Статистическая модель потребительского выбора

Е.Е. Голденок

Красноярский государственный торгово-экономический институт
ул. Лиды Прушинской 2, Красноярск, 660075
email: golde@rambler.ru

Исходные предположения

Ассортимент товаров современного супермаркета составляет 10–15 тысяч наименований. Ежедневная статистика продаж в супермаркете дает в результате случайный вектор из 10–15 тысяч зависимых компонент, имеющих смысл сумм стоимостей соответствующих товаров, проданных в супермаркете за день. Сложная структура зависимостей и взаимодействий компонент подобных случайных векторов представляет собой основную трудность в анализе и моделировании статистических систем потребительского выбора. Применение разработанных случайно – множественных методов позволяет существенно упростить анализ.

Исходные предположения. Статистические наблюдения за продажами товаров в супермаркетах показывают три основные особенности потребительского выбора:

- стоимости покупок любых товаров отделены от нуля некоторым интервалом;
- любой товар не покупается с вероятностью, строго большей нуля;
- частота элементарных покупок обратно пропорциональна их стоимости.

Кроме этого в работе делается общее предположение, что товарный рынок и случайный потребитель находятся в состоянии *статистического равновесия.*

Олег Воробьёв (ред.), Финансово-актуарная математика и смежные вопросы, Труды ФАМ'2002, Красноярск, ИВМ СО РАН

Данные особенности потребительского выбора приводят к трём исходным предположениям, положенным в основу *статистической модели потребительского выбора*:

- гиббсовская случайная величина
  — модель стоимости элементарной покупки,
  а сумма гиббсовских случайных величин
  — модель стоимости покупки;

- гиббсовский случайный вектор
  — модель потребительского выбора;

- гиббсовское случайное множество
  — модель равновесного случайного потребителя.

Методы исследования структур зависимостей и взаимодействий в статистической модели потребительского выбора. Структуру зависимостей и взаимодействий компонент произвольного гиббсовского случайного вектора можно разбить на два уровня. В соответствии с принятыми предположениями на первом уровне в качестве случайно-множественного базиса гиббсовского случайного вектора предлагается использовать гиббсовское случайное множество, которое описывает полную структуру зависимостей и взаимодействий между случайными событиями — покупками товаров, а на втором уровне в качестве количественной надстройки условных распределений предлагается использовать распределение гиббсовских случайных векторов в предположении независимости их компонент. Таким образом, сложная структура зависимостей и взаимодействий компонент исходного гиббсовского случайного вектора разбивается на две: первая учитывает все случайно – множественные зависимости между покупками подмножеств товаров, а вторая предполагает количественную независимость стоимостей покупок (компонент исходного случайного вектора) при условии покупки подмножеств товаров.

Гиббсовский случайный потребитель

Под событиями в статистической системе потребительского выбора понимаются покупки (продажи) тех или иных товаров, обращающихся на рассматриваемом товарном рынке и образующих конечное множество наименований товаров. Обозначение $\mathfrak{X}$ будет использоваться и для множества событий — покупок товаров, и для множества наименований соответствующих товаров, обращающихся на рынке. Считается, что $x \in \mathfrak{X}$ — это событие, заключающееся в покупке товара с наименованием x.

Гиббсовским случайным потребителем на рынке множества товаров $\mathfrak{X}$ с *функцией средней стоимости* подмножеств товаров $H(X)$, с *обратным уровнем среднего дохода* β и с *распределением вероятностей собственных вкусов и предпочтений* $u(X)$ называется гиббсовское случайное множество событий K с функцией статистической энергии $H(X)$, с обратной статистической температурой β и собственным распределением $u(X)$.

Таким образом, гиббсовский случайный потребитель определяется под множеством событий $\mathfrak{X}$ распределением вероятностей:

$$p(X) = P(K = X) = \frac{1}{Z_u} \cdot \exp\{-\beta H(X)\} u(X), \quad X \in 2^{\mathfrak{X}},$$

где

$$Z_u = \sum_{X \in 2^{\mathfrak{X}}} \exp\{-\beta H(X)\} u(X)$$

— нормирующий множитель.

