С. М. Ермаков, В. Б. Мелас, Н. К. Кривулин, В. Г. Курочка

Об одной модели продаж товаров с ограниченным сроком реализации*

В статье предлагается подход к сравнению различных алгоритмов определения оптимальных объемов поставки товаров (продуктов) с ограниченным сроком реализации на пункты продажи. Мы будем рассматривать ситуацию, когда поставка осуществляется ежедневно и продукты, не проданные в течение дня, уничтожаются. При этом могут возникнуть два вида потерь: 1) потери, связанные с уничтожением продуктов и со стоимостью уничтоженных продуктов, 2) потери, связанные с упущенным в результате недопоставки продуктов спросом. В связи с тем, что первый вид потерь может иметь нелинейный характер, а также с тем, что перечисленные потери не полностью характеризуют ситуацию (например, в результате недопоставки можно надолго потерять некоторую часть потенциальных покупателей), мы предположим, что сравнение различных алгоритмов назначения объемов поставки осуществляется экспертным путем. Будем считать, что эксперты могут осуществить такое сравнение, рассматривая зависимость отношения количества поставленного продукта к количеству проданного продукта от доли пунктов продажи, на которых не весь товар был распродан. Задачей настоящей работы является разработка метода построения такой зависимости.

Постановка задачи

Пусть имеется N пунктов продажи и рассматриваются результаты продаж за n дней. Обозначим через L_{ij} количество поставленных товаров, а через V_{ij} – количество проданных товаров в i-й день для j-го пункта, $i=1,\ldots,n,\ j=1,\ldots,N$.

Для каждого $i=1,\ldots,n$ введем средние величины

$$\overline{L}_i=\sum_{j=1}^{N}L_{ij}\Big/N\,;\overline{V}_i=\sum_{j=1}^{N}V_{ij}\Big/N\,;\overline{x}_i=1-\sum_{j=1}^{N}\chi_{\{L_{ij}=V_{ij}\}}\Big/N\,, \tag{1}$$

где $\chi_{\{a=b\}}=1$, если $a=b$, и $\chi_{\{a=b\}}=0$, если $a\neq b$, а также математические ожидания

$$L_i=E\overline{L}_i,\quad V_i=E\overline{V}_i,\quad x_i=E\overline{x}_i.$$

Цель исследования состоит в нахождении зависимости f вида $y_i=f(x_i)$, где $y_i=L_i/V_i$, $i=1,\ldots,n$, на основе величин $\{\overline{x}_i,\overline{L}_i,\overline{V}_i\}$.

Обозначим через Z_{ij} спрос на продукт в i-й день на j-м пункте продажи. Тогда, очевидно, $V_{ij}=\min\{Z_{ij},L_{ij}\}$.

---

* Работа выполнена при поддержке грантов РФФИ-ННИО № 96-01-00015 и РГНФ № 00-02-00228-А.

Регрессионная модель

Будем искать зависимость $\overline{y}_i$ от $\overline{x}_i$ в форме уравнения регрессии

$$\overline{y}_i = f(\overline{x}_i, \theta) + \varepsilon_i, \tag{2}$$

где f – некоторая функция, θ – неизвестный параметр (возможно, векторный), $E\varepsilon_i = 0$, $i = 1, \ldots, n$.

Будем называть $f(\overline{x}_i, \theta)$ моделью отклика, а дисперсионную матрицу W величин ε_i – моделью ошибок.

Пусть $i \in \{1, \ldots, n\}$ фиксировано. Введем обозначения:

$$x = x_i, \quad y = y_i, \quad a = a_i = E\sum_{j=1}^{N} Z_{ij} / N, \quad b = b_i = L_i - a_i$$

Естественно предположить, что усредняемые величины являются однородными. Поэтому будем считать, что вероятность $P\{Z_{ij} < L_{ij}\}$ не зависит от j. Взяв математическое ожидание от обеих частей (1), получим, что эта вероятность равна x_i.

Зависимость y от x определяется только распределениями $\{L_{ij}, Z_{ij}\}$ (а не структурой соответствующих временных рядов) и может быть установлена в общем виде. Однако, для определенности, мы будем предполагать, что эти величины имеют нормальное распределение. Фактически нам потребуется задать распределение разности $L_{ij} - Z_{ij}$.

Пусть $L_{ij} - Z_{ij} \sim N(b, \sigma^2 a^2)$, $j = 1, \ldots, N$, где $N(b, \sigma^2 a^2)$ – нормальное распределение со средним b и дисперсией $\sigma^2 a^2$, $b = b_i$, $a = a_i$, σ – неизвестный параметр.

