

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Смирнова З.М., Крейнина М.В.

1. Общие понятия.
2. Проверка гипотез относительно средних значений. t–критерий Стьюдента.
3. Проверка гипотез для дисперсий. F-критерий Фишера.
4. Ранговый непараметрический U-критерий Манна-Уитни.

Статистические гипотезы – это предположения или допущения о неизвестных генеральных параметрах, выражаемых в терминах вероятности, которые могут быть проверены на основании выборочных показателей с помощью статистических критериев, основанных на использовании статистических распределений.

Статистические гипотезы **различают по виду предположений**, содержащихся в них, при этом рассматривают нулевую и альтернативную гипотезы.

Нулевая гипотеза (H_0) - это гипотеза о том, что две совокупности X и Y, которые сравниваются по одному или нескольким признакам, не различаются между собой. При этом предполагают, что действительное различие сравниваемых величин равно нулю:

$$\bar{X}_r - \bar{Y}_r = 0 \text{ или } \bar{X}_r = \bar{Y}_r$$

Необходимо отметить, что в некоторых случаях, когда расчеты соответствуют нулевой гипотезе, гипотеза H_0 может быть и не принята. Это, скорее всего, указывает на одну из двух возможных причин:

- а) объемы выборок недостаточно велики, чтобы обнаружить имеющиеся различия. Вполне вероятно, что продолжение экспериментов принесет успех.
- б) различия есть, но они настолько малы, что не имеют практического значения. В этом случае продолжение экспериментов не имеет смысла.

Альтернативная гипотеза (H_1) – это предположение, принимаемое в случае отклонения нулевой гипотезы.

Существует **два типа статистических методов** параметрические и непараметрические. Параметрические методы – это количественные методы статистической обработки данных, применение которых требует обязательного знания закона распределения изучаемых признаков в совокупности и вычисления их основных параметров. Например, известно, что выборки извлечены из генеральных совокупностей с нормальным законом распределения и одинаковыми дисперсиями. Требуется выяснить, одинаковы ли генеральные средние этих совокупностей.

Непараметрические методы – это количественные методы статистической обработки данных, применение которых не требует знания закона распределения изучаемых признаков в совокупности и вычисления их основных параметров. Например, *непараметрической* является гипотеза о равенстве генеральных средних значений двух совокупностей, если нет информации о *виде* законов распределения измеряемой величины.

Статистический критерий (К) (критерий значимости)- это некий параметр, вычисленный по определенному алгоритму, который используется для проверки основной гипотезы.

Критерий **К** вычисляется на основе экспериментальных данных (эмпирический критерий) При этом вычисленное эмпирическое значение может оказаться таким, что, действуя по определенному алгоритму, на данном уровне значимости мы примем основную гипотезу, а может оказаться таким, что мы вынуждены будем ее отвергнуть, то есть, принять конкурирующую гипотезу.

Согласно статистическим методам **статистические критерии** делятся на **параметрические** и **непараметрические**.

Параметрические критерии используются в задачах проверки параметрических гипотез и включают в свой расчет показатели распределения, например, средние, дисперсии и т.д. Это такие известные классические критерии, как критерий Стьюдента, критерий Фишера и др. Они позволяют сравнить основные параметры генеральных совокупностей, а также оценить разности средних и различия в дисперсиях. Критерии способны выявить тенденции изменения признака, оценить взаимодействие двух и более факторов в воздействии на изменения признака

Непараметрические критерии проверки гипотез основаны на операциях с другими данными, в частности, частотами, рангами и т.п. Это - критерий Манна-Уитни, критерий Уилкоксона и многие другие. Непараметрические критерии позволяют решить некоторые важные задачи, связанные с выявлением различий исследуемого признака, с оценкой сдвига значений исследуемого признака, выявлением различий в распределениях.

Как параметрические, так и непараметрические методы, используемые для сравнения результатов исследований, т.е. для сравнения выборочных совокупностей, заключаются в применении определенных формул и расчетов определенных показателей в соответствии с предписанными алгоритмами. В конечном результате высчитывается определенная числовая величина **К** (эмпирическая), которую сравнивают с табличными пороговыми значениями **К_{кр}** (критическими).