Утверждение (закаливание и плавление гиббсовского случайного потребителя). *Пусть K — гиббсовский случайный потребитель на рынке товаров $\mathfrak{X}$ с функцией средней стоимости подмножеств товаров $H(X)$, обратным средним уровнем дохода β и вероятностями собственных вкусов и предпочтений $u(X)$. Тогда*

$$\lim_{\beta \to \infty} p^{\beta}(X)/u(X) = \begin{cases} u(X)/u_0, & X \in \mathcal{M}_0, \\ 0, & \textit{иначе}, \end{cases}$$

$$\lim_{\beta\to 0} p^\beta(X) = u(X), \quad X \in 2^{\mathfrak{X}},$$

Здесь

$$\mathcal{M}_0 = \{X : H(X) = \min_Y H(Y)\} \subseteq 2^{\mathfrak{X}},$$

— совокупность подмножеств товаров, на которых функция средней стоимости принимает минимальное значение *(совокупность минимальных потребительских корзин)*, а

$$u_0 = \sum_{X\in\mathcal{M}_0} u(X)$$

— собственная вероятность совокупности минимальных потребительских корзин $\mathcal{M}_0$.

Следствие из теоремы 3 показывает, что, как и для гиббсовского случайного множества можно говорить о "закаливании" и "плавлении" гиббсовского случайного потребителя. Закаливание (снижение статистической температуры гиббсовского случайного множества) в рыночной терминологии означает снижение среднего дохода случайного потребителя, а плавление (повышение статистической температуры) означает повышение его среднего дохода. При снижении среднего дохода потребителя, его предельное распределение сосредоточивается на минимальных потребительских корзинах (с минимальной средней стоимостью) и становится на них пропорциональным вероятностям его вкусов и предпочтений, а при повышении — начинает совпадать с собственными вероятностями вкусов и предпочтений гиббсовского случайного потребителя.

Подмножества с минимальной средней стоимостью представляют собой совокупность подмножеств товаров первой необходимости (минимальных потребительских корзин). Случайный потребитель, придя на рынок с вероятностным распределением своих вкусов и предпочтений $u(X)$, старается привести его в соответствие со средними стоимостями подмножеств товаров $H(X)$, стремясь как можно меньше отклониться от своего собственного распределения, но при этом как можно больше снизить среднюю стоимость покупаемого им подмножества товаров. Стремясь снизить среднюю стоимость купленного подмножества товаров, случайный потребитель как будто максимизирует энтропию своего распределения и поэтому его предельное распределение имеет гиббсовский вид $p(X)$ (по теореме 2).

Гиббсовский случайный потребитель отличается от каждого конкретного потребителя тем, что он, как совокупность всех конкретных потребителей, полностью статистически характеризует данный товарный рынок. Меняя параметры рынка можно моделировать поведение гиббсовского случайного потребителя. В отличие от каждого конкретного потребителя, доход которого не подвластен рынку, доход гиббсовского случайного потребителя может меняться, например, в зависимости от средних стоимостей товаров на рынке.

Применение двухуровневой структуры зависимостей и взаимодействий в алгоритмах статистической оценки распределения гиббсовского случайного потребителя

В силу сделанных предположений в качестве модели гиббсовского случайного потребителя в работе рассматривается гиббсовский случайный вектор. Ключевое свойство гиббсовского случайного вектора заключается в том, что его компоненты обладают так называемыми слепыми интервалами, отделяющими их ненулевые значения от нуля. Это свойство позволяет применить теорему 3 о разложении распределения гиббсовского случайного вектора по случайно – множественному базису. Таким образом, распределение гиббсовского случайного вектора можно представить в виде двухуровневой структуры зависимостей и взаимодействий его компонент. Это позволяет существенно упростить методы статистической оценки распределения гиббсовского случайного вектора. Более того, в силу следствия из теоремы 3 двухуровневая структура может быть использована не только для гиббсовских случайных векторов, а для любых векторов с неотрицательными компонентами, в распределении которых имеется слепой интервал.

