

МЕТОДЫ КОМБИНАТОРНОГО АНАЛИЗА И ТЕОРИИ ВЕРОЯТНОСТЕЙ ДЛЯ ОЦЕНКИ СОСТОЯНИЯ РЫНКА АЛКОГОЛЬНОЙ ПРОДУКЦИИ

**В.Н. БОГДАНОВ, П.С. ВИХЛЯНЦЕВ, С.Н. КАРАХОТИН,
М.В. СИМОНОВ**

В статье рассматривается возможность использования методов комбинаторного анализа и теории вероятностей для оценки объемов нелегальной алкогольной продукции на рынке России.

V.N. Bogdanov, P.S. Vikhlyantsev, S.N. Karahotin, M.V. Simonov. Methods of combinatorial analysis and probability theory for the analysis of the alcohol market

The article considers the possibility of using methods of combinatorial analysis and probability theory to estimate the volumes of illegal alcoholic products in the market in Russia.

Ключевые слова: алкогольная продукция, статистическая выборка, комбинаторный анализ, теория вероятностей

Keywords: alcoholic production, statistical sampling, combinatorial analysis, probability theory

E-mail: p.vikhlyantsev@center-inform.ru

1. Введение

Государственная политика по снижению масштабов злоупотребления алкогольной продукцией и профилактике алкоголизма среди населения Российской Федерации направлена на сокращение объемов потребления населением алкогольной продукции и улучшение демографической ситуации в стране [1]. Важнейшей целью государственной политики в этой области является повышение эффективности регулирования алкогольного рынка и выполнения государственной функции по осуществлению государственного контроля и надзора.

Состояние алкогольного рынка может оцениваться количеством нарушений законодательства и лицензионных требований, а также долей незаконно произведенной алкогольной продукции, находящейся в обороте.

Выявление нарушений осуществляется с использованием сведений, содержащихся в единой государственной автоматизированной информационной системе учета объема производства и оборота этилового спирта, алкогольной и спиртосодержащей продукции (ЕГАИС), анализа деклараций об объемах производства и

оборота, интернет-сервиса проверки сведений, нанесенных на федеральные специальные марки, а также анализа соответствия информации на федеральных специальных (ФСМ) и акцизных марках (АМ) и самой алкогольной продукции, промаркированной этими марками.

Доля незаконно произведенной алкогольной продукции, находящейся в обороте, может оцениваться путем выборочных проверок продукции и обработки статистических данных на основе комбинаторного анализа и основ теории вероятностей [2-4].

Цель настоящей статьи – показать применение комбинаторного анализа и теории вероятностей для обоснования правил формирования выборки и расчета доли незаконно произведенной алкогольной продукции, находящейся в обороте.

2. Постановка задачи

Будем полагать, что на рынке Российской Федерации присутствуют партии одного из видов (водка, коньяк, вино и т.п.) алкогольной продукции различных наименований, изготовленные российскими и зарубежными производителями.

Генеральная совокупность состоит из N единиц алкогольной продукции, из которых M единиц нелегальной алкогольной продукции и $L=N-M$ единиц легальной алкогольной продукции.

В результате контрольных проверок из генеральной совокупности случайным образом извлекается статистическая выборка единиц образцов алкогольной продукции $n < N$. Образцы выборки могут извлекаться одновременно или последовательно без возвращения во множество.

К нелегальной алкогольной продукции относят:

- незаконно произведенную (импортированную) алкогольную продукцию;
- продукцию, незафиксированную в ЕГАИС;
- с поддельными ФСМ (АМ);
- без необходимой сопроводительной документации;
- и т.п.¹

Требуется определить:

а) вероятность нахождения в выборке образцов нелегальной алкогольной продукции при известной доле нелегальной продукции в генеральной совокупности;

б) вероятное количество образцов нелегальной алкогольной продукции и долю нелегальной алкогольной продукции на рынке при выборке заданного объема и известном количестве нелегальной продукции в выборке.

3. Основные расчетные соотношения

Вероятность события $P_{MN}(n, m)$, что в выборке окажется m образцов нелегальной

¹ Необходимые и достаточные признаки отнесения образцов к нелегальной алкогольной продукции определяются Росалкогольрегулированием на основании действующего законодательства Российской Федерации.

алкогольной продукции определяется по формуле (1) [2]:

$$P_{MN}(n, m) = \frac{C_n^m C_{N-n}^{M-m}}{C_N^M}, m = 0, 1, 2, \dots, \min(n, M), \quad (1)$$

где: $C_n^m = \frac{n!}{m!(n-m)!}$.