Используя лемму (см. приложение), получаем, что

$$y = \frac{a + b}{a + b - xb - \sigma a h(x)},$$

где $h(x) = (1/\sqrt{2\pi})\exp\{-(\gamma(x))^2/2\}$, функция $\gamma(x)$ определяется соотношением

$$x = \Phi(x) = \int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du,$$

т. е. $\gamma(x) = \Phi^{-1}(x)$. Кроме того, $\sigma a = b/\gamma(x)$.

Разделив числитель и знаменатель на a и положив $\sigma = \theta$, будем иметь

$$y = \frac{1 + \theta\gamma(x)}{1 + (1 - x)\theta\gamma(x) - \theta h(x)} \tag{3}$$

Таким образом, получена зависимость $y = f(x)$, содержащая один неизвестный параметр σ. Очевидно, что чем меньше этот параметр, тем ближе лежат объемы поставок к реальному спросу.

Для идентификации модели необходимо оценить величину этого параметра по результатам продаж.

Представим модель (2) в виде $\overline{y}_i = f(x_i,\theta)+\varepsilon_i'$, где $f(x_i,\theta)$ равно правой части (3), $\varepsilon_i' = \varepsilon_i + f(\overline{x}_i,\theta) - f(x_i,\theta)$ для всех $i=1,\ldots,n$.

Нетрудно видеть, что $E\varepsilon_i' = E[f(\overline{x}_i,\theta) - f(x_i,\theta)] \neq 0$.

Ясно, что для достаточно больших N величина $E\varepsilon_i'$ будет мала, так как $\overline{x}_i \to x_i$ с вероятностью 1 при $N \to \infty$. Следовательно, обычные оценки метода наименьших квадратов должны обладать хорошими свойствами при больших N. Качество этих оценок целесообразно изучить в имитационном эксперименте, так как их теоретическое изучение (см. [1]) носит лишь асимптотический характер.

Алгоритм и результаты имитационных экспериментов

Поведение величин x_i, y_i может быть изучено с помощью следующего алгоритма.

1. Положим $i=1$.
2. Промоделируем $2N$ независимых случайных величин
$$Z_{i1},\ldots,Z_{iN} \in N(a+b,\sigma_1),\, L_{i1},\ldots,L_{iN} \in N(a,\sigma_2).$$
3. Вычислим $V_{ij} = \min\{Z_{ij},L_{ij}\}$, $j=1,\ldots,N$,
$$\overline{L}_i = \sum_{j=1}^{N}\frac{L_{ij}}{N},\quad \overline{V}_i = \sum_{j=1}^{N}\frac{V_{ij}}{N},\quad \overline{x}_i = 1-\sum_{j=1}^{N}\frac{\chi_{\{L_{ij}=V_{ij}\}}}{N},\quad \overline{y}_i = \frac{\overline{L}_i}{\overline{V}_i}.$$
4. Положим $i=i+1$ и, если $i\le n$, то вернемся к п. 2.
5. Оценим параметр θ по методу наименьших квадратов.

Результаты подгонки модели для $N=100$, $n=50$, $a=100$, $\sigma_1=\sigma_2=\sigma/\sqrt{2}$ представлены в виде графиков на рис. 1–4.

