Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

Методы интеллектуального анализа данных позволяют повысить эффективность работы контролера по обнаружению нестандартных действий (в т.ч. мошеннических). В ходе проведения корпоративных расследований ему могут понадобиться специализированные инструменты для отбора фактов и обнаружения отклонений в объемных и сложных финансовых учетных данных. С помощью предлагаемого в статье метода можно сформировать индикатор риска, позволяющий быстро это сделать.

Евгений ЗВЕРЕВ, *CIA, член НП «Институт внутренних аудиторов»*
Андрей НИКИФОРОВ, *риск-менеджер*

Распределение Бенфорда: выявление нестандартных элементов в больших совокупностях финансовой информации

В предыдущей статье[1] авторы предложили применять кластерный анализ для извлечения и агрегации из больших совокупностей финансовых и (или) иных учетных данных неочевидных, объективных и полезных закономерностей, позволяющих предположить/выявить наличие в них нестандартных (возможно, мошеннических) элементов. Здесь под понятием «элементы совокупности» понимаются транзакции, остатки или обороты по счетам, стоимость активов и т.п.

Метод, рассматриваемый в данной статье, также относится к выявляющим (detective) и (или) компенсирующим (compensative) контролям (если функционирует в режиме реального времени), хотя на его основе можно создавать и предиктивные контроли. По сути это индикатор риска, «красный флаг». «Красные флаги» обычно субъективны (сформированы на основе практического опыта контролера), и поэтому есть вероятность, что какой-то не попадет в поле зрения во время выполнения задания. Подход к формированию «красного

[1] Зверев Е., Никифоров А. Кластерный анализ: формирование индикатора риска для больших совокупностей учетной информации // Внутренний контроль в кредитной организации. 2018. № 3.

флага», рассматриваемый здесь, основан на статистических закономерностях поведения элементов больших совокупностей, поэтому подобный индикатор риска имеет высокую статистическую значимость и является достаточно объективным.

Распределение Бенфорда[1] (частотное распределение значимых цифр в больших *числовых* совокупностях) — подход к анализу финансовых данных, который широко обсуждается в профессиональном сообществе.

Надо отметить, что интернет просто «завален» различными публикациями касательно этого метода. В них есть все: истории появления, попытки объяснить его математическую (вероятностную) основу, словесное описание приемов использования и открывшиеся перспективы применения, но нет основного: как применять и для каких совокупностей, какой результат будет получен[2].

Описание индикатора риска

Распределение Бенфорда (далее — Распределение) как некоторое эмпирическое наблюдение известно сравнительно давно (с третьей четверти XIX в.), но только недавно стало применяться для анализа данных в финансовом секторе:

— банками для выявления фальсификации отчетности заемщиками или мошеннических транзакций по кредитным картам;

— страховыми компаниями для выявления ложных требований застрахованных;

— налоговыми службами для выявления налогового мошенничества;

— аудиторскими фирмами для выявления нарушений бухгалтерского учета.

В чем суть этого распределения?

Берем игральный кубик, на сторонах которого изображены цифры от 1 до 6, и долго бросаем его, например 600 раз. Теория вероятности утверждает, что каждое из шести чисел кубика выпадет приблизительно 100 раз, но чтобы это правило проявилось, нужно бросать кубик многократно, а если бросить его, к примеру, всего шесть раз,

[1] Фрэнк Бенфорд (Frank Benford) — американский физик. См.: Benford F. The law of anomalous numbers // Proceedings of American Philosophical Society. 1938. Vol. 78. No. 4. P. 551-572.