Критическое значение (**К_{кр}**) является функцией уровня статистической значимости (**β**) и числа степеней свободы (**f**): $K_{кр} = \varphi(f, \beta)$.

Степени свободы (f) - это количество значений, способных варьироваться. Число степеней свободы равно числу членов вариационного ряда минус число условий, для которых он был сформирован. К числу таких условий относятся объем выборки, средние значения и дисперсии.

Уровень статистической значимости (β) - это вероятность того, что мы признали различия существенными (приняли альтернативную гипотезу и отклонили нулевую), а они в действительности случайные. Например, если указывается, что различия достоверны на 5%-ном уровне значимости, то вероятность того, что они все же недостоверны, равна 0,05.

Для большинства медико-биологических исследований считается достаточной вероятностью безошибочного прогноза, равная **95% (p=0.95)**, а число объектов генеральной совокупности, в котором могут наблюдаться отклонения от закономерностей, установленных при выборочном исследовании, не будет превышать **5%** (уровень

значимости $\beta=0.05$). При ряде исследований, связанных, например, с применением высокотоксичных веществ, вакцин, оперативного лечения и т.п., в результате чего возможны тяжелые заболевания, осложнения, летальные исходы, применяется вероятность, равная 99% ($p=0.99$) т.е. отклонения от закономерностей в генеральной совокупности возможны не более чем в 1% случаев (уровень значимости ($\beta=0.01$)).

Алгоритм использования любого критерия включает в себя:

- 1) выбор соответствующего статистического метода;
- 2) формулировку нулевой и альтернативной гипотез;
- 3) выбор значения доверительной вероятности (уровня значимости);
- 4) вычисление эмпирического значения критерия;
- 5) нахождение критического значения критерия с помощью таблиц;
- 6) принятие решения на основании сравнения эмпирического и критического значений критерия.

ПАРАМЕТРИЧЕСКИЕ МЕТОДЫ.

Среди параметрических методов оценки достоверности результатов исследования в медицине достаточно широко распространены методы с использованием **t-критерия Стьюдента** и **F-критерия Фишера**.

КРИТЕРИЙ СТЬЮДЕНТА (t – критерий Стьюдента)

Данный критерий был разработан Уильямом Госсетом для оценки качества пива в компании Гиннесс. В связи с обязательствами перед компанией по неразглашению коммерческой тайны, статья Госсета вышла в 1908 году в журнале «Биометрика» под псевдонимом «Student» (Студент).

Задача сравнения *средних* значений двух генеральных совокупностей возникает, когда практическое значение имеет именно *величина* исследуемого признака. Например, когда сравниваются сроки лечения двумя различными методами, или количества осложнений, возникающих при их применении, когда сравнивается эффективность новых лекарственных средств и т.д. В этом случае можно использовать t-критерий Стьюдента.

Особенность использования **t-критерия Стьюдента** при обработке результатов исследования состоит в необходимости выполнения двух условий: **нормального характера распределения изучаемых данных** и **равенства генеральных дисперсий** сравниваемых групп. В противном случае использование t-критерия Стьюдента некорректно.

Данный критерий применим для сравнения больших и средних (объемом более 30 значений) неравновеликих ($n_x \neq n_y$) выборок, а также, для малых (объемом менее 30 значений) равновеликих ($n_x = n_y$) выборок.

t-критерий Стьюдента можно использовать как в случаях **сравнения независимых выборок**, так и при **сравнении зависимых выборок**.

Зависимые выборки - это выборки, представляющие собой параметры одной и той же совокупности до и после воздействия некоторого фактора. Чаще всего зависимые выборки – это измерения **одной и той же группы объектов** в разные моменты времени (например, до и после воздействия какого-либо фактора: *средняя частота пульса у одних и тех же пациентов до и после приема антиаритмического препарата*).

Выборки являются **независимыми**, если набор объектов исследования в каждую из групп осуществлялся независимо от того, какие объекты исследования включены в другую группу. Примером сравнения независимых выборок может служить сопоставление данных анализа крови в группе пациентов с аналогичными показателями в группе здоровых. Допускается, чтобы количество объектов в независимых выборках было различным.

