

Оценка достоверности результатов статистического исследования

Достоверность результатов статистического исследования – это степень уверенности в точности и правильности выводов, сделанных на их основе. Для количественной оценки достоверности используются три ключевых взаимосвязанных критерия.



Доверительная вероятность
(P)



Уровень значимости (p)



Доверительный
коэффициент (t)

Ключевые критерии достоверности

Доверительная вероятность (P)

Характеризует надежность результатов выборочного исследования. Показывает вероятность того, что ошибка рассчитанного показателя не превысит допустимую величину.

- В медицине: обычно 95%, 99% или 99,9%.
- Для большинства задач: достаточно 95%.

Ключевые критерии достоверности

Уровень значимости (p)

Дополняет доверительную вероятность до 100% ($P + p = 100\%$). Отражает вероятность того, что выявленные различия являются случайными.

- Стандартный порог: $p \leq 0,05$.

Ключевые критерии достоверности

Доверительный коэффициент (t)

Соответствует определённому уровню доверительной вероятности и используется для расчёта доверительных границ.

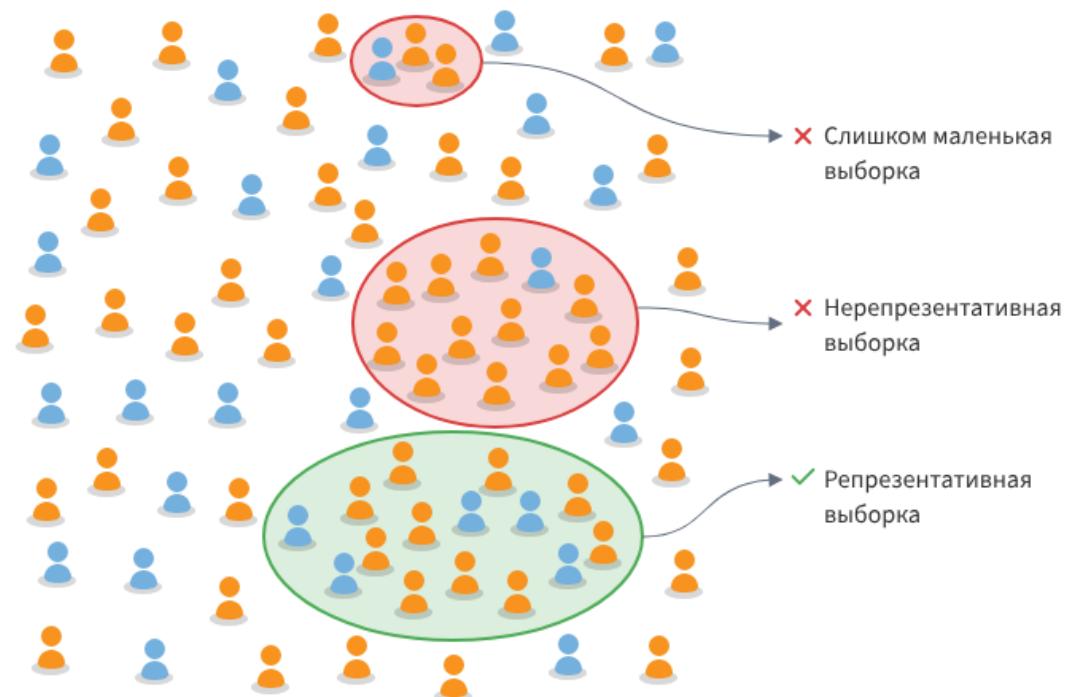
- При $P = 95\%: t = 2$.
- При $P = 99\%: t = 3$.

Репрезентативность выборки

Репрезентативность – это необходимое условие для переноса результатов выборочного исследования на генеральную совокупность. Выборка должна адекватно отражать структуру и свойства всей исследуемой совокупности.

Ошибка репрезентативности (m) характеризует степень возможного отклонения выборочных данных от генеральных параметров и служит мерой точности выборочных оценок.

Генеральная совокупность включает  - 1/3 и  - 2/3



Методика расчёта ошибки репрезентативности (m)

Расчет ошибки репрезентативности (m) критически важен для определения доверительных границ. Формулы различаются для средних и относительных величин.

