

**Лекции по Приложна математика за специалност БИТ,
2009/2010**

Доц. д-р Снежана Гочева-Илиева

Лекция 8. ЕЛЕМЕНТИ НА СТАТИСТИКАТА

- 1. Първична обработка на данни**
- 2. Точкови и интервални оценки**
- 3. Изследване на зависимости – корелация**
- 4. Регресия с метода на най-малките квадрати**
- 5. Многомерен линеен регресионен анализ**

Първична обработка на данни

Вариационен ред е подреден по големина списък на данните от извадката.

Пример 1: Изходни данни:

12, 23, 14, 23, 11, 12, 25, 8, 9, 12, 11, 14, 25, 12, 9.

Вариационен ред:

8, 9, 9, 11, 11, 12, 12, 12, 12, 14, 14, 23, 23, 25, 25.

Статистическият ред или честотната таблица има вида:

x_i	x_1	x_2	x_3	...	x_k
n_i	n_1	n_2	n_3		n_k

където на първия ред се записват във възходящ ред само различните стойности на елементите от извадката, а във втория – тяхната честота, т.е. колко пъти се срещат в извадката. Общият брой данни (наблюдения) n се нарича **обем на извадката**.

Пример 2: От горните данни статистическият ред е:

x_i	8	9	11	12	14	23	25
n_i	1	2	2	4	2	2	2

Относителна честота на даден елемент от статистическия ред е: $v_i = \frac{n_i}{n}$.

x_i	x_1	x_2	x_3	...	x_k
$v_i = \frac{n_i}{n}$	$\frac{n_1}{n}$	$\frac{n_2}{n}$	$\frac{n_3}{n}$		$\frac{n_k}{n}$

Пример 3: На следващата таблица, в третия ред са дадени съответните честоти, брой данни $n=15$.

x_i	8	9	11	12	14	23	25
n_i	1	2	2	4	2	2	2
$v_i = \frac{n_i}{n}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{2}{15}$

Групиран статистически ред. Използва се, когато данните са много. Тогава вариационният ред се разделя на интервали, обикновено с равни дължини, като във всеки интервал трябва да има данни. Броят на подинтервалите k може да се намери от неравенството $2^k > n$. Определят се средите на получените интервали, да ги означим с m_i . Построява се таблица на групирания статистически ред, като на първия ред се записват стойностите на m_i , а на втория – честотите, т.е. броят на данните, които се намират в дадения подинтервал.

M_i	m_1	m_2	m_3	...	m_k
f_i	f_1	f_2	f_3	...	f_k

Пример 4. Нека необработените данни от наблюдение са:

68 71 77 83 79 72 74 57 67 69 50 60 70 66 76 70 84 59 75 94 65 72 85 79 71
83 84 74 82 97 77 73 78 93 95 78 81 79 90 83 80 84 91 101 86 93 92 102 80 69

Размерът на извадката е $n=50$. Приблизително $2^6 > 50$, затова избираме брой подинтервали $k=6$. Дължините на интервалите нека са равни : $h=(x_{\max}-x_{\min})/k=(102-50)/6=8,7 \approx 9$. По-добре е да вземем тук 10 вместо 9, т.е. $h=10$. Групираният статистически ред е даден в долната таблица.

интервал	45-55	55-65	65-75	75-85	85-95	95-105
m_i	50	60	70	80	90	100
f_i	1	4	15	19	8	3
v_i	$\frac{1}{50}$	$\frac{4}{50}$	$\frac{15}{50}$	$\frac{19}{50}$	$\frac{8}{50}$	$\frac{3}{50}$

Точкови и интервални оценки (характеристики на извадка)

I. $\bar{x}$ - **Средноаритметична стойност** на извадката. Тя е ефективна оценка на математическото очакване на генералната съвкупност (ГС):

За вариационен ред:
$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

За статистически ред:
$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n}$$

За групиран статистически ред:
$$\bar{x} = \frac{m_1 f_1 + m_2 f_2 + \dots + m_k f_k}{n}$$

II. $\bar{x}_w$ - **Средноаритметична стойност с тегло w_i за наблюдението x_i :**

$$\bar{x}_w = \frac{x_1 w_1 + x_2 w_2 + \dots + x_n w_n}{w_1 + w_2 + \dots + w_n}$$

III. **Md - Медиана на извадката.** Тя е оценка на медианата на ГС. Нейната стойност се явява среда на извадката, при което броят на елементите по-малки или равни на Md е равен на броя на елементите по-големи или равни на Md.

