коэффициентов $0,8:0,2$. Среднее значение этого бимодального распределения в точности равно 0 , дисперсия σ^2 – примерно $3,56$, значение квантиля уровня 95% составляет порядка 2σ .

Что касается дискретного спроса, то здесь положение дел не столь однозначно. На рис. 9 показаны результаты расчёта вероятности нахождения случайной величины в пределах 3-х сигмовых границ для распределений Пуассона с параметром λ из диапазона $[0,0; 1,0]$. Как видно, эта вероятность опускается ниже уровня 95% (но не ниже гарантированной неравенством Чебышёва величины 89%) только на небольшом интервале значений λ , – приблизительно от $0,052$ до $0,091$, – что соответствует ситуации довольно редкого потребления (примерно один раз за 10-20 дней). При ещё более низких и более высоких значениях λ , включая $\lambda > 1$, уровень 95% обеспечивается с хорошим запасом.

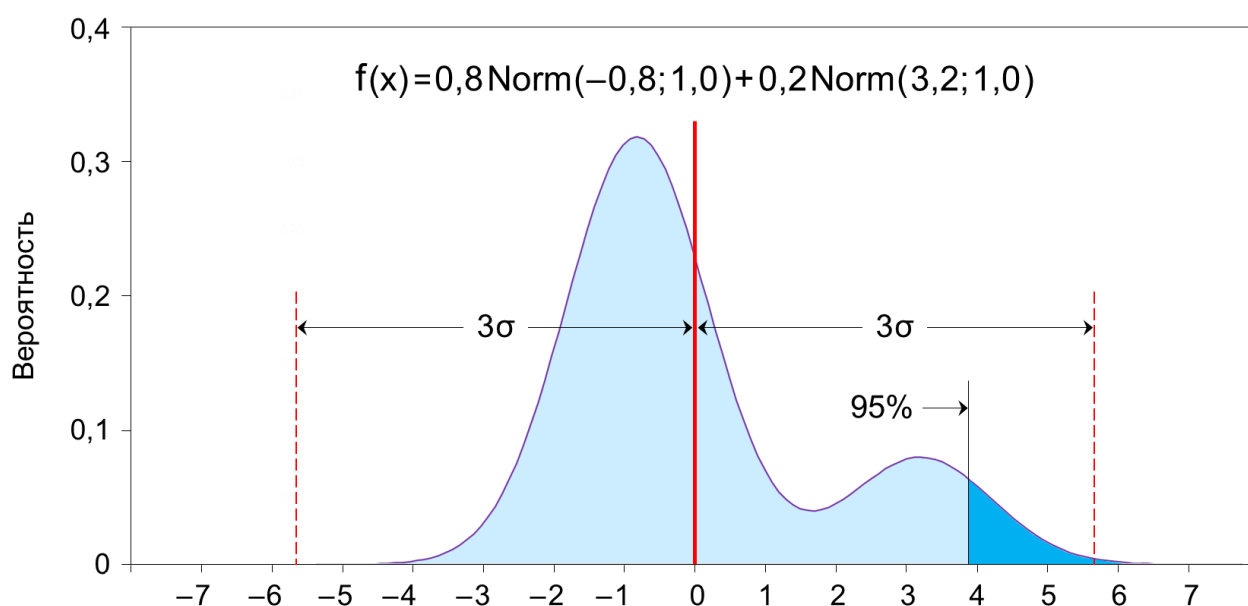


Рис. 8. 3-х сигмовые границы для смеси двух нормальных распределений.

(III) Поскольку для решения сформулированной задачи важны статистические характеристики не самого ежедневного потребления, а суммарного спроса за весь период между пополнениями, то универсальный способ состоит в применении центральной предельной теоремы теории вероятностей, где в качестве двух первых моментов аппроксимирующего нормального распределения используются их выборочные оценки. При этом, очевидно, чем больше время между поставками, тем выше точность получаемых результатов и тем меньше эти результаты зависят от вида исходных распределений (см. примеры на рис. 3 и рис. 5).

Иными словами, если по выборке рассчитаны среднее значение $\bar{x}$ и среднеквадратичное отклонение σ (квадратный корень из дисперсии, с.к.о.), а по ним определён коэффициент вариации $c = \sigma/\bar{x}$, то сумма k *независимых* случайных величин $(x/\bar{x})$ имеет распределение, близкое к нормальному с параметрами:

$$\text{среднее значение } x^*(k) = k; \quad \text{с.к.о. } \sigma^*(k) = c\sqrt{k}$$

Поэтому, взяв значение квантиля $\Omega^0(p)$ стандартного нормального распределения $\text{Norm}(0; 1)$, после соответствующего масштабирования получаем:

$$S_0 = \bar{x} k (1 + \Omega^0(p) c / \sqrt{k}),$$

где (для примера):

$$\Omega^0(80\%) = 0,842; \quad \Omega^0(90\%) = 1,282; \quad \Omega^0(95\%) = 1,645; \quad \Omega^0(99\%) = 2,326.$$

В таком случае появляется возможность ретроспективного анализа статистики запасов с целью грубой оценки их достаточности для обеспечения заданной вероятности отсутствия стокауты при известных характеристиках независи-

