

Применение статистик m^* и d^* для оценки качества данных психологического исследования с помощью закона Бенфорда (на примере измерения времени реакции)

Колачев Н.И.*

Национальный исследовательский университет
«Высшая школа экономики» (ФГБОУ ВО «НИУ ВШЭ»)
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-3214-6675>
e-mail: nkolachev@hse.ru

Целью предлагаемого исследования было рассмотрение свойств статистик, используемых для изучения согласия распределения первой значащей цифры закону Бенфорда, m^* и d^* при достаточно скромных объёмах выборки ($10 \leq n \leq 70$). Для достижения цели было проведено симуляционное исследование. Симулировались данные лог-нормального распределения с параметрами, имитирующими распределение измерений времени реакции. Распределение первой значащей цифры изучалось у стандартизированных значений, возведённых в степень γ ; $5 \leq \gamma \leq 100$. Выяснилось, что статистика m^* не зависит от степени числа в отличие от d^* . Были установлены критические значения для выборок объёмом от 10 до 70 наблюдений с шагом $h = 10$. Оказалось, что критические значения при малых n статистики d^* близки к асимптотическим, в то время как критические значения статистики m^* существенно больше. Функциональность установленных критических значений проверялась в рамках экспериментального исследования: один респондент выполнял когнитивный тест Струпа в соответствии с инструкцией (контрольный случай), другой – нарушал её (экспериментальный случай). Обнаружено, что статистика d^* не позволяет дифференцировать поведение испытуемых. В свою очередь m^* оказалась чувствительной к изменениям в поведении респондентов и в экспериментальном случае существенно чаще позволяла отвергнуть нулевую гипотезу о соответствии распределения первой значащей цифры стандартизированного значения времени реакции закону Бенфорда в сравнении с контрольным. Таким образом, делается предварительный вывод о том, что статистика m^* в сравнении с d^* более функциональна при изучении качества данных о времени реакции при небольших n .

Ключевые слова: закон Бенфорда, время реакции, статистика m^* , статистика d^* .



Для цитаты:

Колачев Н.И. Применение статистик m^* и d^* для оценки качества данных психологического исследования с помощью закона Бенфорда (на примере измерения времени реакции) // Моделирование и анализ данных. 2024. Том 14. № 2. С. 23–44. DOI: <https://doi.org/10.17759/mda.2024140202>

*Колачев Никита Игоревич, кандидат психологических наук, старший преподаватель департамента психологии, Национальный исследовательский университет «Высшая школа экономики» (ФГБОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-3214-6675>, e-mail: nkolachev@hse.ru

1. ВВЕДЕНИЕ

Высокое качество собранных данных имеет ключевое значение для формулирования выводов во многих эмпирических исследованиях. На достоверность собираемых данных могут влиять различные факторы: несовершенства измерительного инструмента, методические погрешности, отсутствие стандартизации процедуры сбора данных и многое другое. Особого внимания требуют результаты измерения нефизических величин (например, мотивации, когнитивных процессов и пр.), которые не поддаются непосредственному, прямому измерению, и, как правило, подвержены влиянию флуктуаций в поведении респондента (испытуемого).

В части исследования качества многое сделано для оценки достоверности данных, собираемых опросно-анкетными методами. К примеру, хорошо известно, как справляться с таким фактором недостоверности самоотчётных данных, как социальная желательность [7]. Более того, разработаны специальные подходы для измерения характеристик личности, которые могут подвергаться искажениям при самоотчёте. К числу таких подходов относятся методы поведенческих сценариев [22], якорных виньеток [23], сценарных шкал [11].

При этом необходимо констатировать, что на сегодняшний день не предложены методы анализа достоверности данных о времени реакции, что делает настоящую работу особенно актуальной, ведь измерение времени реакции является основой когнитивной психологии [9].

Время реакции

Время реакции широко распространено и обычно используется в науках о поведении для оценки и моделирования когнитивных процессов, чаще всего внимания [16]. Под временем реакции понимают временной интервал между появлением какого-либо стимула и ответом (реакцией) на этот стимул со стороны испытуемого [3]. Обычно измеряется в миллисекундах (мс). В исследованиях с использованием показателя времени реакции выделяют три вида возможных реакций [1]:

- Простая сенсомоторная реакция (реакция на появление любого стимула);
- Дифференцировочная реакция простого выбора (реакция на стимулы определенного типа и игнорирование всех остальных);
- Дифференцировочная реакция сложного выбора (ответ на каждый тип стимула должен сопровождаться определенной реакцией).

На протяжении многих лет время реакции служило основой для значимых открытий в экспериментальной, когнитивной психологии, нейрофизиологии, таких как эффект Струпа [26], аффективный прайминг [29] и др. По сей день вопросы работы со временем реакции являются актуальными для исследователей. Так, известно, что распределение времени реакции во многих исследованиях имеет правостороннюю асимметрию [28], поэтому в качестве меры центральной тенденции рекомендуют использовать медиану или моду, а не среднее арифметическое. Кроме того, ряд исследователей говорит о лог-нормальности распределения времени реакции, то есть логарифм времени реакции имеет нормальное распределение [18]. Некоторые исследователи указывают, что распределение времени реакции подчиняется так называемому экс-гауссовскому распределению (смесь гауссовского и экспоненциального) [28].

Основным фактором, влияющим на качество данных о времени реакции, является невыполнение испытуемым поставленной задачи. Если испытуемый хаотично, бессистемно реагирует на стимулы (например, нажимать клавиши на клавиатуре компьютера, ведь большая часть современных исследований проводится в компьютерном формате, даже онлайн [24]), то исследователи получают недостоверные данные, ничего не свидетельствующие об особенностях познавательных процессов испытуемого. Существующие рекомендации по работе с данными о времени реакции лишь предлагают способы «подгонки» данных под существующие статистические методы, как, например, исключение выбросов по правилу « ± 2 стандартных отклонения», или методы работы с пропущенными данными [17]. Все эти рекомендации не позволяют судить о достоверности собранных данных. В связи с этим возникает острая необходимость в поиске способов проверки качества получаемых или генерируемых данных. Одним из вариантов может выступить использование закона Бенфорда.

Закон Бенфорда

Закон Бенфорда – известная закономерность в области математики. Она заключается в том, что в некотором большом наборе числовых данных первая значащая цифра, не включая ноль, встречается с частотой, равной $\log\left(1 + \frac{1}{k}\right)$, где k – первая значащая цифра от 1 до 9 [5]. Закономерность была получена Фрэнком Бенфордом, сотрудником известной компании «General Electric», на примере большого числа данных: длин рек, численности городов США, смертности и пр. [12]. Позднее этот закон нашёл отражение во многих областях человеческой деятельности. К примеру, закономерность Бенфорда используют для выявления подлога в бухгалтерской отчетности и оценки качества отчетности о движении денежных средств [8], фальсификаций на выборах, подсчёта количества числительных в статьях [2], достоверности данных профилактического скрининга в области медицины [10]. Поскольку закономерность Бенфорда воспроизводится на разнообразных данных, некоторые авторы называют её фундаментальным законом природы [6].

Стоит отметить, что не все переменные подчиняются закону Бенфорда. Так, считается, что «почтовые индексы, выигрышные номера в лото и рулетку, номера телефонов и любые объемы данных, размер которых не достаточен для применения



статистических методов» [4, с. 76]. Кроме того, время реакции в пределах от 100 до 200 мс может быть, например, результатом быстрой догадки. Такие быстрые реакции легко выявляются, и их обычно исключают из анализа [28]. При этом такие значения встречаются в распределении реже, чем, допустим, значения в интервале от 400 до 500 мс, что говорит о низкой согласованности первой значащей цифры исходных данных о времени реакции с законом Бенфорда, поскольку, согласно этому закону, цифры 1 и 2 встречаются чаще, чем 4 и 5.

