

УДК 004:331:61

**БЫСТРАЯ ОЦЕНКА ОБЪЕМА ВЫБОРКИ В МЕДИЦИНСКИХ
ИССЛЕДОВАНИЯХ С ПОМОЩЬЮ АНДРОИД ПЛАНШЕТА****Арутюнов С.Д., Перцов С.С., Муслов С.А.***ГБОУ ВПО «Московский государственный медико-стоматологический университет (МГМСУ)
им. А.И. Евдокимова» Минздрава РФ, Москва, e-mail: muslov@mail.ru*

Представлена мини-программа для расчета объема выборки при медицинских исследованиях на основе операционной системы Андроид. Описаны статистические термины, которые используются при описании процесса планирования размера выборки: уровень значимости, p-value, мощность и другие. Расчетная часть программы основана на формуле Лера. В качестве переменных формула содержит стандартное отклонение данных и клинически важный размер эффекта – величина, которая устанавливается врачом-исследователем перед началом исследований. Для данных, стандартное отклонение (дисперсия) которых неизвестно программа рассчитывает этот параметр отдельно на основании серии пробных измерений. Интерфейс программы крайне простой и не требует для работы с ней специальной подготовки. Программа предназначена для студентов, аспирантов и ординаторов медицинских вузов, врачей, исследователей и слушателей курсов переподготовки медицинских работников.

Ключевые слова: объем выборки, статистика, клиническая медицина, программные средства**QUICK EVALUATION OF THE SAMPLE SIZE IN MEDICAL
RESEARCH WITH ANDROID TABLET****Arutyunov S.D., Pertsov S.S., Muslov S.A.***Moscow State Medical Stomatological University (MSMSU) n.a. A.I. Evdokimov,
Moscow, e-mail: muslov@mail.ru*

A mini-program for calculating the sample size for medical research based on the Android operating system is presented. We describe statistical terms that are used in describing the process of planning the sample size: p-value, significance level, power, and others. The calculated part of the program is based on Lehr's formula. As variables, the formula contains the standard deviation of the data and the clinically important size of the effect – the value that is set by the investigator before the studies begin. For data whose standard deviation (variance) is unknown, the program calculates this parameter separately based on a series of trial measurements. The interface of the program is extremely simple and does not require special training to work with it. The program is intended for students, graduate students and residents of medical universities, doctors and researchers and students of retraining courses for medical workers.

Keywords: sample size, statistics, clinic medicine, software

Объем выборки требуется определить перед началом большинства исследований. При этом тема объема выборки в медицинских исследованиях достаточно широко обсуждается в литературе [1, 3, 6–8]. Предложены различные подходы для подсчета размеров выборки в зависимости от типа и параметров исследований, а также ряд компьютерных программ для Windows и Mac (например, GPower и др.) и статистических пакетов (SPSS и т.д.). В данном сообщении представлена программа для быстрой оценки объема выборки с помощью планшетного компьютера или смартфона на базе операционной системы Андроид. Ранее мы сообщали об использовании этих популярных вычислительных средств при решении ряда других статистических задач в медицинских исследованиях [2, 4].

Целью многих статистических изысканий и тестов является сравнение двух методов лечения, лекарственных средств или групп больных. Численное значение, которое

выражает интересующие исследователя различия, называется эффектом. Размер эффекта может характеризовать отношение шансов, относительный или абсолютный риск. Кроме того, весьма популярным в этой области доказательной медицины становится параметр NNT – Number need to treatment (число больных, нуждающихся в лечении) [5]. Затем, как правило, выдвигается нулевая, а также альтернативная гипотезы. По определению нулевая гипотеза постулирует, что эффекта нет (различия между результатами лечения в группах равны нулю, относительный риск равен единице), альтернативная гипотеза предполагает, что эффект есть.

