МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
АВТОМОБИЛЬНО-ДОРОЖНЫЙ ИНСТИТУТ

«УТВЕРЖДАЮ »
ДИРЕКТОР АДИ ГОУВПО «ДонНТУ»
М. Н. ЧАЛЬЦЕВ
21.07.2017

Кафедра «Высшая математика»

УЧЕБНО–МЕТОДИЧЕСКОЕ ПОСОБИЕ К ВЫПОЛНЕНИЮ КОНТРОЛЬНЫХ РАБОТ ПО ДИСЦИПЛИНЕ «МАТЕМАТИЧЕСКАЯ СТАТИСТИКА» ДЛЯ СТУДЕНТОВ НАПРАВЛЕНИЙ ПОДГОТОВКИ: 23.03.01 «ТЕХНОЛОГИЯ ТРАНСПОРТНЫХ ПРОЦЕССОВ», 27.03.04 «УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ», 23.05.01 «НАЗЕМНЫЕ ТРАНСПОРТНО-ТЕХНОЛОГИЧЕСКИЕ СРЕДСТВА», 23.03.03 «ЭКСПЛУАТАЦИЯ ТРАНСПОРТНО-ТЕХНОЛОГИЧЕСКИХ МАШИН И КОМПЛЕКСОВ», 08.03.01 «СТРОИТЕЛЬСТВО», 20.03.01 «ТЕХНОСФЕРНАЯ БЕЗОПАСНОСТЬ», 08.05.02 «СТРОИТЕЛЬСТВО, ЭКСПЛУАТАЦИЯ, ВОССТАНОВЛЕНИЕ И ТЕХНИЧЕСКОЕ ПРИКРЫТИЕ АВТОМОБИЛЬНЫХ ДОРОГ, МОСТОВ И ТОННЕЛЕЙ», 38.03.02 «МЕНЕДЖМЕНТ» (ВСЕХ ФОРМ ОБУЧЕНИЯ)

2/51-2017-01

«РЕКОМЕНДОВАНО»
Учебно-методическая
комиссия факультета
«Транспортные технологии»
Протокол № 1 от 20.01.2017

«РЕКОМЕНДОВАНО»
Учебно-методическая
комиссия факультета
«Автомобильный транспорт»
Протокол № 2 от 16.11.2016

«РЕКОМЕНДОВАНО»
Кафедра «Высшая математика»
Протокол № 4 от 14.01.2016

«РЕКОМЕНДОВАНО»
Учебно-методическая
комиссия факультета
«Экономика и управление»
Протокол № 4 от 24.12.2016

«РЕКОМЕНДОВАНО»
Учебно-методическая
комиссия факультета
«Автомобильные дороги»
Протокол № 5 от 18.01.2017

Горловка – 2017

УДК 519.22

Учебно–методическое пособие к выполнению контрольных работ по дисциплине «Математическая статистика» для студентов направлений подготовки: 23.03.01 «Технология транспортных процессов», 27.03.04 «Управление в технических системах», 23.05.01 «Наземные транспортно-технологические средства», 23.03.03 «Эксплуатация транспортно-технологических машин и комплексов», 08.03.01 «Строительство», 20.03.01 «Техносферная безопасность», 08.05.02 «Строительство, эксплуатация, восстановление и техническое прикрытие автомобильных дорог, мостов и тоннелей», 38.03.02 «Менеджмент» (всех форм обучения), [Электронный ресурс] / составители: Л. П. Вовк, Е. А. Королев. – Горловка: ГОУВПО «ДонНТУ» АДИ, 2017.

Учебно-методическое указание отвечает структуре дисциплины «Высшая математика», приведены примеры решений типовых задач и индивидуальные задания по теме «Математическая статистика» .

Составители: Вовк Л. П, д-р. техн. наук, проф.
Королев Е. А., канд.физ.-мат.наук, доц.

Ответственный за выпуск: Королев Е. А., канд.физ.-мат.наук, доц.