Другой существенной проблемой применения закона Бенфорда является подбор подходящих методов оценки согласия эмпирических данных с этим законом. Тестирование гипотезы о соответствии эмпирического распределения значащих цифр закону Бенфорда обычно происходит методом хи-квадрата или с помощью критерия Колмогорова [21]. Как полагает Джон Морроу, эти тесты являются достаточно консервативными, то есть довольно часто не позволяют отклонить нулевую гипотезу в ситуации, когда данные не согласуются с законом Бенфорда (другими словами, имеют высокую вероятность ошибки II рода). В связи с этим возникла необходимость разработки новых тестов для проверки гипотезы о соответствии данных этому закону.

Модификации статистических тестов согласия с законом Бенфорда

Решением проблемы консервативности статистических тестов занялся Джон Морроу [21]. В своей работе он доказал следующую теорему:

пусть X – непрерывная случайная величина, B – закон Бенфорда, тогда

$$\forall \varepsilon > 0 \exists \gamma^* : \forall \gamma \geq \gamma^* \left| x^\gamma - B \right| < \varepsilon,$$

где x^γ – стандартизированные значения из различных распределений (например, нормального, лог-нормального, экспоненциального и др.), возведённые в степень γ . В оригинальной статье вместо γ фигурирует α , однако, чтобы не путать этот параметр с уровнем значимости, который в статистике обозначается как α , было принято решение использовать другую греческую букву. По итогам анализа Дж. Морроу пришёл к выводу, что 10-е степени ($\gamma = 10$) стандартизированных значений имеют высокое согласие с законом Бенфорда. Однако остаётся неясным, почему выбрана именно 10-я степень.

Кроме того, вместо классических тестов на однородность, таких как хи-квадрат или тест Колмогорова, он предложил критические значения для двух тестов (m^* и d^*), предназначенных для проверки соответствия эмпирического распределения первой значащей цифры закону Бенфорда:

$$m^* = \sqrt{n} \cdot \max_{d \in \{1, \dots, 9\}} \left| P(\text{первая значащая цифра } X = d) - \log_{10} \left(1 + \frac{1}{d} \right) \right|$$

$$d^* = \sqrt{n} \cdot \left[\sum_{d \in \{1, \dots, 9\}} \left(P(\text{первая значащая цифра } X = d) - \log_{10} \left(1 + \frac{1}{d} \right) \right)^2 \right]^{\frac{1}{2}}$$

где n – объём выборки, $\log_{10} \left(1 + \frac{1}{d} \right)$ – закон Бенфорда. В статистической литературе статистику m^* называют расстоянием Чебышёва, а d^* – евклидовым расстоянием [13].

Также учёный показал, что предложенные им статистические тесты достигают достаточной мощности при объёме выборки, превышающем 80 наблюдений (при уровне значимости $\alpha=0.01$).

Необходимо сказать, что психологические исследования, в которых фигурирует время реакции в качестве основной переменной, довольно часто проводятся на скромных по объёму выборках (как правило, $n < 50$). К примеру, Маршалек и коллеги [19] в рамках анализа множества публикаций показали, что средний размер выборки в серьёзном психологическом исследовании составляет 40 человек. Этому есть разумное объяснение. С одной стороны, это обусловлено отсутствием необходимости в больших объёмах, так как небольшой объём выборки компенсируется достаточным количеством производимых измерений у одного человека; с другой стороны, объясняется трудозатратностью самой процедуры сбора данных, так как экспериментальные когнитивные исследования занимают довольно много времени (в среднем 1–1,5 часа) и требуют от участников немалых усилий. Однако несмотря на сложившуюся практику, вопрос проверки качества данных остаётся ключевым, следовательно, даже маленькие по объёму выборки данных должны подвергаться анализу на предмет качества.

В связи с этим целью исследования является изучение свойств статистик m^* и d^* при скромных объёмах выборки ($10 \leq n \leq 70$) на основе данных о времени реакции. Для достижения поставленной цели были сформулированы следующие задачи:

1. Изучить, различаются ли значения m^* и d^* в зависимости от объёма выборки;
2. Исследовать, зависят ли значения статистик m^* и d^* от степени, в которую возводят стандартизированное значение времени реакции, при контроле объёма выборки;
3. Установить критические значения статистик m^* и d^* для небольших по объёму выборок ($10 \leq n \leq 70$) на основе симуляций из лог-нормального распределения;
4. Проверить функциональность установленных критических значений на экспериментальном (эмпирическом) примере.

Для решения поставленных задач сначала проведено симуляционное исследование. Был сгенерирован набор данных о времени реакции следующих объёмов (n): 10, 20, 30, 40, 50, 60 и 70. В качестве функции для генерации данных выбрана лог-нормальная с такими параметрами: $\mu = 5,984$, $\sigma = 0,1245$, что соответствует среднему времени реакции 400 мс и стандартному отклонению 50 мс. Эти значения часто используются в симуляционных исследованиях [28]. Затем сгенерированные значения были стандартизированы по следующей формуле:

$$\frac{x_i - \mu}{\sigma}$$

где x_i – i -е значение переменной из лог-нормального распределения. Далее стандартизированные значения были возведены в k -ю степень с шагом $h = 5$, начиная со второго



элемента; $k \in \{1, 5, 10, \dots, 100\}$. Именно у возведённых в k -ю степень стандартизированных значений изучалось распределение первых значащих цифр с помощью статистик m^* и d^* , описанных выше. В рамках симуляционного исследования изучалось поведение m^* и d^* . Для каждого объёма выборки (n) симулировалось 10^5 репликаций.

Прежде чем произвести поиск критических значений на основе данных симуляций, были проверены различия в средних значениях статистик m^* и d^* в разрезе разных объёмов выборок, а также в зависимости от степени числа. Изучение различий в средних в разрезе разных n проходило с помощью дисперсионного анализа (ANOVA). Модель в матричном виде выглядит следующим образом:

$$Y = X\beta + e$$

где Y – вектор зависимой переменной размера $n \times 1$, X – вектор независимой качественной переменной размера $n \times d$, β – вектор неизвестных параметров размера $d \times 1$, e – вектор остатков (ошибок) размера $n \times 1$.

Дисперсионный анализ основан на вычислении межгрупповой и внутригрупповой дисперсии и расчёте статистики Фишера (F -статистики):

$$D_{\text{меж.}} = \frac{\sum_{i=1}^k n_i \cdot (\bar{x}_i - \bar{x})^2}{k - 1}$$

$$D_{\text{внутр.}} = \frac{\sum_{i=1}^k D_i}{k}$$

$$F = \frac{D_{\text{меж.}}}{D_{\text{внутр.}}}$$

где n_i – объём выборки i -ой группы, \bar{x}_i – среднее i -ой группы, \bar{x} – общее среднее, k – количество групп, D_i – дисперсия i -ой группы.

Зависимость статистик от степени числа изучалась с помощью корреляции Спирмена (ρ):

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

где d_i^2 – разница рангов i -го наблюдения, n – количество наблюдений. Выбор именно этого коэффициента корреляции обоснован тем, что он учитывает не только линейные, но монотонно убывающие или монотонно возрастающие связи или зависимости (в отличие от коэффициента Пирсона).

Затем на основе распределения статистик m^* и d^* произведено отсечение по 90-му, 95-му и 99-му перцентилям, поскольку m^* и d^* являются односторонними критериями, что соответствует критическим значениям для $\alpha = 0,10$, $\alpha = 0,05$ и $\alpha = 0,01$. Полученные критические значения сравнивались с асимптотическими критическими значениями, представленными в работе Джона Морроу [21].