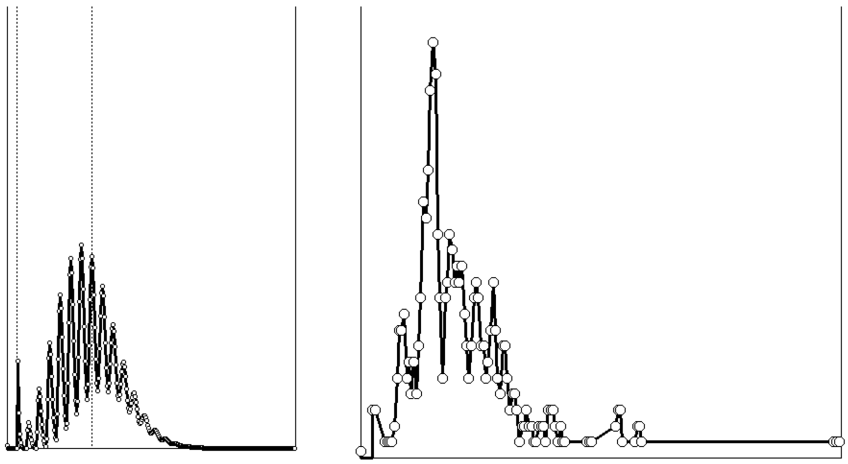

Рис. 1: Сравнение теоретического предельного гиббсовско – пуассоновского распределения пуассоновской суммы гиббсовских случайных величин (слева) и статистической функции плотности распределения, которая была оценена на основе полугодовых данных о суммах ежедневных продаж крепких алкогольных напитков в типичном супермаркете (справа).

Рассмотрим в качестве примера двумерный случайный вектор

$$\boldsymbol{\gamma} = \{\gamma_x, \gamma_y\},$$

описывающий совместную покупку двух товаров x и y случайным потребителем, распределения компонент которого имеют соответствующие слепые интервалы.

Допустим, что имеется статистическая выборка из n наблюдений за значениями такого случайного вектора $\boldsymbol{\gamma}$:

$$\boldsymbol{\gamma}^i = \{\gamma_x^i, \gamma_y^i\}, \quad i = 1, \ldots, n.$$

Это может быть, например, статистика продаж крепких спиртных напитков (рис. 1 справа) и колбасных изделий (рис. 2 слева) Данная обычная статистика позволяет оценить предложенную в работе *двухуровневую структуру* зависимостей и взаимодействий наблюдаемого случайного вектора: *случайно – множественный базис и количественную надстройку.*

1. Случайно – множественный базис. Поскольку рассматривается двумерный случайный вектор, его случайно – множественным базисом служит соответствующее ему случайное множество событий K_γ,

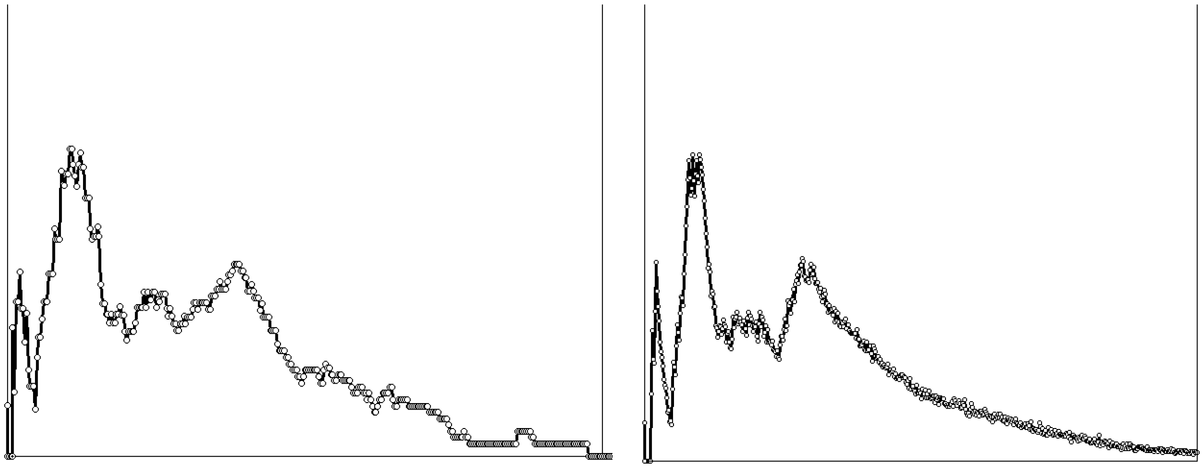

Рис. 2: СЛЕВА: Статистическая оценка плотности суммы стоимостей 262 ежедневных покупок из ассортимента в 209 различных сортов колбасы (полугодовая статистика типичного супермаркета: май – октябрь, 2002, 186 дней). СПРАВА: Монте-Карло генерирование суммы $N = 600$ гиббсовских случайных величин, разбитых на 37 однородных групп, каждая со своими параметрами: (слепой интервал b_i, параметр спуска β_i, вероятность покупки $p_i = \lambda_i/N_i$), $i = 1, \ldots, 37$.