За вариационен и статистически ред: Md = средния елемент на извадката при нечетно n
Md = средно аритметично от двата средни елемента при четно n

За групирани данни:
$$Md = \left(m_{md} - \frac{h}{2}\right) + h \frac{\frac{n}{2} - \sum_{j=1}^{md-1} f_j}{f_{md}},$$

$\frac{n}{2}$ - определя мястото на медианната група $i=$

md , т.е. интервалът на медианата, където

натрупаната честота е $\geq \frac{n}{2}$.

h – ширина на медианната група със среда m_{md} ,

f_{md} – честота в медианната група,

$$\sum_{j=1}^{md-1} f_j - \text{натрупаната честота в предходната група.}$$

IV. M_o - Мода на извадката. Тя е най-често срещаното значение в извадката.

От статистическия ред: **$M_o = x_i$** , за което честотата n_i е най-голяма.

За групирани данни:

$$M_o = (m_{M_o} - \frac{h}{2}) + h \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})},$$

Индексът $i=M_o$ е мястото на модалната група,
т.е. интервалът, в който честотата f_i е най-голяма,

h – ширина на модалната група със среда m_{M_o} ,

f_{M_o} , f_{M_o-1} , f_{M_o+1} – честоти в модалната,
предишната и следващата групи.

V. s^2 - Дисперсия на извадката (извадкова дисперсия). Тя е ефективна оценка на дисперсията на ГС:

За вариационен ред:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

За статистически ред:

$$s^2 = \frac{1}{n_1 + n_2 + \dots + n_k - 1} \sum_{i=1}^k n_i (x_i - \bar{x})^2$$

За групирани статистически ред:

$$s^2 = \frac{1}{f_1 + f_2 + \dots + f_k - 1} \sum_{i=1}^k f_i (m_i - \bar{x})^2$$

VI. Стандартното отклонение на извадката е ефективна оценка на стандартното отклонение на ГС.

$$s = \sqrt{s^2}$$

Забележка. Примери за точкови и интервални оценки ще бъдат разгледани на упражнения.

ИЗСЛЕДВАНЕ НА ЗАВИСИМОСТИ

Корелация

Нека x е представителна извадка на ГС X , а y – представителна извадка на ГС Y .

Корелацията е вид евентуална зависимост между двете генерални съвкупности X и Y . За да установим дали има корелационна зависимост трябва да пресметнем и оценим корелационния коефициент ρ . Негово приближение е т.н. извадков коефициент на корелация r .

Формулите за пресмятане на извадковия коефициент на корелация r са:

$$SSx = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n},$$

$$SSy = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n},$$

$$SSxy = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n},$$

откъдето:

$$r = \frac{SSxy}{\sqrt{SSx} \sqrt{SSy}}.$$

Корелационният коефициент е число между -1 и 1, т.е. $-1 \leq r \leq 1$. Колкото по-голяма е неговата стойност по абсолютна стойност до 1, толкова по-силна е зависимостта между x и y . Когато $r > 0$, наличието на корелационна зависимост означава, че с нарастване на x зависимата променлива y също расте. Когато $r < 0$ с нарастване на x променливата y намалява.

Практическо правило: Ако $|r| > 0,85$, то може да се предполага, че между x и y има зависимост, но това не е достатъчно.

След определянето на корелационния коефициент r се прави проверка за неговата значимост. При предположение за нормално разпределени ГС:

$X \sim N(m_1, \sigma_1)$, $Y \sim N(m_2, \sigma_2)$ се разглеждат хипотезите:

$H_0: \rho = 0$, т.е. няма корелационна зависимост между X и Y ;

$H_1: \rho \neq 0$, т.е. има корелационна зависимост между X и Y .

За проверката на тези хипотези се прилага t -статистиката по разпределението на Стюдънт. Това става по следния начин:

1. Най-напред като се използва r се пресмята

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}.$$

2. След това се намира $t_{\text{критично}}$ - квантил на t -разпределението на Стюдънт с

$n-2$ степени на свобода и порядък $q = 1 - \frac{\alpha}{2}$, където α е избрано ниво на

значимост, напр. $\alpha = 5\%$. Намирането на $t_{\text{кр.}}$ може да стане по таблици или с функцията на Excel $TINV(\alpha; n-2)$. Степените на свобода се определят, като от размера на извадката n се извади броят на участващите променливи. В нашия случай имаме брой променливи равен на 2 (x и y).

