П.А. Карасев

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ЭКОНОМИЧЕСКОГО ПРОГНОЗИРОВАНИЯ

Учебное пособие

Под редакцией
Заслуженного работника высшей школы РФ,
доктора экономических наук,
профессора **М.Н. Кулапова**

ПАЛЕОТИП
Москва
2013

УДК 519.86(075.8)
ББК 65в651я73
К21

Рецензенты:

Б.Т. Кузнецов, д-р тех. наук, проф. (профессор кафедры менеджмента инвестиций и инноваций ФГБОУ ВПО «РЭУ им. Г.В. Плеханова»)

И.С. Кравчук, канд. тех. наук (доцент кафедры «Менеджмент качества» ФГБОУ ВПО «МГУПС (МИИТ)»)

Карасев, П.А.
К21 Математические основы экономического прогнозирования : учебное пособие / Под общ. ред. засл. раб. высш. шк. РФ, д-ра экон. наук, проф. М.Н. Кулапова ; П.А. Карасев. — М. : Издательство «Палеотип», 2013. — 132 с.

ISBN 978-5-94727-662-6

В учебном пособии излагаются наиболее распространенные методы прогнозирования социально-экономических процессов, используемые в аналитическом арсенале управления. Представлены экспертные методы выработки и принятия управленческих решений, раскрыто фундаментальное отличие между детерминированным и вероятностным подходом в прогнозировании. Рассматриваются методы корреляционно-регрессионного анализа и осуществляется их практическое применение в прогнозировании макроэкономических явлений. Рассмотрены основные математические и статистические показатели экономических процессов.

Для студентов бакалавриата и магистратуры, аспирантов экономических специальностей, а также начинающих аналитиков и менеджеров планово-экономических и прогнозно-аналитических отделов организаций различных уровней и сфер деятельности.

УДК 519.86(075.8)
ББК 65в651я73

ISBN 978-5-94727-662-6

Содержание

Предисловие

Предлагаемое Вашему вниманию учебное пособие посвящено исследованию вопросов прогнозирования, не являющегося точным инструментом предсказания поведения экономического объекта в будущем. Основная задача экономического прогнозирования – выявление тенденций и вариантов (то есть альтернатив) развития предприятий и национальной экономики в целом. В этой связи предложенные в настоящем пособии методы – лишь элементы аналитики, обеспечивающей и сопровождающей процессы планирования.

В соответствии с логикой построения учебного процесса в вузе изучение дисциплин профессионального цикла происходит в основном на старших курсах, когда студенты в большей степени готовы к профессиональной деятельности экономиста или менеджера. Однако математические основы закладываются многим ранее. Возникающий разрыв между математическими курсами и специальными дисциплинами в системе подготовки выпускников не способствует активному применению математики в их аналитической деятельности.

Как показывает опыт преподавания автором математических дисциплин в экономическом вузе, студенты не умеют пользоваться количественными методами при оценке экономических явлений. Степень обоснованности вырабатываемых в этих условиях выводов и предложений сильно снижается, поэтому автор надеется, что данная книга станет первым этапом в формировании им цикла учебных и учебно-методических пособий, посвященных количественным методам обоснования управленческих решений, принимаемых в условиях реальной экономики.

Глава 1. Вероятностные методы в прогнозировании

Математический инструментарий, используемый в прогнозировании, крайне обширен. Тем не менее, его можно условно разделить на две обособленные составляющие: детермининированные и вероятностные методы. Детерминированные методы и модели, как правило, предполагают наличие точных формул и жесткие функциональные связи, когда закономерности происходящего и будущего точно известны. Вероятностные методы играют особую роль в математике и прикладных науках, поскольку они, в отличие детерминированных описаний с присущими им ограничениями и допущениями, способны представить фундаментальную картину поведения процессов. Многие явления, события, а тем более прогнозы, носят случайный характер. Это означает, что запланированный результат может как произойти, так и не произойти. В таком случае мы имеем дело с вероятностью наступления того или иного определенного или неопределенного исхода. Блок вероятностных методов в прогнозировании позволяет со всех сторон оценить исследуемый объект и дать оценку вероятности наступления сделанного прогноза в будущем, основываясь не на предположениях, а на расчетах. Детерминированные (функциональные) подходы в математическом прогнозировании ограниченны и они не могут отклониться от тех формул и функций, которые в них заложены. Отличие двух групп методов наглядно изображено на рис. 1.1.

Итак, важнейшее отличие между двумя подходами заключается в том, что первый (детерминированный) отвечает на вопрос: «Если будет событие 1, то совершенно точно произойдет исход А, если ожидается вариант 2, произойдет исход Б», то второму подходу соответствует выражение «если факторы 1, 2 и 3 будут проявлять себя в ходе наблюдения, то событие А наступит с вероятностью около 90 %, а если фактор 3 не проявится, то событие А наступит с вероятностью 62%». Опять же, даже вероятность 99,9 % означает, что событие в 1 из 1000 случаев может и не произойти. Вероятность в математике обозначается $\boldsymbol{p}$ и она не превышает 100%, то есть 1.

Сделанные выше логические заключения позволяют утверждать, что в основе вероятностных методов лежат более сложные и комплексные формулы, расчеты и математические модели, построению, изучению и анализу возможного применения которых будет уделено особое внимание в настоящей главе.