определенное под дуплетом событий $\mathfrak{X} = \{x, y\}$ распределением вероятностей

$$p(X) = \mathbf{P}\left(\bigcap_{x\in X}\{\gamma_x > 0\}\bigcap_{x\in X^{\mathsf{c}}}\{\gamma_x = 0\}\right), \quad X \subseteq \mathfrak{X},$$

содержащим 4 вероятности, которые мы обозначим соответственно:

$$p_{xy} = \mathbf{P}(K_\gamma = \{x, y\})$$

— вероятность покупки двух товаров,

$$p_{xy^{\mathsf{c}}} = \mathbf{P}(K_\gamma = \{x\})$$

— вероятность покупки только товара x,

$$p_{x^{\mathsf{c}}y} = \mathbf{P}(K_\gamma = \{y\})$$

— вероятность покупки только товара y,

$$p_{x^{\mathsf{c}}y^{\mathsf{c}}} = \mathbf{P}(K\gamma = \emptyset).$$

— вероятность не-покупки товаров (вероятность "пустой" покупки под дуплетом товаров $\mathfrak{X} = \{x, y\}$ — вероятность того, что потребитель ушёл из супермаркета без товаров x и y).

2. Количественная надстройка. По определению количественная надстройка распределения случайного вектора, компоненты которого обладают слепыми интервалами, — это совокупность условных распределений

$$\Delta_\gamma = \left\{ F_X(s_x,\ x \in \mathfrak{X}),\ X \in 2^{\mathfrak{X}} \right\}.$$

В данном примере двумерного случайного вектора, эта совокупность состоит из четырёх условных распределений:

1) сосредоточенного в $(0,0)$ нульмерного вырожденного распределения при условии "пустой покупки" (точка символизирует отсутствие аргумента):

$$F_\emptyset\left(\cdot \left| \{\gamma_x = 0\} \cap \{\gamma_y = 0\}\right.\right),$$

2,3) двух одномерных распределений

$$F_x(s_x) = \mathbf{P}\left(\gamma_x \leq s_x \left| \{\gamma_x > 0\} \cap \{\gamma_y = 0\}\right.\right)$$

— при условии покупки только товара x,

$$F_y(s_y) = \mathbf{P}\left(\gamma_y \leq s_y \left| \{\gamma_x = 0\} \cap \{\gamma_y > 0\}\right.\right)$$

— при условии покупки только товара y, и

4) двумерного распределения

$$F_{xy}(s_x, s_y) = \mathbf{P}\left(\gamma_y \leq s_y, \gamma_y \leq s_y \left| \{\gamma_x > 0\} \cap \{\gamma_y > 0\}\right.\right)$$

— при условии покупки двух товаров x и y.

3. Статистическая оценка случайно – множественного базиса. Статистическая оценка случайно – множественного базиса в данном примере наблюдений за двумерным случайным вектором стоимостей покупок двух товаров сводится к оценке распределения случайного множества событий K_γ — четырёх вероятностей

$$p_{xy}, \quad p_{xy^\mathsf{c}}, \quad p_{x^\mathsf{c}y}, \quad p_{x^\mathsf{c}y^\mathsf{c}}.$$

Эту оценку вполне можно провести на основе имеющейся выборки наблюдений по очевидным формулам

$$\widehat{p}_{xy} = \frac{n_{xy}}{n}, \quad \widehat{p}_{xy^\mathsf{c}} = \frac{n_{xy^\mathsf{c}}}{n}, \quad \widehat{p}_{x^\mathsf{c}y} = \frac{n_{x^\mathsf{c}y}}{n}, \quad \widehat{p}_{x^\mathsf{c}y^\mathsf{c}} = \frac{n_{x^\mathsf{c}y^\mathsf{c}}}{n}, \qquad (*)$$

где

$$
\begin{aligned}
n_{xy} &= \left|\{\boldsymbol{\gamma}^i : \ \gamma_x^i > 0, \gamma_y^i > 0\}\right|, \\
n_{xy^{\mathsf{c}}} &= \left|\{\boldsymbol{\gamma}^i : \ \gamma_x^i > 0, \gamma_y^i = 0\}\right|, \\
n_{x^{\mathsf{c}}y} &= \left|\{\boldsymbol{\gamma}^i : \ \gamma_x^i = 0, \gamma_y^i > 0\}\right|, \\
n_{x^{\mathsf{c}}y^{\mathsf{c}}} &= \left|\{\boldsymbol{\gamma}^i : \ \gamma_x^i = 0, \gamma_y^i = 0\}\right|
\end{aligned}
$$

— соответствующие количества наблюдений в рассматриваемой выборке, соответствующие всем четырём комбинациям из двух событий x и y. Очевидно, что

$$n_{xy} + n_{xy^{\mathsf{c}}} + n_{x^{\mathsf{c}}y} + n_{x^{\mathsf{c}}y^{\mathsf{c}}} = n$$

— вся выборка наблюдений разбивается на четыре условные подвыборки, объемы которых равны

$$n_{xy}, \quad n_{xy^{\mathsf{c}}}, \quad n_{x^{\mathsf{c}}y}, \quad n_{x^{\mathsf{c}}y^{\mathsf{c}}},$$

и которые соответствуют четырём комбинациям из двух событий x и y.