Для средней арифметической (M)

При $n > 30$:

$$m = \frac{\sigma}{\sqrt{n}}$$

При $n \leq 30$:

$$m = \frac{\sigma}{\sqrt{n-1}}$$

Для относительной величины (P)

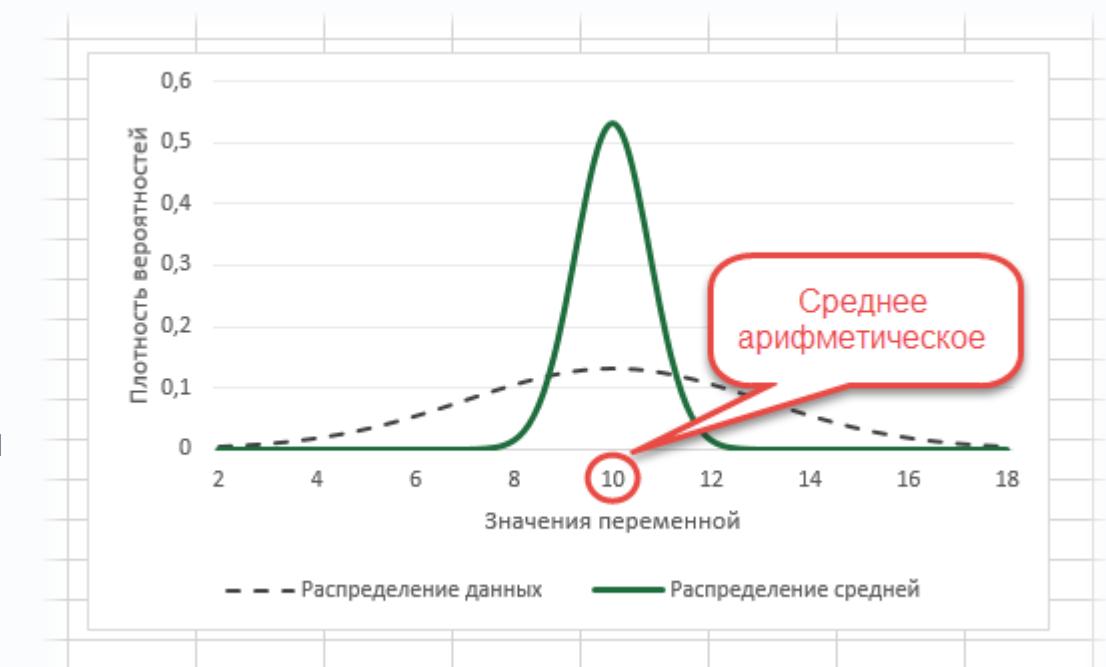
$$m = \sqrt{\frac{P}{n} q}$$

Где $q = 100\% - P$.

Назначение и свойство стандартной ошибки средней арифметической

Числитель – это стандартное отклонение выборки и здесь все понятно. Чем больше разброс данных, тем больше стандартная ошибка средней – прямо пропорциональная зависимость.