Рис. 1.1. Детерминированный и вероятностный подходы в прогнозировании

1.1. Математическая обработка результатов экспертных опросов в прогнозировании

Экспертный опрос применяется в тех случаях, когда невозможно определить с помощью имеющихся моделей развитие какого-либо процесса в будущем, либо эти модели сложны или слишком ограниченны, либо такие модели и формулы отсутствуют. Безусловно, эксперт, как специалист и профессионал, знает и может предсказать больше любой модели, но его оценки чаще качественные, то есть выражены в форме суждений и умозаключений, абсолютно точного ответа, выраженного числом, эксперт дать не в состоянии. Второе важное ограничение экспертных методов состоит в том, что для более точного суждения необходима экспертная группа, то есть мнения нескольких экспертов, а они могут существенно различаться и варьироваться. В этой ситуации на плечи менеджера-аналитика ложится организация процесса обсуждения экспертами заданной проблемы и последующая обработка мнений экспертов. В этом разделе мы поговорим о математической составляющей процесса обработки экспертных оценок и трактовке полученных результатов опроса.

Одной из самых простых процедур проведения экспертных оценок является *балльный метод*, при котором эксперт выставляет по заранее оговоренной шкале балл тому или иному варианту развития ситуации, исходя из собственных предположений.

После того, как все эксперты проставили свои баллы, данные сводятся в итоговую таблицу. Пример ее изображен в табл. 1.1. Затем по данным таблицы проводится вероятностный и статистический анализ, о котором будет сказано ниже.

Таблица 1.1

Пример обобщающей таблицы для занесения данных экспертного опроса

<table>
<tr><th>№ признака / № эксперта</th><th>А</th><th>Б</th><th>В</th><th>Г</th><th>Д</th><th>…</th></tr>
<tr><td>1</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>3</td><td colspan="6" rowspan="2">Баллы, проставленные i-м экспертом j-му признаку</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>6</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>7</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>8</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>9</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>10</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>11</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>12</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Как уже было упомянуто, мнения экспертов часто не совпадают, варьируются. Это означает, что встает вопрос об определении **среднего**, наиболее разделяемого всеми мнения, принимаемого за основу для принятия последующих решений, и об оценке вариации (разброса) мнений, на основе которой можно судить о применимости такого решения.

В математике и статистике приводится множество различных методов расчета среднего в зависимости от ситуации и данных, по которым оно рассчитывается.

Средние бывают *простыми* и *взвешенными*. Взвешиваются индивидуальные значения показателей при расчете средней тогда, когда установлена или подразумевается их различная роль в формировании единой картины, в том числе в случае различной вероятности появления каждого из индивидуальных значений на практике.

Также средние величины могут быть структурными и степенными. К структурным средним относят *моду* (наиболее часто встречающаяся величина в совокупности, вероятность ее среди других значений всегда будет самая высокая) и медиану (величина, находящаяся в середине ряда, то есть справа и слева от которой находится половина единиц совокупности). К степенным средним относят среднюю арифметическую, среднюю гармоническую, среднюю квадратическую, среднюю геометрическую величины.

Средняя арифметическая – самый распространенный вид средних величин. Когда в примере говорится о средней величине без указания ее типа, имеется в виду средняя арифметическая. Она вычисляется в случаях, когда объем усредняемого признака получается как сумма его значений у отдельных элементов изучаемой совокупности.

В зависимости от характера исходных данных средняя арифметическая может быть рассчитана по формуле **простой** или **взвешенной средней**.

Если исходные данные не систематизированы, а просто перечислены, то применяется формула простой средней арифметической:

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (1.1),$$

где n – число значений ряда данных. Значение числового ряда называется *вариантой* (x).

Если исходные данные сгруппированы и представлены *весами* (*частотами*), то есть с числом единиц, имеющих одинаковые значения признака, или вероятностями появления конкретного значения признака в совокупности, то среднюю арифметическую исчисляют по формуле взвешенной средней.

При расчете средней арифметической взвешенной:

- необходимо умножить варианты на частоты / веса / вероятности;
- сложить полученные произведения;
- сложить веса (частоты);
- сумму произведений вариант на веса / частоты / вероятности разделить на сумму последних.

Окончательная формула выглядит следующим образом:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} \quad (1.2)$$

Обычно средняя арифметическая исчисляется по формуле взвешенной средней. Простую среднюю используют только в тех случаях, когда у каждой варианты частота равна единице или если частоты у всех вариант равны друг другу.

Сферы применения других видов средних существенно ограничены. Так, например, средняя геометрическая применяется при исчислении средних темпов динамики, а средняя квадратическая только при исчислении показателей варьируемости признака.

Средняя гармоническая применяется, когда в качестве весов применяются не единицы совокупности, а произведения этих единиц на значения признака. Отсюда приходится взвешивать варианты по объемам признака. Например, даны показатели товарооборота и количество проданного товара по разным дням месяца, и необходимо рассчитать среднемесячную цену продажи. Таких примеров огромное множество, и применение средней гармонической, прежде всего, обусловлено, формой представления данных, из-за которой применение средней арифметической оказывается трудной задачей.

Формулы для расчета средней гармонической имеют следующий вид: для простой средней гармонической: $\overline{x}_{harm} = \frac{n}{\sum_{i=1}^{n} 1/x_i}$ (1.3), для взвешенной: $\overline{x}_{harm} = \frac{\sum_{i=1}^{n} F_i}{\sum_{i=1}^{n} F_i/x_i}$ (1.4).

Средняя геометрическая (простая и взвешенная) рассчитываются по формуле:

$$\overline{x}_{geom} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n} = \sqrt[n]{\prod_{i=1}^{n} x_i} \quad (1.5) \text{ и } \overline{x}_{geom} = \sqrt[\sum_{i=1}^{n} f_i]{\prod_{i=1}^{n} (x_i)^{f_i}} \quad (1.6).$$

Средняя квадратическая находит свое применение при расчете вариации индивидуальных значений признака относительно их среднего (математического ожидания) и рассчитывается по двум различным формулам для простой и взвешенной средней:

$$\overline{x}_{qua} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \quad (1.7) \text{ и } \overline{x}_{qua} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 f_i}{\sum_{i=1}^{n} f_i}} \quad (1.8).$$

Также существует несколько других видов средних, среди которых упоминается средняя хронологическая, но они используются редко и в рассмотрение данной темы не включаются.