Рецензент: Королев М. Е., канд.физ.-мат.наук, доц.

ОГЛАВЛЕНИЕ

1 Общие замечания

При изучении курса высшей математики наибольшие трудности возникают у студентов при решении практических задач. Решение и анализ задач позволяют понять и запомнить основные теоремы и формулы данного курса. Умение решать задачи – лучший критерий оценки глубины изучения программного материала.

Данное учебно-методическое пособие относится к теме «Математическая статистика». Для развития навыков самостоятельного решения практических задач предложены индивидуальные задания по основным разделам темы. Цель заданий – закрепить знание основных теорем и формул, освоить приемы решения задач.

Каждый студент выполняет пять задач, условия которых приведены в разделе 3 настоящей работы. Выбор номеров задач необходимо производить согласно варианту, соответствующему номеру зачетной книжки. Если студент имеет номер зачетной книжки 160987183, то для выбора номеров задач следует принять во внимание лишь три последние цифры и расположить под ними три первые буквы русского алфавита, например,

1	8	3
а	б	в

Затем нужно обратиться к помещенной ниже таблица 1 и из каждого ее вертикального столбца, обозначенного внизу определенной буквой, взять одно число, стоящее в той горизонтальной строке, номер которой соответствует этой букве или разности |б-в|.

Таблица 1

Номер строки	Номера задач контрольных заданий				
1	2	12	25	35	43
2	0	13	21	33	44
3	3	11	26	36	45
4	8	18	24	34	41
5	6	17	30	37	42
6	7	19	22	32	50
7	9	14	27	38	46
8	5	20	28	31	49
9	1	15	23	39	47
0	4	16	29	40	48
	в	|б-в|	б	в	|б-в|

Например, номера задач, которые нужно решить для варианта 160987183, следующие:

3 17 28 36

2 Элементы математической статистики

2.1 Точечные оценки параметров распределения

Пусть для изучения количественного (дискретного или непрерывного) признака X из генеральной совокупности извлечена выборка $x_1, x_2, ...x_k$ объема *n*. Наблюдавшиеся значения x_i признака *X* называют *вариантами*, а последовательность вариант, записанных в возрастающем порядке, – *вариационным рядом*.

Статистическим распределением выборки называют перечень вариант x_i вариационного ряда и соответствующих им частот n_i (сумма всех частот равна объему выборки *n*) или относительных частот ω_i (сумма всех относительных частот равна единице).

Статистическое распределение выборки можно задать также в виде последовательности интервалов и соответствующих им частот (в качестве частоты интервала принимают сумму частот вариант, попавших в этот интервал).

Эмпирической функцией распределения (функцией распределения выборки) называют функцию $F^*(x)$, определяющую для каждого значения *x* относительную частоту события $X < x$,

$$F^*(x) = \frac{n_x}{n}, \tag{1}$$

где n_x – число вариант, меньших *x*;

n – объем выборки.

Эмпирическая функция обладает следующими свойствами:

Свойство 1

Значения эмпирической функции принадлежат отрезку [0;1].

Свойство 2

$F^*(x)$ – неубывающая функция.

Свойство 3

Если x_1 – наименьшая варианта, а x_k – наибольшая, то $F^*(x) = 0$ при $x \le x_1$ и $F^*(x) = 1$ при $x > x_k$.

Полигоном частот называют ломаную, отрезки которой соединяют точки $(x_1,n_1),(x_2,n_2),...,(x_k,n_k)$, где x_i – варианты выборки и n_i – соответствующие им частоты.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки $(x_1,\omega_i),(x_2,\omega_2),...,(x_k,\omega_k)$,
где x_i – варианты выборки;
ω_i – соответствующие им относительные частоты.

При непрерывном распределении признака весь интервал, в котором заключены все наблюденные значения признака, разбивают на ряд частичных интервалов длины h и находят n_i – сумму частот вариант, попавших в интервал.

Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников основаниями которых служат частичные интервалы длины h, а высоты равны отношению $\frac{n_i}{h}$ (плотность частоты). Площадь частичного i-го прямоугольника равна $h\cdot\frac{n_i}{h}=n_i$ – сумме частот вариант, попавших в i-й интервал. Площадь гистограммы частот равна сумме всех частот, т. е. объему выборки n.

Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников основаниями которых служат частичные интервалы длины h, а высоты равны отношению $\frac{\omega_i}{h}$ (плотность относительной частоты). Площадь частичного i-го прямоугольника равна $h\cdot\frac{\omega_i}{h}=\omega_i$ – относительной частоте вариант, попавших в i-й интервал. Площадь гистограммы относительных частот равна сумме всех относительных частот, т. е. единице.

Точечной называют статистическую оценку, которая определяется одним числом.

Несмещенной называют точечную оценку, математическое ожидание которой равно оцениваемому параметру при любом объеме выборки.

Смещенной называют точечную оценку, математическое ожидание которой равно оцениваемому параметру.

Несмещенной оценкой генеральной средней (математического ожидания) служит выборочная средняя

$$\bar{X}_B=\frac{\sum_{i=1}^{k}n_i x_i}{n}, \qquad (2)$$

где x_i – варианта выборки,

n_i – частота варианты n_i ,

$n = \sum_{i=1}^{k} n_i$ – объем выборки.

Если первоначальные варианты x_i – большие числа, то для упрощения расчета целесообразно вычесть из каждой варианты одно и то же число C, т. е. перейти к условным вариантам $u_i = x_i - C$ (в качестве С выгодно принять число, близкое к выборочной средней; поскольку выборочная средняя неизвестна, число С выбирают «на глаз»). Тогда

$$\overline{x}_B = C + \frac{\sum n_i u_i}{n}. \tag{3}$$

<u>*Смещенной оценкой генеральной дисперсии*</u> служит выборочная средняя

$$D_B = \frac{\sum_{i=1}^{k} (x_i - \overline{x}_B)^2}{n}. \tag{4}$$

Эта оценка является смещенной, так как

$$M[D_B] = \frac{n-1}{n} D_\Gamma. \tag{5}$$

Более удобна формула

$$D_B = \overline{x}^2 - [\overline{x}]^2 = \frac{\sum n_i x_i^2}{n} - [\frac{\sum n_i x_i}{n}]^2. \tag{6}$$

Если первоначальные варианты x_i – большие числа, то целесообразно вычесть из всех вариант одно и то же число C, равное выборочной средней или близкое к ней, т. е. перейти к условным вариантам $u_i = x_i - C$ (дисперсия при этом не изменится). Тогда

$$D_B(X) = D_B(u) = \overline{u}^2 - [\overline{u}]^2 = \frac{\sum n_i u_i^2}{n} - [\frac{\sum n_i u_i}{n}]^2. \tag{7}$$

Если первоначальные варианты являются десятичными дробями с k десятичными знаками после запятой, то чтобы избежать действий с дробями, умножают первоначальные варианты на постоянное число $C = 10^k$, т. е. переходят к условным вариантам $u_i = Cx_i$. При этом дисперсия увеличивается в C^2 раз. Поэтому, найдя дисперсию в условных вариантах, надо разделить ее на C^2:

$$D_B(X) = \frac{D_{B_u}}{C^2}. \qquad (8)$$

<u>Несмещенной оценкой генеральной дисперсии</u> служит исправленная выборочная дисперсия

$$S^2 = \frac{n}{n-1} \cdot D_B = \frac{\sum n_i (x_i - x_B)^2}{n-1}. \qquad (9)$$

Более удобна формула

$$S_x^2 = \frac{\sum n_i x_i^2 - \frac{[\sum n_i x_i\,]^2}{n}}{n-1}. \qquad (10)$$

В условных вариантах она имеет вид

$$S_u^2 = \frac{\sum n_i u_i^2 - \frac{[\sum n_i u_i\,]^2}{n}}{n-1}. \qquad (11)$$

Причем, если $u_i = x_i - C$, то $S_X^2 = s_u^2$; если $u_i = C x_i$, то $S_X^2 = \frac{S_u^2}{C^2}$.

