ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 50779.23— 2005 (ИСО 3301:1975)

Статистические методы

СТАТИСТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ДАННЫХ

Сравнение двух средних в парных наблюдениях

ISO 3301 : 1975
Statistical interpretation of data — Comparison of two means in the case of paired observations
(MOD)

Издание официальное

БЗ 2—2004/5

Москва
Стандартинформ
2005

gost.e-ecolog.ru

Предисловие

Задачи, основные принципы и правила проведения работ по государственной стандартизации в Российской Федерации установлены ГОСТ Р 1.0—92 «Государственная система стандартизации Российской Федерации. Основные положения» и ГОСТ Р 1.2—92 «Государственная система стандартизации Российской Федерации. Порядок разработки государственных стандартов»

Сведения о стандарте

1 ПОДГОТОВЛЕН Техническим комитетом по стандартизации ТК 125 «Статистические методы в управлении качеством продукции» и Научно-исследовательским центром контроля и диагностики технических систем на основе собственного аутентичного перевода стандарта, указанного в пункте 4

2 ВНЕСЕН Управлением технического регулирования и стандартизации Федерального агентства по техническому регулированию и метрологии

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 31 мая 2005 г. № 112-ст

4 Настоящий стандарт является модифицированным по отношению к международному стандарту ИСО 3301:1975 «Статистическое представление данных. Сравнение двух средних в парных наблюдениях» (ISO 3301:1975 «Statistical interpretation of data — Comparison of two means in the case of paired observations», MOD) путем включения отдельных фраз, которые выделены в тексте курсивом, с целью гармонизации с национальными стандартами.

Наименование настоящего стандарта изменено относительно наименования указанного международного стандарта для приведения в соответствие с ГОСТ 1.5 (подраздел 3.6)

5 ВВЕДЕН ВПЕРВЫЕ

Информация об изменениях к настоящему стандарту публикуется в указателе «Национальные стандарты», а текст изменений — в информационных указателях «Национальные стандарты». В случае пересмотра или отмены настоящего стандарта соответствующая информация будет опубликована в информационном указателе «Национальные стандарты»

© Стандартинформ, 2005

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Статистические методы

СТАТИСТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ДАННЫХ

Сравнение двух средних в парных наблюдениях

Statistical methods. Statistical interpretation of data. Comparison of two means in the case of paired observations

Дата введения — 2005—07—01

1 Область применения

Настоящий стандарт устанавливает метод проверки статистической гипотезы о равенстве среднего распределения (далее — среднее) разностей парных наблюдений нулю (предположение о несущественности расхождения между рядами наблюдений) или какому-либо другому заданному значению.

2 Нормативные ссылки

В настоящем стандарте использована нормативная ссылка на следующий стандарт:

ГОСТ Р 50779.10—2000 (ИСО 3534-1—93) Статистические методы. Вероятность и основы статистики. Термины и определения (ИСО 3534-1:1993, IDT)

Примечания

1 Настоящий раздел является дополнительным по отношению к содержанию международного стандарта ИСО 3301:1975 (ISO 3301:1975) и включен для учета основополагающих национальных стандартов в области статистических методов.

2 При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов по указателю «Национальные стандарты», составленному по состоянию на 1 января текущего года, и по соответствующим информационным указателям, опубликованным в текущем году. Если ссылочный стандарт заменен (изменен), то при пользовании настоящим стандартом следует руководствоваться замененным (измененным) стандартом. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, применяют в части, не затрагивающей эту ссылку.

3 Термины и определения

В настоящем стандарте применены термины по ГОСТ Р 50779.10, а также следующий термин с соответствующим определением:

парные наблюдения (paired observations): Два результата наблюдений определенных свойств или характеристик объекта x_i и y_i называются парными, если они получены:

- как результаты наблюдений над одним и тем же объектом i (из совокупности), причем данные наблюдения относятся к различным условиям получения этих наблюдений (например, сравнение двух методов анализа свойств одного и того же объекта);
- как результаты наблюдений над объектами, идентичными во всех отношениях, кроме предполагаемого систематического различия в некотором интересующем аспекте; в отношении значимости этого различия проводят проверку статистической гипотезы (например, сравнение урожайности двух соседних участков, засеянных семенами различных сортов).

