

# Введение в медицинскую статистику

$$A(m) = \omega \int_0^{\infty} D(r) \varphi(M) r^2 dr.$$

- **Статистика** — наука, в которой излагаются общие вопросы сбора, измерения и анализа массовых данных.
- **Медицинская статистика** изучает *количественную сторону массовых общественных явлений* в неразрывной связи с их качественной стороной, позволяющая *методом обобщающих показателей* изучить закономерности этих явлений.

# Результаты и выводы

- 1. Результаты получаем, выводы делаем.
- 2. выводы должны опираться на результаты.
- 3. если выводы основать на результатах, просеянных через сито статистики, то они могут подтвердиться и пойти во благо науки.

# Три кита медицинской статистики

- 1. Среднее значение
- 2. Выборка
- 3. Статистические критерии

- **Среднее значение** (числовая характеристика множества чисел или функций) — некоторое число, заключённое между наименьшим и наибольшим из их значений.
- непараметрические средние — мода, медиана.
- среднее значение случайной величины — то же, что математическое ожидание случайной величины.

- **Мода** — значение чаще других встречаемое во множестве наблюдения. *Например: 6, 2, 6, 6, 8, 9, 9, 9, 10; мода — 6 и 9.* В этом случае можно сказать, что совокупность мультимодальна. Из структурных средних величин только мода обладает таким уникальным свойством. Как правило, мультимодальность указывает на то, что набор данных не подчиняется нормальному распределению.

# Расчет размера выборки

- **Генеральная совокупность**
- Суммарная численность объектов наблюдения (люди) обладающих определенным набором признаков (пол, возраст, т.д.), ограниченная в пространстве и времени.
- **Выборка (Выборочная совокупность)**
- Часть объектов из генеральной совокупности, отобранных для изучения, с тем чтобы сделать заключение обо всей генеральной совокупности. Для того чтобы заключение, полученное путем изучения выборки, можно было распространить на всю генеральную совокупность, выборка должна обладать свойством репрезентативности.(Пример с супом)



- **Репрезентативность выборки**

- Свойство выборки корректно отражать генеральную совокупность. Одна и та же выборка может быть репрезентативной и нерепрезентативной для разных генеральных совокупностей.

- **Ошибка выборки (доверительный интервал)**

- Отклонение результатов, полученных с помощью выборочного наблюдения от истинных данных генеральной совокупности.

Ошибка выборки бывает двух видов – статистическая и систематическая.

- **Статистическая ошибка** зависит от размера выборки. Чем больше размер выборки, тем она ниже.

- **Систематическая ошибка отбора** — показывает что выводы, сделанные применительно к какой-либо группе, могут оказаться неточными вследствие неправильного отбора в эту группу, не выполнены условия случайности.
- Типы систематических ошибок:
- **Пространство** «Своевременное» окончание, то есть тогда, когда результаты укладываются в желаемую теорию.
- **Данные** Вычёркивание неких «плохих» данных
- **Участники** Курение и фитнес

# Выбор метода для статистического анализа

- **Шкала измерения**

- При проведении исследования у каждой единицы наблюдения определяются значения различных признаков. В зависимости от того, по какой шкале они измеряются, все признаки делятся на **количественные** и **качественные**. Качественные показатели в исследованиях распределяются по так называемой **номинальной** шкале. Кроме того, показатели могут быть представлены по **ранговой** шкале.

- **Пример:** сравнение показателей сердечной деятельности у спортсменов и лиц, ведущих малоподвижный образ жизни.
- При этом у исследуемых определялись следующие признаки:
- **пол** - является *номинальный* показатель: мужской или женский.
- **возраст** - *количественный* показатель,
- **занятия спортом** - *номинальный* показатель: занимается или не занимается,
- **частота сердечных сокращений** - *количественный* показатель,
- **наличие жалоб на боли в грудной клетке** - является *качественным* показателем:
  - по *номинальной* (есть жалобы - нет жалоб), так и по *ранговой* шкале в зависимости от частоты (например, если боль возникает несколько раз в день - показателю присваивается ранг 3, несколько раз в месяц - ранг 2, несколько раз в год - ранг 1, при отсутствии жалоб на боли в грудной клетке - ставится ранг 0)

# Количество сопоставляемых совокупностей

- В большинстве случаев - *основная и контрольная группы*. **Основной**, или **опытной**, принято считать группу, в которой был применен изучаемый метод диагностики или лечения, или в которой пациенты страдают заболеванием, являющимся предметом данного исследования.
- **Контрольную** группу, напротив, составляют пациенты, получающие обычную медицинскую помощь, плацебо, или лица, у которых отсутствует изучаемое заболевание. Такие совокупности, представленные разными пациентами, называются **несвязанными**.

