

А.В. Копыцкий Е.П. Наумюк В.Н. Хильманович

БИОМЕДИЦИНСКАЯ СТАТИСТИКА

Допущено

*Министерством образования Республики Беларусь
в качестве учебного пособия
для студентов учреждений высшего образования
по специальностям «Лечебное дело», «Педиатрия»,
«Медико-диагностическое дело»,
«Медико-психологическое дело», «Сестринское дело»*



Минск
«Адукацыя і выхаванне»
2024

УДК [57+61]:519.22(075.8)
ББК 51.1я73
К65

Рецензенты: кафедра доказательной медицины и клинической диагностики факультета повышения квалификации и переподготовки кадров учреждения образования «Витебский государственный ордена Дружбы народов медицинский университет» (заведующий кафедрой доктор медицинских наук, профессор *И.В. Жильцов*); профессор кафедры медицинской и биологической физики учреждения образования «Гомельский государственный медицинский университет» доктор биологических наук, доцент *М.Н. Стародубцева*

Копыцкий, А. В.

К65 Биомедицинская статистика: учебное пособие / А.В. Копыцкий, Е.П. Наумюк, В.Н. Хильманович. — Минск : Адукацыя і выхаванне, 2024. — 240 с. : ил.
ISBN 978-985-599-949-3.

Рассматриваются вопросы теории вероятностей, описательной статистики, проверки статистических гипотез, применения параметрических и непараметрических статистических критериев, корреляционного и регрессионного анализа в приложении к задачам в области биологии и медицины. Для лучшего усвоения теоретической информации во все главы издания включены подробные примеры решения конкретных типовых прикладных задач медицинского и биологического характера, требующих применения теории вероятностей и математической статистики.

Предназначено для студентов учреждений высшего образования по специальностям «Лечебное дело», «Педиатрия», «Медико-диагностическое дело», «Медико-психологическое дело», «Сестринское дело».

УДК [57+61]:519.22(075.8)
ББК 51.1я73

Все права на данное издание защищены. Воспроизведение всей книги или любой ее части не может быть осуществлено без разрешения издательства.

ISBN 978-985-599-949-3

© Копыцкий А. В., Наумюк Е. П., Хильманович В. Н., 2024
© Оформление. Республиканское унитарное предприятие «Издательство “Адукацыя і выхаванне”», 2024

ВВЕДЕНИЕ В ДИСЦИПЛИНУ «Биомедицинская статистика»

1.1. ЗАДАЧИ, РЕШАЕМЫЕ С ПОМОЩЬЮ БИМЕДИЦИНСКОЙ СТАТИСТИКИ

Биомедицинская статистика – учебная дисциплина, содержащая систематизированные научные знания о применении методов математической статистики в области биологии и медицины. Биомедицинская статистика изучается в рамках концепции доказательной медицины. Основная идея доказательной медицины, начавшей формироваться в конце 80-х гг. XX в. в медицине, – применение в практике только тех методов диагностики и лечения, эффективность которых доказана результатами клинических исследований. Одним из основных инструментов доказательной медицины является математическая статистика.

Дисциплина «Биомедицинская статистика» занимает важное место в системе подготовки врачей, так как позволяет специалисту описывать группы объектов, достоверно выявлять различия между группами, классифицировать объекты и явления по их числовым характеристикам, по имеющемуся экспериментальному материалу, делать выводы об изучаемых объектах и предсказывать их поведение.

Цель учебной дисциплины «Биомедицинская статистика» – формирование универсальных компетенций для решения задач обработки медико-биологической информации, для самостоятельного принятия решений о выборе методов диагностики, лечения, профилактики заболеваний, исходя из принципов доказательной медицины.

Задачи учебной дисциплины «Биомедицинская статистика» состоят в формировании у обучающихся научных знаний о принципах доказательной медицины, о выборочном методе исследования, о способах организации такого исследования, о методах обработки статистической информации в биологии и медицине, умений и навыков, необходимых для:

- самостоятельной обработки медико-биологической информации;

- интерпретации результатов обработки медико-биологической информации как в рамках своего, так и в рамках сторонних исследований;
- понимания соответствия хода исследования принципам доказательной медицины.