ЗАДАЧА № 1

Найти оценки числовых характеристик генеральной совокупности по данным выборки объема $n = 50$

x_i	1	2	3	4
n_i	20	15	10	5

Решение. Находим среднюю выборочную по формуле (2)

$$\bar{x}_B = \frac{1 \cdot 20 + 2 \cdot 15 + 3 \cdot 10 + 4 \cdot 5}{50} = 2.$$

Смещенная оценка дисперсии генеральной по (3)

$$D_B = 5 - 4 = 1.$$

Исправленная дисперсия по (9) равна

$$S^2 = \frac{50-1}{50}(5 - 2^2) = 0,98.$$

Несмещенная оценка дисперсии генеральной

$$S^2 = 0,98.$$

Несмещенная оценка математического ожидания генеральной совокупности

$$\overline{x}_B = 2.$$

При большом числе данных используется метод произведений вычисления выборочных средней и дисперсии, который заключается в следующем.

Пусть выборка задана в виде распределения равноотстоящих вариант и соответствующих им частот. В этом случае удобно находить выборочные среднюю и дисперсию методом произведений по формулам:

$$\overline{X}_B = M_1^* h + C; \quad D_B = [M_2^* - (M_1^*)^2]h^2, \tag{12}$$

где h – шаг (разность между двумя соседними вариантами);
C – ложный нуль (варианта, которая имеет наибольшую частоту);
$u_i = \dfrac{x_i - C}{h}$ – условная варианта;
$M_1^* \dfrac{\sum n_i u_i}{n}$ – условный момент первого порядка;
$M_2^* \dfrac{\sum n_i u_i^2}{n}$ – условный момент второго порядка.

ЗАДАЧА № 2

Найти методом произведений выборочную среднюю и выборочную дисперсию по заданному распределению выборки объема $n = 100$:

Вариант x_i	12	14	16	18	20	22
Частота n_i	5	15	50	16	10	4

Решение. Составим расчетную таблицу 2. Для этого:

1) запишем варианты в первый столбец;

2) запишем частоты во второй столбец; сумму частот (100) поместим в нижнюю клетку столбца;

3) в качестве ложного нуля C выберем варианту (16), которая имеет наибольшую частоту (в качестве C можно взять любую варианту, расположенную примерно в середине столбца); в клетке третьего столбца, которая принадлежит строке, содержащей ложный нуль, пишем O; над нулем последовательно записываем –1,–2, а под нулем 1, 2, 3;

4) произведения частот n_i на условные варианты u_i запишем в четвертый столбец; отдельно находим сумму (–25) отрицательных чисели отдельно сумму (48) положительных чисел; сложив эти числа, их сумму (23) помещаем в нижнюю клетку четвертого столбца;

5) произведения частот на квадраты условных вариант, т. е. $n_i u_i^2$ запишем в пятый столбец (удобнее перемножить числа каждой строки третьего и четвертого столбцов: $u_i \cdot n_i u_i = n_i u_i^2$) сумму чисел столбца (127) помещаем в нижнюю клетку пятого столбца;

6) произведения частот на квадраты условных вариант, увеличенных на единицу, т. е. $n_i(u_i+1)^1$, запишем в шестой контрольный столбец; сумму чисел столбца (273) помещаем в нижнюю клетку шестого столбца.

В итоге получим расчетную таблицу 2.1.