- **Связанные**, или **парные**, совокупности, когда речь идет об одних и тех же людях, но сравниваются значения какого-либо признака, *до и после* исследования. Число сравниваемых совокупностей при этом также равно 2, однако к ним применяются **другие** методики, нежели к несвязанным.

- Особенностью планирования клинических исследований - исследователь никогда не имеет в своем распоряжении всей популяции (генеральной совокупности).
- Объёма выборки должно хватить для формирования статистически значимого заключения о эффекте по результатам исследования.
- **Нулевая гипотеза** — гипотеза, которая проверяется на согласованность с имеющимися выборочными (эмпирическими) данными.
- ошибка 1 рода - возможность ошибочно отклонить нулевую гипотезу, т.е. найти различия там, где их нет.
- Ошибка 2го рода (ее вероятность обозначается бета) возникает, если мы принимаем нулевую гипотезу, когда она не верна, другими словами, не находим существующее различие.

Таким образом, прежде чем выбрать объем групп для данного исследования, необходимо:

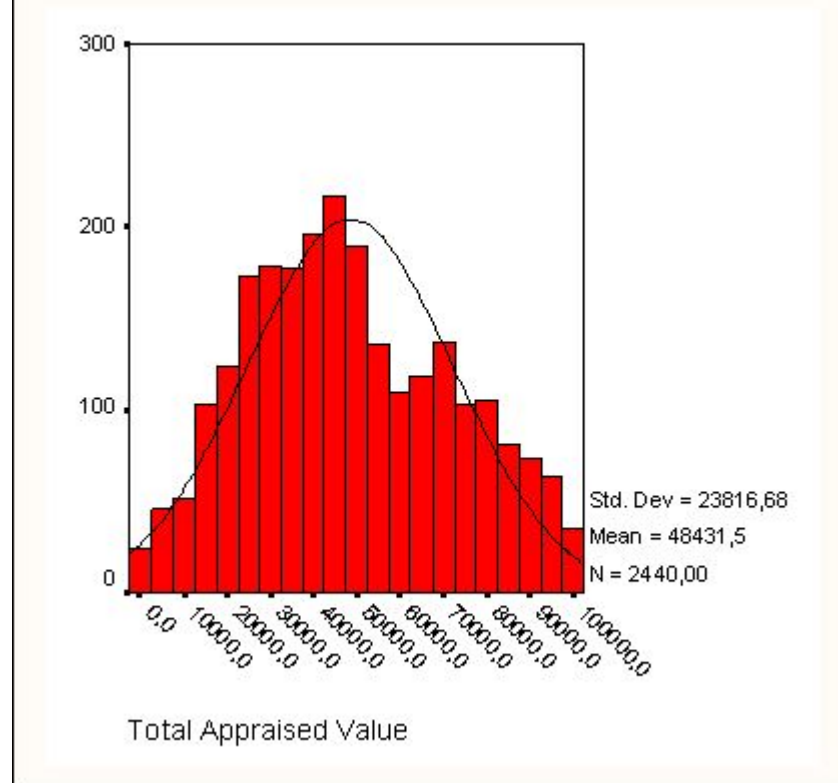
- понять, переменные какого типа (количественные, качественные, и т.д.) будут измерять эффект в данном исследовании
- оценить величину различий между эффектами, приемлемую для данного исследования с клинической точки зрения
- выбрать подходящий статистический критерий для последующего анализа интересующих событий, это определит выбор конкретных формул для расчета
- оценить по данным литературы, пилотному исследованию или результатам сходных исследований величины показателей, входящих в выбранные для расчета формулы
- увеличить рассчитанные по формулами значения объема с учетом возможного исключения в процессе исследования



- Важным также является вопрос **нормальности распределения** изучаемых совокупностей. От этого зависит, можно ли применять методы *параметрического анализа* или только *непараметрического*. Условиями, которые должны соблюдаться в нормально распределенных совокупностях, являются:
  - 1 .максимальная близость или равенство значений средней арифметической, моды и медианы;
  - 2. соблюдение правила "трёх сигм" (в интервале  $M \pm 1\sigma$  находятся не менее 68,3% вариант, в интервале  $M \pm 2\sigma$  - не менее 95,5% вариант, в интервале  $M \pm 3\sigma$  находятся не менее 99,7% вариант;
  - 3. показатели измерены в количественной шкале;
  - 4. положительные результаты проверки на нормальность распределения при помощи специальных критериев - Колмогорова-Смирнова или Шапиро-Уилка.
- <http://medstatistic.ru/theory/change.html>

- **В каком случае использовать параметрический, а в каком - непараметрический метод?**
- Непараметрические методы наиболее приемлемы, когда объем выборок мал. Если данных много (например,  $n > 100$ ), то не имеет смысла использовать непараметрические статистики.
- Дело в том, что когда выборки становятся очень большими, то выборочные средние подчиняются нормальному закону, даже если исходная переменная не является нормальной или измерена с погрешностью.