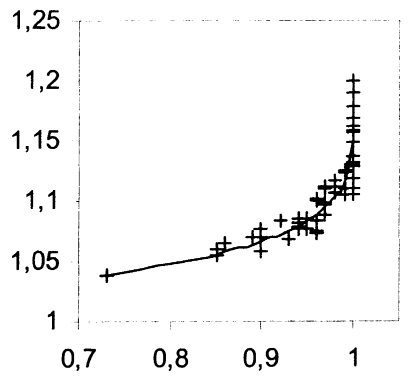

Рис. 1. Результаты подгонки для $\sigma=5$, $b=2\sigma$

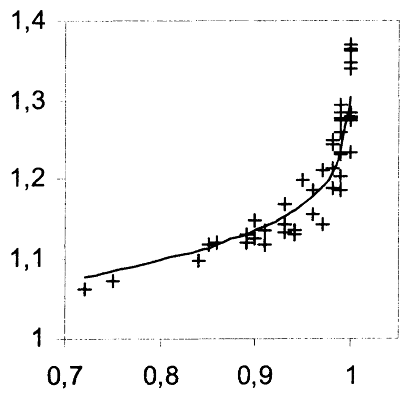

Рис. 2. Результаты подгонки для $\sigma=10$, $b=2\sigma$

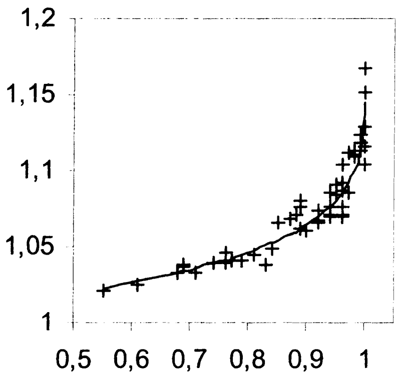

Рис. 3. Результаты подгонки для $\sigma = 5$, $b = 1{,}5\sigma$

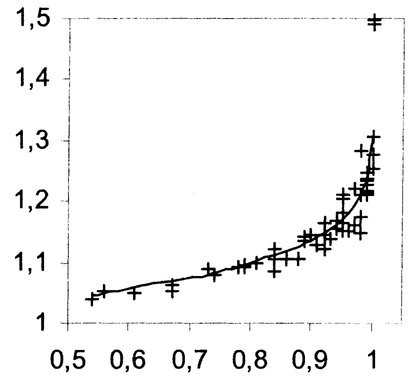

Рис. 4. Результаты подгонки для $\sigma = 10$, $b = 1{,}5\sigma$

Соответствующие оценки $\overline{\theta} = \overline{\sigma}$ по методу наименьших квадратов, полученные по результатам имитационных экспериментов, приведены в следующей таблице:

σ	***b***	
	1,5σ	2σ
5,0	0,04830	0,04981
10,0	0,10183	0,10096

Эти результаты наглядно показывают, что модель хорошо описывает данные.

Приложение

Лемма. Пусть ξ, η – независимые нормально распределенные случайные величины, $E\xi = c$, $E\eta = d$ и $D(\xi - \eta) = \sigma^2$. Тогда

$$E\left(\min\{\xi,\eta\}\right) = cx + d(1-x) - \frac{\sigma}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\gamma(x)\right)^2\right\}, \tag{4}$$

где $x = P\{\xi < \eta\} = \Phi\left((d-c)/\sigma\right)$, $\gamma(x) = \Phi^{-1}(x)$.

Доказательство. Пусть $\zeta = \min\{\xi,\mu\}$. Представим ζ в виде

$$\zeta = \eta + (\xi - \eta)_-,$$

где $(\xi - \eta)_- = \min\{0, \xi - \eta\}$.

Пусть $\Delta = c - d$. Так как $\xi - \eta \in N(c - d, \sigma^2)$, то

$$E\zeta = E\eta + E(\xi - \eta)_- = d + \frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty u\exp\left\{-\frac{1}{2}\left(\frac{u-\Delta}{\sigma}\right)^2\right\}du. \tag{5}$$

Представим второе слагаемое в (5) в виде

$$\frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty u\exp\left\{-\frac{1}{2}\left(\frac{u-\Delta}{\sigma}\right)^2\right\}du-\Delta\int_0^\infty\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{u-\Delta}{\sigma}\right)^2\right\}du+$$

$$+\Delta\int_0^\infty\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{u-\Delta}{\sigma}\right)^2\right\}du=$$

$$=\frac{\sigma}{\sqrt{2\pi}}\int_0^\infty\exp\left\{-\frac{1}{2}\left(\frac{u-\Delta}{\sigma}\right)^2\right\}d\left(\frac{1}{2}\left(\frac{u-\Delta}{\sigma}\right)^2\right)+\Delta\Phi\left(\frac{d-c}{\sigma}\right)=$$

$$=-\frac{\sigma}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(\gamma(x))^2\right\}+(c-d)\Phi\left(\frac{d-c}{\sigma}\right).$$

Кроме того,

$$P\{\xi-\eta<0\}=\frac{1}{2\pi\sigma}\int_0^\infty\exp\left\{-\left(\frac{u-\Delta}{\sigma}\right)^2\right\}du=\Phi\left(\frac{d-c}{\sigma}\right).$$

Подставляя полученные выражения в формулу (5), имеем

$$E\zeta=E(\min\{\xi,\eta\})=cx+d(1-x)-\frac{\sigma}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(\gamma(x))^2\right\}.$$

Литература

1. *Sebe G. A. F., Wil C. J.*, Nonlinear Regression. New York: Wiley, 1989.