Важно помнить, что свое основное свойство **обобщающей характеристики** средняя величина выполняет при сочетании двух обязательных условий: если она получена из однородной совокупности, т.е. из индивидуальных величин одного и того же типа, а также наблюдается большой объем совокупности, то есть число единиц выборки достаточно велико для отражения реальной картины событий.

Вторая группа показателей, используемая в анализе экспертных методов, оценивает вариацию мнений экспертов, то есть насколько мнение каждого отдельно взятого эксперта отличается от усредненного. Таких показателей несколько: это размах вариации, дисперсия, среднее квадратическое отклонение, коэффициент вариации.

Самым простым показателем является *размах выборки* (вариации), равный разности максимального и минимального значения в совокупности $R = x_{\max} - x_{\min}$ (1.9). Он показывает абсолютную величину расхождения «полярных мнений» экспертов. Однако более точную характеристику степени различия во мнениях дают показатели дисперсии, среднего квадратического отклонения и коэффициента вариации, которые учитывают не только разницу индивидуальных (отдельных) значений, но и их связь со средним.

Дисперсия характеризует средний квадрат отклонения отдельно взятого значения от среднего по совокупности, то есть также является средней величиной.

Дисперсия рассчитывается по формуле: $D_x = \overline{x^2} - \left(\overline{x}\right)^2$ (1.10). Чтобы рассчитать дисперсию по данной формуле, нужно найти среднее значение квадрата исследуемого признака. То есть вариантами в данном случае будут $x_1^2, x_2^2, x_3^2, \ldots, x_n^2$, а весами – $f_1, f_2, f_3, \ldots, f_n$ – то есть веса (вероятности) при расчете не изменятся.

Для расчета дисперсии могут быть использованы формулы: для негруппированных данных $D_x = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n}$ (1.11), для сгруппированных данных по весам – $D_x = \frac{\sum_{i=1}^{n}\left(\left(x_i - \overline{x}\right)^2 \cdot f_i\right)}{\sum_{i=1}^{n} f_i}$ (1.12).

В Microsoft Excel для вычисления обычного значения дисперсии без корректировок используется статистическая функция ДИСПР, для среднего – СРЗНАЧ.

Среднее квадратическое отклонение представляет собой арифметический квадратный корень из дисперсии и показывает абсолютную величину отклонения отдельного значения варианты от среднего по совокупности, то есть насколько в среднем отклоняется отдельное значение от среднего. На примере экспертного опроса эта величина показывает среднюю величину расхождения во мнениях экспертов. Здесь опять же ключевое слово – среднее. Поскольку расхождения во мнениях крайне различны, то среднее квадратическое отклонение оценивает **среднюю** величину их различия. Формула для вычисления среднего квадратического отклонения такова:

$$\sigma = \sqrt{D_x} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}\left(\left(x_i - \overline{x}\right)^2 \cdot f_i\right)}{\sum_{i=1}^{n} f_i}} \quad (1.13).$$

В среде Microsoft Excel для определения обыкновенного среднего квадратического отклонения используется функция СТАНДОТКЛОНП.

На рис. 1.2 представлена интерпретация показателя σ. Дан массив значений некоторого признака (x): x_1, x_2, x_3,…,x_i, x_{i+1},…, x_{n-1}, x_n . По данному массиву вычисляется среднее значение ($\overline{x}$) и среднее квадратическое отклонение (σ). Кстати, отмечаем, что $\overline{x}$ не является серединным значением (медианой), оно может сколь угодно близким к границам массива (x_1 и x_n) в случае, если плотность значений около них будет выше, чем в других интервалах.

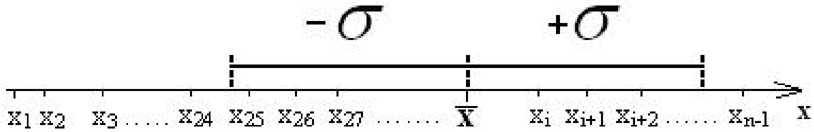

Рис. 1.2. Графическая интерпретация среднего квадратического отклонения

Коэффициент вариации есть мера относительного разброса отдельных значений одного признака; он показывает, какую долю среднего значения этой величины составляет ее средний разброс. Коэффициент вариации равен отношению среднего квадратического отклонения к среднему значению совокупности и выражается в процентах:

$$V = \frac{\sigma}{\overline{x}} \cdot 100\% \quad (1.14).$$

Следует отметить, что коэффициент вариации наряду со средним квадратическим отклонением широко используется в теории рисков для измерения абсолютного и относительного значения того или иного риска, чем он больше, тем больше волатильность (колеблемость) исследуемого признака. В нашем случае высокий процент коэффициента вариации будет свидетельствовать о сильном расхождении во мнениях экспертов и низком качестве управленческого решения, которое будет принято по результатам данного экспертного опроса.

Развитие математического аппарата в области анализа с помощью средних и показателей вариации нашло в теории ранговой корреляции.

Кроме баллов, которые могут проставлять эксперты и которые являются количественной ассоциацией какого-либо явления, процесса или сценария, эксперты могут ранжировать те или иные варианты, прогнозы, модели, проекты по степени их значимости. Это широко используется при выборе того или иного варианта развития, проекта или управленческого решения среди иных других. Ранг – степень значимости и вероятности конкретного объекта в ряду других аналогичных объектов. В теории и практике экспертных оценок, как правило, используется ранжирование по следующей логике: **1, 2, 3, 4, … , *k***, где k – порядковый номер последнего объекта ряда, которому отдается наименьшее предпочтение эксперта, 1 – наиболее значимый и вероятный объект в совокупности. Каждый из экспертов проставляет ранг одному из объектов, причем одно и то же значение ранга не может соответствовать двум и более объектам. Иными словами, вы присваиваете в порядке возрастания числа от 1 до k объектам, значимость которых для вас будет убывать.