кривая нормального распределения  
(Колокол Гаусса).



- Нормальность распределения - при таком распределении большая часть значений группируется около некоторого среднего значения, по обе стороны от которого частота наблюдений равномерно снижается.

- *Непараметрические методы* - когда исследователь ничего не знает о параметрах исследуемой популяции (отсюда и название методов - непараметрические), не основываются на оценке параметров (таких как среднее или стандартное отклонение) при описании выборочного распределения интересующей величины.
- Поэтому эти методы иногда также называются свободными от параметров или свободно распределенными.
- Непараметрические методы позволяют обрабатывать данные "низкого качества" из выборок малого объема с переменными, про распределение которых мало что или вообще ничего не известно.
- По существу, для каждого параметрического критерия имеется, по крайней мере, один непараметрический аналог. Эти критерии можно отнести к одной из следующих групп:
  - 1. Критерии различия между независимыми выборками .
  - 2. Критерии различия между зависимыми выборками.
  - 3. Критерии зависимости между переменными.

- **Различия между независимыми выборками**
- **Две независимые выборки: U-критерий Манна-Уитни и др.**
- Обычно, когда имеются две выборки (например, мужчины и женщины), которые вы хотите сравнить относительно среднего значения некоторой изучаемой переменной, вы используете *t-критерий* для независимых выборок.
- Непараметрическими альтернативами параметрического критерия для двух независимых групп являются:
  - U критерий Манна-Уитни
  - Критерий серий Вальда-Вольфовица
  - Двухвыборочный критерий Колмогорова-Смирнова

- **Несколько независимых групп: критерий Краскела-Уоллиса и др.**
- Если вы имеете несколько групп, то можете использовать *Дисперсионный анализ (ANOVA)*.
- Его непараметрическими аналогами являются:
- Ранговый дисперсионный анализ Краскела-Уоллиса
- Медианный тест

- **Различия между зависимыми выборками**
- **Две зависимые выборки: критерий Вилкоксона и др.**
- Если вы хотите сравнить две переменные, относящиеся к одной и той же выборке (например, математические успехи студентов в начале и в конце семестра), то обычно используется t-критерий для зависимых выборок.
- Альтернативными непараметрическими тестами являются:
  - Критерий Вилкоксона парных сравнений
  - Критерий знаков

**Среднеквадратическое отклонение (сигма)** в теории вероятностей и статистике наиболее распространённый показатель рассеивания значений случайной величины относительно её математического ожидания. (пример с футболом)

**Правило трёх сигм** — практически все значения нормально распределённой случайной величины лежат в интервале ( $x-3$ сигмы;  $x+3$ сигмы). Более строго — приблизительно с вероятностью 0,9973 значение нормально распределённой случайной величины лежит в указанном интервале (при условии, что величина истинная, а не полученная в результате обработки выборки).



# выборочное стандартное отклонение

Если повторить извлечение выборок того же самого объема из популяции, маловероятно, что оценки параметра популяции будут точно такими же в каждой выборке. Однако все оценки должны быть близки к истинному значению параметра (генеральному параметру) в популяции и подобны друг другу.

Определяя величину вариабельности этих оценок, мы поймем, насколько они точны, и таким образом сможем оценить ошибку, обусловленную выборкой.

Обычно берут только одну выборку из популяции. Однако можно использовать знания о теоретическом распределении выборочных оценок для того, чтобы сделать выводы относительно генерального параметра популяции.

*Выборочное стандартное отклонение  $s$  оценивается по наблюдаемой реализации выборки:*

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

# Ошибка среднего

В случае если есть, как обычно, только одна выборка, нашей лучшей оценкой среднего популяции будет выборочное среднее, а так как редко бывает известно стандартное отклонение в популяции (генеральный стандарт), то *стандартную ошибку среднего* оценивают следующим образом:

$$SEM = \frac{s}{\sqrt{n}},$$

где  $s$  – стандартное отклонение в выборке.

Стандартная ошибка среднего отражает точность нашей оценки.

- Большая стандартная ошибка указывает, что оценка неточна;
- Небольшая стандартная ошибка указывает, что оценка точна;
- Стандартная ошибка уменьшится, т.е. мы получим более точную оценку, если:
  - Объем выборки увеличится;
  - Данные имеют небольшое рассеяние.