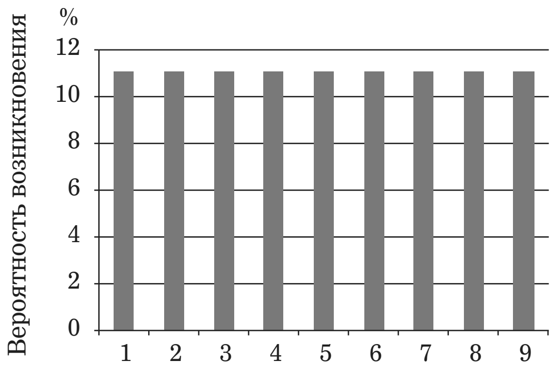
[2] Отметим наиболее значимые публикации: Nigrini M.J. Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2011. P. 330; Nigrini M.J., Wells J.T. Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2012. P. 352; Арнольд В.И. Антинаучная революция и математика // Вестник Российской академии наук. 1999. Том 69. № 6. С. 553-558; Арнольд В.И. Статистика первых цифр степеней двойки и передел мира // Квант. 1998. № 1.

Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

то вовсе не факт, что каждая из цифр выпадет по разу. Математики называют такое выпадение цифр равномерным распределением (рис. 1).

Рисунок 1

Равномерное распределение



Только далеко не во всех случайных событиях вероятность появления каждой из цифр *одинакова*: иногда теория вероятности показывает удивительные фокусы.

Распределение[1] (рис. 2) устанавливает, что для практически любых видов значительных совокупностей данных (длины рек в футах или метрах, население городов и стран, объем торгов на фондовых биржах, количество рейтинговых очков теннисных профи, молекулярные массы химических веществ, высота самых высоких зданий в мире и т.д.) из *естественных источников данных* первая значащая цифра каждого числа будет равна:

— 1 (например, 1, 157, 1812) примерно в 30% случаев;

— 2 примерно в 18% случаев.

Для каждой последующей цифры вероятность возникновения будет уменьшаться, и цифра 9 будет иметь наименьшую частоту — около 5%.

Здесь под термином «естественный источник данных» понимаются требования по отсутствию в совокупности:

[1] В 1938 г. Ф. Бенфорд решил разобраться с такой особенностью чисел. Для этого он изучил более 20 разных списков с числами, такими как площади бассейнов более 300 рек, численность населения стран, показатели смертности в разных государствах, молекулярные массы около 1800 разных веществ и т.д. По всем показателям Ф. Бенфорд составил таблицы, где его интересовала только первая цифра в каждом из чисел. Казалось бы, все эти списки и числа случайны — примерно как результаты бросания кубика, и в этих числах должна равномерно встречаться первая цифра от 1 до 9. Но оказалось, что это не так. Данное распределение Ф. Бенфорд назвал законом аномальных чисел.

Рисунок 2

Распределение Бенфорда для первой значащей цифры числа

— ограничения чисел по максимуму и минимуму. Если есть некая граница (допустим, предельный размер расчетов наличными), то такая совокупность данных уже может не подчиняться идеально Распределению;

— системы нумерации. Числа не должны быть составными системами. Например, набор цифр в ИНН не будет являться Распределением, так как первые две цифры в ИНН — код региона, вторые две — код инспекции, а последняя цифра — контрольная, вычисляется по всем предыдущим.

Существуют определенные условия, которым должны соответствовать совокупности, предполагаемые к анализу на соответствие Распределению:

— экспоненциальное распределение исследуемых данных. Правда, на практике такое соответствие соблюдается редко, но они должны к этому «стремиться»;

— данные должны относиться к одинаковым объектам. Например, нельзя смешивать данные платежных поручений и данные адресов клиентов.