Симуляционное исследование дополнено экспериментальным (эмпирическим). В рамках этого исследования два респондента прошли когнитивный тест (тест Струпа), который заключается в том, что испытуемые нажимают нужные клавиши на клавиатуре компьютера в зависимости от цвета краски появившегося слова по следующему правилу: «←» – красный, «↑» – жёлтый, «↓» – зелёный, «→» – синий. Причём появившееся слово обычно означает цвет, тем самым возникает семантическая интерференция. В ходе выполнения теста фиксируется время реакции и правильность нажатия клавиши. Некоторые стимулы являлись нейтральными (например, вместо слова показывался набор символов «XXXXX»). На рисунке 1 представлен пример стимула.

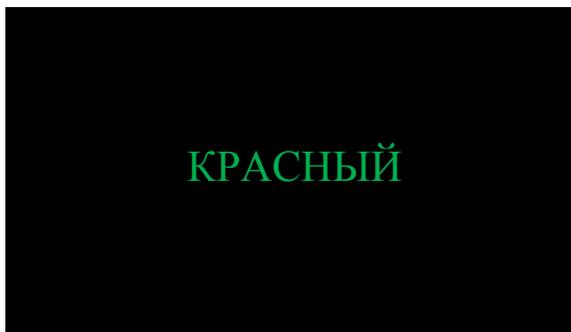


Рис. 1. Пример стимула в тесте Струпа

Первый респондент (контрольный случай) выполнял тест, согласно стандартной инструкции: «Ваша задача – как можно быстрее и точнее реагировать на окраску слова ведущей рукой. Для ответа необходимо использовать стрелки на клавиатуре: ВЛЕВО – для красного, ВВЕРХ – для желтого, ВНИЗ – для зелёного и ВПРАВО – для синего». Второй респондент (экспериментальный случай) был инструктирован выполнять тест в хаотичном, бессистемном режиме, игнорируя инструкцию к тесту.

У каждого респондента было сделано 85 замеров времени реакции, с тем чтобы можно было делать подвыборки измерений следующих размеров (n): 10, 20, 30, 40, 50, 60 и 70 с 105 репликациями для каждого объёма выборки. Данные случайным образом выбирались из всего массива измерений без возвращения. Предварительно данные о времени реакции были стандартизированы и возведены в 10-ю степень, то есть $\gamma = 10$.

Для проверки функциональности установленных критических значений статистик m^* и d^* проведён статистический анализ методом расчёта показателя отношения шансов. Исследовалось, влияет ли поведение респондента при прохождении теста Струпа на вероятность отвергнуть нулевую гипотезу о соответствии распределения первой значащей цифры времени реакции закону Бенфорда при $\alpha = 0,05$, используя установленные критические значения на основе данных симуляций. Такой уровень значимости объясняется тем, что он является конвенциональным в психологической науке [Dixon, 2003]. Сначала рассчитывался простой показатель шанса (*odds*, O) отвергнуть нулевую гипотезу для контрольного и экспериментального случая:



$$O = \frac{p_1}{p_0},$$

где p_1 – вероятность отвергнуть нулевую гипотезу, p_0 – обратное событие ($1 - p_1$).

Затем шанс экспериментального условия делился на шанс контрольного, тем самым получалось отношение шансов (*odds ratio*, *OR*):

$$OR = \frac{O_э}{O_к}$$

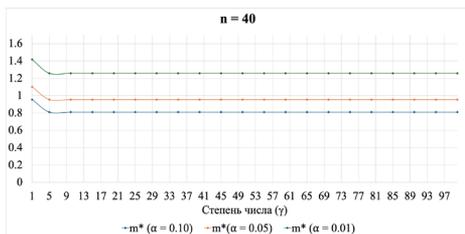
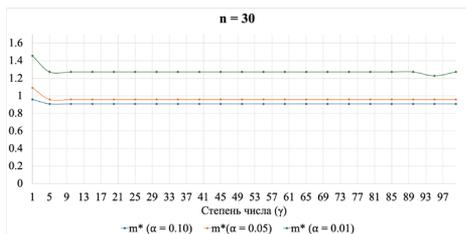
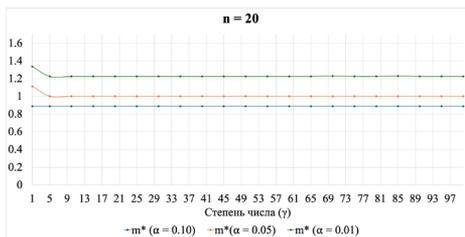
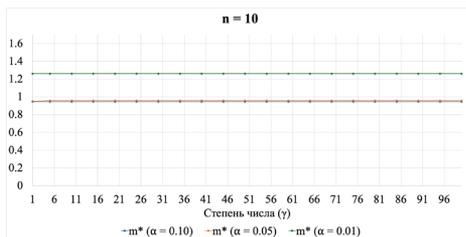
Для интерпретации отношения шансов использованы данные исследования Чена и коллег [14]. Согласно их исследованию, малый эффект наблюдается при отношении шансов, находящемся в полуинтервале [1,52; 2,74), средний – в полуинтервале [2,74; 4,72), большой – в полуинтервале [4,72; +∞).

Симуляция данных выполнена в среде R [24] с использованием базовых функций «*lnorm*» и «*sample*». Статистическая обработка данных проходила в программе *jamovi* [27]. Визуализация выполнена с помощью *Microsoft® Excel* [20].

2. РЕЗУЛЬТАТЫ

Симуляционное исследование

На рисунке 2 представлено поведение статистики m^* в зависимости от объёма выборки и степени числа. Можно заметить, что критические значения отличаются в зависимости от объёма выборки, это подтверждает дисперсионный анализ: $F(6, 140) = 45,709, p < 0,001$. При этом также можно наблюдать, что значения статистики стабилизируются при $\gamma \geq 5$, что даёт возможность предположить, что её значения не зависят от степени числа.



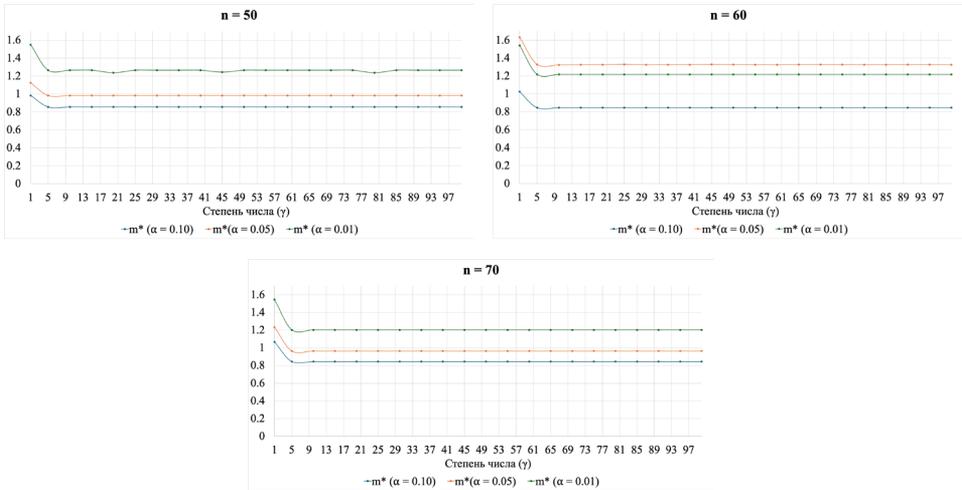


Рис. 2. Поведение статистики m^* в зависимости от объёма выборки и степени числа в симуляционном исследовании

В таблице 1 представлены корреляции Спирмена для разных объёмов выборки между степенью числа и значением статистики m^* с учётом исключения выбросов. Как правило, выбросы наблюдаются при $\gamma = 1$. Отметим, что корреляции не являются значимыми, следовательно, можно заключить, что связи между степенью числа и значением статистики m^* нет. Причём связь отсутствует вне зависимости от объёма выборки.