Клинически важный размер эффекта – это наименьшие различия между средними значениями какой-либо изучаемой величины в группах, которые еще можно рассматривать как диагностические или клинически значимые. Наименьший клинически значимый эффект – минимальные изменения в некоем показателе, которые нельзя игнори-

ровать. Выбор его в основном лежит на исследователе, на его компетентности в сфере решаемой проблемы. Например, изучая реакцию организма пациента на физическую нагрузку, нужно определить, будет ли клинически значимым изменение частоты сердечных сокращений на 5 уд/мин или же на 10 уд/мин, или же какое-то иное значение.

При двухстороннем тесте нулевая гипотеза заключается в отсутствии различий, а альтернативная гипотеза предполагает, что различия между группами могут идти в любом направлении. То есть при одностороннем тесте альтернативная гипотеза определяет предполагаемое направление различий.

Интерпретация результатов теста требует знания и понимания величины p -value (p -значения) – величины, применяемой при статистической проверке гипотез. Уровень значимости или p -value – это наименьшая величина уровня значимости, при которой нулевая гипотеза отвергается для данного значения статистики критерия. Другими словами, уровень значимости – это такое значение вероятности события, при котором событие уже можно считать наступившим не по случайным причинам (а в результате проведенного лечения, например).

Указание точного значения p -value является даже более информативным, чем традиционное оформление результатов проверки нулевой гипотезы в виде неравенства $p < \alpha$. Напомним, что эта форма записи означает, что гипотеза была отклонена на уровне значимости α . Например, очень часто используемая запись $p < 0,05$ равносильна утверждению, что уровень значимости результатов наблюдений менее чем 5%. То есть ложноположительный результат или ошибка I рода при таком уровне значимости возможны не более чем в 5 случаях из 100 (например, при таком уровне значимости мы допускаем по результатам диагностических проб ненужное лечение в действительности 5 здоровых пациентов из 100).

Уровень значимости – это пороговое значение для p -оценки, ниже которого нулевая гипотеза должна быть отвергнута и сделано заключение о том, что имеются доказательства эффекта. Обычно уровень значимости в биологии и медицине устанавливается часто по умолчанию на значении 5%.

Мощность – это вероятность того, что нулевая гипотеза будет адекватно отвергнута, иными словами тогда, когда действительно существуют доказательства реальных различий или взаимосвязей. Ее можно рассматривать как «100 процентов минус вероятность пропуска истинного эффекта» [1]. Поэтому чем выше мощность, тем меньше вероятность пропуска истинного эффекта.

Мощность обычно фиксируется на стандартном уровне 80, 90 или 95%. Согласно большинству авторов, мощность не должна быть меньше 80%. Если крайне важно, чтобы исследование не пропустило существующего эффекта, надо стремиться достичь мощности минимум 90% или даже более.

В медицине мощность (чувствительность) критерия – вероятность не допустить так называемую ошибку II рода. Показатель для характеристики диагностических проб (вероятность положительного результата у больного при наличии болезни). Характеризует способность пробы выявить болезнь.

Итак, в медицинских исследованиях должна быть сформирована выборка такого размера, чтобы если подобные различия действительно существуют, то в исследовании были бы получены статистически значимые результаты.

Обсуждение расчетной формулы

Какие параметры должны учитываться при расчете объема выборки? Для отдельных дискретных величин мощности и уровней значимости необходимые размеры выборки протабулированы, а соответствующие таблицы можно найти в литературе по планированию количественных исследований [8]. Рассчитать размер выборки, необходимый для обеспечения заданной мощности t -критерия, можно также по специальным формулам, среди которых своей простотой выделяется формула Лера [7]. При сравнении двух групп для мощности 80% и двустороннего уровня значимости 0,05 требуемый размер выборки в каждой

группе равен $n = 16 \left(\frac{\sigma}{d} \right)^2$, где σ – стандарт-

ное отклонение наблюдений по выборке, одинаковое в каждой из 2-х групп, d – оценка величины (размера) эффекта (как отмечали выше, наименьшая разность в групповых средних значениях, которая клинически значима). Кстати, научное исследование Р. Лера (1992), в котором опубликована одноименная формула, так и называется: «Sixteen S-squared over D-squared: a relation for crude sample size estimates».