Для контроля вычислений пользуются тождеством

$$\sum n_i(u_i+1)^2 = \sum n_i u_i^2 + 2\sum n_i u_i^2 + n.$$

Контроль:

$$\sum n_i(u_i+1)^2 = 237,$$

$$\sum n_i u_i^2 + 2\sum n_i u_i^2 + n = 12 + 2\cdot 23 + 100 = 237.$$

Совпадение контрольных сумм свидетельствует о правильности вычислений. Вычислим условные моменты первого и второго порядков:

$$M_1^* \frac{\sum n_i u_i}{n} = \frac{23}{100} = 0,23,$$

$$M_2^* \frac{\sum n_i u_i^2}{n} = \frac{127}{100} = 1,27.$$

Найдем шаг (разность между любыми двумя соседними вариантами): $h = 14 - 12 = 2$.

Вычислим искомые выборочные среднюю и дисперсию, учитывая, что ложный нуль (варианта, которая имеет наибольшую частоту) $C = 16$:

$$\overline{x}_B = M_1^* h + C = 0,23\cdot 2 + 16 = 16,46,$$

$$D_B = [M_2^* - (M_1^*)^2]h^2 = [1,27 - 0,23^2]\cdot 2^2 = 4,87.$$

Таблица 2

1	2	3	4	5	6
$x,$	n_i	u_i	$n_i u_i$	$n_i u_i^2$	$n_i(u_i+1)^2$
12	5	–2	–10	20	5
14	15	–1	–15	15	--
16	50	0	–25	–	50
18	16	1	16	16	64
20	10	2	20	40	90
22	4	3	12	36	64
			48		
	$n = 100$		$\sum n_i u_i$	$\sum n_i u_i^2 = 127$	$\sum n_i(u_i+1)^2 = 237$

2.2 Интервальные оценки параметров распределения

Оценку называют интервальной, если она определяется двумя числами – концами интервала.

Интервал (α_1, α_2) называют доверительным, если внутри него с заданной вероятностью (надежностью), близкой к единице, находится оцениваемый параметр.

Надежностью оценки параметра θ называют вероятность γ, с которой осуществляется неравенство $\alpha_1 < \theta < \alpha_2$. Для оценки вероятности этого неравенства в общем случае служит формула

$$P(\alpha_1 < \theta < \alpha_2) = \int_{\alpha_1}^{\alpha 2} f(x)dx, \qquad (16)$$

где $f(x)$ – дифференциальная функция распределения величины.

Для оценки математического ожидания a нормально распределенного количества признака X при известной дисперсии служит доверительный интервал

$$\bar{x}_B - t\frac{\sigma}{\sqrt{n}} < a < \bar{x}_B + t\frac{\sigma}{\sqrt{n}}, \qquad (17)$$

где σ – известное среднеквадратичное отклонение;

n – объем выборки.

Вероятность того, что неизвестное математическое ожидание заключено в указанном интервале, определяется так:

$$P(|X - a| < t\cdot\frac{\sigma}{\sqrt{n}}) = 2\Phi(t) = \gamma,$$

где $\Phi(t)$ – интегральная функция Лапласа.

Из этого равенства находим параметр t по таблице А.1 (приложение А) Если дисперсия генеральной совокупности неизвестна, то при вычислении интервальной оценки для математического ожидания a с заданной надежностью γ используют формулу

$$P\left(\left|\frac{\bar{X} - a}{S}\sqrt{n}\right| < t_\gamma\right| = \gamma. \qquad (18)$$

Доверительный интервал для неизвестного математического ожидания a в этом случае имеет вид

$$\bar{x} - t_\gamma\frac{S}{\sqrt{n}} < a < \bar{x} + t_\gamma\frac{S}{\sqrt{n}}, \qquad (19)$$

где

$$\frac{S}{\sqrt{n}} = \sqrt{\frac{D_B}{n-1}}.$$

ЗАДАЧА № 3

Найти доверительный интервал для оценки с надежностью 0,95 неизвестного математического ожидания a нормально распределенного признака X генеральной совокупности, если даны генеральное среднее квадратическое отклонение σ = 5, выборочная средняя $\bar{x}_B$ = 14 и объем выборки n = 25.