Алгоритм использования t-критерия Стьюдента.

При сравнении двух независимых выборок из различных генеральных совокупностей X и Y, подчиняющихся нормальному закону распределения, проверяются гипотезы:

H₀ - нулевая гипотеза: $\bar{X}_r = \bar{Y}_r, (\bar{X}_r - \bar{Y}_r = 0)$;

H₁ альтернативная гипотеза: $\bar{X}_r \neq \bar{Y}_r, (\bar{X}_r - \bar{Y}_r \neq 0)$

В гипотезе **H₀** принимается предположение о том, что различия не являются статистически достоверными и носят случайный характер.

В гипотезе **H₁** принимается предположение о том, что разность между генеральными средними сравниваемых групп не равна нулю, и исследуемые выборки не принадлежат одной генеральной совокупности, а различия, наблюдаемые между выборочными показателями, носят именно систематический, а не случайный характер.

1. Эмпирическое значение критерия Стьюдента (**t**) вычисляют следующим образом:

$$t = \frac{|\bar{x}_B - \bar{y}_B|}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}, \quad \text{где}$$

$$s_x^2 = \frac{1}{n_x - 1} \sum_{i=1}^n (x_i - \bar{x}_B)^2 \quad s_y^2 = \frac{1}{n_y - 1} \sum_{i=1}^n (y_i - \bar{y}_B)^2$$

\bar{x}_B, \bar{y}_B – выборочные средние значения выборок;

s_x^2, s_y^2 – приведенные (исправленные) оценки выборочных дисперсий для исследуемых выборок;

n_x, n_y – число вариантов в выборках;

2. Теоретический критерий (**t_{кр}**) находят по таблице распределения Стьюдента для заданного уровня значимости (**β**) или доверительной вероятности (**p**) и числа степеней свободы (**f**). (Таблица 1).

Число степеней свободы **f** определяют следующим образом:

если $s_x^2 \approx s_y^2$, то **f = n_x + n_y - 2**;

если s_x^2 и s_y^2 различаются более, чем на порядок, то

f = (n_x + n_y - 2) [1/2 + (s_x² s_y²) / (s_x⁴ + s_y⁴)].

3. Решение о достоверности различий, наблюдаемых между выборочными средними принимают на основании сравнения эмпирического значение критерия (**t**) и критического значение критерия (**t_{кр}**) на определенном уровне значимости (**β**).

Таблица 1.

Решение о принятии гипотезы	Условия принятия гипотезы	Вывод о достоверности различий, наблюдаемых между выборочными средними
Гипотезу H₀ принимают	t ≤ t_{кр}	Различия не являются статистически достоверными и носят случайный характер. ($\bar{X}_r = \bar{Y}_r$)
Гипотезу H₁ принимают (Гипотезу H₀ отвергают)	t > t_{кр}	Различия являются статистически достоверными ($\bar{X}_r - \bar{Y}_r \neq 0$)

КРИТЕРИЙ ФИШЕРА (F–критерий Фишера)

В некоторых клинических исследованиях о положительном эффекте свидетельствует не столько *величина* исследуемого параметра, сколько его *стабилизация*, уменьшение его колебаний. В этом случае возникает вопрос *о сравнении двух генеральных дисперсий* по результатам выборочного обследования. Эта задача может быть решена с помощью параметрического критерия Фишера.

Алгоритм использования F–критерия Фишера.

При сравнении двух независимых выборок из одной или различных генеральных совокупностей, подчиняющихся нормальному закону распределения, проверяются гипотезы:

H₀ - нулевая гипотеза: $\sigma^2_{ГХ} = \sigma^2_{ГУ}$, ($\sigma^2_{ГХ} - \sigma^2_{ГУ} = 0$);

H₁ - альтернативная гипотеза: $\sigma^2_{ГХ} \neq \sigma^2_{ГУ}$, ($\sigma^2_{ГХ} - \sigma^2_{ГУ} \neq 0$)

В гипотезе **H₀** принимается предположение о том, что различия не являются статистически достоверными и носят случайный характер.