Таблица 1

Корреляции Спирмена для разных объёмов выборки между степенью числа и значением статистики m^* с учётом исключения выбросов

Объём выборки (n)	$m^*(\alpha = 0,10)$	$m^*(\alpha = 0,05)$	$m^*(\alpha = 0,01)$
10	0,00	0,00	0,00
20	0,00	0,00	0,29
30	0,00	0,00	-0,34
40	0,00	0,00	0,00
50	0,00	0,00	0,06
60	0,00	0,00	0,00
70	0,00	0,00	0,00

На рисунке 3 представлено поведение статистики d^* в зависимости от объёма выборки и степени числа. Как и в случае со статистикой m^* , критические значения d^* отличаются в зависимости от объёма выборки, это подтверждает дисперсионный анализ: $F(6, 140) = 34,755, p < 0,001$. При этом также можно наблюдать, что значения статистики стабилизируются при $\gamma \geq 5$, что даёт возможность предположить, что её значения не зависят от степени числа.

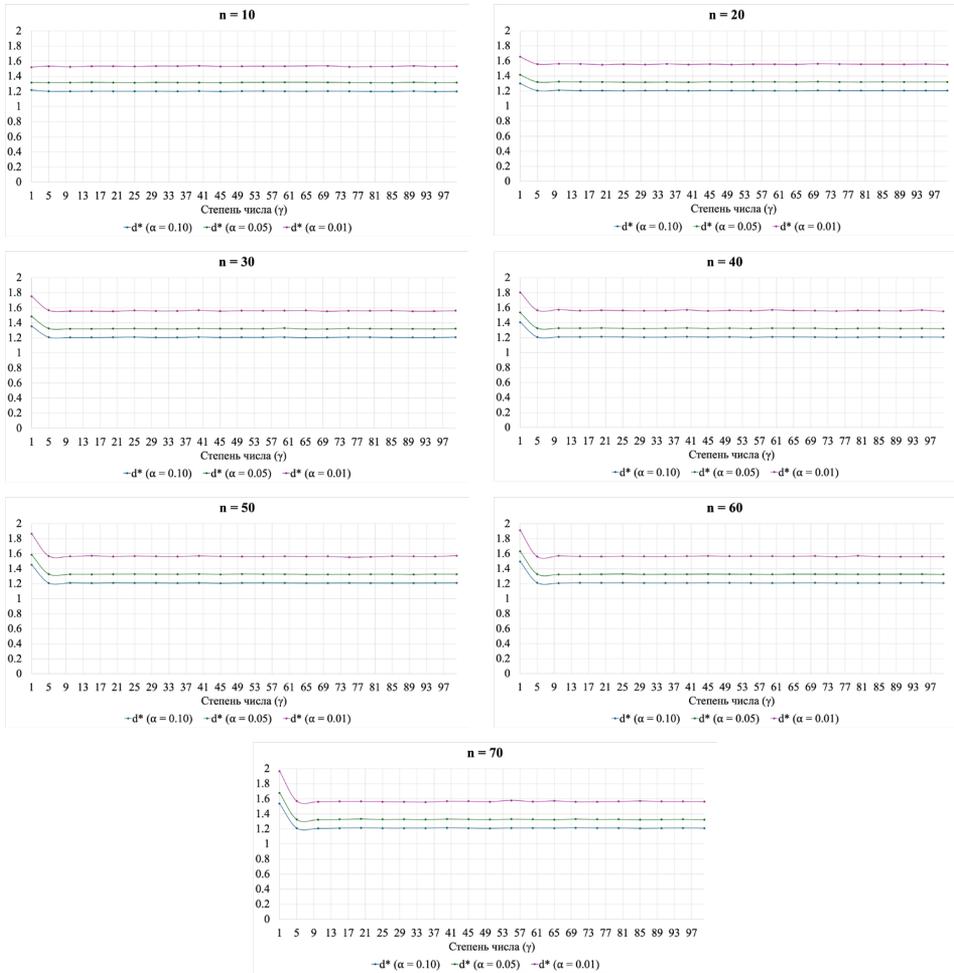


Рис. 3. Поведение статистики d^* в зависимости от объёма выборки и степени числа в симуляционном исследовании

В таблице 2 представлены корреляции Спирмена для разных объёмов выборки между степенью числа и значением статистики d^* с учётом исключения выбросов. Как правило, выбросы наблюдаются при $\gamma = 1$. Отметим, что корреляции не являются значимыми, однако наблюдается большая зависимость от степени числа, чем у статистики m^* .

Таблица 2

Корреляции Спирмена для разных объёмов выборки между степенью числа и значением статистики d^* с учётом исключения выбросов

Объём выборки (n)	$d^*(\alpha = 0,10)$	$d^*(\alpha = 0,05)$	$d^*(\alpha = 0,01)$
10	-0,17	0,03	-0,08
20	-0,20	0,01	-0,23
30	-0,01	-0,18	-0,06
40	-0,37	-0,55*	-0,38
50	-0,11	-0,33	-0,24
60	-0,40	0,01	-0,29
70	0,05	-0,15	0,04
Прим. * $p < 0,05$			

В таблице 3 представлены критические значения для статистики m^* , полученные в рамках симуляционного исследования. Отдельно указаны асимптотические значения статистики, полученные в исследовании Джона Морроу [21]. Можно заметить, что симуляционные и асимптотические критические значения становятся близки при увеличении объёма выборки.

Таблица 3

Критические значения для m^* при разных уровнях значимости

Объём выборки	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
10	0,946	0,958	1,262
20	0,890	1,002	1,225
30	0,907	0,958	1,273
40	0,809	0,955	1,259
50	0,856	0,983	1,266
60	0,847	0,968	1,219
70	0,845	0,965	1,204
Дж. Морроу (n > 80)	0,851	0,967	1,212

В таблице 4 представлены критические значения для статистики d^* , полученные в рамках симуляционного исследования. Отдельно указаны асимптотические значения статистики, полученные в исследовании Джона Морроу [21]. В отличие от m^* , симуляционные критические значения d^* достаточно близки к асимптотическим при любом значении α .

Таблица 4

Критические значения для d^* при разных уровнях значимости

Объём выборки	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
10	1,204	1,321	1,538
20	1,208	1,326	1,563
30	1,211	1,331	1,569
40	1,213	1,330	1,574

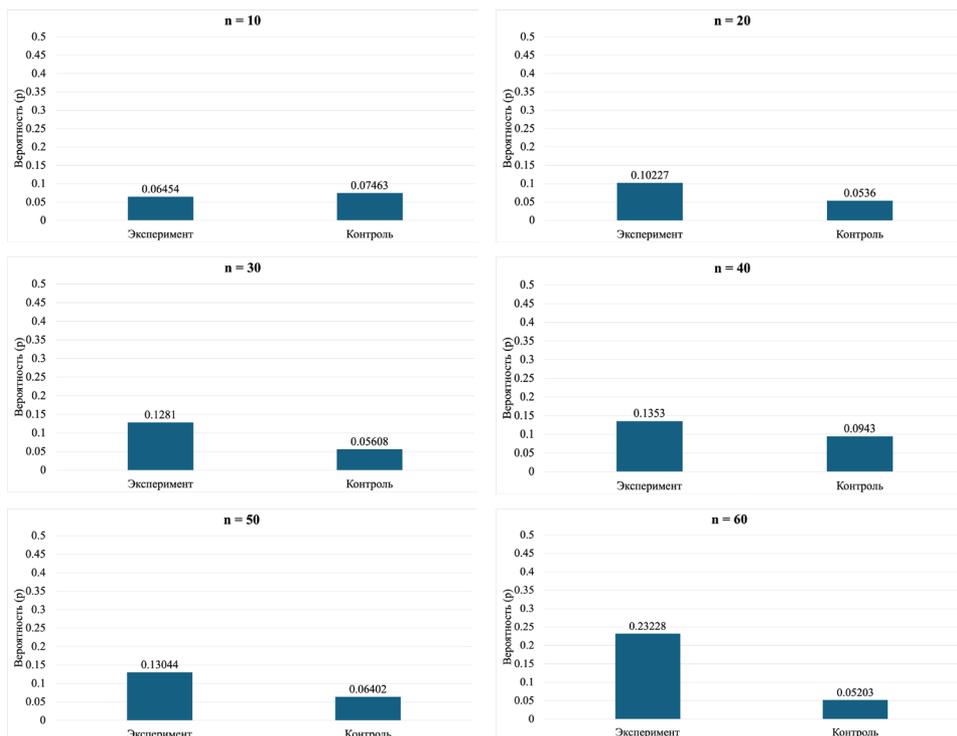


Объём выборки	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
50	1,211	1,329	1,572
60	1,212	1,330	1,571
70	1,214	1,333	1,578
Дж. Морроу (n > 80)	1,212	1,330	1,569