Решение. Требуется найти доверительный интервал по (17)

$$\bar{x}_B - t\frac{\sigma}{\sqrt{n}} < a < \bar{x}_B + t\frac{\sigma}{\sqrt{n}}.$$

Здесь все величины, кроме t, известны. Найдем t. Из соотношения $2\Phi(t) = 0{,}95$ получим $\Phi(t) = 0{,}475$. По таблице А.1 (приложение А) находим $t = 1{,}96$. Подставив $t = 1{,}96$, $\bar{x}_B = 14$, σ = 5, n = 25 в (17), окончательно получим искомый доверительный интервал.

$$12{,}04 < a < 15{,}96.$$

ЗАДАЧА № 4

В результате наблюдения за 115 автомобилями ЗИЛ-130 был получен средний пробег до первого капитального ремонта рулевого механизма 145,0 тыс. км при $\sigma = 37{,}5$ тыс. км. Какова точность полученного результата при надежности, равной 0,95.

Решение. Считая, что величина пробега до первого капитального ремонта рулевого механизма распределена по нормальному закону, точность результата можно получить с помощью доверительного интервала:

$$\bar{x}_B - t\frac{\sigma}{\sqrt{n}} < a < \bar{x}_B + t\frac{\sigma}{\sqrt{n}}.$$

По условию задачи γ =0,95. Тогда $\Phi(t) = 0{,}475$, а $t = 1{,}96$. При построении доверительного интервала получили:

$$\sigma = \frac{1{,}96 \cdot 37{,}6}{\sqrt{115}} \approx 6{,}88.$$

Отсюда в 95 случаях из 100 для автомобилей ЗИЛ-130 средний пробег до первого капитального ремонта рулевого механизма составляет $145{,}0 \pm 6{,}9$ тыс. км. Абсолютная погрешность 6,9 тыс. км, относительная погрешность результата

$$\Delta = \frac{6,88}{145,0} \cdot 100\% = 4\,\%.$$

ЗАДАЧА № 5

Измерения твердости 16 образцов легированной стали (в условных единицах) дали следующие результаты: 13,1; 12,8; 11,9; 12,4; 13,5; 13,7; 12,0; 13,8; 10,6; 12,4; 13,5; 11,7; 13,9; 11,5; 12,5; 11,9. В предположении, что выборка измерений получена из нормально распределенной генеральной совокупности, найти доверительные интервалы для среднего и дисперсии при доверительной вероятности $\gamma = 0,95$.

Решение. По выборке найдем оценки среднего, дисперсии и среднеквадратичного отклонения:

$$\overline{x}_B = 15,57; S^2 = 0,91 \text{ или } S{=}0,95.$$

По таблице Б.1 (приложение Б) находим $t_\gamma = 2,131$, при n = 16 и $\gamma = 0,95$. Тогда доверительный интервал для математического ожидания

$$\overline{x} - t_\gamma \frac{\text{S}}{\sqrt{n}} < a < \overline{x} + t_\gamma \frac{\text{S}}{\sqrt{n}}$$

$$12,57 - 2,131 \cdot \frac{0,95}{16} < a < 12,57 + 2,131 \frac{0,95}{16}$$

$$12,06 < a < 13,08.$$

2.3 Элементы теории линейной корреляции

Если обе линии регрессии *Y* на *X* и *X* на *Y* – прямые, то корреляцию называют линейной.

Уравнения линейной регрессии с учетом коэффициента корреляции имеют вид:

$$y - \overline{y} = r_B \frac{\sigma_y}{\sigma_x}(x - \overline{x}) \text{ при} Y \text{ на } X; \tag{20}$$

$$x - \overline{x} = r_B \frac{\sigma_x}{\sigma_y}(y - \overline{y}) \text{ при} Y \text{ на } X, \tag{21}$$

где σ_x и σ_y – выборочные среднеквадратичные отклонения;

$\overline{x}$ и $\overline{y}$ – выборочные соединения признаков *X* и *У*;

r_B – эмпирический (выборочный) коэффициент корреляции величин Х и У.