Анализ экспертного опроса с использованием рангов по показателям среднего и вариации не отличается от анализа, проведенного с помощью баллов. Преимущество использования рангов заключается в том, что в данном случае можно получить так называемую «ранговую корреляцию», по результатам которой проводится наиболее объективная оценка согласованности мнений экспертов. Если устанавливается низкая согласованность мнений экспертов, то экспертный опрос проводится заново с изменением технологии и(ли) процедуры опроса, и(ли) состава участников.

Простейшим методом анализа и определения приоритетности объекта является *метод средних рангов*. Для этого нужно рассчитать сумму рангов по каждому объекту и разделить ее на число экспертов, принявших участие в опросе, в результате будет получен средний арифметический ранг. По средним рангам строится итоговое ранжирование, исходя из принципа - чем меньше средний ранг, чем лучше объект. Объект с наименьшим средним рангом в итоговой ранжировке получает ранг 1. В случае, если объекты имеют одинаковый средний ранг (например, 4,53), то они занимают места в ранжировке с i-го по (i+1)-е, но поскольку все эксперты отдали им одинаковое предпочтение, их ранги в итоговой таблице будут одинаковыми и равными $\frac{i+(i+1)}{2}=\frac{2i+1}{2}=i+0,5$. Иными словами, если эквивалентные объекты в итоговой ранжировке должны занимать 5 и 6 место, то их ранг будет равен 5,5.

Другим методом, более адекватным ранжированию, является *метод медиан рангов*. В качестве «золотой середины» для каждого из объектов оценки выбирается не среднее арифметическое рангов, а их медиана, поскольку не всегда правильно усреднять порядковые номера. Следовательно, можно использовать то самое серединное значение, которым является *медиана*.

Существует несколько методов расчета медианы. В нашем случае, когда значения представлены дискретно (то есть, не сгруппированы в интервалы), можно использовать такой порядок: необходимо расставить в порядке возрастания ранги и простым счетом найти центральную величину в ряде или определить ее порядковый номер по формуле:

$$I_{Me}=\frac{1}{2}(n+1) \quad (1.15).$$

Для определения медианы в дискретном ряду при наличии частот, сначала исчисляется полусумма частот, а затем определяется какое значение варьируемого признака ей соответствует.

После того, как медианы для каждого из объектов найдены, строится новая ранжировка по принципу, аналогичному тому, который используется в методе средних рангов.

Ранговая корреляция экспертных оценок – установление связи между парой мнений экспертов об одном и том же объекте оценки. Если ранговая корреляция парная, то она позволяет установить, насколько мнения **двух** экспертов согласуются между собой и какова степень их расхождения по поводу, например, проекта А среди других аналогичных проектов.

Также в общей и математической статистике есть несколько показателей ранговой корреляции. Мы рассмотрим два коэффициента парной ранговой корреляции **Спирмена** и **Кендалла** и коэффициент **конкордации** (множественный коэффициент ранговой корреляции).

Простейшим видом корреляционной связи является парная связь, когда рассматривается взаимодействие и взаимовлияние двух факторов друг на друга. В нашем случае мы должны изучить сходство во мнениях двух любых экспертов по всей группе объектов оценки.

Ранговый коэффициент корреляции Спирмена можно рассчитать по следующей формуле:

$$k_s = 1 - \frac{6}{n\left(n^2-1\right)} \sum_{i=1}^{n} \left(R_2\left(x_i\right) - R_1\left(x_i\right)\right)^2 \quad (1.16),$$

где $R_1\left(x_i\right)$ – последовательность рангов, присвоенных 1-ым экспертом всем объектам, $R_1\left(x_i\right)$ – последовательность рангов, присвоенных 2-ым экспертом всем объектам.

Данный коэффициент, как и все прочие корреляционные коэффициенты, обладает следующими **свойствами**:

1^0. Коэффициент корреляции Спирмена k_s показывает только меру связи между двумя позициями экспертов, выраженными в числах. На деле качественные оценки экспертов могут расходиться.

2^0. Значение коэффициента Спирмена заключено в следующих пределах: $-1 \le k_s \le 1$.

3^0. Если значение коэффициента по модулю близко к 1 ($|k_s| \to 1$), то связь между *числовыми последовательностями* рангов тесная, а мнения экспертов обладают следующими характеристиками:

3.1^0 если $k_s = -1$, то мнения экспертов прямо противоположны,

3.2^0 если $k_s = +1$, то мнения экспертов полностью совпадают,

4^0. Если $k_s = 0$, то какой-либо связи во мнениях двух экспертов нет.

Можно также использовать *коэффициент ранговой корреляции Кендалла*. Для его нахождения рассмотрим ряд рангов $\zeta_1, \zeta_2, \ldots, \zeta_n$, введенный по следующему принципу:

• значения Υ (первого признака) ранжируются в порядке возрастания или убывания;

• значения ξ (второго признака) располагаются в порядке, соответствующем значениям Υ;

То есть значения второго признака получают новые порядковые номера из соответствующих им значений первого признака. Таким образом получаем новый ряд ξ.