1.2. ТИПЫ ДАННЫХ. КАЧЕСТВЕННЫЕ, ПОРЯДКОВЫЕ И КОЛИЧЕСТВЕННЫЕ ПОКАЗАТЕЛИ. ШКАЛЫ ИЗМЕРЕНИЯ ДАННЫХ

Измеряемые показатели называются величинами, переменными, признаками (*variables*). Набор результатов измерений от одного или нескольких объектов называется *данными* (*data*). В самом простом случае данные делятся на качественные (категориальные и ранговые) и количественные (численные). В зависимости от природы численной измеряемой величины она может быть дискретной или непрерывной.

При проведении измерений объекту или явлению приписывают по определенному закону числовое или буквенное значение. Говорят, что измерение однозначно отображает характеристику объекта или явления в область символьных обозначений. Закон такого отображения называется *шкалой измерения*. Наиболее распространенная классификация шкал (в силу простоты) – классификация Стивенса. По Стивенсу, выделяют 2 типа шкал:

1) *метрические шкалы*. В этих шкалах есть так называемые метрики – расстояния между объектами, заданные в определенных единицах. Если в такой шкале нет абсолютного 0 (т.е. полного отсутствия показателя), то говорят об *интервальной шкале*. При наличии абсолютного 0 говорят о *шкале отношений*. Приведем *примеры показателей*, заданных в разных метрических шкалах:

- *интервальная шкала*: уровень IQ испытуемого по Айзенку, температура по Цельсию, концентрация глюкозы в крови пациента, гематокрит крови испытуемого;
- *шкала отношений*: вирусная нагрузка, интенсивность окраски раствора при реакции Вассермана, температура по Кельвину;

2) *неметрические шкалы*. В них уже нет метрик – естественных расстояний между объектами. Выделяют 2 неметрических шкалы:

- *шкалу наименований* (синонимы: номинативная шкала, шкала имен, категориальная шкала, шкала факторная). В данной шкале объекты различаются только названиями, их нельзя упорядочить по степени выраженности показателя. *Примеры показателей*, заданных в этой шкале: пол (м/ж), диагноз (подтвержден / не подтвержден),

полиморфизм гена (доминантная / рецессивная гомозигота / гетерозигота), тип темперамента;

- *шкалы рангов* (синонимы: порядковая шкала, ординарная шкала). В данной шкале объекты можно упорядочить по степени выраженности показателя, однако нельзя указать дистанцию между ними. *Примеры признаков*, заданных в этой шкале: группа инвалидности, стадия онкологического заболевания.

1.3. ФОРМИРОВАНИЕ ВЫБОРОК, РАНДОМИЗАЦИЯ

Один из главных методов в математической статистике – выборочный метод, к которому прибегают, когда не могут изучить генеральную совокупность. *Генеральная совокупность (популяция)* – множество всех объектов, признаки которых должны быть измерены. В выборочном методе изучают только часть генеральной совокупности – выборку. *Выборка (выборочная совокупность)* – специальным образом отобранная из генеральной совокупности группа объектов для проведения измерений, ограниченная по численности. В общем случае результаты изучения выборки обобщают на всю генеральную совокупность. Для этого выборка должна быть репрезентативной, валидной и надежной. *Репрезентативность* (представительность) – соответствие основных параметров выборки параметрам генеральной совокупности (т. е. выборка должна полно передавать свойства генеральной совокупности). *Валидность* – соответствие выбранного способа выборочного исследования изучаемому явлению в генеральной совокупности. *Надежность* – выборка и метод исследования должны быть такими, чтобы при повторении эксперимента другой исследователь получал близкие результаты. *Рандомизация* – способ случайного отбора объектов в выборку из генеральной совокупности или распределения объектов по группам в зависимости от способа организации эксперимента (дизайна).

Способы формирования выборок делят на вероятностные (рандомизированные) и невероятностные. Для вероятностных способов есть специальные формулы для оценивания возможных ошибок выборки, для невероятностных таких формул нет.

1.4. ВЕРОЯТНОСТНЫЙ ПОДХОД

1.4.1. Простая случайная выборка

Это простейший тип организации выборки: элементы в выборку отбираются случайным образом, и формально каждый элемент генеральной совокупности (ГС) имеет одинаковую вероятность

попасть в выборку, причем эти вероятности для каждого объекта независимы. Схематически простую случайную выборку можно изобразить следующим образом (рис. 1.1).