Статистическая оценка $(*)$ распределения случайного множества событий K_γ позволяет получить оценку ковариаций зависимости и силы статистического взаимодействия случайных событий x и y по формулам, основанным на определениях этих понятий:

$$\widehat{\mathsf{Kov}}_{xy} = \widehat{\mathbf{P}}(x \cap y) - \widehat{\mathbf{P}}(x)\widehat{\mathbf{P}}(y), \quad \widehat{\mathsf{For}}_{xy} = \frac{\widehat{\mathbf{P}}(x)\widehat{\mathbf{P}}(y)}{\widehat{\mathbf{P}}(x \triangle y)}$$

с учетом следующих соотношений, связывающих вероятности событий:

$$\mathbf{P}(x) = p_{xy} + p_{xy^{\mathsf{c}}}, \quad \mathbf{P}(y) = p_{xy} + p_{x^{\mathsf{c}}y},$$

$$\mathbf{P}(x \cap y) = p_{xy},$$

$$\mathbf{P}(x \triangle y) = \mathbf{P}(x) + \mathbf{P}(y) - 2\mathbf{P}(x \cap y) = p_{xy^{\mathsf{c}}} + p_{x^{\mathsf{c}}y}.$$

Таким образом, окончательные статистические оценки ковариации зависимости и силы статистического взаимодействия событий x и y через оценки распределения $(*)$ принимают вид:

$$\widehat{\mathsf{Kov}}_{xy} = \widehat{p}_{xy} - (\widehat{p}_{xy} + \widehat{p}_{xy^{\mathsf{c}}})(\widehat{p}_{xy} + \widehat{p}_{x^{\mathsf{c}}y}),$$

$$\widehat{\mathsf{For}}_{xy} = \frac{(\widehat{p}_{xy} + \widehat{p}_{xy^c})(\widehat{p}_{xy} + \widehat{p}_{x^cy})}{\widehat{p}_{x^cy} + \widehat{p}_{xy^c}}.$$

Напомним, что ковариация зависимости событий x и y измеряет насколько статистически притягиваются ($\mathsf{Kov}_{xy} > 0$) или отталкиваются ($\mathsf{Kov}_{xy} < 0$) эти события. Оценка силы статистического взаимодействия For_{xy} двух событий x и y позволяет построить их орбитальную структуру, визуализирующую взаимодействия случайных событий на плоскости, которая в данном примере является частным решением задачи 2 случайных событий — частным решением следующей системы обыкновенных дифференциальных уравнений второго порядка:

$$\mathbf{P}(x)\frac{d^2r_x}{dt^2} = -\mathsf{For}_{xy}[r_y - r_x] - \mathsf{For}_{xy^c}[r_{y^c} - r_x] - \mathsf{For}_{xx^c}[r_{x^c} - r_x],$$

$$\mathbf{P}(y)\frac{d^2r_y}{dt^2} = -\mathsf{For}_{xy}[r_x - r_y] - \mathsf{For}_{x^cy}[r_{x^c} - r_y] - \mathsf{For}_{y^cy}[r_{y^c} - r_y],$$

$$\mathbf{P}(x^c)\frac{d^2r_{x^c}}{dt^2} = -\mathsf{For}_{x^cy}[r_y - r_{x^c}] - \mathsf{For}_{x^cy^c}[r_{y^c} - r_{x^c}] - \mathsf{For}_{x^cx}[r_x - r_{x^c}],$$

$$\mathbf{P}(y^c)\frac{d^2r_{y^c}}{dt^2} = -\mathsf{For}_{xy^c}[r_x - r_{y^c}] - \mathsf{For}_{x^cy^c}[r_{x^c} - r_{y^c}] - \mathsf{For}_{y^cy}[r_y - r_{y^c}].$$

4. Статистическая оценка количественной надстройки. Условные функции распределения, составляющие количественную надстройку наблюдаемого случайного вектора также несложно оценить стандартными методами математической статистики по имеющейся статистической выборке наблюдений за его значениями. Каждая условная функция распределения из количественной надстройки оценивается по соответствующей подвыборке наблюдений. В итоге получаем