Зададим величины R_i следующим образом: пусть правее ζ_1 имеется R_1 рангов, больших ζ_1; правее ζ_2 имеется R_2 рангов, больших ζ_2 и т.д. Тогда, если обозначить $R = R_1 + R_2 + \ldots + R_{n-1}$, то выборочный коэффициент ранговой корреляции Кендалла («тау» Кендалла) будет определяться формулой:

$$\tau_k = \frac{4R}{n(n-1)} - 1 \quad (1.17).$$

Нетрудно видеть, что коэффициент Кендалла обладает теми же свойствами, что коэффициент Спирмена. Считается, что коэффициент Кендалла лучше использовать для выборок и совокупностей небольшого объема. Дело в том, что при подсчете $R_1, R_2, \ldots, R_{n-1}$ в ситуации равномерного возрастания количества альтернатив (по арифметической прогрессии) объем вычислений будет возрастать в геометрической прогрессии.

Коэффициент Спирмена всегда больше коэффициента Кендалла для одинаковой совокупности данных. Следовательно, оценка связи с применением коэффициента ранговой корреляции Кендалла получается более «осторожной».

Для измерения тесноты связи между любым произвольным числом ранжированных последовательностей используется коэффициент множественной ранговой корреляции – *коэффициент конкордации* (W), который вычисляется по формуле:

$$W = \frac{12S}{m^2\left(n^3 - n\right)} \quad (1.18),$$

где S – сумма квадратов отклонений суммы рангов, присвоенных каждым экспертом i-ому объекту оценки, от средней суммы; m – число экспертов, n – число объектов экспертных оценок.

Коэффициент конкордации изменяется в пределах от 0 до 1.

Значение коэффициента, равное 0, свидетельствует о полном отсутствии согласия между экспертами, значение W, равное 1, говорит о полной согласованности мнений экспертов.

Рассмотрим пример расчета и интерпретации рассмотренных показателей.

ПРИМЕР. Крупной компании XYZ требуется принять коллективное решение по новой стратегии бизнеса в меняющихся рыночных условиях. Для выработки стратегических альтернатив было приглашено 15 экспертов (обозначены по номерам от 1 до 15) из различных областей деятельности. Методом мозгового штурма были определены 10 альтернатив стратегического развития (обозначены как А, Б, В, Г, Д, Е, Ж, З, И, К). Затем экспертам предложили проставить ранги выбранным альтернативам, начиная с 1 (самый приоритетный, наиболее вероятный) и заканчивая 10 (наименее возможный, последняя очередь). Требуется выбрать три наиболее приоритетных и оценить качество проведенного экспертного опроса с использованием математических методов.

Мнения экспертов по итогам опроса представлены в табл. 1.2.

Таблица 1.2

Сводная таблица экспертного опроса, проведенного в компании XYZ

Альтернативы (i) / *Эксперты (j)*	***А***	***Б***	***В***	***Г***	***Д***	***Е***	***Ж***	***З***	***И***	***К***	*Вспом. столб.*
1	1	4	3	5	8	7	10	2	9	6	
2	3	7	1	6	4	5	8	2	10	9	
3	8	1	2	4	3	5	7	9	10	6	
4	7	10	6	1	2	3	8	9	5	4	
5	7	1	3	4	2	10	6	5	9	8	
6	2	8	6	1	9	5	3	7	10	4	
7	4	8	3	2	1	10	9	7	5	6	
8	7	6	1	8	4	5	2	3	10	9	
9	4	3	1	2	6	10	9	5	7	8	
10	2	3	9	7	5	8	10	1	4	6	
11	9	1	4	7	8	6	10	2	3	5	
12	2	5	1	10	3	4	9	8	6	7	
13	3	9	5	2	1	8	10	4	7	6	
14	10	6	8	2	1	5	9	7	4	3	
15	3	10	2	6	1	7	4	5	9	8	
Сумма рангов Σ	72	82	55	67	58	98	114	76	108	95	
Ср. арифмет. рангов $\overline{r}$	4,80	5,47	3,67	4,47	3,87	6,53	7,60	5,07	7,20	6,33	
Медиана рангов	4	6	3	4	3	6	9	5	7	6	
Дисперсия	8,457	10,27	6,810	7,838	7,695	5,124	7,114	7,210	6,457	3,381	

Альтернативы (i) / *Эксперты (j)*	***А***	***Б***	***В***	***Г***	***Д***	***Е***	***Ж***	***З***	***И***	***К***	*Вспом. столб.*
Ср. квадратическое отклонение	2,9081	3,2042	2,6095	2,7997	2,7740	2,2636	2,6673	2,6851	2,5411	1,8387	*Сумма квадратов разностей рангов*
Коэффициент вариации	60,59 %	58,61 %	71,17 %	62,68 %	71,74 %	34,65 %	35,10 %	52,99 %	35,29 %	29,03 %	
Ранги 7 эксперта	**4**	**8**	**3**	**2**	**1**	**10**	**9**	**7**	**5**	**6**	
Ранги 13 эксперта	**3**	**9**	**5**	**2**	**1**	**8**	**10**	**4**	**7**	**6**	
Квадрат разности рангов 7 и 13 эксперт. $R_7(x_i)$ *и* $R_{13}(x_i)$	1	1	4	0	0	4	1	9	4	0	***24***
Коэффициент Спирмена[1]	***0,85455***										**Средняя** сумма рангов
Сумма рангов (повторно)	72	82	55	67	58	98	114	76	108	95	*82,5*
Отклонения суммы рангов от их средней	-10,5	-0,5	-27,5	-15,5	-24,5	15,5	31,5	-6,5	25,5	12,5	***Сумма (S)***
Квадраты отклонений сумм рангов от средней	110,25	0,25	756,25	240,25	600,25	240,25	992,25	42,25	650,25	156,25	**3788,5**
Коэффициент конкордации	***0,204094***										

Для расчета коэффициента Кендалла необходимо сформировать новый ряд значений для рангов 13-го эксперта, взяв порядковые номера нового ряда из значений рангов 7-го эксперта.