Издание официальное

Во втором случае эффективность проверки гипотез зависит от степени уверенности в отсутствии каких-либо других систематических различий между объектами, кроме некоторого возможного различия, в отношении которого проверяют гипотезу.

4 Применение метода для сравнения двух способов обработки экспериментальных данных

Установленный в настоящем стандарте метод проверки статистической гипотезы может быть применен с целью подтвердить различие двух способов обработки. В этом случае можно считать, что результаты наблюдений x_i получены одним способом обработки, а результаты наблюдений y_i — некоторым другим способом. Две серии результатов наблюдений не являются независимыми, поскольку каждому x_i первой серии (первый способ обработки) ставится в соответствие вполне определенное y_i второй серии (второй способ обработки). Термин «способ обработки» понимают в широком смысле. При выявлении возможного систематического расхождения два сравниваемых способа обработки могут относиться к двум методам испытаний, к двум измерительным устройствам или к двум лабораториям. Два способа обработки, выполненные над одним и тем же экспериментальным материалом, могут влиять друг на друга, и полученное значение может зависеть от последовательности обработки. Оптимальный план эксперимента должен устранять возможные систематические смещения. В качестве альтернативы сравнению способов обработки можно рассмотреть воздействие одного способа обработки по сравнению со случаем полного отсутствия обработки.

5 Условия применения метода

Для корректного применения метода необходимо выполнение следующих двух условий:

- последовательность разностей $d_i = x_i - y_i$ является выборкой независимой случайной величины;
- распределение величин $d_i = x_i - y_i$ является нормальным или близким к нормальному.

Если распределение разностей отклоняется от нормального, метод проверки гипотез применим при условии, что объем выборки достаточно велик. При больших отклонениях от нормальности требуется соответственно бóльшая выборка. Однако даже в предельных случаях отклонения от нормальности выборки объемом 100 могут быть достаточными для корректного применения метода в большинстве прикладных задач.

6 Расчетные формулы и правила принятия решения

Изучаемая проблема _____

Условия проведения эксперимента _____

Статистические данные	**Расчетные формулы**
Объем выборки $n =$ Суммы результатов наблюдений: $\sum x_i =$ $\sum y_i =$ Сумма разностей: $\sum d_i =$ Сумма квадратов разностей: $\sum d_i^2 =$	$\bar{d} = \frac{1}{n}\left(\sum x_i - \sum y_i\right) = \frac{1}{n}\sum d_i =$ $s_d^2 = \frac{1}{n-1}\left[\sum d_i^2 - \frac{1}{n}\left(\sum d_i\right)^2\right] =$ $\sigma_d = \sqrt{s_d^2} =$

Окончание

Заданное значение (среднее случайных разностей парных наблюдений): d_0 = Число степеней свободы: $\nu = n - 1$ = Выбранный уровень значимости: α =	$A_1 = [t_{1-\alpha}(\nu) / \sqrt{n}]\,\sigma_d$ = $A_2 = \left[t_{1-\alpha/2}(\nu) / \sqrt{n}\right]\sigma_d$ =

Правило принятия решения

1 Двусторонняя критическая область:

Гипотезу о том, что среднее совокупности разностей равно d_0 (нулевая гипотеза H_0), отвергают, если

$$|\bar{d} - d_0| > A_2 .$$

2 Односторонняя критическая область:

а) Гипотезу о том, что среднее совокупности разностей равно d_0 (нулевая гипотеза H_0), отвергают, если

$$\bar{d} < d_0 - A_1 .$$

б) Гипотезу о том, что среднее совокупности разностей равно d_0 (нулевая гипотеза H_0), отвергают, если

$$\bar{d} > d_0 + A_1 .$$

Примечание — $t_{1-\alpha}(\nu)$ — квантиль уровня 1 — α статистики Стьюдента с (ν) степенями свободы. Значения $t_{1-\alpha}(\nu)/\sqrt{n}$ приведены в таблице 1.