3. Прилага се решаващото правило:

Решаващо правило: Ако $|t| > t_{\text{кр.}}$, то се отхвърля основната хипотеза H_0 и се приема, че X и Y са корелационно зависими. Ако $|t| < t_{\text{кр.}}$, то се приема, че X и Y са корелационно независими.

Пример 1. Дадени са данни от две извадки x и y , представители на две нормално разпределени ГС. Да се изследва има ли зависимост между тях и да се оцени значимостта на извадковия корелационен коефициент със степен на доверие $\alpha = 5\%$. При наличие на зависимост да се построи уравнението на линейна регресия по метода на най-малките квадрати.

x	5	4	7	10	11	8	9	11	8	9
y	30	27	38	48	59	54	42	63	52	47

Решение: Построяваме таблицата, изчисляваме сумите и заместваем във формулите за r .

i	x_i	y_i	x_i^2	y_i^2	x_i*y_i
1	5	30	25	900	150
2	4	27	16	729	108
3	7	38	49	1444	266
4	10	48	100	2304	480
5	11	59	121	3481	649
6	8	54	64	2916	432
7	9	42	81	1764	378
8	11	63	121	3969	693
9	8	52	64	2704	416
10	9	47	81	2209	423
Суми	82	460	722	22420	3995

По формулите: $n = 10$,

$$SSx = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 722 - \frac{82 \cdot 82}{10} = 49,6$$

$$SSy = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 22420 - \frac{460 \cdot 460}{10} = 1260,$$

$$SSxy = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} = 3995 - \frac{82 \cdot 460}{10} = 223,$$

$$r = \frac{SSxy}{\sqrt{SSx} \sqrt{SSy}} = \frac{223}{\sqrt{49,6} \sqrt{1260}} = 0,892029.$$

Тъй като $|r| > 0,85$ заключаваме, че има корелационна зависимост, но това не е достатъчно. Необходимо е да направим проверка за значимостта на получения извадков корелационен коефициент r . При предположение за нормално разпределени ГС: $X \sim N(m_1, \sigma_1)$, $Y \sim N(m_2, \sigma_2)$ пресмятаме t по формулите:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0,892029}{\sqrt{\frac{1-0,892029^2}{8}}} \approx 5,5822.$$

Така имаме: $r = 0,892029$, $t = 5,5822$

За да направим оценката на хипотезите от таблицата на t разпределението на Стюдънт при даденото ниво на значимост $\alpha = 5\%$, за $n-2=8$ степени на свобода и порядък на квантила $q = 1 - \frac{\alpha}{2} = 0,975$, намираме $t_{кр.} = 2,306$.

Тъй като $t > t_{кр.}$, заключаваме, че съществува корелационна зависимост между x и y .

Регресия с метода на най-малките квадрати

Ако с t -статистиката се установи наличието на зависимост, тя може да се определи явно като формула $y = \hat{y}(x)$ с различни методи за приближаване на функции. Един от най-използваните е методът на най-малките квадрати (МНМК). Това се нарича регресия, а графиката на получената зависимост – регресионна крива.

Линейна регресия. Когато графиката на данните прилича на отсечка, приближаваща функция се търси като полином от първа степен във вида $\hat{y}(x) = a_0 + a_1 x$. Коефициентите a_0 , a_1 се определят от линейната система

$$\begin{cases} na_0 + \sum_{i=1}^n x_i a_1 = \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i a_0 + \sum_{i=1}^n x_i^2 a_1 = \sum_{i=1}^n x_i y_i \end{cases}.$$

Квадратична регресия. Търси се приближаваща функция като полином от втора степен:

$\hat{y}(x) = a_0 + a_1 x + a_2 x^2$. Коефициентите a_0 , a_1 и a_2 се намират от линейната система

$$\begin{cases} na_0 + \sum_{i=1}^n x_i a_1 + \sum_{i=1}^n x_i^2 a_2 = \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i a_0 + \sum_{i=1}^n x_i^2 a_1 + \sum_{i=1}^n x_i^3 a_2 = \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i^2 a_0 + \sum_{i=1}^n x_i^3 a_1 + \sum_{i=1}^n x_i^4 a_2 = \sum_{i=1}^n x_i^2 y_i \end{cases} .$$

Формулата за грешката на приближенията по МНМК е в средноквадратичен смисъл:

$$R = \sqrt{\sum_{i=1}^n (y_i - \hat{y}(x_i))^2} .$$

Пример 2. По данните от Пример 1, (Корелация) да се определи уравнението на:
а) линейната регресия, б) квадратична регресия.