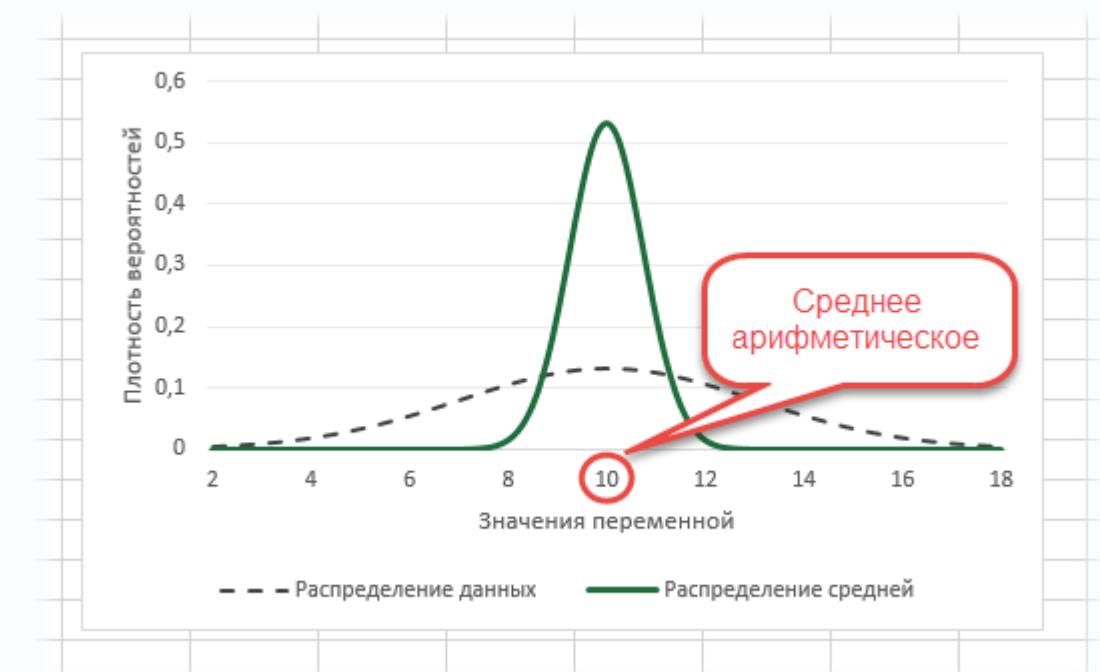
Посмотрим на знаменатель. Здесь находится квадратный корень из объема выборки. Соответственно, чем больше объем выборки, тем меньше стандартная ошибка средней. Для наглядности изобразим на одной диаграмме график нормально распределенной переменной со средней равной 10, сигмой – 3, и второй график – распределение средней арифметической этой же переменной, полученной по 16-ти наблюдениям (которое также будет нормальным)..



Назначение и свойство стандартной ошибки средней арифметической

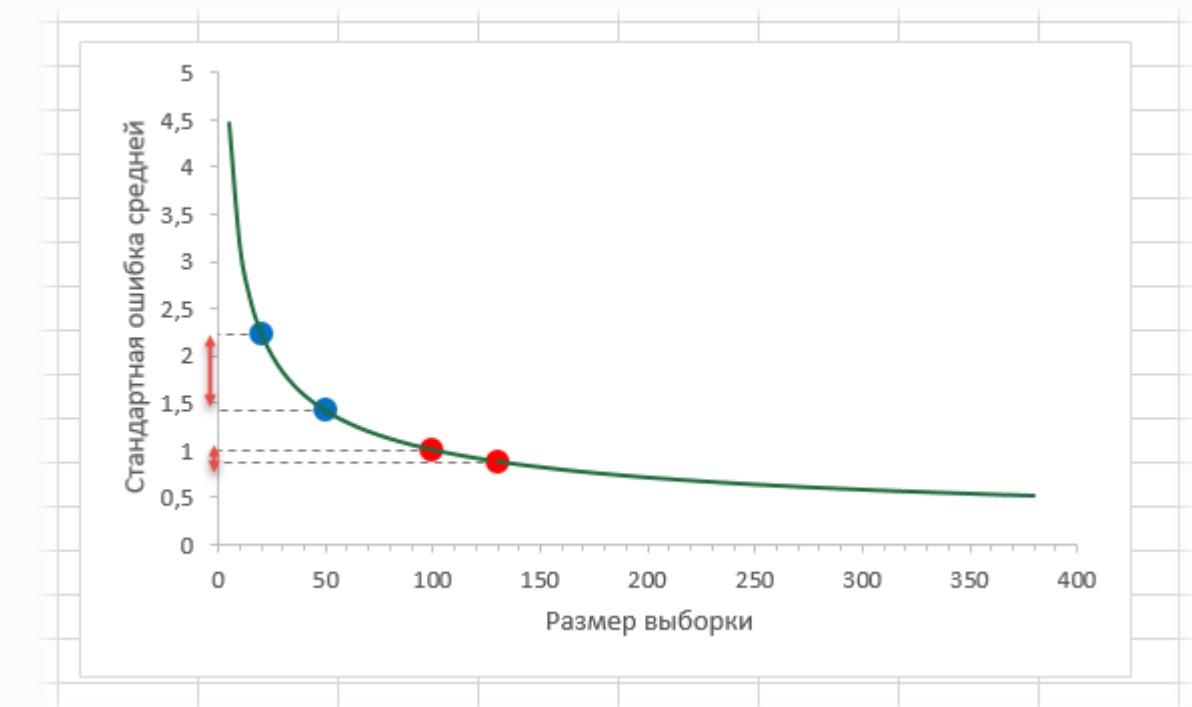
Судя по формуле, разброс стандартной ошибки средней должен быть в 4 раза (корень из 16) меньше, чем разброс исходных данных, что и видно на рисунке выше. Чем больше наблюдений, тем меньше разброс средней.