Стандартное отклонение (корень из дисперсии) – это мера того, насколько сильно разбросаны опытные данные относительно их среднего значения. Анализ величины стандартного отклонения призван исключить из результатов статистического анализа ошибку «среднего значения по больнице». Из формулы видно, что чем больше разброс данных и чем меньше размер эффекта, тем больше необходимый объем выборки. Это, конечно, ожидаемый результат.

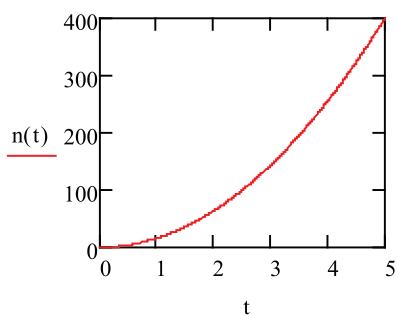


Рис. 1. Объем выборки n в зависимости от отношения $t = \frac{\sigma}{d}$

Формула Лера может быть использована для построения графика (рис. 1), однако точность такого представления невысока и ценность его использования на практике в ряде случаев весьма сомнительна.

Нами данная формула была положена в основу программы, которая легко производит расчет объема выборки с известной и неизвестной дисперсией данных. Программа выполнена на базе операционной системы Android и может быть реализована с помощью планшетного компьютера или смартфона, работающих под управлением данной системы.



Рис. 2. Окно программы SampleSize, в котором пользователь определяет, известна дисперсия данных или нет

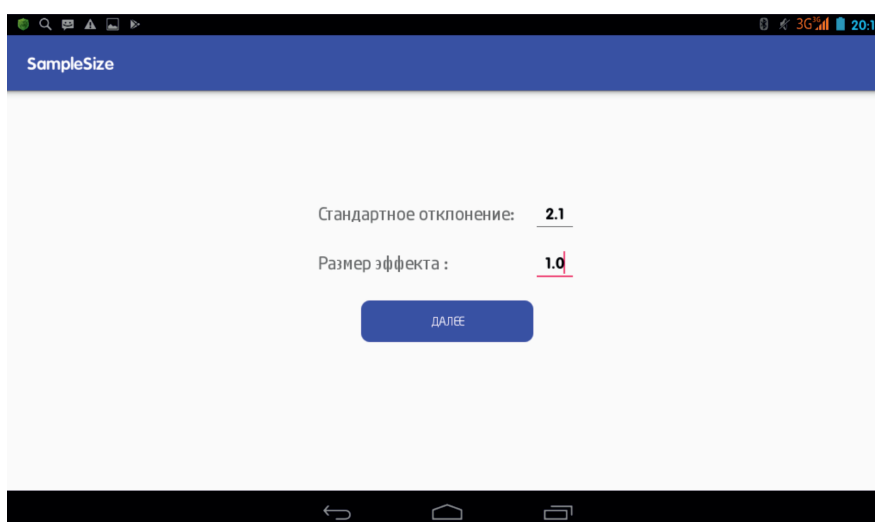


Рис. 3. Окно программы SampleSize для ввода численных значений стандартного отклонения данных и размера эффекта