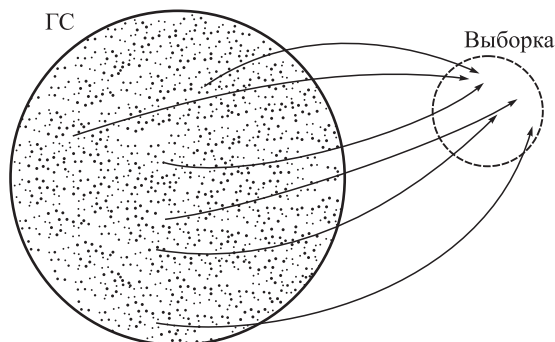


Рис. 1.1. Схема простой случайной выборки

Данный метод хорош тем, что для него можно легко рассчитать величины ошибок, он прост для реализации и объяснения. Однако в некоторых ситуациях его реализация может оказаться неэффективной (по соотношению затраты / информация).

Рассмотрим следующий пример. При оценке качества таблеток, выходящих с конвейера в течение одной смены, данный подход, при котором в произвольное время смены отбирается произвольное количество таблеток для контрольных испытаний, будет хорошим решением. Однако если есть основания полагать, что в разные смены качество продукции не будет существенно отличаться (например, при высоком уровне автоматизации), то такой подход может оказаться избыточным, требующим дополнительных затрат на контроль качества.

При обследовании пациентов, например, различных возрастных групп, данный метод может оказаться недостаточным и дать сильно смещенные оценки показателей здоровья, если распределение обследуемых по возрастам неравномерное. Пусть, условно говоря, вся генеральная совокупность состоит из 100 объектов: 90 объектов класса *A* и 10 объектов класса *B*. Очевидно, что объекты класса *A* будут попадать в выборку при случайном отборе чаще, чем объекты класса *B*. Обобщение результатов на генеральную совокупность будет происходить в сторону, соответствующую объектам класса *A*, и чем меньше объем выборки, тем сильнее будет этот перекос. Если не учесть распределения объектов по классам, то есть риск не заметить этот перекос.

Для организации простого случайного отбора могут потребоваться существенные объемы выборок, учитывающие специфику исследуемого явления. В медицине, например, нужно обследовать паци-

ентов мужского и женского пола, различных социальных слоев, регионов проживания, возрастов, сопутствующих патологий и т. д., в противном случае выборка может оказаться нерепрезентативной, или оценки, построенные на ней, окажутся сильно смещенными.

Обобщая, можно заключить, что такой подход к формированию выборки уместен тогда, когда изучаемая генеральная совокупность однородна по всем показателям, например все обследуемые пациенты одного пола, одной социальной группы, одного возрастного периода и т. д.

1.4.2. Систематическая случайная выборка (механический отбор)

Здесь выборка формируется следующим образом: предварительно отбираются случайным образом N объектов из генеральной совокупности, далее они упорядочиваются по заранее известному показателю (например, по возрасту, по времени). И наконец, из упорядоченного ряда отбираются n объектов через постоянный интервал Δn .

Схематически данную процедуру можно представить следующим образом (рис. 1.2).

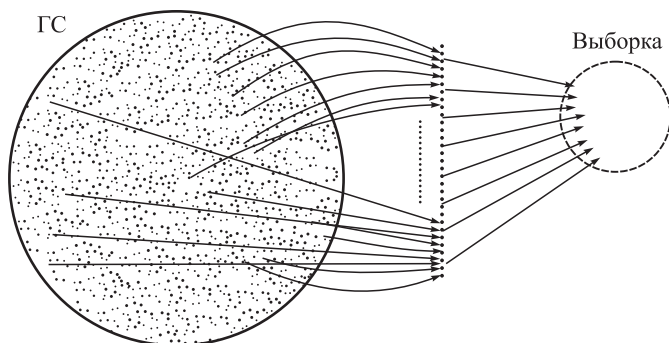


Рис. 1.2. Схема систематической случайной выборки

Если заранее известно, что выборка будет неоднородной по одному или нескольким показателям (например, по возрасту), то этот подход будет эффективным, так как неоднородность предварительной выборки будет компенсирована одинаковыми вероятностями попадания объектов из предварительной выборки в изучаемую. Также данный метод широко используется, например, для оценки качества выпускаемой продукции на предприятиях, когда через равные промежутки времени отбираются образцы для испытаний.