Казалось бы, что для получения наиболее точной средней достаточно использовать максимально большую выборку и тогда стандартная ошибка средней будет стремиться к нулю, а сама средняя, соответственно, к математическому ожиданию. Однако квадратный корень объема выборки в знаменателе говорит о том, что связь между точностью выборочной средней и размером выборки не является линейной. Например, увеличение выборки с 20-ти до 50-ти наблюдений, то есть на 30 значений или в 2,5 раза, уменьшает стандартную ошибку средней только на 36%, а со 100-а до 130-ти наблюдений (на те же 30 значений), снижает разброс данных лишь на 12%.



Назначение и свойство стандартной ошибки средней арифметической

Лучше всего изобразить эту мысль в виде графика зависимости стандартной ошибки средней от размера выборки. Пусть стандартное отклонение равно 10 (на форму графика это не влияет). Таким образом, при достижении некоторого размера выборки ее дальнейшее увеличение уже почти не оказывается на точности средней. Этот факт имеет далеко идущие последствия. Например, при проведении выборочного обследования населения (опроса) чрезмерное увеличение выборки ведет к неоправданным затратам, т.к. точность почти не меняется. Именно поэтому количество опрошенных редко превышает 1,5 тысячи человек. Точность при таком размере выборки часто является достаточной, а дальнейшее увеличение выборки – нецелесообразным. Дисперсия и стандартная ошибка средней имеют большое практическое значение. Они используются в проверке гипотез и расчете доверительных интервалов.

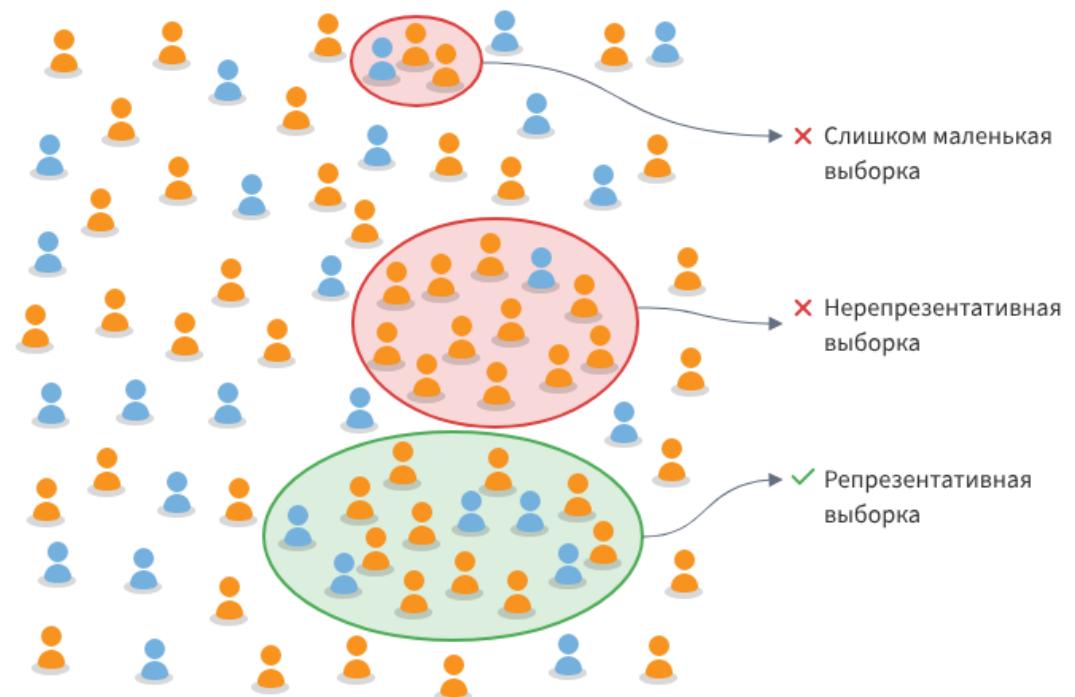


Репрезентативность Выборки

Репрезентативность – это необходимое условие для переноса результатов выборочного исследования на генеральную совокупность. Выборка должна адекватно отражать структуру и свойства всей исследуемой совокупности.