1) оценку сосредоточенного в $(0,0)$ нульмерного вырожденного распределения при условии "пустой покупки":

$$\widehat{F}_{\emptyset}\left(\cdot\,|\{\gamma_x = 0\} \cap \{\gamma_y = 0\}\right) \equiv 1,$$

на основе подвыборки объемом $n_{x^cy^c}$. (Процедура оценки здесь, разумеется, "вырождена" потому что значение двумерного вектора детерминировано и равно $(0,0)$);

2,3) оценки двух одномерных распределений

$$\widehat{F}_x(s_x) = \widehat{\mathbf{P}}\left(\gamma_x \le s_x\,|\{\gamma_x > 0\} \cap \{\gamma_y = 0\}\right)$$

— при условии покупки только товара x,

$$\widehat{F}_y(s_y) = \widehat{\mathbf{P}}\left(\gamma_y \le s_y \,|\{\gamma_x = 0\} \cap \{\gamma_y > 0\}\right)$$

— при условии покупки только товара y на основе подвыборок объемами n_{xy^c} и $n_{x^c y}$ соответственно;

4) оценку двумерного распределения

$$\widehat{F}_{xy}(s_x, s_y) = \widehat{\mathbf{P}}\left(\gamma_y \le s_y, \gamma_y \le s_y \,|\{\gamma_x > 0\} \cap \{\gamma_y > 0\}\right)$$

— при условии покупки двух товаров x и y на основе подвыборки объемом n_{xy}.

5. Разложение распределения случайного вектора по случайно – множественному базису. Теперь всё готово для построения оценки распределения наблюдаемого случайного вектора $\boldsymbol{\gamma}$, которая конструируется на основе теоремы 3 о разложении распределения по случайно – множественному базису K_γ. В силу этой теоремы формула оценки распределения для рассматриваемого двумерного случайного вектора стоимостей покупок двух товаров имеет вид:

$$\widehat{F_\gamma}(s_x, s_y) = \hat{p}_{x^c y^c} + \widehat{F_x}(s_x) \cdot (\hat{p}_{xy} + \hat{p}_{xy^c}) +$$
$$+ \widehat{F_x}(s_x) \cdot (\hat{p}_{xy} + \hat{p}_{x^c y}) + \widehat{F_{xy}}(s_x, s_y) \cdot \hat{p}_{xy}.$$

6. Предположение независимости условных распределений из количественной надстройки. Можно ввести упрощающее предположение о независимости условного двумерного распределения F_{xy}, которое приведет к тому, что для оценки распределения исходного случайного вектора окажется достаточным оценок не более, чем одномерных условных распределений:

$$F_{x|xy}(s_x) = \mathbf{P}\left(\gamma_x \le s_x \,|\{\gamma_x > 0\} \cap \{\gamma_y > 0\}\right),$$
$$F_{y|xy}(s_y) = \mathbf{P}\left(\gamma_y \le s_y \,|\{\gamma_x > 0\} \cap \{\gamma_y > 0\}\right),$$

поскольку в силу предположения независимости

$$F_{xy}(s_x, s_y) = F_{x|xy}(s_x) \cdot F_{x|xy}(s_x).$$

Аналогичные упрощающие предположения могут быть сделаны и при статистических оценках случайных векторов большей размерности, когда предполагается независимость условных распределений, размерность которых больше двух, трёх и т.д. в зависимости от конкретной задачи.

Таким образом, предложенная в работе двухуровневая структура зависимостей и взаимодействий случайных событий — это продвижение вперед в анализе статистических зависимостей, так как она позволяет "выдавить" часть учитываемых зависимостей на случайно множественный уровень, чтобы "ослабить" предположение независимости на количественном уровне и, в конечном итоге, более полно описать общую структуру зависимостей и взаимодействий компонент случайного вектора.

Обоснованность статистической модели гиббсовского потребительского выбора

Предлагаемая в работе статистическая модель потребительского выбора разрабатывалась при многократном сравнении с реальными данными о статистике продаж типичного городского супермаркета. Все сделанные предположения, на которых основана разработанная ст атистическая модель, проверялись на большом объеме реальной статистики. Практическая идентичность модельных расчетов и многочисленных статистических данных, которая проиллюстрирована на ряде примеров, позволила ограничиться качественным сравнением статистической модели потребительского выбора и результатов наблюдений за продажами товаров в супермаркете.