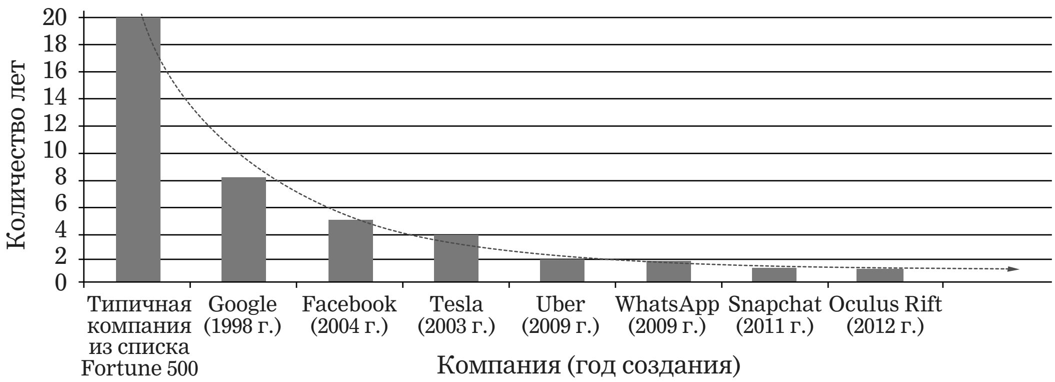
Естественные финансовые учетные данные (остатки на счетах кредитных карт и (или) транзакции по ним, суммы бухгалтерских проводок, страховых выплат, гарантийного ремонта, выставленных счетов к оплате, объемы поставок, суммы в налоговых декларациях, списки стоимости покупок или денежных поступлений и пр.) также подчиняются Распределению, причем оно фактически выполняется, даже если значения конвертируются из одной валюты в другую.

Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

Применимость распределения Бенфорда для таких данных, по мнению авторов, объясняется тем, что при отсутствии ограничений развитие бизнеса порождает экспоненциальный рост, когда прирост пропорционален достигнутому результату (рис. 3).

Рисунок 3

Достижение рыночной капитализации $1 млрд[1]



Несмотря на эмпирическое подтверждение эффективности Распределения, оно не получало практического применения до тех пор, пока у финансистов не возникла необходимость разработать какие-то действенные инструменты, используя которые можно было бы выявить сфальсифицированные данные.

В 1990-х годах американский математик Марк Нигрини[2] обратил внимание, что Распределению подчиняются числа в налоговых и бухгалтерских документах. Поэтому он составил специальную программу для проверки налоговых данных. Дело в том, что налоговые службы не могут проверить все компании на правильность уплаты налогов — это физически невозможно из-за очень большого объема данных. Применение Распределения дает возможность «показать пальцем» (сформировать индикатор риска), какие данные нужно проверять, поскольку несоответствие данных распределению первой цифры вызывает подозрение в том, что они были кем-то «нарисованы», а не получены в процессе *естественной* финансово-хозяйственной деятельности. Иначе говоря, какая-то фирма попыталась

[1] Исмаил С., Мэлоун М., ван Геест Ю. Взрывной рост: почему экспоненциальные организации в десятки раз продуктивнее вашей (и что с этим делать). М.: Альпина Паблишер, 2017. С. 440.

[2] См.: Nigrini M.J. Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2011. P. 330; Nigrini M.J., Wells J.T. Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2012. P. 352.

исказить суммы налоговых платежей и таким образом укрыться от обязательных выплат государству.

На данный момент известно около десяти видов тестов, основанных на Распределении, из них наиболее распространены шесть (рис. 4).

Рисунок 4

Перечень тестов[1], основанных на распределении Бенфорда

1. Анализ частоты первой значащей цифры

- В данном случае используется само Распределение

2. Совместный анализ частот первой и второй цифр

- При использовании данного теста отдельно проверяется частота цифры от 1 до 9 на первой позиции и частота цифры от 0 до 9 на второй. Затем составляется таблица соответствий, которая анализируется на отличие частоты цифр фактического распределения от самого Распределения

3. Анализ дублей

- Данный метод опирается только на методологию Бенфорда, а не на само Распределение. Проверка выявляет частоту числовых повторов в большом количестве документации. Все повторяющиеся числа в исследуемых данных сортируются по частотности повторов, а затем проверяются уплотнения повторов ряда чисел

4. Анализ первой пары цифр

- Этот метод фактически представляет собой усовершенствованный второй тест, так как он исследует частоту появления цифр в начале числа от 10 до 99

5. Анализ первой тройки цифр

- Метод, более точный в сравнении с первым, вторым и четвертым тестами. Анализируется частота первой тройки цифр от 100 до 999 в изучаемой числовой последовательности. Данный метод используют при проверке большого объема информации (от 10 000 значений)

6. Анализ округлений

- Тест проводится для проверки частоты последних значащих цифр анализируемой числовой последовательности и позволяет выявить не соответствующую Распределению частоту постоянного округления в большую или меньшую сторону

[1] Гусев И.Ю. Методы, с помощью которых аудиторы выявляют мошенничество внутри компании // Российский налоговый курьер. 2012. № 21.

Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

Описание модели индикатора риска

Распределению Бенфорда подчиняются цифры в числах, образующих геометрические прогрессии или имеющих экспоненциальную шкалу измерения[1]. Нетипичные транзакции, «выпадающие» из логики бизнеса какой-либо производственной компании или финансовой организации, с большой вероятностью будут отклоняться от Распределения. Определение частот выпадения значащих цифр в таких данных и сравнение их с теоретическим распределением позволят выявить подозрительные (нетипичные) транзакции для дальнейшего детального анализа их внутренним аудитором.

Распределение Бенфорда для первой значащей цифры рассчитывается следующим образом:

$$P(n) = \log_{10}\left(1 + \frac{1}{n}\right),$$

где n — первая значащая цифра числа, от 1 до 9;
P — вероятность.

Графически оно представлено выше (см. рис. 2). Следует отметить, что данная формула с небольшими дополнениями справедлива для формирования Распределения любой по порядку значащей цифры числа, а также для набора цифр, например для первой пары.

Сравнение распределений для первой, второй, третьей и четвертой значащих цифр представлено далее (рис. 5). Из данного сравнения видно, что осуществлять тестирование следует по первой и второй цифрам, поскольку третья и четвертая цифры (на рис. 5 их графики совпали), а также все последующие распределяются практически равномерно, что не позволит выявлять нетипичные для бизнеса (подозрительные) транзакции.

Весьма полезно тестировать распределение первой пары цифр числа на соответствие распределению Бенфорда (рис. 6). Марк Нигрини[2] считает это более предпочтительным, чем тестирование распределения только одной цифры.

Кроме того, при анализе распределения первых двух цифр получается существенно большее количество значений, что позволяет более корректно применять статистические тесты (см. далее) для оценки совпадения фактической частоты распределения цифр и теоретического Распределения.

[1] См.: Арнольд В.И. Статистика первых цифр степеней двойки и передел мира // Квант. 1998. № 1.

[2] См.: Nigrini M.J. Benford’s Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2011. P. 330; Nigrini M.J., Wells J.T. Benford’s Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2012. P. 352.

Рисунок 5

Распределение Бенфорда для первой, второй, третьей и четвертой цифр

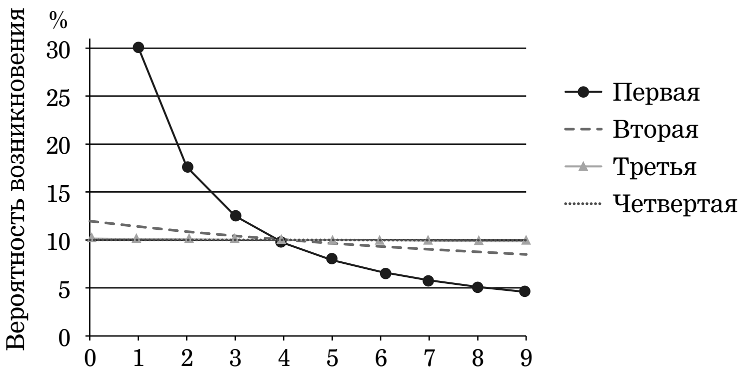
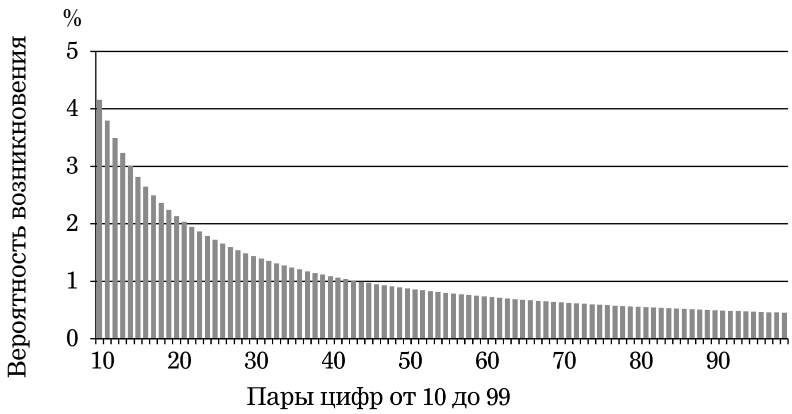