Ошибка репрезентативности (m) характеризует степень возможного отклонения выборочных данных от генеральных параметров и служит мерой точности выборочных оценок.

Генеральная совокупность включает  - 1/3 и  - 2/3



Доверительные границы: интервальная оценка

Доверительные границы определяют интервал, в пределах которого с заданной вероятностью находится истинное значение параметра в генеральной совокупности. Это позволяет перейти от точечной оценки к более надежной интервальной. При числе наблюдений больше 30 и стандартной доверительной вероятности 95% коэффициент t принимается равным 2. Например, при $M_{выб} = 45.6$ кг и $m = \pm 0.14$ кг, доверительные границы рассчитываются как $45.6 \pm 2 \cdot 0.14$, что дает интервал от 45.32 до 45.88 кг. Это означает, что с вероятностью 95% средний вес во всей генеральной совокупности находится в этих пределах.

Для средней величины (M)

$$M_{\text{ген}} = M_{\text{выб}} \pm t \cdot m$$

Особенности малых выборок

При работе с малыми выборками ($n \leq 30$) необходимо применять специальные поправки для обеспечения достоверности результатов.

Поправка на степени свободы

Используется поправка на число степеней свободы ($n-1$) при расчете ошибки средней арифметической.

Коэффициент Стьюдента

Доверительный коэффициент t определяется по специальным таблицам Стьюдента, а не принимается равным 2 или 3.

Это позволяет сохранить надежность статистических выводов даже при ограниченном объеме данных.

Доверительные границы: интервальная оценка

Доверительные границы определяют интервал, в пределах которого с заданной вероятностью находится истинное значение параметра в генеральной совокупности. Это позволяет перейти от точечной оценки к более надежной интервальной. Для большой выборки, при $P_{выб} = 70\%$ и $m = \pm 1.5\%$, доверительные границы при $t=2$ составят $70 \pm 3\%$, то есть от 67% до 73%. В случае малой выборки, например, 25 человек с показателем распространенности 60% и ошибкой $m = \pm 10\%$, при табличном значении $t=2.0$ для 24 степеней свободы, границы будут равны $60 \pm 20\%$, что дает очень широкий интервал от 40% до 80%.

Для относительной величины (P)

$$P_{ген} = P_{выб} \pm t \cdot m$$

Точность оценки

Точность статистической оценки напрямую связана с объемом выборки, уровнем доверия и вариабельностью признака.

- 1 Объем выборки (n)
- 2 Доверительная вероятность (P)
- 3 Вариабельность признака (σ)

1

Чем **больше** выборка, тем **уже** доверительный интервал, что **повышает** точность.

2

Доверительная вероятность (P)

Чем **выше** доверительная вероятность (например, 99% вместо 95%), тем **шире** интервал.

3

Вариабельность признака (σ)

Чем **меньше** вариабельность признака, тем **точнее** оценка.

Оценка достоверности различий

Для сравнения двух выборок и определения статистической значимости различий между ними используется расчет критерия t.

Разность средних величин

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

Разность относительных величин

$$t = \frac{P_1 - P_2}{\sqrt{m_1^2 + m_2^2}}$$

Оценка достоверности различий

Разность средних величин

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

Разность относительных величин

$$t = \frac{P_1 - P_2}{\sqrt{m_1^2 + m_2^2}}$$

Этот метод относится к параметрическим, что предполагает обязательное знание закона распределения изучаемых признаков и вычисление их основных параметров – средней арифметической величины и ее ошибки для количественных признаков или относительного показателя и его ошибки репрезентативности для атрибутивных данных. Ключевым условием для его корректного применения является репрезентативность обеих независимых сравниемых выборочных совокупностей.

Интерпретация критерия t

Значение критерия t позволяет количественно определить вероятность случайности выявленных различий и сделать научно обоснованный вывод.

$$t \geq 2 (p \leq 0,05)$$

Различия считаются **статистически достоверными**. Вероятность случайности не превышает 5%.

$$t < 2$$

Различия, скорее всего, объясняются **случайными причинами** или ошибкой выборки. Вывод о значимости не делается.

$$t \geq 3 (p \leq 0,01)$$

Достоверность выводов **превышает 99%**. Различия высоко значимы.