[1] Расчет и интерпретация парных коэффициентов ранговой корреляции проведем для двух любых экспертов (например, для 7-го и 13-го экспертов). Для других пар они рассчитываются аналогично.

Таблица 1.3

Расчет коэффициента Кендалла для экспертного опроса компании XYZ.

Старые ранги										
Ранги 7 эксперта	**4**	**8**	**3**	**2**	**1**	**10**	**9**	**7**	**5**	***6***
Ранги 13 эксперта	**3**	**9**	**5**	**2**	**1**	**8**	**10**	**4**	**7**	***6***
Новые ранги										
Ранги 7 эксперта – порядковые номера рангов 13-го эксперта	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	***10***
Сами новые ранги 13 эксперта	***1***	***2***	***5***	***3***	***7***	***6***	***4***	***9***	***10***	***8***
Кол-во рангов, находящихся правее и превосходящих данный ранг	*9 (все)*	*8 (все нах. правее)*	*5 (кр. 3 и 4)*	*6 (все вправо)*	*3 (только 9,10,8)*	*3 (только 9,10,8)*	*3 (только 9,10,8)*	*1 (только 10)*	*0 (ничего)*	-
Сумма R	*Σ = 9+8+5+6+3+3+3+1+0=**38***									
Коэффициент Кендалла	**0,68889**									

Для ответа на вопрос, какие из альтернатив являются более приоритетными и какие следует выбрать, создадим сводную таблицу (табл. 1.2), применив к ней метод средних арифметических рангов и метод медиан.

По методу средних рангов выберем три альтернативы, среднее значение рангов по которым меньше остальных: это альтернативы В, Д, Г.

По методу медиан рангов это альтернативы В, Д, Г, А (Г и А имеют одинаковую медиану, равную 4).

Интерпретируем полученные результаты: по выбранным альтернативам наблюдается высокое расхождение во мнениях экспертов, что подтверждается высокими значениями средних квадратических отклонений и коэффициентов вариации. Следовательно, концентрация вокруг средних рангов является невысокой (мнения экспертов разнятся), средние значения рангов получаются низкими (для нас это очень важно) не из-за схожих мнений экспертов, а за счет такого сочетания чисел.

Невысокое значение коэффициента конкордации свидетельствует о низкой согласованности экспертов относительно путей стратегического развития компании.

В данном случае рекомендуется провести повторно мозговой штурм и экспертный опрос, изменив обстоятельства и процедуру их проведения, и сформировать новую таблицу. Тем не менее, полученные результаты экспертного опроса являются значимыми, на них можно опираться при принятии решений, а с помощью парных коэффициентов ранговой корреляции Спирмена и Кендалла было установлено, что между мнениями седьмого и тринадцатого экспертов существует тесная связь, позиции данных экспертов являются согласованными.

1.2. Теоретические распределения в изучении фундаментальных взаимосвязей социально-экономических процессов

Ключевой проблемой в данной главе является проблема взаимоотношения случайности и закономерности. Практически все происходящее носит случайный характер. Ведь мы не знаем, что завтра фондовый индекс Доу-Джонса (Dow Jones Industrial Average) примет значение 7952,55, а значит, дневное значение данного индекса можно считать случайной величиной. С другой стороны, все прогнозирование опирается в основном на определенные тенденции, сложившиеся в настоящий момент в исследуемой области. Ответ на данное противоречие оказывается не таким сложным: для связи между случайными процессами и закономерностями (тенденциями) должно пройти определенное время, чтобы данных за временной промежуток было достаточно для утверждения о наличии четкого тренда (тенденции).

Математическая и статистическая науки, неразрывно связанные между собой и одновременно с этим, самостоятельные и самодостаточные, обобщили данную ситуацию на примере генеральной совокупности и выборки.

Большинство вероятностных и статистических подходов сводится к поиску и применению методов и инструментов оценки выборки на предмет ее соответствия генеральной совокупности. Дело в том, что всю совокупность данных невозможно получить и отследить по многим причинам (ограниченность финансовых и человеческих ресурсов, технические и технологические трудности, нехватка времени и пр.). Поэтому ученые и специалисты проводят выборочные наблюдения (выборки) из общего массива данных и на основе инструментов, часть из которых будет представлена ниже, проводят оценку соответствия полученных выборочных данных общим законам поведения главного (генерального) процесса.

Рассмотрим понятия генеральной совокупности и выборки. Генеральная совокупность – массив данных, детальным и подробным образом характеризующий процесс, объект или признак. Это степень детализации достигается благодаря тому, что генеральная совокупность содержит все возможные значения, которые может принимать *случайная величина*. Все возможные значения признака, все исходы в серии наблюдений, все качественные описательные характеристики включаются в генеральную совокупность. В теории вероятностей понятия случайной величины и генеральной совокупности эквивалентны.

На практике чаще встречается и используется генеральная совокупность, содержащая числа, а не суждения. Иначе сфера и состав показателей для анализа данных, представленных не в количественной форме, заранее резко сужается. В дальнейшем мы будем рассматривать числовые генеральные совокупности.

Выборка – однородный массив данных, полученный из генеральной совокупности при прочих равных условиях и представляющий ее малую часть (чаще всего от 5 до 30 %). Важным свойством и целью выборки является ее репрезентативность, то есть степень отражения реальной картины событий. Также выборка должна быть правильно осуществлена. На процесс проведения выборочных наблюдений не должны оказывать влияния другие факторы, как прямого, так и косвенного воздействия. Выборка должна проводиться строго из генеральной совокупности и не допускать никаких предположений.