Решение:

а) Като използваме намерените суми

$$\sum_{i=1}^n x_i = 82, \quad \sum_{i=1}^n x_i^2 = 722, \quad \sum_{i=1}^n y_i = 460, \quad \sum_{i=1}^n x_i y_i = 3995 \text{ и заместим в системата за}$$

линейна регресия, получаваме системата:

$$\begin{cases} 10a_0 + 82a_1 = 460 \\ 82a_0 + 722a_1 = 3996 \end{cases} .$$

Можем да решим системата например по метода на Крамер. Пресмятаме детерминантите

$$\Delta = \begin{vmatrix} 10 & 82 \\ 82 & 722 \end{vmatrix} = 10 \cdot 722 - 82 \cdot 82 = 496 ,$$

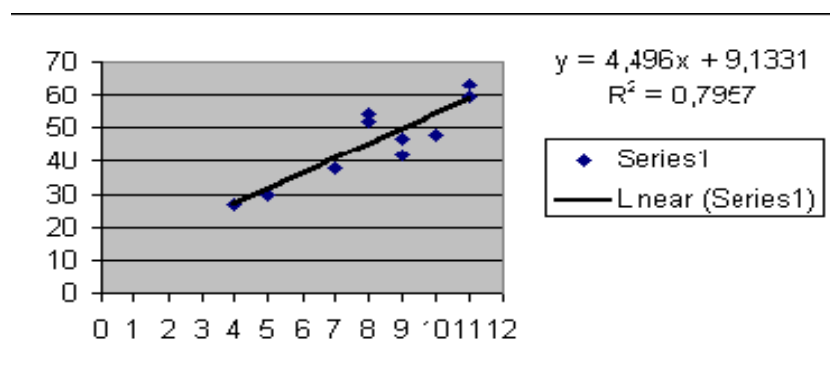
$$\Delta_1 = \begin{vmatrix} 460 & 82 \\ 3995 & 722 \end{vmatrix} = 460 \cdot 722 - 82 \cdot 3995 = 4530 ,$$

$$\Delta_2 = \begin{vmatrix} 10 & 460 \\ 82 & 3995 \end{vmatrix} = 10 \cdot 3995 - 82 \cdot 460 = 2230 .$$

Тогава $a_0 = \frac{\Delta_1}{\Delta} = \frac{4530}{496} = 9,1331$, $a_1 = \frac{\Delta_2}{\Delta} = \frac{2230}{496} = 4,4960$.

Уравнението на линейна регресия е: $\hat{y}(x) = 9,1331 + 4,4960x$.

На графиката е построена линейната зависимост по МНМК за данните от примера.



Б) За да намерим квадратичната крива на регресия пресмятаме още три стълба към таблицата:

	x_i^3	x_i^4	$x_i^2 y_i$
	125	625	750
	64	256	432
	343	2401	1862
	1000	10000	4800
	1331	14641	7139
	512	4096	3456
	729	6561	3402
	1331	14641	7623
	512	4096	3328
	729	6561	3807
суми	6676	63878	36599

Съставяме системата:

$$\begin{cases} 10a_0 + 82a_1 + 722a_2 = 460 \\ 82a_0 + 722a_1 + 6676a_2 = 3996 \\ 722a_0 + 6676a_1 + 63878a_2 = 36599 \end{cases}$$

Нейното решение е: $a_0=6,8924$, $a_1=5,1434$ и $a_2=-0,0425$. Следователно уравнението на приближаващата крива е:

$$\hat{y}(x) = 6,8924 + 5,1434x - 0,0425x^2.$$

Многомерна линейна регресия

Разгледаните случаи на регресия всъщност са частен случай на линейна регресия с една зависима и една независима променлива. Линейността се разглежда спрямо коефициентите на регресия, по-горе означени с $a_0, a_1, a_2 \dots$.

Накратко ще приведем необходимите теоретични постановки за многомерен РА.