Если частоты совместного появления признаков X и Y равны единице, то характеристики, входящие в (20), (21), записываются в виде:

$$\overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i, \quad \overline{y} = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i, \tag{22}$$

$$\overline{x}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2, \quad \overline{y}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i^2, \tag{23}$$

$$\sigma_x = \sqrt{\overline{x}^2} - (\overline{x})^2, \quad \sigma_y = \sqrt{\overline{y}^2} - (\overline{y})^2, \tag{24}$$

$$\overline{xy} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i \cdot y_i), \tag{25}$$

$$r_B = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \sigma_y}. \tag{26}$$

ЗАДАЧА № 6

В результате образования группы однотипных предприятий получены данные (таблица 2.2) о зависимости продукции Y, руб., от объема основных фондов *X*, млн. руб.

Таблица № 2.2

x_i	8,1	12,3	17,6	21,5	27,4
y_i	3,4	2,6	1,8	1,5	1,2

Требуется сопоставить эмпирическое линейное уравнение регрессии Y на X и изобразить в системе координат xOy опытные данные и прямую регрессии.

Решение. Вычислим эмпирические оценки соответствующих параметров линейного уравнения регрессии по приведенным, не сгруппированным данным:

$$\overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \frac{1}{5}(8{,}1 + 12{,}3 + 17{,}6 + 2{,}5 + 27{,}4) = 17{,}38;$$

$$\overline{y} = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i = \frac{1}{5}(3{,}4 + 2{,}6 + 1{,}8 + 1{,}5 + 1{,}2) = 2{,}1;$$

$$\overline{x}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 = \frac{1}{5}(8{,}1^2 + 12{,}3^2 + 17{,}6^2 + 2{,}5^2 + 27{,}4^2) = 347{,}9;$$

$$\overline{y}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i^2 = \frac{1}{5}(3{,}4^2 + 2{,}6^2 + 1{,}8^2 + 1{,}5^2 + 1{,}2^2) = 5{,}05;$$

$$\sigma_x = \sqrt{\overline{x}^2 - (\overline{x}^2)} = \sqrt{347{,}9 - 17{,}38^2} = 6{,}77;$$

$$\sigma_y = \sqrt{\overline{y^2} - (\overline{y})^2} = \sqrt{5,05 - 2,1^2} = 0,8;$$

$$\overline{xy} = \frac{1}{5}(8,1 \cdot 3,4 + 12,3 \cdot 2,6 + 17,6 \cdot 1,8 + 21,5 \cdot 1,5 + 27,4 \cdot 1,2) = 31,27.$$

Эмпирический коэффициент корреляции

$$r_B = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \sigma_y} = \frac{31,27 - 17,38 \cdot 2,1}{6,77 \cdot 0,8} - 0,96.$$

Запишем эмпирическое уравнение регрессии

$$y - \overline{y} = r_B \cdot \frac{\sigma_y}{\sigma_x} \cdot (x - \overline{x})$$

Подставив вычисленные значения параметров в уравнение, получим

$$y = -0,96 \cdot \frac{0,8}{6,77} \cdot (x - 17,38) + 2,1 \text{ или } y = -0,113x + 4,07.$$

Начертим график (рисунок 2.1)