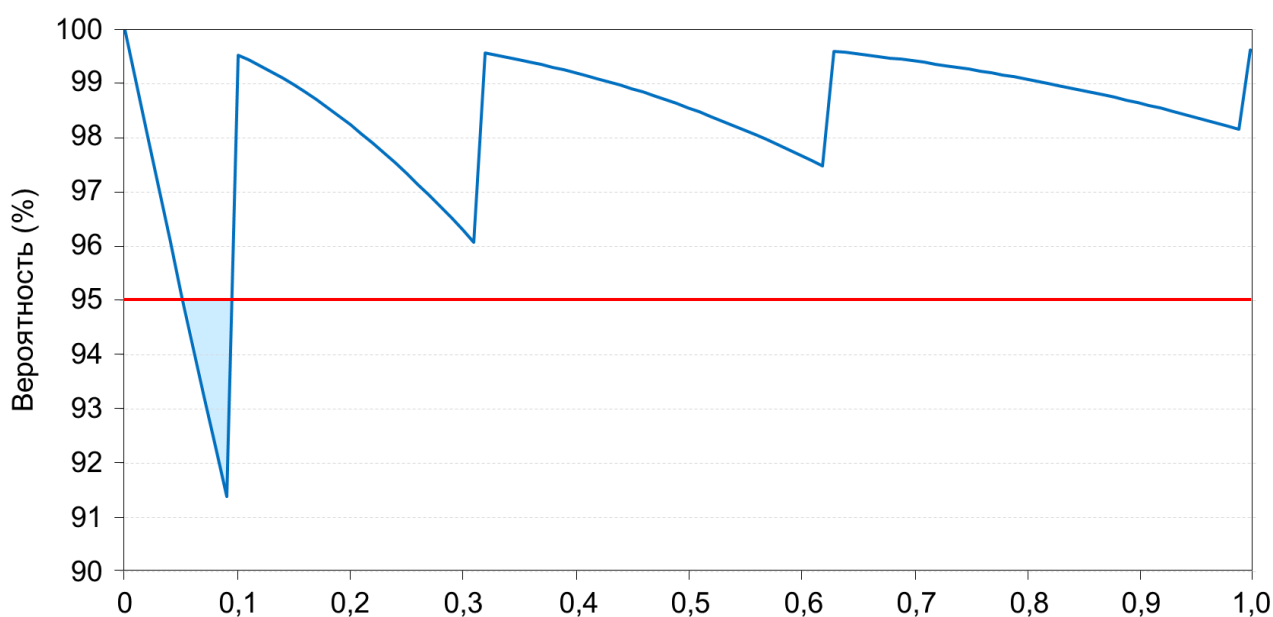


Рис. 9. Вероятность нахождения в пределах 3-х сигмовых границ для распределений Пуассона в зависимости от значения параметра λ .

мого спроса и интервалах между поставками. При этом «оптимальный» средний размер запасов определяется из соотношения:

$$^0S^* = \dot{x} k (1 + \Omega^0(p) c / \sqrt{k}) / 2,$$

Следовательно, если значение безразмерного «индекса оптимальности запасов»

$$I_s(k; c; p) \equiv S^* / ^0S^*$$

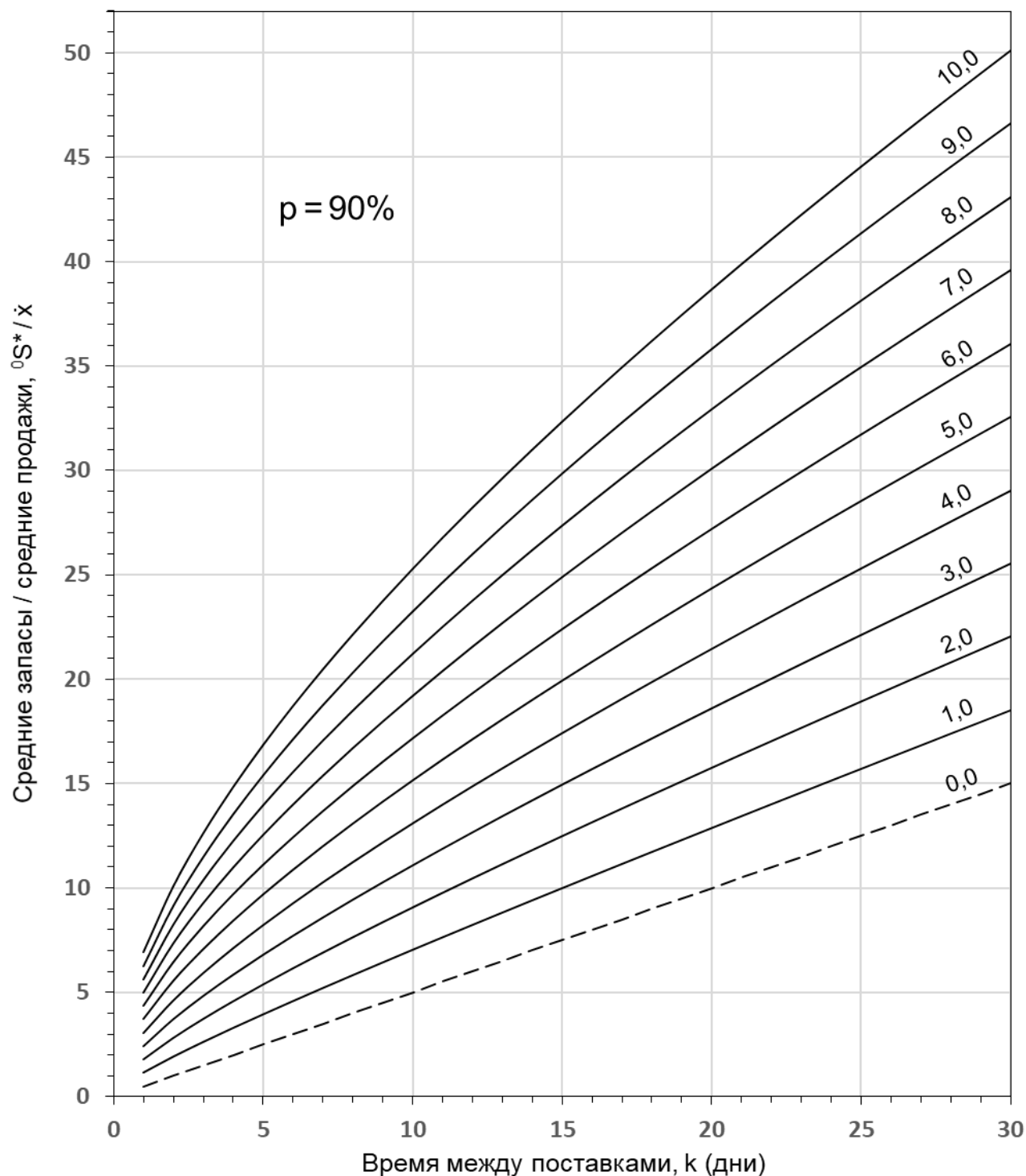


Рис. 10. «Оптимальный» средний размер запасов для $p = 90\%$ в зависимости от времени между поставками k и коэффициента вариации спроса c .