Задачата е:

Нека са известни измервания (данни) за някаква променлива величина y , зависеща от p на брой независими величини $x_1, x_2, \dots, x_p$. Всяка променлива величина $y, x_1, x_2, \dots, x_p$ има по n на брой измерени стойности (наблюдения). Търсим линейна зависимост от вида

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p, \quad (1)$$

където $b_0, b_1, b_2 \dots, b_p$ са неизвестните коефициенти на регресия. Най-стандартният метод за тяхното определяне е методът на най-малките квадрати, разгледан в лекция 6 и в примерите по-горе. Получените стойности на коефициентите $b_0, b_1, b_2 \dots, b_p$ в (1) се наричат оценки на регресията, а $\hat{y}$ - предсказана стойност или оценка на зависимата изходна променлива y .

Разликите $\varepsilon_i = \hat{y}_i - y_i$, $i = 1, 2, \dots, n$ са грешките (остатъци, отклонения) на регресията.

За прилагането както на едномерен, така и многомерен РА се изисква изпълнението на следните основни условия:

- Изходните данни трябва да имат случаен характер, да бъдат от интервален или относителен тип. Категорийните данни трябва да бъдат предварително трансформирани в двоични.
- Наблюденията трябва да бъдат независими.
- За всяка независима променлива, зависимата променлива трябва да има нормално разпределение или близко до нормалното.
- При линейна регресия се предполага, че общата зависимост има линеен или близък до линеен характер.

Съществуват голям брой статистически тестове за валидност на регресионния модел. Когато те са изпълнени, регресионният модел се счита за статистически значим и пригоден за практическо използване и прогнози. Ще приведем само най-основните изисквания за валидност на РА (при ниво 95%):

- Нивото на значимост на целия модел, оценяван с F статистика (в таблица ANOVA) трябва да бъде < 0.05
- Коефициентът на детерминация R^2 и коефициентът на многомерна корелация R трябва да са близки до 1 и желателно по-големи от 0.5
- Изследването за мултиколинеарност между независимите променливи да е приемливо (за SPSS тестът е $VIF < 10$)
- Да се направи анализ на остатъците от регресията: разпределението на остатъците да е нормално (напр. с хистограма, P-P, Q-Q тест, тест на Колмогоров-Смирнов, тест на Шапиро-Уилкс), разстоянията на Махалобис и Кук да са малки, да се изследват отдалечените точки и др.
- Препоръчва се и провеждане на процедура на крос-валидация. Тя се състои в разделяне на дадената извадка на две непресичащи се подмножества, наречени тренировъчно и тестово. С тренировъчното множество се построява и изследва нов модел. С този модел се предсказват

стойностите на зависимата променлива за наблюденията от тестовото множество. Ако резултатите са близки до модела с тренировъчното подмножество и с изходния модел, се приема че изходния модел е валиден.

Подробно решени примери с SPSS са разгледани по-нататък в учебния курс, както и на упражнения.

ТАБЛИЦА НА НОРМАЛНОТО СТАНДАРТНО РАЗПРЕДЕЛЕНИЕ N(0,1)

За $x > 0$: $\Phi(-x) = 1 - \Phi(x)$, $P\{0 < \xi < x\} = P\{-x < \xi < 0\} = \Phi(x) - 0,5$ $P\{-x < \xi < x\} = 2\Phi(x) - 1$
 $P\{\alpha < \xi < \beta\} = \Phi(\beta) - \Phi(\alpha)$

Нормиране на произволно нормално разпределение $N(m, \sigma)$:

$F(x) = \Phi((x-m)/\sigma)$

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,50000	0,50	0,69146	1,00	0,84134	1,50	0,9332	2,00	0,97725	2,50	0,9938	3,00	0,9987
0,01	0,50399	0,51	0,69497	1,01	0,84375	1,51	0,93448	2,01	0,97778	2,51	0,99396	3,01	0,99869
0,02	0,50798	0,52	0,69847	1,02	0,84614	1,52	0,93574	2,02	0,97831	2,52	0,99413	3,02	0,99874
0,03	0,51197	0,53	0,70194	1,03	0,84849	1,53	0,93699	2,03	0,97882	2,53	0,99430	3,03	0,99878
0,04	0,51595	0,54	0,70540	1,04	0,85083	1,54	0,93822	2,04	0,97932	2,54	0,99446	3,04	0,99882
0,05	0,51994	0,55	0,70884	1,05	0,85314	1,55	0,93943	2,05	0,97982	2,55	0,99461	3,05	0,99886
0,06	0,52392	0,56	0,71226	1,06	0,85543	1,56	0,94062	2,06	0,98030	2,56	0,99477	3,06	0,99889
0,07	0,52790	0,57	0,71566	1,07	0,85769	1,57	0,94179	2,07	0,98077	2,57	0,99492	3,07	0,99893
0,08	0,53188	0,58	0,71904	1,08	0,85993	1,58	0,94295	2,08	0,98124	2,58	0,99506	3,08	0,99896
0,09	0,53586	0,59	0,72240	1,09	0,86214	1,59	0,94408	2,09	0,98169	2,59	0,99520	3,09	0,99900
0,10	0,53983	0,60	0,72575	1,10	0,86433	1,60	0,94520	2,10	0,98214	2,60	0,99534	3,10	0,99903
0,11	0,54380	0,61	0,72907	1,11	0,86650	1,61	0,94630	2,11	0,98257	2,61	0,99547	3,11	0,99906
0,12	0,54776	0,62	0,73237	1,12	0,86864	1,62	0,94738	2,12	0,98300	2,62	0,99560	3,12	0,99910
0,13	0,55172	0,63	0,73565	1,13	0,87076	1,63	0,94845	2,13	0,98341	2,63	0,99573	3,13	0,99913
0,14	0,55567	0,64	0,73891	1,14	0,87286	1,64	0,94950	2,14	0,98382	2,64	0,99585	3,14	0,99916
0,15	0,55962	0,65	0,74215	1,15	0,87493	1,65	0,95053	2,15	0,98422	2,65	0,99598	3,15	0,99918
0,16	0,56356	0,66	0,74537	1,16	0,87698	1,66	0,95154	2,16	0,98461	2,66	0,99609	3,16	0,99921
0,17	0,56749	0,67	0,74857	1,17	0,87900	1,67	0,95254	2,17	0,98500	2,67	0,99621	3,17	0,99924
0,18	0,57142	0,68	0,75175	1,18	0,88100	1,68	0,95352	2,18	0,98537	2,68	0,99632	3,18	0,99926
0,19	0,57535	0,69	0,75490	1,19	0,88298	1,69	0,95449	2,19	0,98574	2,69	0,99643	3,19	0,99929
0,20	0,57926	0,70	0,75804	1,20	0,88493	1,70	0,95543	2,20	0,98610	2,70	0,99653	3,20	0,99931
0,21	0,58317	0,71	0,76115	1,21	0,88686	1,71	0,95637	2,21	0,98645	2,71	0,99664	3,21	0,99934
0,22	0,58706	0,72	0,76424	1,22	0,88877	1,72	0,95728	2,22	0,98679	2,72	0,99674	3,22	0,99936
0,23	0,59095	0,73	0,76730	1,23	0,89065	1,73	0,95818	2,23	0,98713	2,73	0,99683	3,23	0,99938
0,24	0,59483	0,74	0,77035	1,24	0,89251	1,74	0,95907	2,24	0,98745	2,74	0,99693	3,24	0,99940
0,25	0,59871	0,75	0,77337	1,25	0,89435	1,75	0,95994	2,25	0,98778	2,75	0,99702	3,25	0,99942
0,26	0,60257	0,76	0,77637	1,26	0,89617	1,76	0,96080	2,26	0,98809	2,76	0,99711	3,26	0,99944
0,27	0,60642	0,77	0,77935	1,27	0,89796	1,77	0,96164	2,27	0,98840	2,77	0,99720	3,27	0,99946
0,28	0,61026	0,78	0,78230	1,28	0,89973	1,78	0,96246	2,28	0,98870	2,78	0,99728	3,28	0,99948