Таблица 1 — Значения отношения $t_{1-\alpha}(\nu)/\sqrt{n}$ для $\nu = n-1$

$\nu = n-1$	Двусторонняя гипотеза		Односторонняя гипотеза		$\nu = n-1$	Двусторонняя гипотеза		Односторонняя гипотеза	
	$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,995}}{\sqrt{n}}$	$\frac{t_{0,95}}{\sqrt{n}}$	$\frac{t_{0,99}}{\sqrt{n}}$		$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,995}}{\sqrt{n}}$	$\frac{t_{0,95}}{\sqrt{n}}$	$\frac{t_{0,99}}{\sqrt{n}}$
1	8,985	45,013	4,465	22,501	21	0,443	0,604	0,367	0,537
2	2,434	5,730	1,686	4,021	22	0,432	0,588	0,358	0,523
3	1,591	2,920	1,177	2,270	23	0,422	0,573	0,350	0,510
4	1,242	2,059	0,953	1,676	24	0,413	0,559	0,342	0,498
5	1,049	1,646	0,823	1,374	25	0,404	0,547	0,335	0,487
6	0,925	1,401	0,734	1,188	26	0,396	0,535	0,328	0,477
7	0,836	1,237	0,670	1,060	27	0,388	0,524	0,322	0,467
8	0,769	1,118	0,620	0,966	28	0,380	0,513	0,316	0,458
9	0,715	1,028	0,580	0,892	29	0,373	0,503	0,310	0,449
10	0,672	0,956	0,546	0,833	30	0,367	0,494	0,305	0,441
11	0,635	0,897	0,518	0,785	40	0,316	0,422	0,263	0,378
12	0,604	0,847	0,494	0,744	50	0,281	0,375	0,235	0,337
13	0,577	0,805	0,473	0,708	60	0,256	0,341	0,214	0,306
14	0,554	0,769	0,455	0,678	70	0,237	0,314	0,198	0,283
15	0,533	0,737	0,438	0,651	80	0,221	0,293	0,185	0,264
16	0,514	0,708	0,423	0,626	90	0,208	0,276	0,174	0,248
17	0,497	0,683	0,410	0,605	100	0,197	0,261	0,165	0,235
18	0,482	0,660	0,398	0,586	200	0,139	0,183	0,117	0,165
19	0,468	0,640	0,387	0,568	500	0,088	0,116	0,074	0,104
20	0,455	0,621	0,376	0,552	∞	0	0	0	0

Пример — Приведенные в таблице 2 данные собраны в процессе исследований, спланированных с целью определить, зависит ли скорость изнашивания шеек коленчатого вала двигателя внутреннего сгорания от типа материала металлических вкладышей подшипников скольжения этого вала.

Таблица 2 — Износ шеек коленчатых валов в течение заданной наработки, измеренный в стотысячных долях дюйма

Номер вала i	Износ вала		Разность $d_i = x_i - y_i$
	Медно-свинцовые вкладыши x_i	Алюминиевые вкладыши y_i	
1	3,5	1,5	2,0
2	2,0	1,3	0,7
3	4,7	4,5	0,2
4	2,8	2,5	0,3
5	6,5	4,5	2,0
6	2,2	1,7	0,5
7	2,5	1,8	0,7
8	5,8	3,3	2,5
9	4,2	2,3	1,9
Сумма	34,2	23,4	10,8