Рисунок 6

Распределение Бенфорда для первых двух цифр



Для получения лучшего результата рекомендуется сочетать несколько методов и обращать внимание на совпадающие результаты. Например, можно проверить распределение первой значащей цифры, второй значащей цифры и сравнить результат с распределением по первым двум цифрам.

Методология применения модели

Модель применяется следующим образом:

1) выбираются данные для анализа;

2) рассчитывается частота выпадения цифр;

Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

3) полученное распределение сравнивается с теоретическим Распределением;

4) определяются транзакции с максимальным отклонением от Распределения;

5) по отобранным транзакциям проводится детальная проверка.

Культура статистического исследования предполагает оценку статистической значимости[1] как соответствия фактического распределения теоретическому, так и достоверности полученных результатов. Для этой цели авторы предлагают:

— во-первых, для оценки соответствия фактического (рассчитанного) распределения теоретическому использовать критерий χ^2 (хи-квадрат), чтобы определить, подтверждается ли гипотеза экспериментом, и (или) оценить, насколько полученное распределение в целом соответствует теоретическому. С этой целью можно использовать встроенную функцию Excel: = ХИ2.ТЕСТ(фактические_данные; теоретические_данные), где фактические_данные — данные, содержащие фактические частоты выпадения цифр в исследуемой совокупности, теоретические_данные — данные, содержащие теоретическую (в соответствии с Распределением) частоту выпадения цифр;

— во-вторых, для подтверждения достоверности выявленных нестандартных транзакций использовать Z-тест с заданным уровнем доверительной вероятности (предполагается, что отклонение фактического распределения от теоретического возможно в результате воздействия большого количества случайных факторов, т.е. распределяется по нормальному закону).

Значение для Z-теста рассчитывается следующим образом[2]:

$$Z = \frac{|f - p| - \left(1/2n\right)}{\sqrt{\dfrac{p(1-p)}{n}}},$$

где f — фактическая частота появления цифры;

p — теоретическая вероятность появления цифры (распределение Бенфорда);

n — количество рассматриваемых цифр (или комбинаций цифр).

[1] Статистическая значимость (statistical significance) — статистические критерии для оценки получаемых результатов. Эти критерии позволяют оценить вероятность того, что такие результаты могли появиться чисто случайно. Термин «статистическая значимость» употребляется как раз в связи с использованием таких критериев. Характеристика статистически значимых дается результатам, вероятность случайного появления которых равна или ниже некоторого общепринятого уровня. Традиционно статистически значимый уровень 5% (или ниже) — вероятность случайного получения результата; это обычно выглядит как $p < 0{,}05$ (или $p < 0{,}01$) и означает, что если бы данное исследование повторили 100 раз, случайного появления таких результатов можно было бы ожидать менее чем в 5 случаях (или менее чем в 1 случае соответственно).

[2] См.: Nigrini M.J. Benford’s Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2011. P. 330; Nigrini M.J., Wells J.T. Benford’s Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons, 2012. P. 352.

Критическое значение Z определяется по доверительной вероятности:

$$Z_c = \mathrm{ABS}\left(N^{-1}\left(\frac{1-P_d}{2}\right)\right),$$

где N^{-1} — обратная функция стандартного нормального распределения, встроенная функция Excel: = НОРМ.СТ.ОБР();

P_d — доверительная вероятность.