0,29	0,61409	0,79	0,78524	1,29	0,90147	1,79	0,96327	2,29	0,98899	2,79	0,99736	3,29	0,99950
0,30	0,61791	0,80	0,78814	1,30	0,90320	1,80	0,96407	2,30	0,98928	2,80	0,99744	3,30	0,99952
0,31	0,62172	0,81	0,79103	1,31	0,90490	1,81	0,96485	2,31	0,98956	2,81	0,99752	3,31	0,99953
0,32	0,62552	0,82	0,79389	1,32	0,90658	1,82	0,96562	2,32	0,98983	2,82	0,99760	3,32	0,99955
0,33	0,62930	0,83	0,79673	1,33	0,90824	1,83	0,96638	2,33	0,99010	2,83	0,99767	3,33	0,99957
0,34	0,63307	0,84	0,79955	1,34	0,90988	1,84	0,96712	2,34	0,99036	2,84	0,99774	3,34	0,99958
0,35	0,63683	0,85	0,80234	1,35	0,91149	1,85	0,96784	2,35	0,99061	2,85	0,99781	3,35	0,99960
0,36	0,64058	0,86	0,80511	1,36	0,91308	1,86	0,96856	2,36	0,99086	2,86	0,99788	3,36	0,99961
0,37	0,64431	0,87	0,80785	1,37	0,91466	1,87	0,96926	2,37	0,99111	2,87	0,99795	3,37	0,99962
0,38	0,64803	0,88	0,81057	1,38	0,91621	1,88	0,96995	2,38	0,99134	2,88	0,99801	3,38	0,99964
0,39	0,65173	0,89	0,81327	1,39	0,91774	1,89	0,97062	2,39	0,99158	2,89	0,99807	3,39	0,99965
0,40	0,65542	0,90	0,81594	1,40	0,91924	1,90	0,97128	2,40	0,99180	2,90	0,99813	3,40	0,99966
0,41	0,65910	0,91	0,81859	1,41	0,92073	1,91	0,97193	2,41	0,99202	2,91	0,99819	3,41	0,99968
0,42	0,66276	0,92	0,82121	1,42	0,92220	1,92	0,97257	2,42	0,99224	2,92	0,99825	3,42	0,99969
0,43	0,66640	0,93	0,82381	1,43	0,92364	1,93	0,97320	2,43	0,99245	2,93	0,99831	3,43	0,99970
0,44	0,67003	0,94	0,82639	1,44	0,92507	1,94	0,97381	2,44	0,99266	2,94	0,99836	3,44	0,99971
0,45	0,67364	0,95	0,82894	1,45	0,92647	1,95	0,97441	2,45	0,99286	2,95	0,99841	3,45	0,99972
0,46	0,67724	0,96	0,83147	1,46	0,92785	1,96	0,97500	2,46	0,99305	2,96	0,99846	3,46	0,99973
0,47	0,68082	0,97	0,83398	1,47	0,92922	1,97	0,97558	2,47	0,99324	2,97	0,99851	3,47	0,99974
0,48	0,68439	0,98	0,83646	1,48	0,93056	1,98	0,97615	2,48	0,99343	2,98	0,99856	3,48	0,99975
0,49	0,68793	0,99	0,83891	1,49	0,93189	1,99	0,97670	2,49	0,99361	2,99	0,99861	3,49	0,99976

Таблица на t-разпределение на Стюдънт

двустранно

α	0,01	0,02	0,05	0,10
q	0,99	0,98	0,95	0,90

n - степени на свобода

3	5,8408	4,5407	3,1824	2,3534
4	4,6041	3,7469	2,7765	2,1318
5	4,0321	3,3649	2,5706	2,0150
6	3,7074	3,1427	2,4469	1,9432
7	3,4995	2,9979	2,3646	1,8946
8	3,3554	2,8965	2,3060	1,8595
9	3,2498	2,8214	2,2622	1,8331
10	3,1693	2,7638	2,2281	1,8125
11	3,1058	2,7181	2,2010	1,7959
12	3,0545	2,6810	2,1788	1,7823
13	3,0123	2,6503	2,1604	1,7709
14	2,9768	2,6245	2,1448	1,7613
15	2,9467	2,6025	2,1315	1,7531
16	2,9208	2,5835	2,1199	1,7459
17	2,8982	2,5669	2,1098	1,7396
18	2,8784	2,5524	2,1009	1,7341
19	2,8609	2,5395	2,0930	1,7291
20	2,8453	2,5280	2,0860	1,7247
21	2,8314	2,5176	2,0796	1,7207
22	2,8188	2,5083	2,0739	1,7171
23	2,8073	2,4999	2,0687	1,7139
24	2,7970	2,4922	2,0639	1,7109

25	2,7874	2,4851	2,0595	1,7081
26	2,7787	2,4786	2,0555	1,7056
27	2,7707	2,4727	2,0518	1,7033
28	2,7633	2,4671	2,0484	1,7011
29	2,7564	2,4620	2,0452	1,6991
30	2,7500	2,4573	2,0423	1,6973
40	2,7045	2,4233	2,0211	1,6839
60	2,6603	2,3901	2,0003	1,6706
120	2,6174	2,3578	1,9799	1,6576