Итак, стандартная ошибка отображает точность выборочного среднего и должна быть указана, если интересуется среднее значение набора данных.

- *t-критерий Стьюдента* – общее название для класса методов статистической проверки гипотез (статистических критериев), основанных на распределении Стьюдента. Наиболее частые случаи применения *t-критерия* связаны с проверкой равенства средних значений в двух выборках.
- **Для чего используется *t-критерий Стьюдента*?**
- *t-критерий Стьюдента* используется для определения статистической значимости различий средних величин. Может применяться как в случаях сравнения независимых выборок (*например, группы больных сахарным диабетом и группы здоровых*), так и при сравнении связанных совокупностей (*например, средняя частота пульса у одних и тех же пациентов до и после приема антиаритмического препарата*)

### Как рассчитать t-критерий Стьюдента?

Для сравнения средних величин t-критерий Стьюдента рассчитывается по следующей формуле:

где  $M_1$  - средняя арифметическая первой сравниваемой совокупности (группы),  $M_2$  - средняя арифметическая второй сравниваемой совокупности (группы),  $m_1$  - средняя ошибка первой средней арифметической,  $m_2$  - средняя ошибка второй средней арифметической.

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}}$$

### Как интерпретировать значение t-критерия Стьюдента?

Полученное значение t-критерия Стьюдента необходимо правильно интерпретировать. Для этого нам необходимо знать количество исследуемых в каждой группе ( $n_1$  и  $n_2$ ). Находим число степеней свободы  $f$  по следующей формуле:

$$f = (n_1 + n_2) - 2$$

После этого определяем критическое значение t-критерия Стьюдента для требуемого уровня значимости (например,  $p=0,05$ ) и при данном числе степеней свободы  $f$  по таблице

#### **Сравниваем критическое и рассчитанное значения критерия:**

Если рассчитанное значение t-критерия Стьюдента *равно или больше* критического, найденного по таблице, делаем вывод о статистической значимости различий между сравниваемыми величинами.

Если значение рассчитанного t-критерия Стьюдента *меньше* табличного, значит различия сравниваемых величин статистически не значимы.

Степень свободы	Границы значения		Степень свободы	Границы значения	
	$p=0,05$	$p=0,01$		$p=0,05$	$p=0,05$
1	12,71	63,60	21	2,08	2,82
2	4,30	9,93	22	2,07	2,82
3	3,18	5,84	23	2,07	2,81
4	2,78	4,60	24	2,06	2,80
5	2,57	4,03	25	2,06	2,79
6	2,45	3,71	26	2,06	2,78
7	2,37	3,50	27	2,05	2,77
8	2,31	3,36	28	2,05	2,76
9	2,26	3,25	29	2,04	2,76
10	2,23	3,17	30	2,04	2,75
11	2,20	3,11	40	2,02	2,70
12	2,18	3,06	50	2,01	2,68
13	2,16	3,01	60	2,00	2,66
14	2,15	2,98	80	1,99	2,64
15	2,13	2,95	100	1,98	2,63
16	2,12	2,92	120	1,98	2,62
17	2,11	2,90	200	1,97	2,60
18	2,10	2,88	500	1,96	2,59

### Пример расчета t-критерия Стьюдента

Для изучения эффективности нового препарата железа были выбраны две группы пациентов с анемией. В первой группе пациенты в течение двух недель получали новый препарат, а во второй группе - получали плацебо. После этого было проведено измерение уровня гемоглобина в периферической крови. В первой группе средний уровень гемоглобина составил  $115,4 \pm 1,2$  г/л, а во второй -  $103,7 \pm 2,3$  г/л (данные представлены в формате  $M \pm m$ ), сравниваемые совокупности имеют нормальное распределение. При этом численность первой группы составила 34, а второй - 40 пациентов. Необходимо сделать вывод о статистической значимости полученных различий и эффективности нового препарата железа.

**Решение:** Для оценки значимости различий используем t-критерий Стьюдента, рассчитываемый как разность средних значений, поделенная на сумму квадратов ошибок:

$$t = \frac{115,4 - 103,7}{\sqrt{1,2^2 + 2,3^2}}$$

После выполнения расчетов, значение t-критерия оказалось равным 4,51. Находим число степеней свободы как  $(34 + 40) - 2 = 72$ . Сравниваем полученное значение t-критерия Стьюдента 4,51 с критическим при  $p=0,05$  значением, указанным в таблице: 1,993. Так как рассчитанное значение критерия больше критического, делаем вывод о том, что наблюдаемые различия статистически значимы (уровень значимости  $p < 0,05$ ).