По итогам симуляционного исследования можно заключить, что, по нашим данным, статистика m^* в сравнении с d^* более предпочтительна при изучении соответствия распределения первой значащей цифры стандартизированного значения времени реакции закону Бенфорда, поскольку она не зависит от степени, в которую возводят стандартизированное значение (в пределах $5 \leq \gamma \leq 100$).

Экспериментальное исследование

На рисунке 4 представлены вероятности отвержения нулевой гипотезы о соответствии распределения первой значащей цифры закону Бенфорда для экспериментального и контрольного случаев в разрезе разных объёмов выборки. Можно заметить, что распределение вероятности для контрольного случая флуктуирует в районе значения 0,06, при этом для экспериментального увеличивается при увеличении объёма выборки. Также на основе представленных вероятностей можно рассчитать отношение шансов при сравнении экспериментального и контрольного случаев.



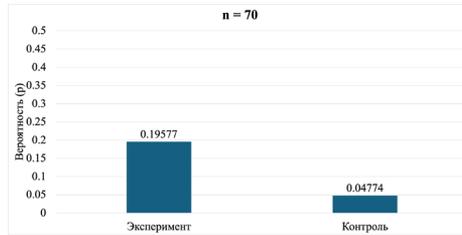


Рис. 4. Вероятность отвержения нулевой гипотезы о соответствии распределения первой значащей цифры закону Бенфорда для экспериментального и контрольного случаев на основе статистики t^*

При $n = 10$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,06454}{1-0,06454}=0,068992795$, в контрольном – $\frac{0,07463}{1-0,07463}=0,080648822$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,068992795}{0,080648822}=0,85547183 \approx 14\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу примерно на 14% ниже, чем в контрольном. Или, если перевернуть отношение шансов, то получится, что в контрольном случае на 17% выше шанс отвергнуть нулевую гипотезу, чем в экспериментальном: $\frac{0,080648822}{0,068992795}=1,168945563 \approx 17\%$. При $n = 20$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,10227}{1-0,10227}=0,113920667$, в контрольном – $\frac{0,0536}{1-0,0536}=0,056635672$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,113920667}{0,056635672}=2,011464904 \approx 101\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу примерно в 2 раза (или на 101%) выше, чем в контрольном. При $n = 30$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,1281}{1-0,1281}=0,146920518$, в контрольном – $\frac{0,05608}{1-0,05608}=0,059411815$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,146920518}{0,059411815}=2,472917542 \approx 147\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу почти в 2.5 раза (или на 147%) выше, чем в контрольном. При $n = 40$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,1353}{1-0,1353}=0,156470452$, в контрольном – $\frac{0,0943}{1-0,0943}=0,104118361$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,156470452}{0,104118361}=1,50281324 \approx 50\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу в 1.5 раза (или на 50%) выше, чем в контрольном. При $n = 50$ шанс отвергнуть нулевую



гипотезу в экспериментальном случае равен $\frac{0,13044}{1-0,13044}=0,1500069$, в контрольном – $\frac{0,06402}{1-0,06402}=0,068398897$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,1500069}{0,068398897}=2,193118686 \approx 119\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу почти в 2.2 раза (или на 119%) выше, чем в контрольном. При $n = 60$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,23228}{1-0,23228}=0,302558224$, в контрольном – $\frac{0,05203}{1-0,05203}=0,054885703$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,302558224}{0,054885703}=5,512514318 \approx 451\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу в 5.5 раза (или на 451%) выше, чем в контрольном. При $n = 70$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,19577}{1-0,19577}=0,243425388$, в контрольном – $\frac{0,04774}{1-0,04774}=0,050133367$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,243425388}{0,050133367}=4,855556352 \approx 386\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу почти в 4.9 раза (или на 386%) выше, чем в контрольном.

На основе границ размеров эффекта можно сделать вывод, что при объёме выборки от 10 до 50 различия между экспериментальным и контрольным случаем довольно малы. В то же время при $n = 60$ и $n = 70$ различия являются большими.

На рисунке 5 представлены вероятности отвержения нулевой гипотезы о соответствии распределения первой значащей цифры закону Бенфорда на основе статистики d^* для экспериментального и контрольного случаев в разрезе разных объёмов выборки. Можно заметить, что распределение вероятности обоих случаев нестабильно, изменяется в зависимости от объёма выборки – увеличивается при увеличении n .

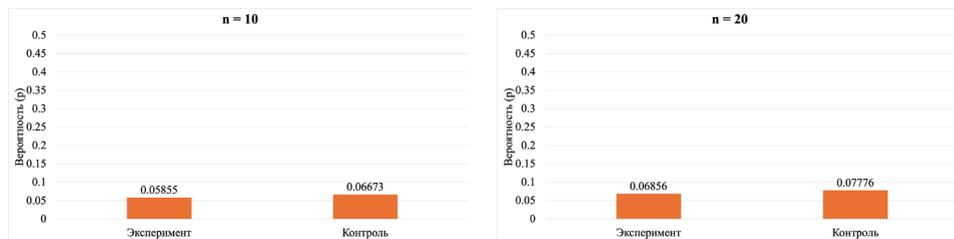


Рис. 5. Вероятность отвержения нулевой гипотезы о соответствии распределения первой значащей цифры закону Бенфорда для экспериментального и контрольного случаев на основе статистики d^*