Математическое ожидание и дисперсия случайной величины попадания в выборку m образцов нелегальной алкогольной продукции определяются по формулам:

$$M(n, m) = n \frac{M}{N} = np, \quad (2)$$

$$D(n, m) = np(1-p) \left(1 - \frac{n-1}{N-1}\right). \quad (3)$$

Распределение вероятности $P_{MN}(n, m)$ зависит от трех параметров: $n, p=M/N$ и N . Если N неограниченно возрастает, а p остается постоянным, то члены гипергеометрического распределения стремятся к соответствующим членам биномиального распределения:

$$P_{MN}(n, m) \rightarrow C_n^m p^m (1-p)^{n-m}. \quad (4)$$

Вероятность события, что в выборке из n образцов не окажется ни одного нелегального, может быть найдена по формуле (1) при $m=0$ и составляет:

$$P_{MN}(n, 0) = \left(1 - \frac{M}{N}\right) \left(1 - \frac{M}{N-1}\right) \dots \left(1 - \frac{M}{N-n+1}\right). \quad (5)$$

Вычисления вероятностей гипергеометрического распределения можно вести последовательно, используя два соседних члена распределения [4]:

$$\frac{P_{MN}(n, m+1)}{P_{MN}(n, m)} = \frac{(n-m)(M-m)}{(m+1)(N-M-n+m+1)}.$$

Для каждого фиксированного набора значений M, n, m существует значение объема \hat{N} генеральной совокупности, при котором вероятность $P_{MN}(n, m)$ принимает максимальное значение. Данное значение принято называть оценкой максимального правдоподобия N [1].

Для поиска значения максимального правдоподобия \hat{N} рассмотрим отношение вероятности $P_{MN}(n, m)$ при объеме генеральной

совокупности N и вероятности $P_{MN-1}(n, m)$ при объеме генеральной совокупности $N-1$:

$$\frac{P_{MN}(n, m)}{P_{MN-1}(n, m)} = \frac{N^2 - NM - Nn - Mn}{N^2 - NM - Nn - Nm}.$$

Данное соотношение больше единицы при $Nm > Mn$.

С возрастанием N при постоянных заданных значениях M , n и m значения вероятностей $P_{MN}(n, m)$ вначале возрастают, достигая максимума, когда N является целым числом, не превосходящим n/m , а затем убывают.

Таким образом, величина максимального правдоподобия приблизительно равна $\hat{N} = M n/m$.

Во всей генеральной совокупности на практике чаще всего известно количество легальной продукции L , а количество нелегальных единиц алкогольной продукции неизвестно.

Вероятность события, что в выборке объемом n из генеральной совокупности с параметрами L и N будет содержаться l образцов легальной продукции, определяется по формуле:

$$P_{LN}(n, l) = \frac{C_n^l C_{N-m}^{L-l}}{C_N^l}, l = 1, 2, \dots, \min(n, L).$$

Значение вероятности $P_{LN}(n, l)$ достигает максимума, когда N является целым числом, которое не превосходит величину Ln/l , т.е. величина максимального правдоподобия равна $\hat{N} = Ln/l$.

4. О формировании случайной выборки

Понятие случайной выборки является одним из основных в математической статистике. Не всякий выбор алкогольной продукции дает случайную выборку в том смысле, как это понимается в математической статистике.

В случайную выборку должна отбираться алкогольная продукция одного вида, различных наименований и разных производителей. Из отбора должна исключаться алкогольная

продукция одного вида, одного наименования, одного производителя и одной даты розлива, так как данная продукция относится к одной партии алкогольной продукции, фиксируется в ЕГАИС по одной заявке о фиксации и для ее маркирования используются ФСМ (АМ) одной серии.

5. Примеры использования методов комбинаторного анализа и теории вероятности для анализа состояния рынка алкогольной продукции

Рассмотрим на примерах возможности использования приведенных выражений комбинаторного анализа и теории вероятности для расчета показателей состояния рынка алкогольной продукции.

Пример 1. Пусть доля нелегальной алкогольной продукции, находящейся в обороте, составляет от 30%. Из генеральной совокупности алкогольной продукции ($N=100$, $M=30$) одного вида сформирована случайная выборка $n=10$ единиц разных наименований от различных изготовителей.

Требуется определить вероятность того, что в этой случайной выборке не будет ни одной единицы нелегальной алкогольной продукции.

Подставляя в (5) значения $N=100$, $M=30$, $n=10$ получаем вероятность того, что в эту случайную выборку не попадет ни одна единица нелегальной алкогольной продукции, равную $P_{NM}(n, m) = P_{100, 30}(10, 0) = 0,0288$.

Вероятность попадания в выборку хотя бы одной единицы нелегальной алкогольной продукции составляет $P_{NM}(n, m \geq 1) = 1 - P_{NM}(n, 0) = 0,9772$.

При увеличении генеральной совокупности полученные значения вероятности изменяются незначительно. Так, при $N=200$, $M=60$ вероятность того, что в этой

Таблица 1

Значения вероятностей попадания в выборку объема n хотя бы одного образца нелегальной продукции при доле p нелегальной продукции в генеральной совокупности

n	p=M/N				n	p=M/N	
	0,1	0,2	0,3	0,4		0,1	0,2
10	0.651	0.893	0.972	0.994	21	0.891	0.991
11	0.686	0.914	0.98	0.996	22	0.902	0.993
12	0.718	0.931	0.986	0.998	23	0.911	0.994
13	0.746	0.945	0.990	0.999	24	0.920	0.995
14	0.771	0.956	0.993		25	0.928	0.996
15	0.794	0.965	0.995		26	0.935	0.997
16	0.815	0.972	0.997		27	0.942	0.998
17	0.833	0.977	0.998		28	0.948	0.998
18	0.850	0.982	0.998		29	0.953	0.998
19	0.865	0.986	0.999		30	0.958	0.999
20	0.878	0.988					

случайную выборку не попадет ни одна единица нелегальной алкогольной продукции, равна $P_{NM}(n,m)=P_{200,60}(10,0)=0,0256$, а при $N=3000$, $M=900$ данная вероятность равна $P_{NM}(n,m)=P_{3000,900}(10,0)=0,0281$.