1.4.3. Стратифицированная выборка (типический отбор)

Если заранее известно, что выборка будет неоднородной по одному или нескольким показателям, то имеет смысл разбить выборку на страты, в каждой из которых будет соблюдаться однородность, при этом в страте должны оказаться как можно более похожие элементы. Например, выборку пациентов можно предварительно разбить на страты по возрастным группам, по полу, месту проживания, социальному статусу и т. д. так, чтобы по побочным показателям внутри страты явных различий между пациентами не наблюдалось. Схематически отбор в стратифицированную выборку можно представить следующим образом (рис. 1.3).

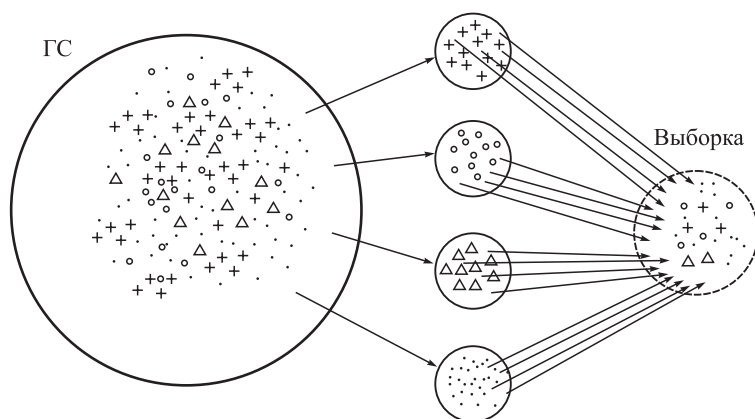


Рис. 1.3. Схема стратифицированной выборки

Последующее извлечение из страты в выборку осуществляется случайно или систематически.

1.4.4. Кластеризованная выборка (серийный отбор)

Если заранее известно, что выборка будет неоднородной по одному или нескольким показателям, то можно использовать другой, отличный от стратифицированного подхода, способ. При этом из выборки предварительно случайно выбираются целые области (регионы) – кластеры. При этом в отличие от стратификации в каждом кластере объекты не должны быть максимально похожи, наоборот, кластер должен напоминать по своему многообразию срез генеральной совокупности. Схематически кластеризация выглядит следующим образом (рис. 1.4).

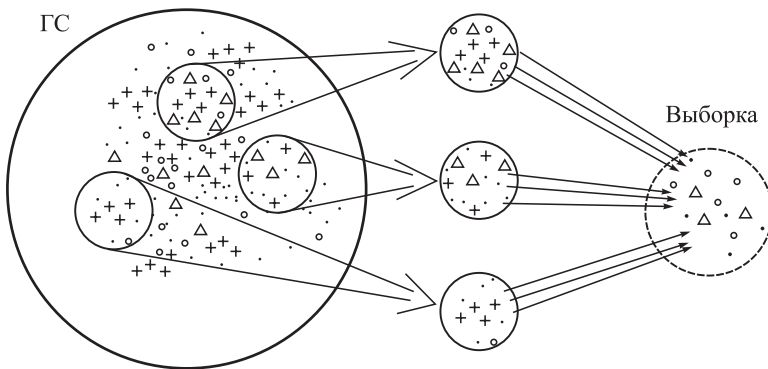


Рис. 1.4. Схема кластеризованной выборки

Кластеризация в медицинских исследованиях часто носит географический характер, когда кластерами считаются случайно выбранные учреждения здравоохранения или их подразделения. Далее из этих учреждений случайно или систематически выбираются испытуемые.

1.5. НЕВЕРОЯТНОСТНЫЙ ПОДХОД

При данном подходе отбор объектов в выборку осуществляется без использования необходимости равной вероятности включения объекта в выборку из генеральной совокупности. Отбор осуществляется произвольно, по усмотрению экспериментатора. Для данного подхода невыполнимо корректное обобщение результатов исследования на популяцию, так как оценить ошибку результатов не представляется возможным. Такие выборки чаще всего заведомо нерепрезентативные. Несмотря на это, данный подход может использоваться на начальных этапах исследований, в поисковых и пилотных исследованиях, для грубой оценки каких-либо эффектов или величин.