Для проверки полученных в работе выводов и положений был выбран типичный супермаркет города Красноярска и исследована статистика его продаж в течение 6-ти месяцев 2001 года (май–октябрь, 186 дней). Кроме этого рассматривалась также статистика продаж типичной фирмы, торгующей строительными товарами.

Общие суммарные характеристики статистики продаж типичного супермаркета за полгода.

Общая сумма розничной торговли — 7 760 180 руб.

Объем розничной торговли в день — 41 721 руб.

Общее количество покупок — 116 177

Количество покупок в день — 625

Средний объем одной покупки — 67 руб.

Общая сумма оптовых закупок — 5 815 806 руб.

Общий доход 7760180 – 5815806 = 1 944 374 руб.

Доходность за полгода — 1944374/5815806 = 0.33 (33 %)

Средний доход в день — 10 453 руб.

На рис. 3 показаны статистические оценки плотностей стоимостей соответствующих групп товаров. График плотности суммы стоимостей покупок крепких алкогольных напитков, построенный на основе статистических наблюдений продаж типичного супермаркета за 186 дней качественно близок к теоретическому графику плотности стоимостей. Имеющиеся отличия объясняются, главным образом, тем, что теоретический график суммы плотностей построен для гиббсовских случайных величин, имеющих одинаковые параметры b, β, λ (теорема о предельном распределении пуассоновской суммы взаимно независимых одинаково распределенных гиббсовских случайных величин [3]).

На рис. 1 показано сравнение теоретического предельного гиббсовско – пуассоновского распределения пуассоновской суммы гиббсовских случайных величин (в центре) и статистической функции плотности распределения, которая была оценена на основе полугодовых данных о суммах ежедневных продаж крепких алкогольных напитков в типичном супермаркете (справа). Сравнение графика плотности суммы стоимостей реальных покупок 209 различных сортов колбасы, взятых из статистики того же супермаркета, с графиком плотности суммы стоимостей специальным образом генерированных гиббсовских случайных величин с различными параметрами b, β, λ, также показывает их качественное сходство. Это подтверждает адекватность гиббсовской модели потребительского выбора реальной статистике продаж.