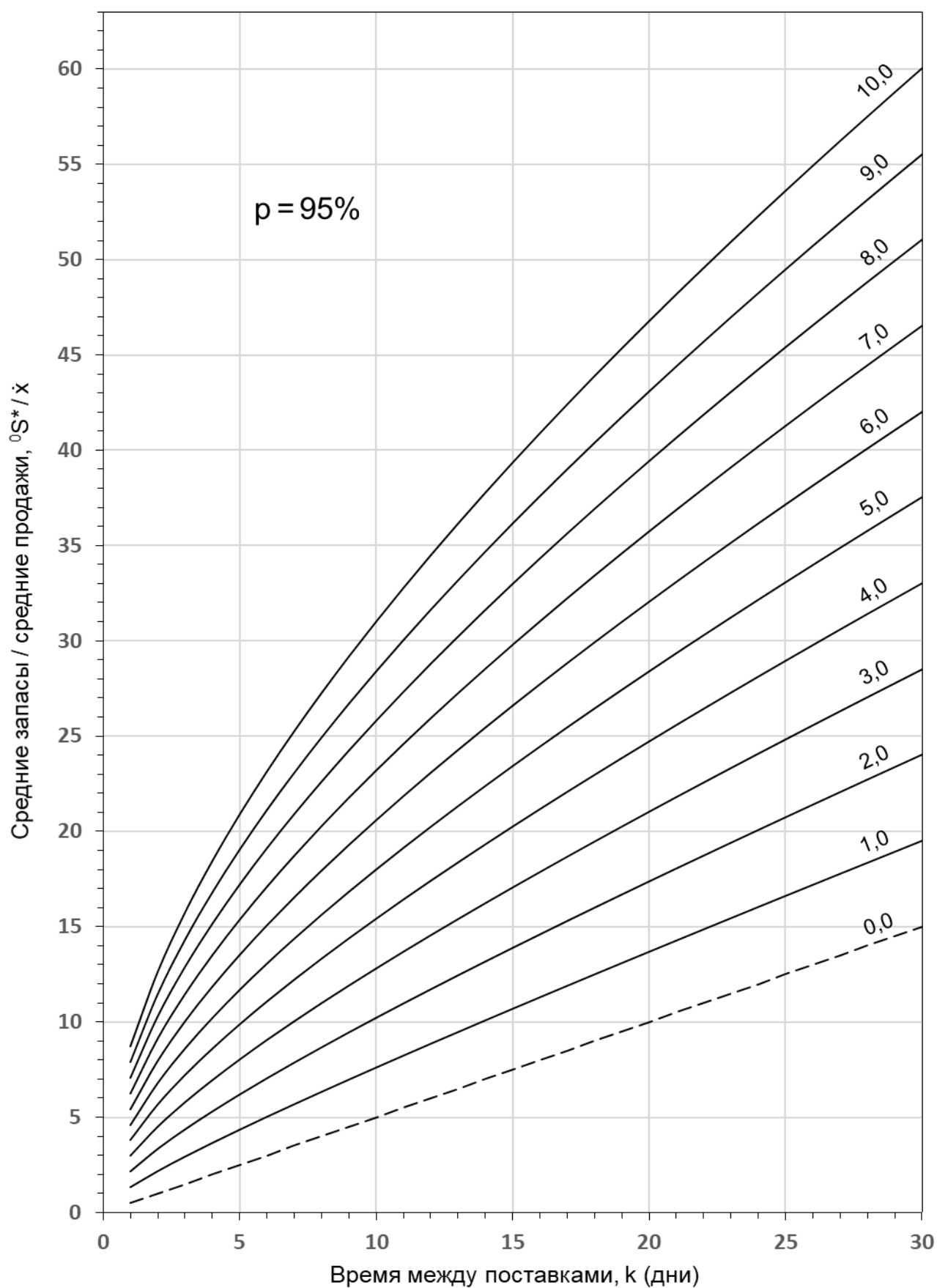


Рис. 11. «Оптимальный» средний размер запасов для $p = 95\%$ в зависимости от времени между поставками k и коэффициента вариации спроса s .

превышает единицу (хотя бы на 5-10%), то поддерживаемый средний размер запасов при заданном наборе исходных параметров, скорее всего, является избыточным. А если значение индекса меньше единицы (хотя бы на 5-10%), то для обеспечения заданной вероятности удовлетворения текущих потребностей таких запасов, скорее всего, недостаточно.

На рис. 10 и рис. 11 показаны результаты расчётов по полученной выше формуле «оптимального» относительного размера запасов ($^0S^*/\dot{x}$) как функции времени между поставками k (от 1 до 30 дней), – для широкого диапазона коэффициентов вариации c (от 0 до 10 с шагом 1; значения указаны рядом с соответствующими линиями на диаграммах) и двух вариантов доверительной вероятности: $p = 90\%$ (рис. 10) и $p = 95\%$ (рис. 11).

Кроме того, используя формулу «оптимального» среднего размера запасов, нетрудно решать обратные задачи, а именно, для любого случайного набора потребностей с известным средним $\dot{x}$ и коэффициентом вариации c при заданном значении S^* оценивать *среднее время между поставками* k , обеспечивающее желаемую доверительную вероятность p :

$$k = \left[\sqrt{\frac{[\Omega^0(p) c]^2}{4} + \frac{2S^*}{\dot{x}}} - \frac{\Omega^0(p) c}{2} \right]^2$$

либо для принятого среднего времени между поставками k оценивать величину квантиля $\Omega^0(p)$ стандартного нормального распределения $\text{Norm}(0; 1)$ и, следовательно, *вероятность* p удовлетворения потребностей, которую может обеспечивать заданное значение S^* :

$$\Omega^0(p) = \frac{\sqrt{k}}{c} \left(\frac{2S^*}{\dot{x}k} - 1 \right)$$

ЧТО В СУХОМ ОСТАТКЕ?