Если критерий $Z > Z_c$, то считается, что отклонение частоты появления цифры превышает допустимую случайную величину, то есть полученный результат не случаен (статистически значим). Транзакции с такими цифрами следует отобрать для более детальной проверки. Рекомендуется выбирать значение доверительной вероятности в диапазоне от 80 до 99%, поскольку *увеличение* доверительной вероятности ограничивает количество транзакций, подлежащих более детальному анализу.

Описание тестируемой совокупности

Предлагаемая авторами модель была применена для анализа обязательств контрагентов производственной компании, иными словами, рассматривалась совокупность *естественных финансовых учетных данных*, отражающих реальный бизнес-процесс. При этом было выполнено три теста: анализ распределения частоты выпадения первой цифры, второй цифры и пары первых двух цифр.

Анализ частоты первой значащей цифры

Рассчитывалась частота выпадения первой значащей цифры (от 1 до 9) у каждого элемента (числа) из совокупности.

Результат теста по первой цифре представлен в табл. 1 и на рис. 7. Доверительная вероятность для Z-теста выбрана на уровне 95%, величина Z_c равна 1,960.

Таблица 1

Тестирование совокупности по первой значащей цифре

Первая цифра	Распределение, %		Разность, %	Z-тест	Z_c
	фактич.	теоретич.			
1	30,0	30,1	–0,1	0,257	1,960
2	18,3	17,6	0,7	1,774	
3	12,4	12,5	–0,1	0,208	

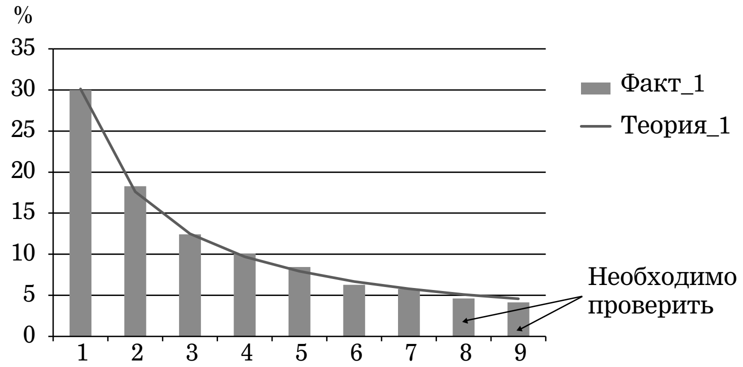
Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

Окончание табл. 1

Первая цифра	Распределение, %		Разность, %	Z-тест	Z_c
	фактич.	теоретич.			
4	10,1	9,7	0,4	1,264	
5	8,5	7,9	0,5	1,951	
6	6,3	6,7	–0,4	1,679	
7	5,8	5,8	0,0	0,189	
8	4,6	5,1	–0,5	**2,135**	
9	4,1	4,6	–0,4	**2,109**	

Рисунок 7

Графическое представление теста по первой цифре



При проведении данного теста были отобраны транзакции, содержащие среди значащих цифр на первом месте цифры 8 и 9.

Анализ частоты второй цифры

Рассчитывалась частота выпадения второй цифры (от 0 до 9) у каждого элемента (числа) из совокупности.

Результат теста по второй цифре представлен в табл. 2 и на рис. 8. Доверительная вероятность для Z-теста выбрана на уровне 80%, величина Z_c равна 1,282.

Таблица 2

Тестирование совокупности по второй цифре

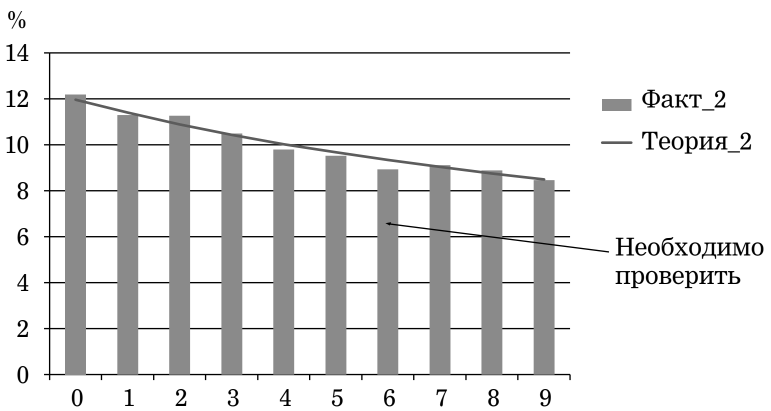
Вторая цифра	Распределение, %		Разность, %	Z-тест	Z_c
	фактич.	теоретич.			
0	12,2	12,0	0,2	0,669	1,282
1	11,3	11,4	–0,1	0,293	