Можно выделить несколько распространенных схем, характерных для данного подхода:

- исследование согласных. Как следует из названия, в данном способе организации в выборку попадают только те испытуемые, которые согласились принимать участие в исследовании;
- исследование по принципу удобства. При таком подходе в исследование включаются объекты, исследование которых не требует каких-либо дополнительных условий. Например, отбор студентов-добровольцев, отбор пациентов, данные карточек которых легко получить, работа с пациентами только своего отделения и т. д.;

- метод снежного кома. При этом подходе в выборку случайно отбирается испытуемый. Далее устанавливаются связи испытуемого, через которые находятся другие испытуемые, и т.д. Данный метод может использоваться при исследованиях инфекционных заболеваний, при исследованиях различных сообществ и т.д.;

- квотная выборка. Это подвид предыдущих способов; при данном подходе исследователь следит за тем, чтобы пропорции между количествами объектов в выборке примерно соответствовали пропорциям между объектами в генеральной совокупности. При правильном квотировании данный подход может дать результаты, близкие к вероятностному стратифицированному подходу;

- отбор типичных случаев. При этом подходе в выборку включаются только типичные испытуемые, соответствующие целому подвиду объектов. Например, включение в выборку только двух женщин пенсионного возраста, страдающих от гипертонии, включение только двух мужчин пенсионного возраста, также страдающих от гипертонии, включение двух женщин пожилого возраста и т.д.

1.6. ОПРЕДЕЛЕНИЕ ОШИБОК ВЫБОРОК ДЛЯ ДОЛЕЙ, ОПРЕДЕЛЕНИЕ НЕОБХОДИМЫХ ОБЪЕМОВ ВЫБОРОК

При вероятностном подходе, как уже было сказано выше, можно оценить ошибки выборки. *Ошибкой выборки* называется разность между характеристикой изучаемого показателя в выборке и истинным значением данной характеристики в генеральной совокупности, например выборочным средним арифметическим и средним арифметическим в генеральной совокупности. Другой пример – разность между долей, вычисленной по выборке, и истинной долей в генеральной совокупности. Поскольку при выборочном методе истинные значения характеристик интересующего исследователя показателя в генеральной совокупности точно не известны, то можно их оценить по выборочным значениям. Например, можно указать так называемую предельную ошибку доли Δ_p , задающую интервал $p \pm \Delta_p$, за который с заданной вероятностью γ не выходит истинное значение доли в генеральной совокупности. Предельная ошибка выборки позволяет исследователю оценивать, в каком диапазоне в генеральной совокупности находится изучаемая характеристика показателя. Также предельную ошибку выборки можно использовать для вычисления необходимого объема выборки, при котором с вероятностью γ эта ошибка не будет превосходить некоторого заранее выбранного исследователем значения.

Предельная ошибка выборки для доли и соответствующий необходимый объем для простой и систематической выборок могут быть определены по следующим формулам.

Предельная ошибка доли:

- бесповторная выборка:

$$\Delta_p = t \sqrt{\frac{p(1-p)}{n} \left(1 - \frac{n}{N}\right)}; \quad (1.1)$$

- повторная выборка:

$$\Delta_p = t \sqrt{\frac{p(1-p)}{n}}. \quad (1.2)$$

Необходимый объем выборки:

- бесповторная выборка:

$$n = \frac{t^2 N p(1-p)}{(\Delta_p)^2 N + t^2 p(1-p)}; \quad (1.3)$$

- повторная выборка:

$$n = \frac{t^2 p(1-p)}{(\Delta_p)^2}, \quad (1.4)$$

где t – коэффициент доверия, определяемый для заданной вероятности γ по специальным таблицам; p – измеренная доля (относительная частота встречаемости) признака в выборке; n – объем выборки; N – объем генеральной совокупности.

Для определения коэффициента доверия t можно использовать следующую таблицу:

Вероятность γ , %	68,27	86,64	95,00	95,45	98,76	99,01	99,73	99,95
t	1	1,5	1,96	2	2,5	2,58	3	3,5

Задача 1.1. В городе с населением $N = 300\,000$ человек было определено по результатам изучения случайно отобранных $n = 1000$ историй болезней, что диагноз «Артериальная гипертензия» встречается с частотой $p = 15\%$. Оценить предельную ошибку доли. Коэффициент доверия принять равным $t = 2$. Указать долю встречаемости диагноза с учетом ошибки выборки.