Например, нужно получить выборку по данным продаж компании Х. В первую очередь, эти данные являются случайной величиной, так как на спрос потребителей компания не может оказать стопроцентного влияния (за исключением продаж, осуществляемых строго по заранее заключенным сделкам с полностью хеджированными рисками с обеих сторон, что тоже в практике ведения бизнеса представляет редкость). В генеральную совокупность будут входить данные по продажам за каждый день работы компании, и данная совокупность будет неограниченной, пока компания будет осуществлять реализацию своих товаров. Безусловно, работать с неограниченным массивом данных не представляется реальной задачей, поэтому можно поставить цель о взятии искомых данных за период или по группе (одному или нескольким видам) товаров, что будет являться выборкой. В данном случае выборкой **не будут** являться данные, взятые в разные дни по разным видам товаров (в день 1 – товар А, в день 2 – товар А и Б), значения продаж, взятые с большим временным разрывом (нерегулярно, например, в дни 3, 4, 5, 10, 25, 26, 27, 28). Также в нашем случае к выборке сложно отнести данные за 1992, 1993, 1994, 1996 и 2005, 2006, 2007 годы, так как цены по указанным годам не сопоставимы даже, если произвести их пересчет на уровень инфляции. И уж точно выборкой нельзя назвать данные, взятые из продаж другой компании для восполнения пробелов в данных исследуемой компании.

После того, как проведена выборка, она представляется в удобном для анализа виде, производится группировка данных в интервальные ряды и начинается анализ.

Но прежде чем рассматривать инструментарий выборочного наблюдения, нужно изучить закономерности генеральных совокупностей, которые описывают полученные выборки.

Как уже отмечалось, генеральная совокупность и случайная величина – понятия эквивалентные. В теории вероятностей *случайной величиной* принято называть функцию, определенную на множестве событий, точный исход которых неизвестен, но известны варианты и вероятности наступления каждого из этих исходов.

Случайные величины бывают двух видов – *дискретные* и *непрерывные*. Дискретные случайные величины могут принимать конечное счетное множество значений (то есть значения данных величин можно пересчитать), а непрерывная случайная величина определена на интервале своих значений (их количество бесконечно). Дискретной случайной величиной можно считать последовательность значений индекса потребительских цен в России, поскольку данный индекс рассчитывается по итогам месяца и является фиксированной величиной. Непрерывной случайной величиной, к примеру, можно считать цены на продовольственные и непродовольственные товары в оптовых и розничных сетях, которые гораздо чаще подвергаются изменениям, и их уж точно не представляется возможным записать в виде последовательности с точным числом значений.

Соответственно, и свойства и закономерности распределения случайных величин дифференцируются на дискретные и непрерывные.

Что же есть вероятность? В теории вероятностей определений вероятности несколько: классическое, статистическое, геометрическое. Можно сформулировать и другие определения. Наиболее понятно определение вероятности события как отношения числа успехов (числа наступлений данного события в ходе опыта) к общему числу испытаний (т.е. сумме успешных и неуспешных исходов). Успехом в испытании признается случай, когда событие произошло, или наступило.

Так, например, общая вероятность всех исходов одного события соответствует 100 % (или 1, что является общепринятым в математике). Такие исходы образуют *полную группу событий*, и их одновременное совместное наступление невозможно.

Важно подчеркнуть, что недостаточно знать, какие значения будет принимать случайная величина (например, поднимет ли Банк России ставку рефинансирования на 0,5 процентных пункта, 1 п.п. или снизит на 1,5 п.п.). Для принятия упреждающих решений наряду с самими событиями важно знать и вероятности их наступления. Поэтому в математической науке при изучении вероятностей большое внимание уделяется законам распределения вероятностей.

Законом распределения случайной величины (СВ) называют соответствие значений случайной величины и вероятностей их появления в серии испытаний. Помимо просто значений случайной величины, в законе распределения должно быть прописано и установлено правило, по которым можно рассчитать, какова вероятность того, что случайная величина примет определенное значение, или просто вместе со значениями случайной величины указаны (уже найдены) данные вероятности. Закон распределения может быть задан аналитически, в виде таблицы или графически.

Таблица соответствия значений случайной величины и их вероятностей называется *рядом распределения*. Ряд распределения также бывает двух видов – дискретный и интервальный (непрерывный). Пример дискретного ряда дан в табл. 1.3 Случайная величина, представленная в таблице, заключается в действии Центробанка страны W по снижению (–) или увеличению (+) ставки рефинансирования.

Таблица 1.3

Табличное представление дискретного ряда распределения

Значение (x_i) случайной величины (СВ), п.п.	**-1.5**	**-1**	**-0.5**	**0**	**0.5**	**1**	**1.5**
Вероятность того, что СВ примет значение x_i	*0.05*	*0.095*	*0.18*	*0.16*	*0.27*	*0.21*	*0.035*

Графическое представление этой таблицы называется *многоугольником распределения* (см. рис. 1.3). При этом сумма всех ординат многоугольника распределения представляет собой сумму вероятностей всех возможных значений случайной величины, а, следовательно, равна единице.

На рис. 1.3 видно, что наиболее вероятно действие Центробанка страны W, направленное на повышение ставки рефинансирования на 0,5 процентных пункта.