Окончание табл. 2

Вторая цифра	**Распределение, %**		**Разность, %**	**Z-тест**	**Z_c**
	фактич.	**теоретич.**			
2	11,3	10,9	0,4	1,230	
3	10,5	10,4	0,1	0,203	
4	9,8	10,0	–0,2	0,752	
5	9,5	9,7	–0,1	0,484	
6	8,9	9,3	–0,4	**1,347**	
7	9,1	9,0	0,1	0,279	
8	8,9	8,8	0,1	0,453	
9	8,5	8,5	0,0	0,090	

Рисунок 8

Графическое представление теста по второй цифре



При проведении данного теста были отобраны транзакции, содержащие среди значащих цифр на втором месте цифру 6.

Анализ частоты первой пары цифр

Рассчитывалась частота выпадения первой пары цифр (от 10 до 99) у каждого элемента (числа) из совокупности.

Результат теста по первым двум цифрам представлен в табл. 3 и на рис. 9. Доверительная вероятность для Z-теста была выбрана на уровне 95%, величина Z_c равна 1,960.

При проведении данного теста были отобраны транзакции, содержащие среди первой пары цифр следующие комбинации: 43, 55, 72, 96, 98.

Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

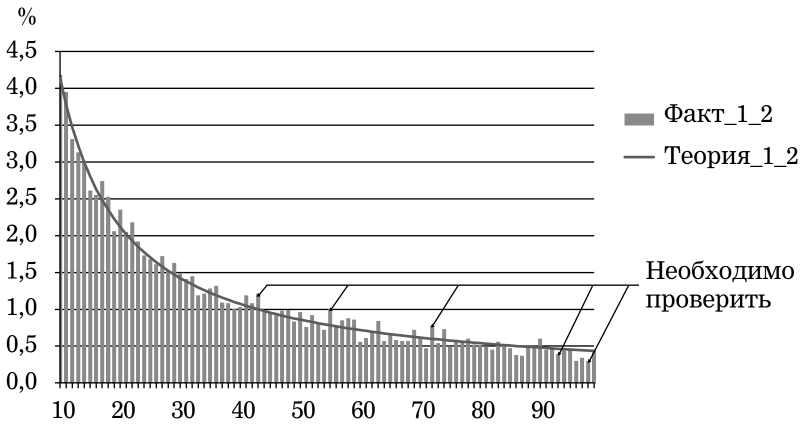
Таблица 3

Тестирование совокупности по первой паре цифр

Первая пара цифр	Распределение, %		Разность, %	Z-тест	Z_c
	фактич.	теоретич.			
10	4,2	4,1	0,0	0,13	1,960
11	3,9	3,8	0,2	0,82	
…	…	…	…	…	
42	1,1	1,0	0,0	0,43	
43	1,2	1,0	0,2	**1,98**	
44	1,0	1,0	0,0	0,21	
…	…	…	…	…	
54	0,7	0,8	–0,1	0,92	
55	1,0	0,8	0,2	**2,41**	
56	0,8	0,8	0,0	–0,04	
…	…	…	…	…	
71	0,5	0,6	–0,1	1,83	
72	0,8	0,6	0,2	**2,02**	
…	…	…	…	…	
96	0,3	0,5	–0,2	**2,32**	
…	…	…	…	…	
98	0,3	0,4	–0,2	**2,20**	
99	0,4	0,4	0,0	–0,02	

Рисунок 9

Графическое представление теста по первой паре цифр



Распределение Бенфорда: выявление нестандартных элементов в больших совокупностях финансовой информации

Обращаем внимание, что по определенным транзакциям результаты теста по первой и второй цифрам и по первой паре цифр совпали. При тесте по первой цифре была выбрана цифра 9, при тесте по второй цифре — 6. При тесте по первой паре цифр была выбрана цифра 96. Таким образом, в обоих случаях были отобраны транзакции с суммами, начинающимися на 96.