Представленные выше соображения теоретического характера позволяют осмысленно отвечать на некоторые важные практические вопросы управления запасами. Например, из рис. 10 и рис. 11 следует, что *в предположении о независимости ежедневных продаж* при среднем времени между поставками 15 дней и коэффициенте вариации спроса около 1,0 (ситуация типа экспоненциального распределения) для удовлетворения потребностей с вероятностью 90% вполне должно хватать 10-дневного среднего размера запасов, а с вероятностью 95% – 11-дневного. Однако при коэффициенте вариации 8,0 (ситуация типа распределения Пуассона с параметром $\lambda = 0,016$, что соответствует положению дел, когда одна единица товара продаётся в среднем один раз за два месяца) этого будет уже недостаточно. В таком случае необходимо поддерживать как минимум 27-дневный (90%) либо почти 33-дневный (95%) запас.

Если же вернуться к приведенным в начале заметки данным о продажах и запасах двух аптек, то по ним нетрудно получить следующие расчётные характеристики:

➤ Аптека 1 (диаграмма А рис. 1):

- текущий обеспечиваемый уровень сервиса (при $S^* = 46,2$ и $K = 10$): $\Omega^0(p) \approx 2,97$, то есть $p = 99,85\%$;
- «оптимальный» размер запасов (при $p = 95\%$ и $K = 10$): ${}^0S^* = 37,3$, то есть $I_s(95\%) = 1,24$;
- количество поставок за период (при $p = 95\%$ и $S^* = 46,2$): $k = 28,4$ дней, то есть $K = 243/k - 1 = 7,6$.

➤ Аптека 2 (диаграмма Б рис. 1):

- текущий обеспечиваемый уровень сервиса (при $S^* = 42,3$ и $K = 9$): $\Omega^0(p) \approx 1,11$, то есть $p = 86,7\%$;
- «оптимальный» размер запасов (при $p = 95\%$ и $K = 9$): ${}^0S^* = 46,8$, то есть $I_s(95\%) = 0,90$;
- количество поставок за период (при $p = 95\%$ и $S^* = 42,3$): $k = 21,6$ дней, то есть $K = 243/k - 1 = 10,3$.

Иными словами, подтверждаются первоначальные качественные выводы о завышенных запасах в первом случае и заниженных – во втором. Но теперь это можно обсуждать с конкретными цифрами в руках и количественно оценивать параметры альтернативных сценариев развития событий.

Прежде всего, поддерживаемый в первой аптеке размер запасов соответствует неимоверно высокому (почти 99,9%) уровню сервиса, тогда как для второй аптеки он составляет меньше 90%.

Возможные варианты изменения ситуации для обеспечения, например, 95-процентного уровня сервиса:

- в первом случае – при той же частоте поставок примерно на 20% сократить запасы (с 46,2 до 37,3 штук), либо при тех же запасах почти на 30% (с 22,1 до 28,4 дня) увеличить среднее время между поставками;
- во втором случае – при той же частоте поставок почти на 10% *увеличить* запасы (с 42,3 до 46,8 штук), либо при тех же запасах примерно на 10% (с 24,3 до 21,6 дня) сократить среднее время между поставками.

Напомню, что в рассматриваемом кейсе значительное увеличение частоты поставок было бы нежелательно, однако 10-процентный рост – вполне приемлемая цена за заметное повышение уровня сервиса. Что и требовалось доказать! Осталось только научиться всем этим управлять на практике.

ОТ ТЕОРИИ К ПРАКТИКЕ И ОБРАТНО

Как известно, – перефразируя классику, – «100-процентную гарантию даёт только страховой полис», но к управлению запасами это, похоже, не относится. В любой момент времени может нагрянуть клиент, который заберёт весь сток, ... и вот у вас уже стокаут. Поэтому о каких-либо гарантиях уместно говорить только в статистическом смысле.

Практические методы управления запасами в условиях неопределённости (типа Мин-макс, «точки перезаказа» или МТА), как правило, представляют собой различные версии и модификации традиционной «системы полных запасов» (Base Stock System)² и имеют набор настраиваемых параметров, так или иначе связывающих между собой срок, размер партий пополнения и величину страховки. При этом для учёта последней в явном или неявном виде (а также в меру допустимости предположения о независимости спроса) используется приведенное выше решение задачи теории вероятностей.

Например, в варианте МТА (Make-To-Availability, см. рис. 12) теории ограничений ключевыми элементами методики являются «буфер» и «целевой уровень»⁹. Буфер для SKU определяется суммой всех имеющихся в наличии складских запасов в пункте продаж и товаров «в пути» по сделанным ранее и находящимся на разных стадиях исполнения заказам $O_1, \dots, O_M$. Состояние системы регулярно контролируется, и если в какой-то момент времени фактический размер буфера оказывается ниже установленного значения целевого уровня, то формируется потребность на пополнение запасов в количестве не менее образовав-