Решение

Так как мы имеем дело с простой случайной выборкой с бесповторным отбором, то для определения ошибки доли используем формулу (1.1):

$$\Delta = t \sqrt{\frac{p(1-p)}{n} \left(1 - \frac{n}{N}\right)}.$$

Подставляя числа, получаем:

$$\begin{aligned} \Delta &= t \sqrt{\frac{p(1-p)}{n} \left(1 - \frac{n}{N}\right)} = 2 \sqrt{\frac{0,15(1-0,15)}{1000} \left(1 - \frac{1000}{300000}\right)} = \\ &= 2 \cdot 0,0113 = 0,0225 = 2,25\%. \end{aligned}$$

С учетом предельной ошибки выборки долю можно указать как

$$p = 15,00\% \pm 15,00\% \cdot 0,0225 = (15,00 \pm 0,34)\%.$$

Задача 1.2. Какое количество респондентов необходимо для того, чтобы оценить долю курящих в популяции лиц с предельной ошибкой выборки не более $\Delta_p = 1\%$, если размер популяции составляет $N = 1000000$ человек, а по данным пилотного исследования доля курящих оценивается в $p = 33\%$? Коэффициент доверия принять равным $t = 2$. Решить задачу для случаев повторных и бесповторных выборок.

Решение

Для случая повторной выборки будем использовать формулу (1.4):

$$n = \frac{t^2 p(1-p)}{(\Delta_p)^2}.$$

Подставляя числа, получаем:

$$n = \frac{t^2 p(1-p)}{(\Delta_p)^2} = \frac{2^2 \cdot 0,33(1-0,33)}{0,01^2} = 8844.$$

Для случая бесповторной выборки используем формулу (1.3):

$$n = \frac{t^2 N p(1-p)}{(\Delta_p)^2 N + t^2 p(1-p)}.$$

Подставляя числа, получаем:

$$n = \frac{t^2 N p(1-p)}{(\Delta_p)^2 N + t^2 p(1-p)} = \frac{2^2 \cdot 1000000 \cdot 0,33(1-0,33)}{0,01^2 \cdot 1000000 + 2^2 \cdot 0,33(1-0,33)} = 8766.$$

1.7. ПОНЯТИЕ О ЦЕНЗУРИРОВАННЫХ (НЕПОЛНЫХ) ДАННЫХ

Неполные данные (пропущенные данные) — такие наборы результатов измерений, при которых для некоторых объектов некоторые показатели не измерены либо исключены:

Объект	Хар. 1	Хар. 2	Хар. 3	Хар. 4	Хар. 5	Хар. 6	Хар. 7
1	43	125	0,7	2	А	0	12,8
2	39	140	0,7	2	А	0	
3	24	160		0	Б	0	10,1
4	62	133	0,6	3	А	1	16,9
...

Неполные данные могут быть результатом как ошибок измерений, нарушения протокола исследования, так и выбывания объекта из исследования (переезд в другое место, гибель и т.д.) — в последнем случае говорят о цензурированных данных.

1.8. ЭТАПЫ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ

Можно выделить следующие основные этапы исследования, связанного с анализом статистических данных:

- 1) выдвижение научных гипотез;
- 2) выбор дизайна исследования, формулирование статистических гипотез;
- 3) определение объемов выборок, формирование выборок;
- 4) проведение выборочных измерений;
- 5) сведение полученных данных;
- 6) получение описательных статистик, валидация данных;
- 7) проверка статистических гипотез;
- 8) интерпретация результатов — подтверждение или опровержение ранее выдвинутых научных гипотез;
- 9) публикация результатов (отчеты, научные статьи).

1.9. ВИДЫ СБОРА ДАННЫХ

Выделяют следующие виды (способы) сбора данных:

- 1) непосредственное наблюдение (получение данных от самого объекта с помощью устройств или наблюдателя);
- 2) документальное наблюдение (источники данных — документы: истории болезни, протоколы, рецепты, отчеты и т.д.);
- 3) опрос (получение ответов от респондентов: опросные листы, анкеты и пр.).