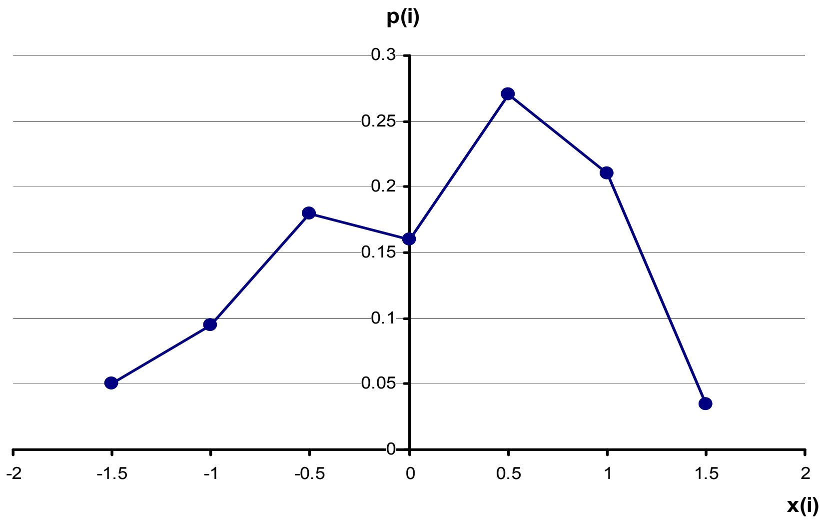

Рис. 1.3. Многоугольник распределения

Однако нерешенным остается вопрос о том, как были рассчитаны вероятности этих событий. Один из способов – экспертные оценки – мы уже рассмотрели. Но в математических дисциплинах вопрос расчета вероятностей и прогнозирования решается и с помощью более продвинутых и сложных математических методов. Один из них – установление фундаментальных характеристик случайной величины, на основании которых можно утверждать о ее неизвестных значениях.

К фундаментальным характеристикам случайных величин помимо закона распределения вероятностей, относятся такие их характеристики, как математическое ожидание, дисперсия, среднее квадратическое отклонение, функция и плотность распределения.

Закон распределения полностью характеризует случайную величину. Однако когда невозможно найти закон распределения, или этого не требуется, можно ограничиться нахождением значений, называемых числовыми характеристиками случайной величины. Эти величины определяют некоторое среднее значение, вокруг которого группируются значения случайной величины, и степень их разброса вокруг этого среднего значения. Рассмотрим каждое из них.

Математическим ожиданием дискретной случайной величины называется среднее арифметическое значение возможных значений случайной величины, взвешенное на их вероятности:

$$M(X) = x_1 p_1 + x_2 p_2 + ... + x_n p_n = \sum_{i=1}^{n} x_i p_i \quad (1.19).$$

Математическое ожидание характеризуется следующими <u>свойствами:</u>

1. Математическое ожидание константы (постоянной величины) равно самой константе:

$$M(C) = C.$$

2. Константу можно вынести вперед за знак от математического ожидания:

$$M(Cx) = CM(x).$$

3. Математическое ожидание совместного наступления двух независимых случайных величин есть произведение математических ожиданий этих величин:

$$M(XY) = M(X)M(Y).$$

Это свойство справедливо для произвольного числа случайных величин.

4. Математическое ожидание суммы двух случайных величин есть сумма математических ожиданий двух этих величин:

$$M(X+Y) = M(X) + M(Y).$$

Это свойство также справедливо для произвольного числа случайных величин.

5. Математическое ожидание M(X) числа появлений некоего события в n независимых испытаниях равно произведению числа испытаний на вероятность p появления события в каждом испытании:

$$M(X) = np.$$

Однако математическое ожидание не может полностью охарактеризовать случайный процесс. Кроме математического ожидания есть величина, которая характеризует отклонение значений случайной величины от математического ожидания.

Это отклонение равно разнице между случайной величиной и математическим ожиданием. При этом математическое ожидание данного отклонения обращается в ноль. Это объясняется тем, что одни возможные отклонения положительны, другие отрицательны, и в результате их взаимного погашения получается ноль.

Дисперсией (рассеиванием) дискретной случайной величины называется средний квадрат отклонения значений случайной величины от их математического ожидания:

$$D(X) = M\left(X - M(X)\right)^2 \quad (1.20).$$

Дисперсия равна разности между математическим ожиданием квадрата случайной величины X и квадратом ее математического ожидания:

$$D(X) = M\left(X^2\right) - \left(M(X)\right)^2 \quad (1.21).$$

Свойства дисперсии следующие (они понадобятся для нахождения дисперсий случайных величин, распределенных по разным законам):

1. Дисперсия константы равна нулю:

$$D(C) = 0 .$$

2. Константа выносится за знак дисперсии в квадрате:

$$D(CX) = C^2 D(X) .$$

3. Дисперсия суммы двух независимых случайных величин[1] есть сумма дисперсий данных величин:

$$D(X+Y) = D(X) + D(Y) .$$

4. Дисперсия разности двух независимых случайных величин есть также сумма дисперсий данных величин:

$$D(X-Y) = D(X) + D(Y) .$$

Средним квадратическим отклонением (мы рассматривали этот показатель раньше) случайной величины X называется квадратный корень из дисперсии:

$$\sigma(X) = \sqrt{D(X)} .$$

Среднее квадратичное отклонение суммы конечного числа взаимно независимых случайных величин равно квадратному корню из суммы квадратов средних квадратических отклонений этих величин:

$$\sigma(X_1 + X_2 + \ldots + X_n) = \sqrt{\sigma^2(X_1) + \sigma^2(X_2) + \ldots + \sigma^2(X_n)} \quad (1.22).$$

Во всех рассмотренных выше случаях случайная величина определялась путем задания значений самой величины и вероятностей этих значений. Однако такой метод применим далеко не всегда. Например, в случае непрерывной случайной величины, ее значения могут заполнять некоторый произвольный интервал. Очевидно, что в этом случае задать все значения случайной величины просто нереально.

Даже в случае, когда это сделать можно, зачастую задача решается чрезвычайно сложно, особенно, когда множество значений слу-

---

[1] Две и более случайные величины считаются независимыми, если одно из них никак не влияет на исход другого (например, две компании из разных отраслей, не связанные друг с другом экономическими и хозяйственными отношениями, терпят банкротство).