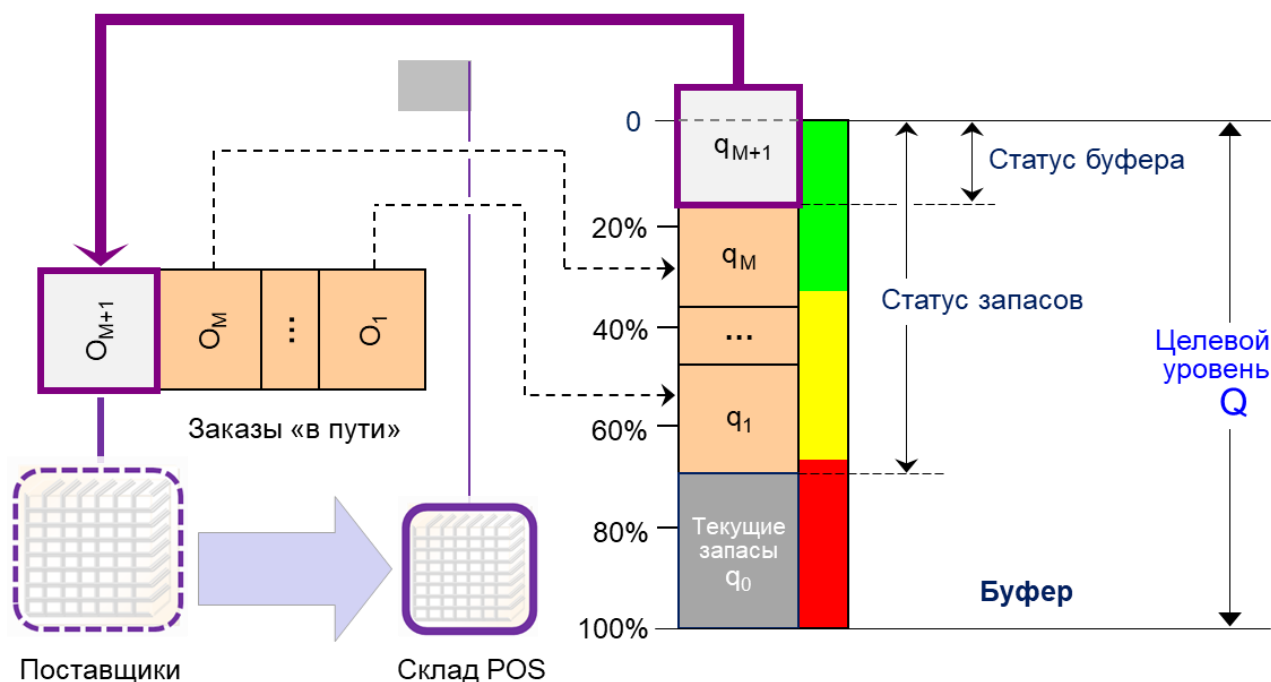


Рис. 12. Общая схема управления запасами по методике МТА.

шегося расхождения. В таком случае «статус буфера» или приоритет соответствующего заказа определяется следующим образом:

$$pr < O_{M+1} > = (1 - \sum_{i=0}^M q_i / Q) \times 100\%$$

либо принимается равным нулю при отрицательном расчётном значении, где Q – целевой уровень, q_0 – текущий размер запасов, q_i – размер заказа O_i , $i = 1, \dots, M$, а с целью визуализации приоритетов вводятся цветовые коды:

Зона	«Голубая»	«Зелёная»	«Жёлтая»	«Красная»	«Чёрная»
Приоритет (%)	0	1 – 33	34 – 66	67 – 99	100

В качестве дополнительного индикатора состояния системы иногда используется так называемый «статус запасов», рассчитываемый по формуле:

$$St = (1 - q_0 / Q) \times 100\%$$

(либо принимаемый равным нулю при отрицательном расчётном значении) и показывающий долю имеющихся в наличии запасов относительно целевого уровня.

Статическая схема МТА с постоянным целевым уровнем и стандартным размером партии пополнения обычно позволяет весьма эффективно управлять запасами и поставками в условиях стабильного независимого спроса. При этом целевой уровень определяется статистическими свойствами потока потребностей, – исходя из условия их удовлетворения на заданном уровне сервиса p в течение всего срока пополнения запасов. Иными словами, если $\dot{x}$ и c – среднее и коэффициент вариации спроса, k^* – срок пополнения, $\Omega^0(p)$ – квантиль уровня p стандартного нормального распределения, то из приведенных ранее теоретических выкладок следует, что в качестве «оптимальной» оценки целевого уровня разумно принять величину:

$$Q_0 = \dot{x} k^* (1 + \Omega^0(p) c / \sqrt{k^*})$$

В этой связи крайне странно выглядят приводимые даже в солидных источниках⁹ рекомендации об использовании вместо выражения в скобках фиксированного фактора в размере 1,5 или 2. (Очевидно, что почти при любой вариативности спроса для больших значений k^* заданный подобным образом целевой уровень становится избыточным, а для небольших – недостаточным.) Не говоря уже про литературу из серии «управление запасами для чайников»¹⁰ с предложениями брать вместо средних выборочных значений максимальные и измышлениями относительно применимости строгих теоретических подходов только для нормальных распределений. Короче, как ни крути, всё равно получается «пальцем в небо». А уж про «коэффициент паранойи», – пребывая в здравом уме и твёрдой памяти, – я вообще не берусь что-либо всерьёз обсуж-

дать. Обычно тут приводится встречный аргумент о возможности последующей корректировки начального (пусть и не вполне адекватного) целевого уровня посредством специальных эвристических приёмов. Но об этом чуть позже.

Другим важным параметром при управлении запасами является размер партии пополнения. Дело в том, что (как показано в теоретической части заметки) при известных статистических характеристиках спроса все составляющие вероятностной модели взаимосвязаны. В частности, по заданному уровню сервиса ρ и среднему времени между поставками k_0 однозначно определяются «оптимальные» средние запасы. Поэтому если в качестве стандартного размера партии пополнения для статической схемы установить, например, величину $L_0 = \bar{x}k_0$, то ситуация становится полностью замкнутой и контролируемой. Безусловно, лучше всего было бы формировать заказы на допоставку сразу же после и в размере каждой текущей продажи, но в реальной жизни, – прошу прощения за каламбур, – это часто нереально. Зато если имеются технологические либо организационные ограничения по частоте заказов, то их легко заложить в схему управления запасами в явном виде.

К сожалению, стабильность спроса также не слишком характерна для реальной жизни. И при значительных отклонениях от стационарности с неизвестными заранее трендами (включая аномальные выбросы и сезонные колебания) требуется более гибкий подход с адаптацией к изменениям условий внешнего окружения. Грубо говоря, при выявлении заметной тенденции к росту потребностей величина целевого уровня должна повышаться и, следовательно, при тех же размерах текущих запасов приоритеты формирования новых заказов на пополнение увеличиваться. А при проявлении признаков устойчивого снижения спроса должен действовать обратный эффект.

Динамическая схема MTA / DBM (Dynamic Buffer Management) предполагает реализацию подобного механизма с помощью средств и инструментов различной степени сложности и обоснованности. Чаще всего в этой связи произносятся сакраментальные фразы типа “too much green” (слишком много зелёного) или “too much red” (слишком много красного) и звучат предложения, соответственно, уменьшать либо увеличивать текущую величину целевого уровня, как правило, на треть⁹. Конечно, за неимением лучшего можно пользоваться и такими рекомендациями. Правда, всё равно придётся хоть как-то – более или менее формально – определяться с термином «слишком много». Естественно, здесь открывается широкий простор для всевозможных фантазий и вербальных интервенций, хотя встречаются и ссылки на практический опыт¹⁰.