В гипотезе **H₁** принимается предположение о том, что разность между генеральными дисперсиями сравниваемых групп не равна нулю, и исследуемые выборки не принадлежат одной генеральной совокупности, а различия, наблюдаемые между выборочными показателями, носят именно систематический, а не случайный характер.

1. Эмпирическое значение критерия вычисляют по формуле:

$$F = \frac{s_x^2}{s_y^2},$$

s_x^2, s_y^2 – приведенные (исправленные) оценки выборочных дисперсий для исследуемых выборок x и y при условии, что $s_x^2 > s_y^2$ и:

$$s_x^2 = \frac{1}{n_x - 1} \sum_{i=1}^n (x_i - \bar{x}_B)^2 \quad s_y^2 = \frac{1}{n_y - 1} \sum_{i=1}^n (y_i - \bar{y}_B)^2$$

2. По таблице определяют критическое значение критерия Фишера (**F_{кр}**) для выбранного уровня значимости (**β**) или доверительной вероятности (**p**) и числа степеней свободы (**f_x** и **f_y**): **f_x = n_x - 1** и **f_y = n_y - 1** (Таблица 2).

3. Решение о достоверности различий, наблюдаемых между выборочными дисперсиями, принимают на основании сравнения двух величин **F** и **F_{кр}** на определенном уровне значимости. (Например, **β = 5%**).

Таблица 2.

Решение о принятии гипотезы	Условия принятия гипотезы	Вывод о достоверности различий, наблюдаемых между выборочными дисперсиями
Гипотезу H₀ принимают	F ≤ F_{кр}	Различия не являются статистически достоверными и носят случайный характер, т.е. выборки взяты из одной и той же генеральной совокупности или из разных генеральных совокупностей, при условии, что $\sigma^2_{ГХ} = \sigma^2_{ГУ}$
Гипотезу H₁ принимают (Гипотезу H₀ отвергают)	F > F_{кр}	Различия являются статистически достоверными, т.е. выборки взяты из разных генеральных совокупностей и $\sigma^2_{ГХ} \neq \sigma^2_{ГУ}$

НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ.

Применение параметрических методов для проверки статистических гипотез основано на представлении о нормальном распределении совокупностей, из которых взяты сравниваемые выборки, а также, о равенстве генеральных параметров сравниваемых выборок. Однако эти условия выполняются не всегда, так как не все биологические признаки распределяются нормально и об их законе распределения часто мало что известно. Следует учитывать еще и то, что исследователю приходится подвергать анализу не только количественные, но и качественные (атрибутивные) признаки, которые могут выражаться не только абсолютными или относительными числами, но и порядковыми номерами, индексами, условными знаками и т.д. В таких случаях параметрические методы не применимы. Поэтому в медико-биологических исследованиях наряду с параметрическими используют наиболее простые **непараметрические (порядковые или ранговые)** методы, основанные на использовании соответствующих статических критериев. Непараметрические критерии предназначены для обнаружения различий в законах распределения двух генеральных совокупностей с неизвестными законами распределения.

КРИТЕРИЙ МАННА-УИТНИ (U-критерий Манна-Уитни)

Данный метод выявления различий между выборками был предложен в 1945 году Фрэнком Уилкоксоном (*F. Wilcoxon*). В 1947 году он был существенно переработан и расширен Х.Б.Манном (*H. B. Mann*) и Д.Р.Уитни (*D. R. Whitney*), по именам которых и называется.

U-критерий Манна-Уитни – непараметрический статистический критерий, используется для оценки различий между двумя независимыми и несвязанными малыми выборками по уровню какого-либо признака, измеренного количественно, например, для обнаружения различий в средних значениях двух генеральных совокупности (*оценка эффективности новых лекарственных препаратов, новых методов физиотерапевтического лечения, сравнение результатов биохимического исследования в двух группах лиц и т.д.*), выборки из которых представлены в *ранговой шкале*. Измеренные признаки располагаются на этой шкале в порядке возрастания, а затем нумеруются целыми числами 1, 2, ... Эти числа и называются *рангами*. Значение имеет не сама величина признака, а лишь *порядковое место*, который она занимает среди других величин. Мерой отличия является число **T** - сумма рангов для каждой из групп.