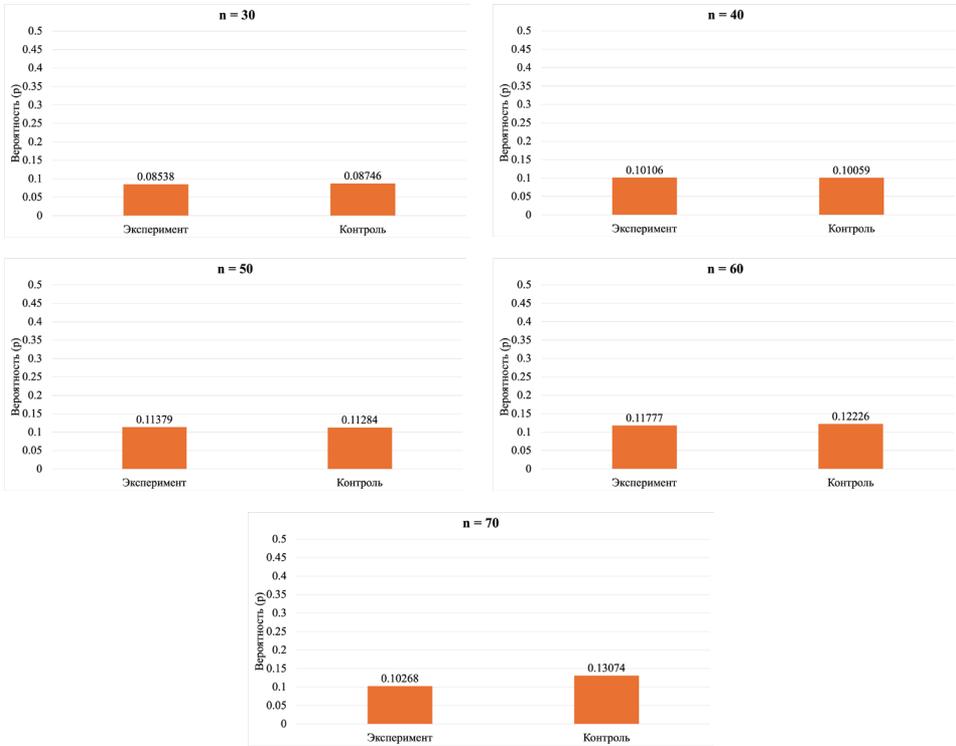


Рис. 5. Вероятность отвержения нулевой гипотезы о соответствии распределения первой значащей цифры закону Бенфорда для экспериментального и контрольного случаев на основе статистики d^*

При $n = 10$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,05855}{1 - 0,05855} = 0,062191301$, в контрольном – $\frac{0,06673}{1 - 0,06673} = 0,07150128$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,062191301}{0,07150128} = 0,869792824 \approx 13\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу примерно на 13% ниже, чем в контрольном. При $n = 20$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,06856}{1 - 0,06856} = 0,073606459$, в контрольном – $\frac{0,07776}{1 - 0,07776} = 0,084316447$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,073606459}{0,084316447} = 0,87297866 \approx 13\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу примерно на 13% ниже, чем в контрольном. При $n = 30$ шанс отвергнуть нулевую гипотезу в экспериментальном



случае равен $\frac{0,08538}{1-0,08538}=0,093350244$, в контрольном – $\frac{0,08746}{1-0,08746}=0,095842374$.

При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,093350244}{0,095842374}=0,973997616 \approx 3\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу на 3% ниже, чем в контрольном.

При $n = 40$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,10106}{1-0,10106}=0,112421296$,

в контрольном – $\frac{0,10059}{1-0,10059}=0,111839984$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,112421296}{0,111839984}=1,005197714 \approx 0,5\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу в 1,005 раза (или на 0,5%) выше, чем в контрольном.

При $n = 50$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,11379}{1-0,11379}=0,128400718$,

в контрольном – $\frac{0,11284}{1-0,11284}=0,127192389$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,128400718}{0,127192389}=1,009500006 \approx 1\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу почти в 1,01 раза (или на 1%) выше, чем в контрольном.

При $n = 60$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,11777}{1-0,11777}=0,133491266$,

в контрольном – $\frac{0,12226}{1-0,12226}=0,139289539$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,133491266}{0,139289539}=0,958372519 \approx 4\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу на 4% ниже, чем в контрольном.

При $n = 70$ шанс отвергнуть нулевую гипотезу в экспериментальном случае равен $\frac{0,10268}{1-0,10268}=0,114429635$,

в контрольном – $\frac{0,13074}{1-0,13074}=0,150403792$. При этом отношение шансов (эксперимент vs. контроль) составляет $\frac{0,114429635}{0,150403792}=0,760816158 \approx 24\%$. То есть в экспериментальном случае шанс отвергнуть нулевую гипотезу почти на 24% ниже, чем в контрольном.

В случае с использованием статистики d^* получается парадоксальная ситуация – шансы отвергнуть нулевую гипотезу оказываются немного выше в контрольном случае, нежели в экспериментальном. Правда шансы достаточно малы по размеру, это означает, что они практически не отличаются в обоих случаях.

3. ЗАКЛЮЧЕНИЕ

В предлагаемом исследовании на основе симуляций из лог-нормального распределения, отражающего распределение времени реакции, с параметрами $\mu = 5.984$, $\sigma = 0.1245$ изучалось поведение статистик согласия m^* и d^* . Выяснилось, что значения m^* не зависят от степени, в которую возводят стандартизированное значение

времени реакции, в то время как у d^* наблюдается небольшая отрицательная зависимость. Обе статистики зависят от объёма выборки (n), поэтому для каждого из анализируемых n были установлены отдельные критические значения при $\alpha = 0,10$, $\alpha = 0,05$ и $\alpha = 0,01$.

Необходимо отметить, что результаты наших симуляций относительно статистики d^* совпадают с заключениями Л. Кампанелли [13], который вывел линейные закономерности между n и d^* . Так, для $26 \leq n \leq 1000$ зависимость выглядит следующим образом:

$$d_{N,1-\alpha}^* = a + bn^{-1}$$

К примеру, при $\alpha = 0,05$ для $50 \leq n \leq 1000$ коэффициенты его формулы принимают вид: $a = 1,3286, b = -0,1151$. Если подставить $n=50$, то получится 1,326. В нашем случае критическое значение равняется 1,329, что довольно близко к результату по формуле Кампанелли (естественно, в рамках погрешности метода). По такой же логике можно проверить остальные критические значения. При этом важно учесть, что при $26 \leq n \leq 49$ коэффициенты a и b другие, а для $5 \leq n \leq 25$ значения представлены в табличном виде. Однако несмотря на сходство полученных результатов, статистика d^* оказалась нефункциональна при сравнении данных экспериментального исследования. Она оказалась нечувствительна к изменениям в поведении испытуемых, что отразилось в практически одинаковых шансах отвергнуть нулевую гипотезу о соответствии распределения первой значащей цифры закону Бенфорда.

Если говорить о статистике m^* , то как уже было отмечено, она оказалась менее чувствительной к степени, в которую мы возводили стандартизированное значение времени реакции, чем d^* . Кроме того, она оказалась более функциональной при сравнении двух разных паттернов поведения в экспериментальном исследовании. Выяснилось, что в эмпирическом исследовании при $20 \leq n \leq 70$ в экспериментальном случае шансы отвергнуть нулевую гипотезу о соответствии распределения первой значащей цифры стандартизированного значения времени реакции закону Бенфорда выше, чем в контрольном. Причём при $60 \leq n \leq 70$ эти шансы в 4–5 раз выше. Это даёт нам основание предположить, что статистика m^* позволяет лучше дифференцировать достоверные и недостоверные данные при измерении времени реакции. Тем не менее, требуются другие исследования (например, с симуляциями из отличных от лог-нормального распределений), чтобы увереннее говорить о большей полезности статистики m^* в сравнении с d^* при анализе данных о времени реакции.

Ограничения исследования

Предлагаемое исследование имеет ряд ограничений. Во-первых, в части симуляций был сделан упор на лог-нормальное распределение с конкретными параметрами. В связи с этим полученные закономерности неправомерно распространять на другие виды распределений или даже на лог-нормальное с иными значениями μ и σ . Требуются дальнейшие исследования по изучению свойств m^* и d^* . Во-вторых, функциональность статистик m^* и d^* изучалась в рамках эксперимента всего с двумя



испытуемыми, что ограничивает возможности обобщения выводов на более широкие случаи. Более того, экспериментальные условия и методика измерений были настроены под специфику дизайна и идеи исследования, что может влиять на воспроизводимость результатов в иных контекстах. Следовательно, крайне важно продолжить исследовательскую работу, направленную на проверку полученных результатов в более разнообразных условиях и с использованием различных методологических подходов, что позволит укрепить достоверность выводов и их применимость в более широком спектре ситуаций.