Однако, если понимать суть вопроса и основы статистического прогнозирования, то (с известными оговорками относительно стационарности на характерных временах и эргодичности рассматриваемых случайных процессов) с пос-

тавленной задачей должны хорошо справляться обычные линейные фильтры. При этом выражение для целевого уровня легко модифицируется:

$$Q(t) = \dot{x}(t) k^* (1 + \Omega^0(p) c(t) / \sqrt{k^*})$$

где $\dot{x}(t)$ и $c(t)$ – оценки среднего значения и коэффициента вариации спроса в момент времени t . Первая из них в простейшем варианте может быть получена из разностных уравнений:

$$\dot{x}(0) = x_0, \dot{x}(t) = \alpha \dot{x}(t-1) + \beta x_t, t > 0,$$

где x_t – значение спроса в момент t , x_0 – начальное значение спроса (например, среднее за некоторый период времени в прошлом), а весовые коэффициенты α и β представляют собой дополнительные параметры настройки фильтра с соблюдением условия их нормировки $(\alpha + \beta = 1)$ ¹¹. Вторую величину в нулевом приближении можно зафиксировать в качестве постоянной характеристики вариабельности спроса $C(t) = C_0$.

В таком случае появляется возможность обойтись без спекуляций на содержании термина «слишком много (зелёного или красного)» и пересчитывать значение целевого уровня при каждом очередном сеансе анализа потребностей в пополнении запасов (например, ежедневно) на основании всей имеющейся к данному моменту информации. Что же касается встречающихся в литературе высказываний о нежелательности слишком частой корректировки, то – с моей точки зрения, – их следует рассматривать исключительно в контексте оправдания собственного неумения применять другие, теоретически более осмысленные подходы. А ссылки⁹ на эксперименты Деминга с воронкой и tampering (эффект дестабилизации системы при её перестройке по текущим выборочным значениям) здесь вообще не уместны, поскольку использование низкочастотных фильтров практически исключает резкие колебания управляющих параметров. Как известно, подобные статистические инструменты обычно достаточно уверенно «ловят» сезонность и «режут» выбросы.

На рис. 13 показаны результаты работы описанного алгоритма типа МТА / DBM для статистического управления запасами на приведенных в начале заметки данных по продажам двух аптек. Напомню, что для первой из них (диаграмма А рис. 13) «оптимальный» средний размер запасов (при $p = 95\%$ и $K = 10$) составляет 37,3 шт. и фактически превышен почти на четверть. При фиксированном значении параметра $k_0 = 22,1$ модель также без стокаута «опускает» его практически до оптимальной величины 37,4, увеличив количество поставок за период до $K_M = 11$ (оранжевая линия на диаграммах демонстрирует модельную динамику запасов). А для второй аптеки (диаграмма Б рис. 13) текущие запасы по сравнению с «оптимальным» значением 46,8 шт. (при $p = 95\%$ и $K = 9$) занижены примерно на 10%, и модель при $k_0 = 24,3$ «поднимает» их до величины 46,3, также на единицу увеличив количество поставок, но теперь уже при одном стокауте вместо 14 фактических.

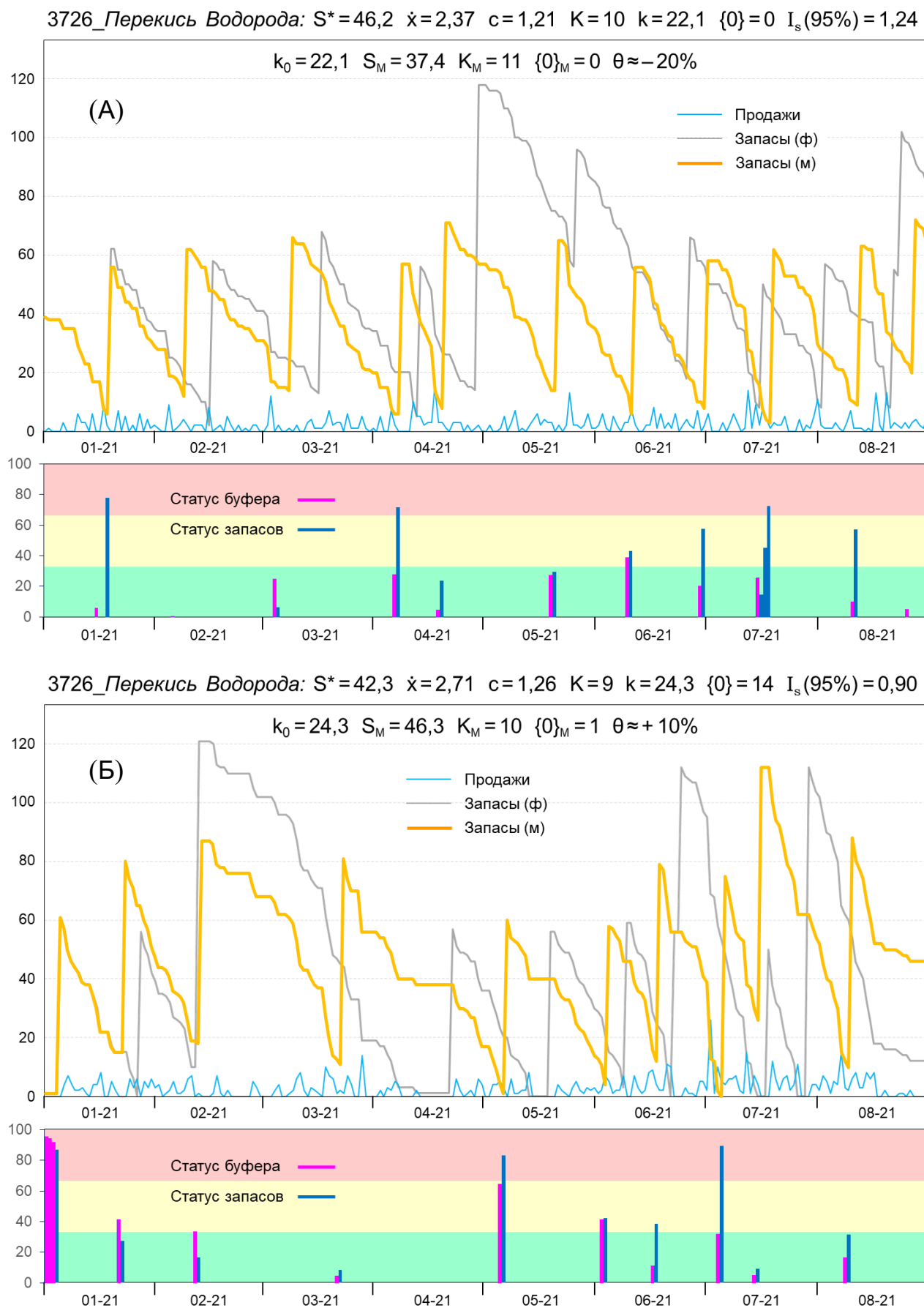


Рис. 13. Динамика продаж и запасов (факт и модель, две аптеки, одна SKU).