Технические характеристики

Статистические данные

Объем выборки:
$n = 9$
Суммы результатов наблюдений:
$\sum x_i = 34{,}2 \qquad \sum y_i = 23{,}4$
Сумма разностей:
$\sum d_i = 10{,}8$
Сумма квадратов разностей:
$\sum d_i^2 = 19{,}22$
Заданное значение:
$d_0 = 0$
Число степеней свободы:
$\nu = 8$
Выбранный уровень значимости:
$\alpha = 0{,}01$

Вычисления

$$\bar{d} = \frac{1}{9}(34{,}2 - 23{,}4) = 1{,}2$$

$$s_d^2 = \frac{1}{8}19{,}22 - \frac{10{,}8^2}{9} = 0{,}7825$$

$$\sigma_d = \sqrt{0{,}7825} = 0{,}8846$$

$$t_{0{,}995} / \sqrt{9} = 1{,}118$$

$$A_2 = 1{,}118 \times 0{,}8846 = 0{,}99$$

Выводы

Сравнение среднего совокупности с заданным значением 0:
Проверка двусторонней гипотезы:

$$|\bar{d} - d_0| = 1{,}2 > 0{,}99.$$

Гипотезу равенства скоростей изнашивания шеек коленчатого вала двигателя в случае подшипников с различными вкладышами (медно-свинцовые и алюминиевые) отвергают на 1 %-м уровне значимости.

7 Ошибка второго рода

Вероятность отклонения нулевой гипотезы, когда она верна, равна уровню значимости α. Отклонение нулевой гипотезы, когда она верна, называется ошибкой первого рода, и поэтому выбор значения α ограничивает риск такой ошибки.

С другой стороны, можно совершить ошибку второго рода, то есть проверить нулевую гипотезу, когда она неверна. Вероятность 1 — β отклонения нулевой гипотезы, когда она неверна, называется мощностью критерия статистической проверки гипотезы. Вероятность ошибки второго рода в таком случае равна β.

Для заданной выборки n и ошибки первого рода вероятность β зависит не только от истинного среднего D наблюдаемых разностей $d_i = x_i - y_i$, для которых устанавливают различные альтернативные гипотезы, но также и от стандартного отклонения σ_d этих разностей. Это стандартное отклонение неизвестно, и, если n мало, выборка может обеспечить только плохую оценку. В результате невозможно устанавливать верхний предел вероятности ошибки второго рода.

На графиках (рисунки 1 и 2) показаны зависимости между мощностью 1 — β критерия проверки гипотез и истинным средним совокупности, деленным на соответствующее стандартное отклонение (D / σ_d) для случая односторонней гипотезы H_0: $D \leq 0$, различных значений n и уровней значимости 0,05 и 0,01 соответственно.

На основании этих графиков можно сделать следующие заключения:

1) Мощность критерия проверки гипотез однозначно определяется истинным средним D совокупности разностей, измеренных в единицах стандартного отклонения σ_d, уровнем значимости α и объемом выборки.

2) Функция мощности является монотонно возрастающей функцией истинного среднего совокупности разностей. Она также монотонно возрастает с ростом объема выборки и уровня значимости α при условии, что $D > 0$ и α отличается от 0 и 1.

3) Для уровня значимости 0,05 и при объеме выборки 50 достигается мощность 0,95, если истинное среднее разностей превышает половину стандартного отклонения разностей. Для $n = 20$ такая мощность достигается для $D / \sigma_d = 0{,}78$ и бóльших значений.