U-критерий Манна-Уитни представляет непараметрическую альтернативу t-критерию Стьюдента для независимых выборок и вычисляется по формуле:

$$U = n_x \cdot n_y + \frac{n(n+1)}{2} - T, \quad (1), \text{ где:}$$

n_x и n_y - объемы выборок;

n - объем выборки, имеющей большую ранговую сумму;

T - большая сумма рангов из выборок X и Y .

Алгоритм использования U-критерия Манна-Уитни.

При сравнении двух независимых малых (число вариантов в выборках менее 30) выборок за «рабочую» гипотезу принимается **альтернативная гипотеза: $H_1 (\bar{X}_r - \bar{Y}_r \neq 0)$** т.е. признается *статистическая значимость различий между уровнями признака в рассматриваемых выборках.*

1. Составить единый ранжированный ряд из обеих сопоставляемых выборок, расставив их элементы по степени нарастания признака и приписав меньшему значению меньший номер (ранг). Если значения совпадают, им присваивают один и тот же **средний ранг** (например, если два значения поделили 3-е и 4-е номера обоим присваивают ранг 3,5).
2. Подсчитать **отдельно** сумму рангов для первой и второй выборок.
3. Определить **наибольшую** из двух ранговых сумм **T**.
4. Вычислить эмпирическое значение **U-критерия (U)** (по формуле (1)).
5. Определить по таблице для избранного уровня статистической значимости (**$\beta = 5\%$** или **$\beta = 1\%$**) или доверительной вероятности (**$p = 0.95$** или **$p = 0.99$**) **критическое значение ($U_{кр}$)** при заданной численности групп. (Таблица 3 приложения).
6. Решение о достоверности различий, наблюдаемых между уровнем признака в рассматриваемых выборках, принимают на основании сравнения полученных эмпирического (**U**) и критического (**$U_{кр}$**) значений критерия Манна-Уитни.

Таблица 3

Решение о принятии гипотезы	Условия принятия гипотезы	Вывод о достоверности различий между уровнем признака в рассматриваемых выборках
гипотезу H_0 принимают	$U > U_{кр}$	Различия не являются статистически достоверными и носят случайный характер ($\bar{X}_r = \bar{Y}_r$)
Гипотезу H_1 принимают (Гипотезу H_0 отвергают)	$U \leq U_{кр}$	Различия являются статистически достоверными. ($\bar{X}_r - \bar{Y}_r \neq 0$)

Пример. Результаты тестирования по 30-бальной шкале для группы X и группы Y представлены в таблице 4. Сравнить эффективность двух методов обучения студентов в двух группах для уровня статистической значимости $\beta = 5\%$.

Таблица 4.

X	18	10	7	15	14	11	13				
Y	15	20	10	8	16	10	19	7	15	14	29

Составим объединенную ранговую таблицу 5, расположив во второй строке значения вариант из обеих выборок в порядке возрастания

Таблица 5.

Принадлежность к выборке	X	Y	Y	X	Y	Y	X	X	X	Y	X	Y	Y	Y	X	Y	Y	Y
Значения	7	7	8	10	10	10	11	13	14	14	15	15	15	16	18	19	20	29
Номер	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Ранг	1,5	1,5	3	5	5	5	7	8	9,5	9,5	12	12	12	14	15	16	17	18

Найдем суммы рангов каждой выборки и выберем **большую** из них: $T = 113$

Рассчитаем эмпирическое значение критерия по формуле (1):

$$U = 7 \cdot 11 + [11 \cdot (11+1)] / 2 - 113 = 30.$$

Определим по таблице 3 приложения критическое значение критерия при уровне значимости $\beta = 5\%$: $U_{кр} = 16$.

Вывод: так как эмпирическое значение U -критерия больше критического ($U > U_{кр}$) при уровне значимости $\beta = 5\%$, то гипотеза H_0 о равенстве средних принимается и различия в методиках обучения не существенны.

Список сокращений.