Приведенный пример свидетельствует о том, что при доле нелегальной продукции в 30% даже в относительно небольшой выборке окажется хотя бы один образец нелегальной продукции.

Пример 2. На складах организации оптовой торговли находится 20 ($N=20$) партий алкогольной продукции одного вида, разных наименований и различных производителей. Предположительно доля нелегальной алкогольной продукции составляет около 30% ($M=6$). Сформирована случайная выборка из 10 единиц ($n=10$) алкогольной продукции по одной единице из каждой партии. Определить вероятность появления в данной выборке ровно m единиц нелегальной алкогольной продукции.

По формуле (1) рассчитываем значения вероятностей при различных отношениях m от нуля до 6: $P_{MN}(n,0)=0,0054$, $P_{MN}(n,1)=0,0650$,

$$P_{MN}(n,2)=0,2438, P_{MN}(n,3)=0,3715, P_{MN}(n,4)=0,2438, P_{MN}(n,5)=0,0650, P_{MN}(n,6)=0,0054.$$

Максимальное значение вероятности $P_{MN}(n,m)$ достигает при $m=3$.

Вероятность непопадания в выборку ни одной единицы нелегальной алкогольной продукции весьма мала и составляет всего 0,0054.

Следует отметить, что математическое ожидание случайной величины количества единиц нелегальной продукции в выборке из данной генеральной совокупности по формуле (2) составляет $M(10,6)=10 \bullet 6/20=3$.

Если в результате обследования отобранных образцов в выборке окажется другое количество единиц нелегальной алкогольной продукции, отличное от 3, то это означает, что в исследуемой генеральной совокупности, скорее, всего другая процентная доля нелегальной алкогольной продукции.

Пример 3. По данным из ЕГАИС, в течение нескольких месяцев российскими производителями алкогольной продукции на рынок поставлено 5000 партий водки ($L=5000$). Примем гипотезу, что на рынке алкогольной продукции

может находиться 10, 20, 30 или 40% нелегальной водки, не зафиксированной в ЕГАИС. Определить вероятность события, что в случайной выборке объемом n ($n=10,12, \dots, 30$) окажется хотя бы один образец продукции, не зафиксированной в ЕГАИС.

С учетом выражения (4) вероятность события, что в выборке n окажется хотя бы один образец нелегальной продукции, можно определить по формуле $P_{NM}(n, m \geq 1) = 1 - (1-p)^n$.

Результаты расчета сведены в табл.1.

Из анализа результатов расчетов следует, что при снижении доли нелегальной продукции в общем объеме продукции для достижения заданной вероятности выявления образцов нелегальной продукции необходимо увеличивать объем выборки.

Так, для достижения вероятности обнаружения не ниже 0,95 в выборке хотя бы одной единицы нелегальной продукции при доле нелегальной продукции 20% объем выборки должен составлять не менее 14 единиц продукции, а при доле нелегальной продукции 10% – не менее 29 единиц.

6. Выводы

1. Методы комбинаторного анализа и теории вероятностей могут оказать существенную помощь при обработке статистических данных по результатам проведения проверок легальности алкогольной продукции, находящейся в обороте.

2. С помощью данного математического аппарата можно, с одной стороны, получить статистическую оценку объема нелегальной алкогольной продукции по каждому ее виду или по всем видам продукции, а с другой, – проследить динамику уменьшения объема нелегальной продукции по ее видам по времени и

дать оценку эффективности принимаемых мер по регулированию алкогольного рынка.

3. Достоверность получаемых данных в значительной степени зависит от правильного формирования случайной выборки. В случайную выборку должна отбираться алкогольная продукция одного вида, различных наименований и разных производителей.

4. Рассмотренные методы справедливы при условии однородности распределения нелегальной продукции на анализируемой территории субъекта РФ, федерального округа или всей страны.

Литература

1. Концепция реализации государственной политики по снижению масштабов злоупотребления алкогольной продукцией и профилактике алкоголизма среди населения Российской Федерации на период до 2020 года. Одобрена распоряжением Правительства Российской Федерации от 30.12.2009 N 2128-р.

2. Феллер В. Введение в теорию вероятностей и ее приложений. Том 1. – М.: Мир, 1984. – 528с.

3. Справочник по вероятностным расчетам. – М.: Воениздат, 1970. – 536 с.

4. Справочник по теории вероятностей и математической статистике. – М.: Наука, 1985. – 640 с.