Во всех расчётах в качестве исходных значений X_0 и C_0 приняты соответствующие выборочные средние и коэффициенты вариации, для фильтров установлены нормировочные параметры $\alpha = 0,9$, $\beta = 0,1$, а начальные запасы подобраны таким образом, чтобы в течение первых 10 дней их динамика полностью совпадала с фактической. Если из-за возникшего стокаута размер продаж текущего дня оказывался меньше величины спроса, то «дефицит» добавлялся к потребностям последующих дней¹². Кроме того, с целью визуализации текущей ситуации и дополнительного обоснования необходимости и своевременности принятия модельных решений о пополнении запасов для каждого из двух примеров на рис. 13 добавлены карты изменения во времени статуса буфера и статуса запасов, определяемых на начало каждого операционного дня.

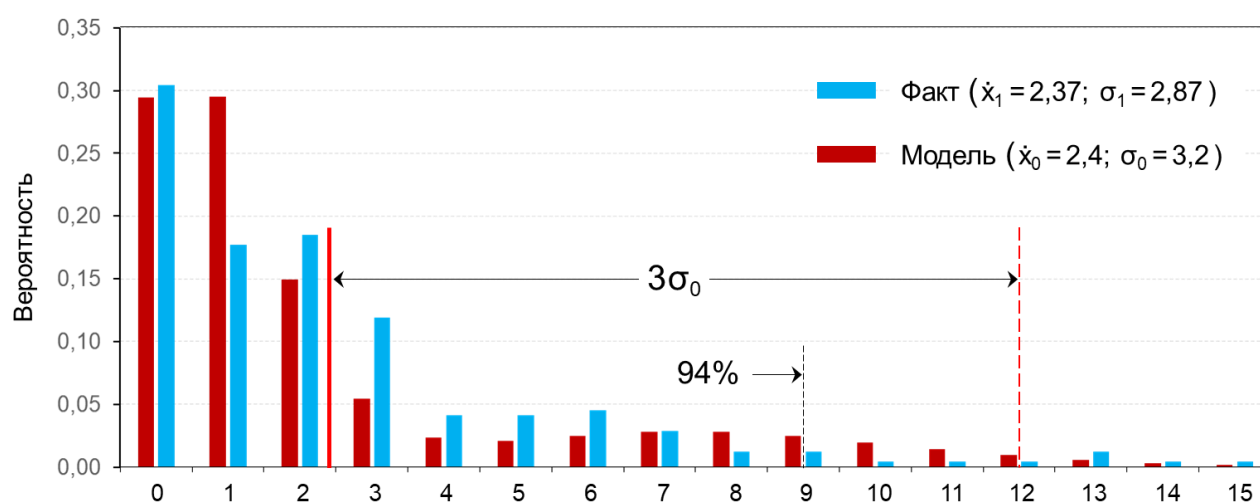


Рис. 14. Фактическое (для первой аптеки) и модельное распределение спроса.

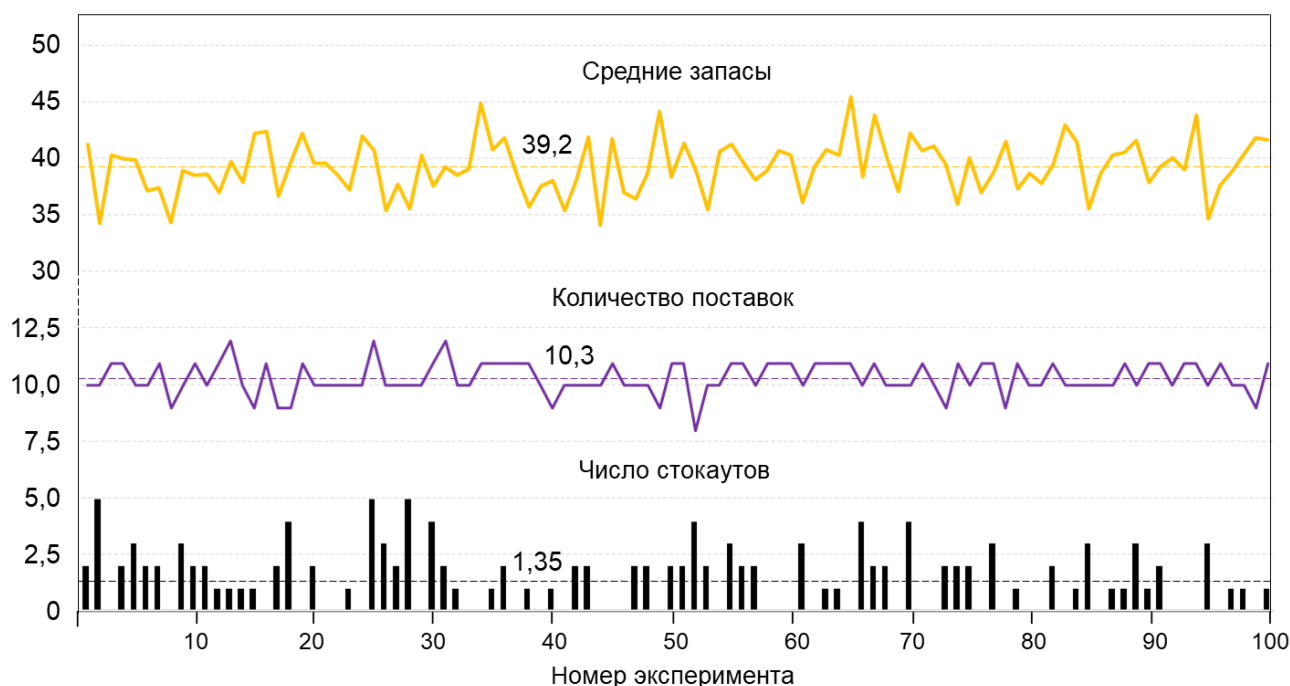


Рис. 15. Результаты имитационного моделирования.