- x и y - совокупность значений вариант для выборок;
- n_x и n_y - объемы выборок x и y ;
- $\bar{x}_в$ и $\bar{y}_в$ - выборочные средние значения для выборок x и y ;
- $\bar{X}_Г$ и $\bar{Y}_Г$ - генеральные средние значения для генеральных совокупностей, из которых получены выборки x и y ;
- $\sigma^2_{Гx}$ и $\sigma^2_{Гy}$ - генеральные дисперсии для генеральных совокупностей X и Y ;
- s_x^2 и s_y^2 - приведенные (исправленные) оценки выборочных дисперсий для выборок x и y ;
- p - доверительная вероятность;
- β - уровень значимости; f - число степеней свободы

ПРИЛОЖЕНИЕ

ТАБЛИЦА 1. Критические значения t- критерия Стьюдента ($t_{кр}$)

f	p				f	p				f	p			
	0,9	0,95	0,99	0,999		0,9	0,95	0,99	0,999		0,9	0,95	0,99	0,999
1	6,314	12,70	63,65	636,61	31	1,696	2,040	2,744	3,633	61	1,670	2,000	2,659	3,457
2	2,920	4,303	9,925	31,602	32	1,694	2,037	2,738	3,622	62	1,670	1,999	2,657	3,454
3	2,353	3,182	5,841	12,923	33	1,692	2,035	2,733	3,611	63	1,669	1,998	2,656	3,452
4	2,132	2,776	4,604	8,610	34	1,691	2,032	2,728	3,601	64	1,669	1,998	2,655	3,449
5	2,015	2,571	4,032	6,869	35	1,690	2,030	2,724	3,591	65	1,669	1,997	2,654	3,447
6	1,943	2,447	3,707	5,959	36	1,688	2,028	2,719	3,582	66	1,668	1,997	2,652	3,444
7	1,895	2,365	3,499	5,408	37	1,687	2,026	2,715	3,574	67	1,668	1,996	2,651	3,442
8	1,860	2,306	3,355	5,041	38	1,686	2,024	2,712	3,566	68	1,668	1,995	2,650	3,439
9	1,833	2,262	3,250	4,781	39	1,685	2,023	2,708	3,558	69	1,667	1,995	2,649	3,437
10	1,812	2,228	3,169	4,587	40	1,684	2,021	2,704	3,551	70	1,667	1,994	2,648	3,435
11	1,796	2,201	3,106	4,437	41	1,683	2,020	2,701	3,544	71	1,667	1,994	2,647	3,433
12	1,782	2,179	3,055	4,318	42	1,682	2,018	2,698	3,538	72	1,666	1,993	2,646	3,431
13	1,771	2,160	3,012	4,221	43	1,681	2,017	2,695	3,532	73	1,666	1,993	2,645	3,429
14	1,761	2,145	2,977	4,140	44	1,680	2,015	2,692	3,526	74	1,666	1,993	2,644	3,427
15	1,753	2,131	2,947	4,073	45	1,679	2,014	2,690	3,520	75	1,665	1,992	2,643	3,425
16	1,746	2,120	2,921	4,015	46	1,679	2,013	2,687	3,515	76	1,665	1,992	2,642	3,423
17	1,740	2,110	2,898	3,965	47	1,678	2,012	2,685	3,510	78	1,665	1,991	2,640	3,420
18	1,734	2,101	2,878	3,922	48	1,677	2,011	2,682	3,505	79	1,664	1,990	2,639	3,418
19	1,729	2,093	2,861	3,883	49	1,677	2,010	2,680	3,500	80	1,664	1,990	2,639	3,416
20	1,725	2,086	2,845	3,850	50	1,676	2,009	2,678	3,496	90	1,662	1,987	2,632	3,402
21	1,721	2,080	2,831	3,819	51	1,675	2,008	2,676	3,492	100	1,660	1,984	2,626	3,390
22	1,717	2,074	2,819	3,792	52	1,675	2,007	2,674	3,488	110	1,659	1,982	2,621	3,381
23	1,714	2,069	2,807	3,768	53	1,674	2,006	2,672	3,484	120	1,658	1,980	2,617	3,373
24	1,711	2,064	2,797	3,745	54	1,674	2,005	2,670	3,480					
25	1,708	2,060	2,787	3,725	55	1,673	2,004	2,668	3,476					
26	1,706	2,056	2,779	3,707	56	1,673	2,003	2,667	3,473					
27	1,703	2,052	2,771	3,690	57	1,672	2,002	2,665	3,470					
28	1,701	2,049	2,763	3,674	58	1,672	2,002	2,663	3,466					
29	1,699	2,045	2,756	3,659	59	1,671	2,001	2,662	3,463					
30	1,697	2,042	2,750	3,646	60	1,671	2,000	2,660	3,460					