ОГЛАВЛЕНИЕ

Глава 1. Введение в дисциплину «Биомедицинская статистика»	3
1.1. Задачи, решаемые с помощью биомедицинской статистики	3
1.2. Типы данных. Качественные, порядковые и количественные показатели. Шкалы измерения данных	4
1.3. Формирование выборок, рандомизация	5
1.4. Вероятностный подход	5
1.4.1. Простая случайная выборка	5
1.4.2. Систематическая случайная выборка (механический отбор)	7
1.4.3. Стратифицированная выборка (типический отбор)	8
1.4.4. Кластеризованная выборка (серийный отбор).	8
1.5. Невероятностный подход	9
1.6. Определение ошибок выборок для долей, определение необходимых объемов выборок	10
1.7. Понятие о цензурированных (неполных) данных	13
1.8. Этапы статистического исследования	13
1.9. Виды сбора данных.	13
Глава 2. Основные понятия и теоремы теории вероятностей.	14
2.1. Случайные события. Виды событий.	14
2.2. Классическое и статистическое определение вероятности случайного события	15
2.2.1. Классическое определение вероятности случайного события.	15
2.2.2. Статистическое определение вероятности случайного события	15
2.3. Теоремы теории вероятностей	18
2.3.1. Теорема сложения вероятностей для несовместных событий	18
2.3.2. Теорема умножения вероятностей независимых событий.	19
2.3.3. Условная вероятность. Теорема умножения вероятностей зависимых событий.	22
2.3.4. Формула полной вероятности. Формула Байеса	23
Глава 3. Повторные независимые испытания. Формула Бернулли. Предельные случаи формулы Бернулли	26
3.1. Повторные независимые испытания. Формула Бернулли	26
3.2. Локальная теорема Муавра – Лапласа	29
3.3. Интегральная теорема Муавра – Лапласа	30
3.4. Формула Пуассона	33
Глава 4. Дискретные и непрерывные случайные величины	35
4.1. Случайные величины: дискретные и непрерывные	35
4.2. Закон распределения дискретной случайной величины	35
4.3. Функция распределения дискретной случайной величины	37

4.4. Характеристики дискретных случайных величин	39
4.4.1. Характеристики положения	39
4.4.2. Характеристики рассеяния	40

Глава 5. Некоторые теоретические законы распределений случайных величин 42

5.1. Биномиальное распределение	42
5.2. Распределение Пуассона	44
5.3. Равномерное распределение	47
5.4. Нормальное распределение	48

Глава 6. Описательные статистики количественных и качественных показателей. 55

6.1. Описательные статистики количественных и качественных показателей.	55
6.2. Выборочные описательные статистики и оценки описательных статистик в генеральной совокупности	55
6.3. Способы представления результатов измерений (вариационные ряды распределений)	57
6.4. Среднее арифметическое значение, дисперсия, среднее квадратическое отклонение, стандартное отклонение, стандартная ошибка среднего.	58
6.5. Квантили распределения: квартили, процентиля	62
6.6. Мода	67
6.7. Графическое представление экспериментальных данных.	68
6.8. Доверительный интервал для среднего значения. Диаграмма размаха для среднего значения	73
6.9. Доверительный интервал для долей (процентов)	74

Глава 7. Статистические аспекты доказательной медицины 76

7.1. Базовые понятия в областях диагностики, лечения и литературного анализа.	76
7.1.1. Базовые понятия в области диагностики	76
7.1.2. Базовые понятия в области лечения	80
7.1.3. Базовые понятия в области литературного анализа	82
7.2. Фазы клинических испытаний.	83
7.3. Понятие конечной точки в клиническом испытании	85
7.4. Таблицы сопряженности 2×2 в описательных и экспериментальных исследованиях	85

Глава 8. Проверка статистических гипотез 90

8.1. Статистические гипотезы.	90
8.2. Статистика критерия. Распределение статистики критерия.	92
8.3. Принцип проверки статистических гипотез	93
8.4. Ошибки при принятии статистических гипотез	95