Литература

1. *Зайцев А.В., Лунандин В.И., Сурнина О.Е.* Время реакции в теоретических и прикладных исследованиях // Психологический вестник Уральского государственного университета. Вып. 3. 2002. Электронный ресурс: <https://elar.urfu.ru/bitstream/10995/3964/3/pv-01-03.pdf> (дата обращения: 03.03.2024).
2. *Зенков А.В.* Отклонения от закона Бенфорда и распознавание авторских особенностей в текстах // Компьютерные исследования и моделирование. 2015. Том 7. № 1. С. 197–201. DOI:10.20537/2076-7633-2015-7-1-197-201
3. *Карпенко Л.А., Петровский А.В., Ярошевский М.Г.* Краткий психологический словарь. Ростов-на-Дону, «ФЕНИКС», 1998. 505 с.
4. *Кувакина Л.В., Долгополова А.Ф.* Закон Бенфорда: сущность и применение // Современные наукоемкие технологии. 2013. № 6. С. 74–76.
5. *Куликова А.А., Прохоров Ю.В.* Односторонние устойчивые распределения и закон Бенфорда // Теория вероятностей и ее применения. 2004. Том 49. № 1. С. 178–184. DOI:10.4213/tvp244
6. *Лонэ М.* Теорема зонтика, или искусство правильно смотреть на мир через призму математики. Бомбора, 2022. 352 с.
7. *Осин Е.Н.* Проблема социальной желательности в исследованиях личностного потенциала. 2011. Электронный ресурс: <https://www.hse.ru/data/2012/08/28/1242770673/Осин%202011.pdf?ysclid=lt46qcsj468944400> (дата обращения: 27.02.2024)
8. *Попина О.Ю., Савельева М.Ю., Бородина Ю.Б.* Оценка качества отчетов о движении денежных средств российских организаций с использованием закона Бенфорда // Научные исследования: от теории к практике. 2016. № 4–2. С. 187–190.
9. Психология высших когнитивных процессов / Под ред. Т. Н. Ушаковой, Н. И. Чуприковой. М.: Ин-т психологии РАН, 2004. 303 с.
10. *Старунова О.А., Руднев С.Г., Иванова А.Е., Семенова В.Г., Стародубов В.И.* Применение закона Бенфорда для оценки качества данных профилактического скрининга // Математическая биология и биоинформатика. 2022. Том 17. № 2. С. 230–249. DOI:10.17537/2022.17.230
11. *Antipkina I., Ludlow L.H.* Parental involvement as a holistic concept using Rasch/Guttman scenario scales // Journal of Psychoeducational Assessment. 2020. Vol. 38. № 7. P. 846–865. DOI:10.1177/0734282920903164
12. *Benford F.* The law of anomalous numbers // Proceedings of the American philosophical society. 1938. P. 551–572.
13. *Campanelli L.* On the Euclidean distance statistic of Benford's law // Communications in Statistics-Theory and Methods. 2024. Vol. 53. № 2. P. 451–474. DOI:10.1080/03610926.2022.2082480

14. *Chen H., Cohen P., Chen S.* How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies // *Communications in Statistics – simulation and Computation*®. 2010. Vol. 39. № 4. P. 860–864. DOI:10.1080/03610911003650383
15. *Dixon P.* The p-value fallacy and how to avoid it // *Canadian Journal of Experimental Psychology/ Revue canadienne de psychologie expérimentale*. 2003. – Vol. 57. № 3. P. 189–202. DOI:10.1037/h0087425
16. *Dutilh G. et al.* The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models // *Psychonomic bulletin & review*. 2019. Vol. 26. P. 1051–1069. DOI:10.3758/s13423-017-1417-2
17. *Lachaud C.M., Renaud O.* A tutorial for analyzing human reaction times: How to filter data, manage missing values, and choose a statistical model // *Applied Psycholinguistics*. 2011. Vol. 32. № 2. P. 389–416. DOI:10.1017/S0142716410000457
18. *Marmolejo-Ramos F. et al.* On the efficacy of procedures to normalize Ex-Gaussian distributions // *Frontiers in psychology*. 2015. Vol. 5. DOI:10.3389/fpsyg.2014.01548
19. *Marszalek J. M., Barber C., Kohlhart J., Cooper B. H.* Sample size in psychological research over the past 30 years // *Perceptual and motor skills*. 2011. Vol. 112. № 2. P. 331–348. DOI:10.2466/03.11.PMS.112.2.331-348
20. *Microsoft Corporation.* Microsoft Excel. 2018. URL: <https://office.microsoft.com/excel> (Accessed: 02.03.2024)
21. *Morrow J.* Benford's Law, Families of Distributions and a Test Basis // *CEP Discussion Papers dp1291*, Centre for Economic Performance, LSE. 2014. URL: <https://cep.lse.ac.uk/pubs/download/dp1291.pdf> (Accessed: 12.03.2024)
22. *Peng K., Nisbett R.E., Wong N.Y.C.* Validity problems comparing values across cultures and possible solutions // *Psychological methods*. 1997. Vol. 2. № 4. P. 329–344. DOI:10.1037/1082-989X.2.4.329
23. *Primi R. et al.* Anchoring Vignettes: Can They Make Adolescent Self-Reports of Social-Emotional Skills More Reliable, Discriminant, and Criterion-Valid? // *European Journal of Psychological Assessment*. 2016. Vol. 32. № 1. P. 39–51. DOI:10.1027/1015-5759/a000336
24. *R Core Team.* R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021. URL: <https://www.R-project.org/> (Accessed: 02.03.2024)
25. *Stoet G.* PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments // *Teaching of Psychology*. 2017. Vol. 44. № 1. P. 24–31. DOI:10.1177/0098628316677643
26. *Stroop J.R.* Studies of interference in serial verbal reactions // *Journal of experimental psychology*. 1935. Vol. 18. № 6. P. 643–662.
27. *The jamovi project.* jamovi (Version 2.3) [Computer Software]. 2023. URL: <https://www.jamovi.org> (Accessed: 02.03.2024)
28. *Whelan R.* Effective analysis of reaction time data // *The psychological record*. 2008. Vol. 58. P. 475–482. DOI:10.1007/BF03395630
29. *Zhang Q., Kong L., Jiang Y.* The interaction of arousal and valence in affective priming: behavioral and electrophysiological evidence // *Brain research*. 2012. Vol. 1474. P. 60–72. DOI:10.1016/j.brainres.2012.07.023



Application of the m^* and d^* Statistics for Assessing the Quality of Data in Psychological Research Using Benford's Law (Illustrated with Reaction Time Measurements)

Nikita I. Kolachev*

HSE University, Moscow, Russia

ORCID: <https://orcid.org/0000-0002-3214-6675>

e-mail: nkolachev@hse.ru

The objective of the proposed study was to examine the properties of statistics used to assess the conformity of the distribution of the first significant digit to Benford's Law, m^* and d^* , with relatively modest sample sizes ($10 \leq n \leq 70$). A simulation study was conducted to achieve this goal. Data were simulated following a log-normal distribution with parameters that mimic the distribution of reaction time measurements. The distribution of the first significant digit was examined in standardized values raised to the power of γ ; $5 \leq \gamma \leq 100$. It was found that the statistic m^* does not depend on the power of the number, unlike d^* . Critical values were established for samples ranging from 10 to 70 observations with an increment of $h = 10$. It turned out that for small n , the critical values of the statistic d^* are close to asymptotic, while the critical values of the statistic m^* are significantly larger. The functionality of the established critical values was tested within the framework of an experimental study: one respondent performed the Stroop cognitive test in accordance with the instructions (control case), while another violated them (experimental case). It was discovered that the statistic d^* does not allow for differentiation in the behavior of subjects. Conversely, m^* proved sensitive to changes in respondent behavior, and in the experimental case, it significantly more often allowed for the rejection of the null hypothesis regarding the conformity of the distribution of the first significant digit of the standardized value of reaction time to Benford's Law compared to the control. Thus, a preliminary conclusion is made that the statistic m^* is more functional compared to d^* in studying the quality of data on reaction time with small n .