Дальнейший подробный анализ результатов тестирования подтвердил наличие транзакций, нетипичных для деятельности конкретно этой производственной компании: отсутствующие в предыдущие отчетные периоды авансовые платежи на небольшие суммы, не характерные для деятельности компании, ссуды физическим лицам и прочие операции, требующие детальной проверки со стороны службы внутреннего аудита.

Рекомендации

Для более эффективного тестирования на основе Распределения авторы рекомендуют:

1. Выбирать данные в денежном выражении, которые отражают специфику бизнеса (*естественные финансовые учетные данные*). Нельзя смешивать данные, относящиеся к разным учетным системам (например, данные производственного и бухгалтерского учета), организационным структурам (например, данные по филиалу и головной компании) или бизнес-процессам (например, данные закупок и продаж).

2. Сочетать и комбинировать результаты разных тестов: по первой цифре, по второй цифре, совместный анализ частот первой и второй цифр, по первым двум цифрам. Нет смысла проводить тесты по третьей, четвертой и последующим цифрам, так как распределение вероятности появления таких цифр близко к равномерному (см. рис. 5).

3. Доверительную вероятность для Z-теста выбирать в пределах от 80 до 99%; варьируя доверительную вероятность, можно определять количество транзакций для более детального анализа, исходя из наличия ресурсов и трудоемкости его проведения.

4. Использовать отклонения от распределения Бенфорда, чтобы судить по ним о возможном проведении нетипичных (подозрительных) транзакций, которые нарушают экспоненциальную шкалу или логику масштабирования бизнеса организации, близкую к геометрической прогрессии. Например: разбиение крупных платежей на мелкие, ограничение платежей, проведение транзакций, не характерных для бизнеса компании.

Евгений ЗВЕРЕВ
Андрей НИКИФОРОВ

Заключение

Несмотря на широту применения распределения Бенфорда, не следует забывать, что существуют данные, не подчиняющиеся ему, а также объемы данных, размер которых недостаточен для применения статистических методов. Тем не менее, тесты, предложенные М. Нигрини, вполне справедливо на нем основываются. Они совершили переворот в аудите, поскольку раньше данные в налоговых декларациях проверялись лишь выборочно, а теперь можно осуществить тестирование практически любого количества информации. Естественно, результаты тестирования не всегда верны и могут указывать на ложные элементы, но нельзя отрицать, что они являются важным дополнительным инструментом в делах, связанных с махинациями.

Однако даже при использовании столь «продвинутых» методов контролеру необходимо скрупулезно проверять указанные им элементы совокупности для формирования достоверного заключения, поскольку при вынесении профессионального суждения он всегда должен руководствоваться качеством доказательств.

Послесловие ко всему циклу статей

Периодически авторам задают вопрос: «Сможет ли мошенник, зная методы, предлагаемые вами, подобрать величины элементов финансовой совокупности таким образом, что эти методы их “не увидят”?» Если теоретическое распределение достоверно отображает фактическую совокупность, то внедренные в нее мошеннические элементы скрыть невозможно, поскольку они будут «выпадать» из естественного поведения финансовых учетных данных, а критерием указанного соответствия являются рекомендуемые авторами коэффициенты статистической значимости[1].

[1] См.: Зверев Е., Никифоров А. Сравнение данных учетных систем: как выявить манипуляции с бухгалтерской отчетностью? // Внутренний контроль в кредитной организации. 2018. № 2; Кластерный анализ: формирование индикатора риска для больших совокупностей учетной информации // Внутренний контроль в кредитной организации. 2018. № 3.