Глава 9. Некоторые параметрические критерии	98
9.1. Критерий Фишера сравнения дисперсий	98
9.2. Критерий Стьюдента сравнения средних при неизвестных и равных дисперсиях	101
9.3. Критерий Уэлча	103
9.4. <i>t</i> -Критерий для повторных измерений	105
9.5. Одновыборочный критерий Стьюдента	107
Глава 10. Дисперсионный анализ: основные идеи и методы	109
10.1. Основные идеи и математические основания дисперсионного анализа	109
10.2. Однофакторный дисперсионный анализ для независимых выборок	111
10.3. Однофакторный дисперсионный анализ для повторных измерений	116
10.4. Многофакторный дисперсионный анализ для независимых выборок	121
Глава 11. Некоторые непараметрические критерии. Апостериорные сравнения	129
11.1. Сравнение групп. Алгоритм выбора непараметрического статистического критерия сравнения групп	129
11.2. Критерий Манна – Уитни	131
11.3. Критерий Краскела – Уоллиса	136
11.4. Критерий Вилкоксона	139
11.5. Критерий Фридмана	144
11.6. Апостериорные попарные сравнения. Проблема попарных сравнений. Поправки Бонферрони и Холма	146
11.6.1. Поправка Бонферрони	147
11.6.2. Поправка Холма	147
Глава 12. Критерии согласия и сравнения распределений. Критерии нормальности	149
12.1. Некоторые критерии сравнения распределений непрерывных величин: критерий Лемана – Розенблатта, критерий Смирнова	149
12.1.1. Критерий Лемана – Розенблатта	149
12.1.2. Критерий Смирнова	152
12.2. Критерии нормальности распределения: критерий Шапиро – Уилка, критерий Лиллиефорса	158
12.2.1. Критерии нормальности	158
12.2.2. Критерий Шапиро – Уилка	159
12.2.3. Критерий Лиллиефорса	161
Глава 13. Статистическая и корреляционная зависимости. Показатели корреляционной связи. Статистическая значимость корреляционной связи в генеральной совокупности	164
13.1. Функциональная и статистическая зависимости	164
13.1.1. Понятие корреляционной связи	165

13.1.2. Параметрические показатели корреляции	166
13.2. Показатели корреляции в генеральной совокупности.	168
13.3. Непараметрические коэффициенты корреляции.	171
13.3.1. Коэффициент корреляции τ -Кендалла.	172
13.3.2. Рангово-бисериальный коэффициент корреляции	175
13.3.3. Коэффициент ранговой корреляции ρ -Спирмена	177
13.3.4. Коэффициент ассоциации ϕ	179
Глава 14. Основы регрессионного анализа	182
14.1. Метод наименьших квадратов в линейной регрессии	182
14.2. Определение статистической значимости коэффициентов регрессионного уравнения для случая линейной регрессии. Проверка адекватности регрессионного уравнения	184
14.3. Нелинейные регрессионные уравнения и множественная регрессия	190
Глава 15. Анализ качественных признаков в группах.	191
15.1. Принципы описания и сравнения качественных признаков. Схема выбора статистического критерия сравнения распределений.	191
15.2. Критерий согласия Пирсона	192
15.3. Критерий однородности Пирсона.	195
15.4. Поправка Йейтса для критерия однородности Пирсона.	198
15.5. Точный критерий Фишера	200
15.6. Критерий Мак-Немара для повторных измерений. Биномиальный критерий для повторных измерений	203
Приложение	205
Список использованных источников.	236

Учебное издание

Копыцкий Андрей Витальевич
Наумюк Елена Петровна
Хильманович Валентина Николаевна
БИОМЕДИЦИНСКАЯ СТАТИСТИКА

Учебное пособие

Редактор *Л.Н. Макейчик*. Художественный редактор *Т.Ю. Таран*.
Компьютерная верстка *Н.В. Шабуни*. Корректор *Л.Н. Макейчик*.

Подписано в печать 01.11.2024. Формат 60×90/16. Бумага офсетная. Печать офсетная. Усл. печ. л. 15,0.
Уч.-изд. л. 12,0. Тираж 200 экз. Заказ 235.

Республиканское унитарное предприятие «Издательство “Адукацыя і выхаванне”».
Свидетельство о государственной регистрации издателя, изготовителя, распространителя
печатных изданий №1/19 от 02.08.2013. Ул. Будённого, 21, 220070, г. Минск.

Республиканское унитарное предприятие «Издательский центр Белорусского государственного
университета». Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 2/63 от 19.03.2014.
Ул. Красноармейская, 6, 220030, г. Минск.