Естественно, приведенные примеры следует рассматривать с поправкой на случайный характер фактических данных. Возможно, при других реализациях из той же генеральной совокупности показатели были бы хуже. Поэтому для подтверждения релевантности подхода проводился большой объём имитационного моделирования. Так, в одной из серий компьютерных экспериментов для описания потока потребностей, похожего на реальную динамику продаж первой аптеки, использовалась смесь двух распределений Пуассона с параметрами (средними значениями) 1,0 и 8,0 и соотношением весовых коэффициентов 0,8:0,2. Среднее значение $\bar{X}_0$ для такой смеси, очевидно, равно 2,4, среднеквадратичное отклонение $\sigma_0 = 3,2$. Модельное распределение в сравнении с эталонным показано на рис. 14 и, как видно, весьма далеко от нормального. В каждом из 100 экспериментов серии при одних и тех же управляющих параметрах (включая $p = 95\%$ и $K = 10$) генерировались 233 квазислучайных независимых величины. Представленные на рис. 15 результаты (по каждому эксперименту – количество поставок K_m , средние запасы S_m и число стокаутов $\{O\}_m$) свидетельствуют об устойчивой и стабильной работе алгоритма. В частности, количество поставок менялось от 8 до 12 со средним 10,3. Средний размер запасов в основном находился в диапазоне 35–45 штук со средним 39,2 (при теоретически «оптимальном» для данного набора параметров значении 38,9), а среднее число стокаутов составило 1,35 и ни в одном из экспериментов за весь период из 243 дней не превысило 5 штук.

ВЫВОД

Альберту Эйнштейну приписывают следующий афоризм:

«Теория — это когда всё известно, но ничего не работает. Практика — это когда всё работает, но никто не знает почему. Мы же объединяем теорию и практику: ничего не работает и никто не знает почему!»

Надеюсь, мне удалось показать, что по крайней мере в области управления запасами связка теории с практикой может работать вполне успешно.

ССЫЛКИ И КОММЕНТАРИИ

- ¹ Талёб Н. *О секретах устойчивости: Эссе*. – М.: КоЛибри, Азбука-Аттикус, 2012; с. 139.
- ² См., например: Hopp W., Spearman M. *Factory Physics*. – McGraw-Hill / Irwin, 2007 (для читателей с хорошей математической подготовкой), или Hopp W.J. *Supply Chain Science*. – Waveland Press Inc., 2011 (для читателей со средней математической подготовкой).
- ³ Жаринов С. *Об учётных и управляющих системах*. – http://www.lipro.ru/Load/8_Об_учётных_и_управляющих_системах.pdf

- 4 При этом следует иметь в виду, что если запасы по факту учитываются в конце каждого операционного дня, то их средний размер за k дней составляет $S^* = (k-1)/2$. А если запасы по факту учитываются в начале дня, то их средний размер за k дней составляет $S^* = (k+1)/2$.
- 5 Квантиль уровня p – значение, которое случайная величина не превышает с вероятностью p . Если p задана в процентах, то такую характеристику часто называют p -процентилем, но я и в этом случае пользуюсь термином “квантиль”.
- 6 Доказательство неравенства Чебышёва можно найти в учебниках, например: **Вентцель Е.С.** *Теория вероятностей*. – М.: Высш. шк., 2006; с. 287-290.
- 7 **Высочанский Д.Ф., Петунин Ю.И.** *Об одном неравенстве Гаусса для одновершинных распределений*. – Теория вероятностей и её применения, 1982, том 27, выпуск 2, с. 339-341.
- 8 Кроме того, при сравнении с фактическими сведениями о запасах, полученными по данным на конец операционного дня, необходимо учитывать поправку на особенность, указанную выше в комментарии (4), а именно, скорректировать теоретическое значение $^0S^*$ на величину $\Delta = \dot{x}/2$ в сторону уменьшения. А при использовании данных, полученных на начало операционного дня, целесообразно скорректировать $^0S^*$ на ту же величину Δ , но в сторону увеличения. В общем случае, когда не известно либо нет полной уверенности или однозначности в способе получения фактических данных, важно понимать, что «оптимальный» размер запасов может отклоняться от теоретического значения $^0S^*$ на величину $\pm \Delta$ и что при коротких интервалах времени между поставками такое отклонение может быть существенным.
- 9 См., например: ***Theory of Constraints Handbook*** (Edited by James F. Cox III & John G. Schleier, Jr.) – The McGraw-Hill Companies, Inc., 2010; p. 239-264.
- 10 См., например: **Егоров Д.** *Управление запасами в цепях поставок: Как обеспечить наличие и не морозить деньги на складах?* – [б. м.]: Издательские решения, 2021.
- 11 Описан вариант дискретного IIR-фильтра 1-го порядка (с бесконечной импульсной характеристикой), то есть действующего только на основании данных о потребностях в прошлом. Если же имеются сведения о будущем спросе (например, предварительные заказы либо надёжные прогнозы), то аналогичным образом можно применять комбинацию IIR и FIR-фильтров (с конечной импульсной характеристикой).
- 12 Дополнительная специфика рассматриваемого аптечного кейса определяется относительно короткими и календарно-зависимыми сроками поставок. Так, если заказ на пополнение сделан, например, в понедельник, то его, как правило, доставят на следующий день во вторник ($k^* = 2$ дня). Однако пятничный заказ, скорее всего, будет выполнен не раньше следующего понедельника ($k^* = 4$ дня). А если ещё добавятся праздники, то k^* может увеличиться до 5, 6 и более дней. При этом основные статистические свойства решения остаются неизменными.

Баошичи (ЧГ), декабрь 2021 г.