f - число степеней свободы, **p** - доверительная вероятность

ТАБЛИЦА 2. Критические значения F-критерия Фишера для уровня значимости $\beta = 5\%$ и доверительной вероятности $p=0.95$.

f_y	1	2	3	4	5	f_x 6	7	8	12	24	∞
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.68	4.53	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.79	2.61	2.41
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.69	2.51	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.20	2.01	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.18	1.98	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.16	1.96	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.12	1.91	1.66
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	1.92	1.70	1.39
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	1.88	1.65	1.33
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.85	1.63	1.28
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.83	1.61	1.26
∞	3.84	3.00	2.61	2.37	2.22	2.10	2.01	1.94	1.75	1.52	1.00

f_x и f_y - число степеней свободы большей и меньшей несмещенных оценок выборочных дисперсий.

ТАБЛИЦА 3. Критические значения U-критерия Манна-Уитни для уровня значимости $\beta = 5\%$ и доверительной вероятности $p=0.95$.

n_x	n_y													
	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	12	15	17	20	23	26	28	30	34	37	39	42	45	48
10	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	16	19	23	26	30	33	37	40	44	48	51	55	58	62
12	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	22	26	30	36	40	45	50	55	59	64	67	74	78	83
15	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	26	31	37	42	48	53	59	64	70	75	81	86	92	98
17	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	34	41	48	55	62	69	76	83	90	98	105	112	119	127

n_x и **n_y** - объем выборок.

ЛИТЕРАТУРА

1. Гланц С. Медико-биологическая статистика. – М., 1998.
2. Леман.Э. Проверка статистических гипотез. - 2 изд., пер. с англ., М. 1979.
3. Большев Л. Н., Смирнов Н. В., Таблицы математической статистики, 2 изд., М., 1968.
4. Крамер Г. Математические методы статистики, 2 изд., пер. с англ., М., 1975.

Задачи на тему "Оценка статистической значимости различий средних и относительных величин при помощи t-критерия Стьюдента"

1. Было проведено исследование уровня лейкоцитов в периферической крови при остром аппендиците и остром аднексите. В группе больных острым аппендицитом ($n=60$) наблюдался лейкоцитоз $\bar{x}=10,1$, среднее квадратическое отклонение $\sigma_x=2,5$; в группе больных острым аднекситом ($n=85$) наблюдался лейкоцитоз $\bar{y}=9,5$, среднее квадратическое отклонение $\sigma_y=1,8$.

Сделайте вывод о диагностической значимости уровня лейкоцитоза для дифференцировки острого аппендицита и острого аднексита.

2. Учебной частью одной из кафедр медицинского университета было проведено исследование успеваемости студентов в зависимости от посещаемости лекций. Для студентов, посетивших менее половины лекционного курса ($n=36$), средняя оценка на экзамене составила 3,2, $\sigma_1=0,2$. Для студентов, посетивших более 90% лекций по предмету ($n=150$), средняя оценка на экзамене составила 4,5, $\sigma_2=0,5$.

Сделайте вывод о достоверности различий успеваемости студентов в зависимости от посещаемости лекций по предмету.

3. Одна из фармацевтических компаний провела исследование эффективности нового препарата из группы нестероидных противовоспалительных средств. Для этого была измерена температура тела у больных острым тонзиллитом до приема и через 30 минут после приема нового препарата. Объем выборки составил 400 человек. Температура тела до приема препарата составила: $x_1=39,2$, $\sigma_1=0,6$; а после приема: $x_2=37,6$, $\sigma_2=0,8$.

Сделайте вывод об эффективности нового препарата для купирования лихорадки.