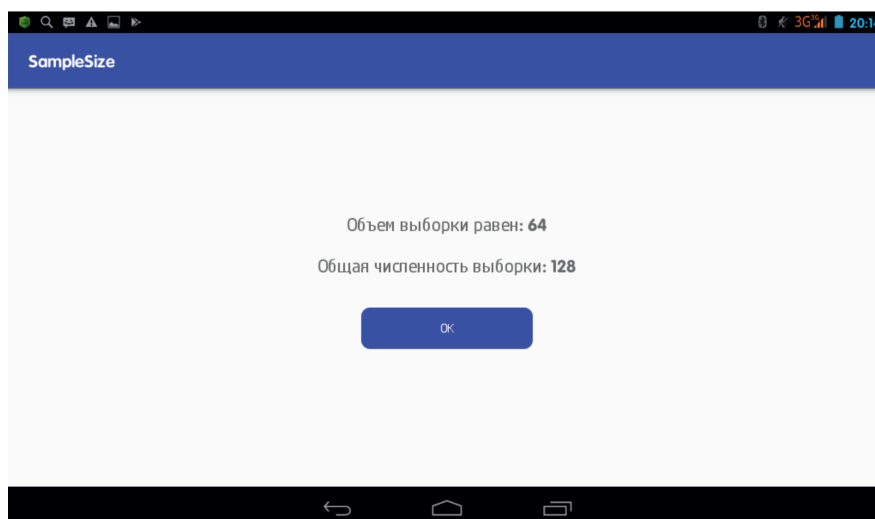


Рис. 4. Окно программы SampleSize оценки объема выборки с решением



Рис. 5. Окно программы SampleSize для ввода данных пробной выборки и размера эффекта в случае, если дисперсия данных заранее не известна

Алгоритм работы программы

Первоначально пользователь определяет, известно или нет стандартное отклонение (дисперсия) исследуемых данных (рис. 2). В 1-м случае в новом открывшемся окне предлагается заполнить поле для величины стандартного отклонения и поле, характеризующее размер эффекта (рис. 3). После этого программа вычисляет искомый объем выборки по формуле Лера (рис. 4). Во 2-м случае в новом окне программа запрашивает объем серии пробных измерений, затем предлагает для заполнения соответствующий массив полей для численных значений этой серии и поле для

размера эффекта (рис. 5). После осуществления данных манипуляций вычисляется стандартное отклонение (не отображается) и объем выборки для данных, разброс которых исследователю заранее неизвестен. Округление числа n до целых производится по формуле $n = \{n\} + 1$, где $\{n\}$ – целая часть числа n .

Таким образом, независимо от того, известно ли стандартное отклонение данных или нет, мы получаем количество больных, необходимое для включения в каждую из групп при выполнении двухвыборочного t -теста. Общая численность выборки, подлежащей исследованию, будет в два раза выше.

Список литературы

1. Артюхов И.П., Поликарпов Л.С., Хамнагадаев И.И., Деревянных Е.В. Методы эпидемиологического изучения неинфекционных заболеваний. Учебное пособие для студентов медицинских вузов по специальности 060101 – лечебное дело. – Красноярск, 2010. – 145 с.
2. Арутюнов С.Д., Перцов С.С., Муслев С.А., Шанидзе З.Л. Анализ гигиенического состояния челюстных протезов-обтураторов и программа анализа для Андроид планшета. В печати.
3. Койчубеков Б.К., Сорокина М.А., Мхитарян К.Э. Определение размера выборки при планировании научного исследования // Международный журнал прикладных и фундаментальных исследований. – 2014. – № 4. – С. 71–74.
4. Перцов С.С., Стюрева Г.М., Муслев С.А., Корнев А.А., Зайцева Н.В., Синицын А.А., Подымов В.Ф., Сидорова Л.Н. Программа анализа медицинских данных (представленных таблицами сопряженности 2x2) для ANDROID планшета / Сборник научных трудов по итогам международной научно-практической конференции (10 марта 2016г.) «Актуальные вопросы современных математических и естественных наук». Екатеринбург, 2016. – С. 12–13.
5. Ярема И.В., Муслев С.А., Еремеев В.А., Стюрева Г.М., Сухочев П.Ю. Применение методов доказательной медицины при оценке эффективности вмешательств // Медицинский алфавит. Серия Больница. – 2009. – № 2. – С. 20–23.
6. Bland. M. Расчет размера выборки [Электронный ресурс]. – Режим доступа: <http://www.pubhealth.spb.ru/COPC/STATSH/ssize.htm> (дата обращения: 01.07.17).
7. Lehr R. (1992). Sixteen S-squared over D-squared: a relation for crude sample size estimates. *Statistics in Medicine*, 11, 1099-1102.
8. Machin D. & Campbell M.J. (1995). *Statistical Tables for the Design of Clinical Trials*, 2nd ed. Blackwell Scientific Publications, Oxford.