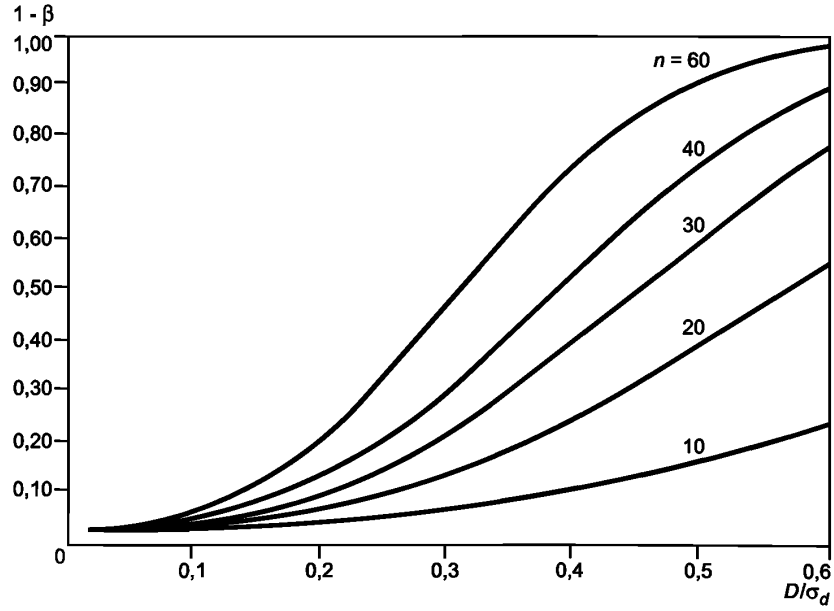


Рисунок 1 — Мощность критерия для одной выборки (односторонний критерий), $\alpha = 0{,}01$

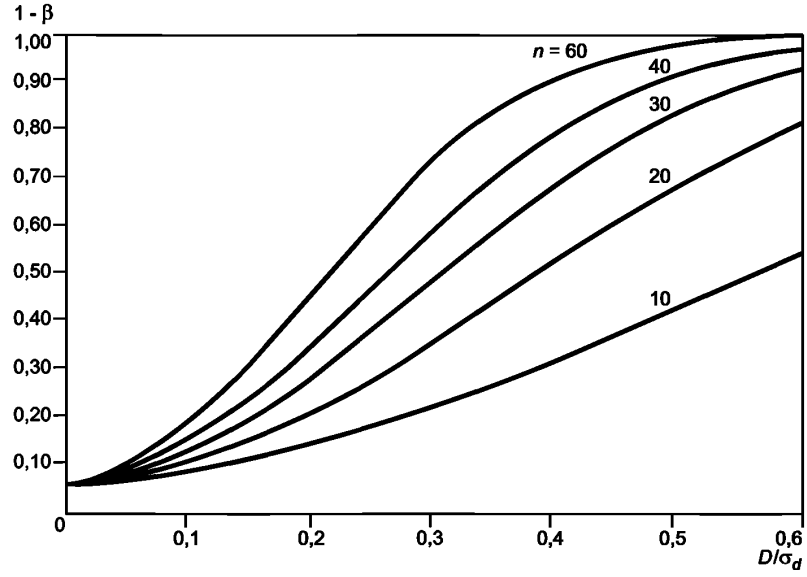


Рисунок 2 — Мощность критерия для одной выборки (односторонний критерий), $\alpha = 0,05$

УДК 519.25:620.113:006.354 ОКС 03.120.10 Т59

Ключевые слова: проверка статистической гипотезы, результаты наблюдений, ошибка второго рода, стандартное отклонение, мощность критерия, случайный отбор, выборочное среднее

Редактор *Л.В. Афанасенко*
Технический редактор *Н.С. Гришанова*
Корректор *В.Е. Нестерова*
Компьютерная верстка *Е.Н. Мартемьяновой*

Сдано в набор 06.06.2005. Подписано в печать 21.06.2005. Формат 60×84 1/8. Бумага офсетная. Гарнитура Ариал. Печать офсетная. Усл. печ. л. 0,93. Уч.-изд. л. 0,65. Тираж 664 экз. Зак. 381. С 1416.

ФГУП «Стандартинформ», 123995 Москва, Гранатный пер., 4.
gost.e-ecolog.ru gost.e-ecolog.ru
Набрано во ФГУП «Стандартинформ» на ПЭВМ.
Отпечатано в филиале ФГУП «Стандартинформ» — тип. «Московский печатник», 105062 Москва, Лялин пер., 6.

gost.e-ecolog.ru