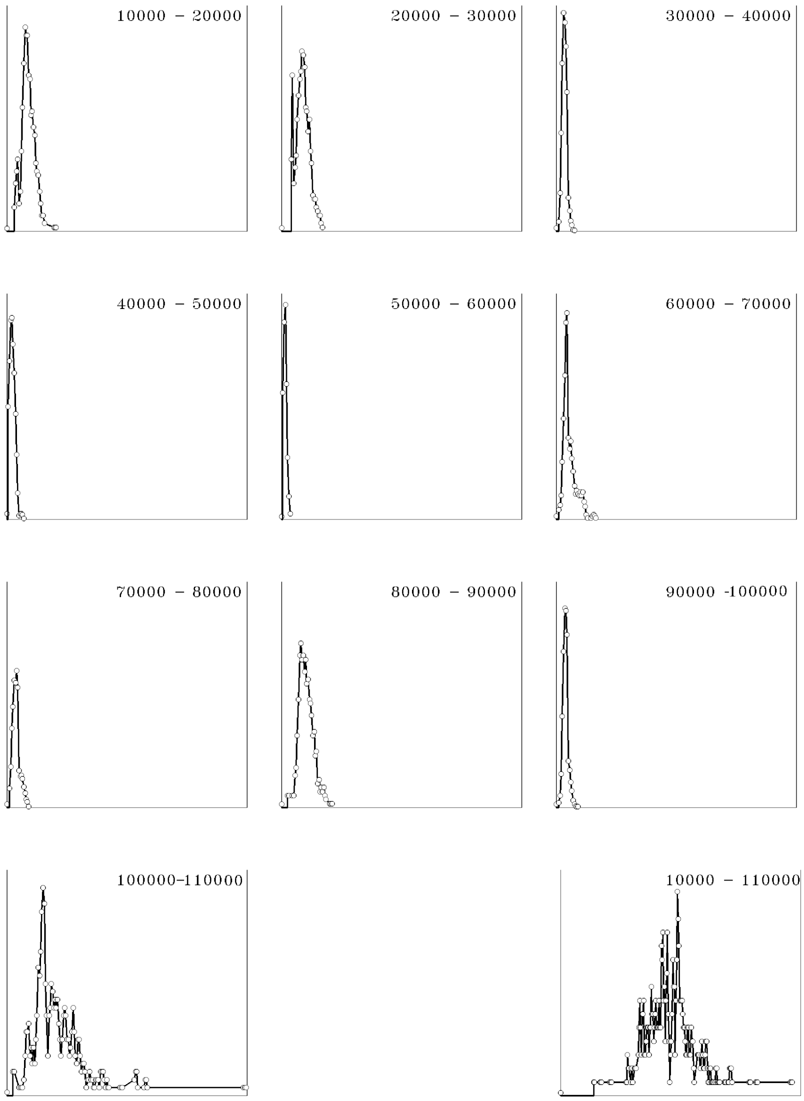

Рис. 3: Статистические оценки плотности стоимостей продаж различных групп товаров в супермаркете: 10-молочные, 20-колбасные, 30-кулинария, 40-птица, 50-бакалея, 60-рыба, 70-кондитерские, 80-овощи-фрукты, 90-масло-сыр, 100-алкоголь, 10–110-все группы (справа внизу).

На рис. 2 слева показана статистическая оценка плотности суммы стоимостей 262 ежедневных покупок из ассортимента в 209 различных сортов колбасы (полугодовая статистика типичного супермаркета: май – октябрь, 2002, 186 дней).

На рис. 2 справа показано генерирование методом Монте-Карло суммы 600 разнородных гиббсовских случайных величин, разбитых на 37 однородных групп, каждая со своими параметрами: (слепой интервал b_i, параметр спуска β_i, вероятность покупки $p_i = \lambda_i / N_i$), $i = 1, \ldots, 37$.

Рекомендации по использованию статистической модели случайного потребителя

Предложенную статистическую модель гиббсовского потребительского выбора можно использовать для решения многих практических задач в социально – экономической области. В качестве возможных приложений и рекомендаций по использованию статистической модели формулируются три задачи из области экономики товарных рынков.

Простейшим примером возможной области приложений результатов работы может служить естественность и простота основанного на понятии ковариации случайных событий нового статистического определения известных экономических характеристик спроса и предложения на товарных рынках — *взаимозаменяемости* и *взаимодополняемости* некоторой пары товаров x и y. Ещё две рекомендации основаны на разработанной статистической модели потребительского выбора.

1. **Взаимозависимые и взаимодополняемые товары.** Два товара x и y на рынке товаров $\mathfrak{X}$ случайного потребителя K называются *статистически взаимодополняемыми* всякий раз, когда $\mathsf{Kov}_{xy} > 0$, и *статистически взаимозаменяемыми* всякий раз когда $\mathsf{Kov}_{xy} < 0$.

Это лаконичное статистическое определение понятий *взаимозаменяемости* и *взаимодополняемости* полностью согласуется с существующим в экономической теории потребительского выбора представлением, согласно которому взаимозаменяемыми товары считаются тогда, когда один товар вытесняет другой при повышении цен на один из них; а взаимодополняемыми считаются товары с противоположным свойством — товары вместе вытесняются в рынка при повышении цен на один из них. Статистически это означает, что покупка и тех и других товаров — это зависимые случайные события, причем взаимозаменяемые товары реже (чем при независимом спросе) покупаются вместе, а взаимодополняемые товары наоборот — реже покупаются отдельно. Данное статистическое свойство, записанное на языке ковариации случайных событий, и заложено в приведенном выше статистическом определении. Это определение можно использовать на практике — оно позволяет разрабатывать методы статистической оценки взаимозаменяемости и взаимодополняемости товаров или услуг на различных рынках и системах потребительского выбора.

2. **Статистический анализ распределения дохода случайного потребителя.** На основе статистической выборки наблюдений за ежедневными суммами стоимостей проданных товаров можно оценить распределение уровня дохода случайного потребителя и стратифицировав выборку по слоям, соответствующим различным уровням его дохода, изучать распределения случайных потребителей с разным уровнем среднего дохода.

3. **Статистическая оценка распределения вероятностей собственных вкусов и предпочтений случайного потребителя.** После статистической оценки распределения уровня дохода случайного потребителя (и его стратификации по слоям) для каждого слоя уровня дохода может быть решена задача статистической оценки распределения собственных вероятностей случайных потребителей с разными уровнями среднего дохода. В результате появляется возможность строить и анализировать статистические модели вкусов и предпочтений потребителей с разными уровнями дохода — очень важная для практики товарных рынков задача.

Литература

[1] Воробьев О.Ю. (2001) *Предметный указатель шести томов персональных Записок ФАМ семинара (1–6, 1997–2001).* Красноярск: ИВМ СО РАН, 11 с.

[2] Воробьев О.Ю. (2002, в печати) *Теория случайных событий и её применение.* Красноярск: ИВМ СО РАН, 340 с.