Keywords: Benford's Law, reaction time, m^* statistics, d^* statistics.

For citation:

Kolachev N.I. Using m^* and d^* Statistics to Evaluate the Quality of Data in Psychological Research via Benford's Law (Illustrated by Reaction Time Measurements). *Modelirovanie i analiz dannykh = Modelling and Data Analysis*, 2024. Vol. 14, no. 2, pp. 23–44. DOI: <https://doi.org/10.17759/mda.2024140202> (In Russ., abstr. in Engl.).

***Nikita I. Kolachev**, PhD (Psychology), Senior Lecturer, Department of Psychology, HSE University, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-3214-6675>, e-mail: nkolachev@hse.ru

References

1. Zajcev A.V., Lupandin V.I., Surnina O.E. Vremya reakcii v teoreticheskikh i prikladnykh issledovaniyakh. *Psihologicheskij vestnik Ural'skogo gosudarstvennogo universiteta*. Vyp. 3, 2002. Available at: <https://elar.urfu.ru/bitstream/10995/3964/3/pv-01-03.pdf> (Accessed: 03.03.2024). (In Russ.).
2. Zenkov A.V. Otkloneniya ot zakona Benforda i raspoznavanie avtorskih osobennostej v tekstah // *Kompyuternye issledovaniya i modelirovanie=Computer Research and Modeling*, 2015. Vol. 7, no. 1, pp. 197–201. DOI:10.20537/2076-7633-2015-7-1-197-201 (In Russ.).
3. Karpenko L.A., Petrovskij A.V., YArOshevskij M.G. Kratkij psihologicheskij slovar'. Rostov-na-Donu, «FENIKS», 1998. 505 p. (In Russ.).
4. Kuvakina L.V., Dolgopolova A.F. Zakon Benforda: sushchnost' i primenenie. *Sovremennye naukoemkie tekhnologii*, 2013. no. 6, pp. 74–76. (In Russ.).
5. Kulikova A. A., Prohorov YU. V. Odnostoronnie ustojchivye raspredeleniya i zakon Benforda. *Teoriya veroyatnostej i ee primeneniya*, 2004. Vol. 49, no. 1, pp. 178–184. DOI:10.4213/tpv244 (In Russ.).
6. Lone M. Teorema zontika, ili iskusstvo pravil'no smotret' na mir cherez prizmu matematiki. *Bombora*, 2022. 352 p. (In Russ.).
7. Osin E.N. Problema social'noj zhelatel'nosti v issledovaniyakh lichnostnogo potentsiala. 2011. Available at: <https://www.hse.ru/data/2012/08/28/1242770673/Osin%202011.pdf?ysclid=lt46qcsj1j468944400> (Accessed: 27.02.2024) (In Russ.).
8. Popina O.YU., Savel'eva M.YU., Borodina YU.B. Ocenka kachestva otchetov o dvizhenii den-zhnyh sredstv rossijskikh organizacij s ispol'zovaniem zakona Benforda. *Nauchnye issledovaniya: ot teorii k praktike*, 2016. no. 4–2, pp. 187–190. (In Russ.).
9. Ushakova T.N., CHuprikova N.I. (Eds.) Psihologiya vysshih kognitivnyh processov. Moscow: In-t psihologii RAN, 2004. 303 p. (In Russ.).
10. Starunova O.A., Rudnev S.G., Ivanova A.E., Semenova V.G., Starodubov V.I. Primenenie zakona Benforda dlya ocenki kachestva dannyh profilakticheskogo skrininga. *Matematicheskaya biologiya i bioinformatika=Mathematical biology and bioinformatics*, 2022. Vol 17, no. 2. pp. 230–249. DOI:10.17537/2022.17.230 (In Russ.).
11. Antipkina I., Ludlow L. H. Parental involvement as a holistic concept using Rasch/Guttman scenario scales. *Journal of Psychoeducational Assessment*, 2020. Vol. 38, no. 7, pp. 846–865. DOI:10.1177/0734282920903164
12. Benford F. The law of anomalous numbers. *Proceedings of the American philosophical society*, 1938. pp. 551–572.
13. Campanelli L. On the Euclidean distance statistic of Benford's law. *Communications in Statistics-Theory and Methods*, 2024. Vol. 53, no. 2. pp. 451–474. DOI:10.1080/03610926.2022.2082480
14. Chen H., Cohen P., Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics – simulation and Computation®*, 2010. Vol. 39, no. 4, pp. 860–864. DOI:10.1080/03610911003650383
15. Dixon P. The p-value fallacy and how to avoid it // *Canadian Journal of Experimental Psychology. Revue canadienne de psychologie expérimentale*, 2003. Vol. 57, no. 3, pp. 189–202. DOI:10.1037/h0087425
16. Dutilh G. et al. The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models. *Psychonomic Bulletin & Review*, 2019. Vol. 26, pp. 1051–1069. DOI:10.3758/s13423-017-1417-2
17. Lachaud C. M., Renaud O. A tutorial for analyzing human reaction times: How to filter data, manage missing values, and choose a statistical model. *Applied Psycholinguistics*, 2011. Vol. 32, no. 2, pp. 389–416. DOI:10.1017/S0142716410000457



18. Marmolejo-Ramos F. et al. On the efficacy of procedures to normalize Ex-Gaussian distributions. *Frontiers in Psychology*, 2015. Vol. 5. DOI:10.3389/fpsyg.2014.01548
19. Marszalek J. M., Barber C., Kohlhart J., & Cooper B. H. Sample size in psychological research over the past 30 years. *Perceptual and Motor Skills*, 2011. Vol. 112, no. 2, pp. 331–348. DOI:10.2466/03.11.PMS.112.2.331-348
20. Microsoft Corporation. Microsoft Excel. 2018. Available at: <https://office.microsoft.com/excel> (Accessed: 02.03.2024)
21. Morrow J. Benford's Law, Families of Distributions and a Test Basis. *CEP Discussion Papers dp1291*, Centre for Economic Performance, LSE, 2014. Available at: <https://cep.lse.ac.uk/pubs/download/dp1291.pdf> (Accessed: 12.03.2024)
22. Peng K., Nisbett R. E., Wong N. Y. C. Validity problems comparing values across cultures and possible solutions. *Psychological Methods*, 1997. Vol. 2, no. 4, pp. 329–344. DOI:10.1037/1082-989X.2.4.329
23. Primi R. et al. Anchoring Vignettes: Can They Make Adolescent Self-Reports of Social-Emotional Skills More Reliable, Discriminant, and Criterion-Valid? *European Journal of Psychological Assessment*, 2016. Vol. 32, no. 1, pp. 39–51. DOI:10.1027/1015-5759/a000336
24. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021. Available at: <https://www.R-project.org/> (Accessed: 02.03.2024)
25. Stoet G. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 2017. Vol. 44, no. 1, pp. 24–31. DOI:10.1177/0098628316677643
26. Stroop J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 1935. Vol. 18, no. 6, pp. 643–662.
27. The jamovi project. jamovi (Version 2.3) [Computer Software]. 2023. Available at: <https://www.jamovi.org> (Accessed: 02.03.2024)
28. Whelan R. Effective analysis of reaction time data. *The Psychological Record*, 2008. Vol. 58, pp. 475–482. DOI:10.1007/BF03395630
29. Zhang Q., Kong L., Jiang Y. The interaction of arousal and valence in affective priming: behavioral and electrophysiological evidence. *Brain Research*, 2012. Vol. 1474, pp. 60–72. DOI:10.1016/j.brainres.2012.07.023

Получена 13.03.2024

Принята в печать 28.03.2024

Received 13.03.2024

Accepted 28.03.2024