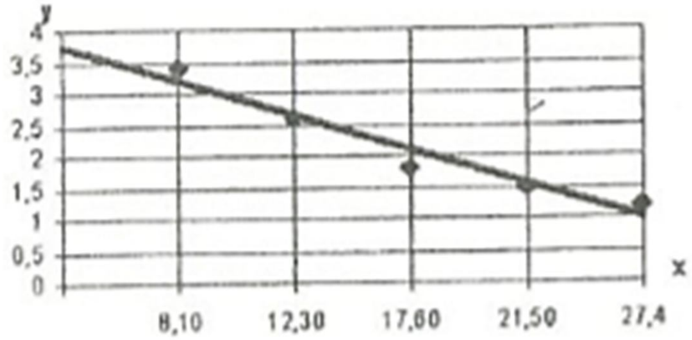

Рисунок 2.1

ЗАДАЧА № 7

Предел выносливости стали при изгибе Y (Н/мм^2) оценивается на основании другой ее характеристики – предела упругости при кручении X (Н/мм^2). По опытным данным для 12 марок стали (таблица 2.3) найти уравнения линейной регрессии Y на X и X на Y и вычислить коэффициент корреляции между этими характеристиками.

Таблица 2.3

x_i	51	67	84	81	101	109	71	97	109	51	105	89
y_i	25	30	43	44	57	58	43	46	62	45	55	45

$$\sum x_i = 1015; \ldots \sum y_i = 553;$$

$$\sum x_i^2 = 90667; \ldots \sum y_i^2 = 26807; \sum x_i y_i = 48888.$$

Ответ: $r \approx 0,837; y = 0,44x + 8,96; x = 1,6y + 10,96.$

Если данные наблюдений над признаками *X* и *Y* заданы в виде корреляционной таблицы с равностоящими вариантами, то целесообразно перейти к условным вариантам

$$u_i = \frac{X - C_1}{h_1}, \nu_i = \frac{y_1 - C_2}{h_2},$$

где C_1 – «ложный нуль» вариант *x* (новое начало отсчета); в качестве ложного нуля выгодно принять варианту, которая расположена примерно в середине вариационного ряда (условимся принимать в качестве ложного нуля варианту, имеющую наибольшую частоту);

h_1 – шаг, т. е. разность между двумя соседними вариантами *X*;

C_2 – ложный нуль вариант *Y*;

h_2 – шаг вариант *Y*.

В этом случае выборочный коэффициент корреляции

$$r_B = \frac{\sum n_{u\nu} u\nu - n\overline{u\nu}}{n\sigma_u \sigma_\nu},$$

причем слагаемое $\sum n_{u\nu} u\nu$ удобно вычислять, используя расчетную таблицу 2.6.

Величины $u, \nu, \sigma_u, \sigma_\nu$ могут быть найдены либо методом произведений (при большом числе данных), либо непосредственно по формулам:

$$\overline{u} = \frac{\sum n_u u}{n}, \overline{\nu} = \frac{\sum n_\nu \nu}{n}, \sigma_u = \sqrt{\overline{u}^2 - (\overline{u})^2}, \sigma_\nu = \sqrt{\overline{\nu}^2 - (\nu)^2}.$$

Зная эти величины, можно определить входящие в уравнения в регрессии (20) и (21) величины по формулам:

$$\overline{x} = \overline{u}h_1 + C_1, \overline{y} = \overline{\nu}h_2 + C_2, \sigma_x = \sigma_u h_1, \sigma_y = \sigma_\nu h_2.$$

Для оценки силы линейной корреляционной связи служит выборочный коэффициент корреляции r_B; чем ближе $|r_B|$ к единице, тем связь сильнее; чем ближе $|r_B|$ к нулю, тем связь слабее.

ЗАДАЧА № 8

Найти выборочное уравнение прямой линии регрессии Y на X по данным, приведенным в корреляционной таблице 2.4.

Решение. Составим корреляционную таблицу 5 в условных вариантах, выбрав в качестве ложных нулей $C_1 = 30$ и $C_2 = 36$ (каждая из этих вариант расположена в середине соответствующего вариационного ряда)

Найдем $\overline{u}$ и $\overline{\nu}$

$$\overline{u} = \frac{\sum n_u u}{n} = \frac{4\cdot(-2)+14\cdot(-1)+46\cdot 0+16\cdot 1+20\cdot 2}{100} = 0,34,$$

$$\overline{\nu} = \frac{\sum n_\nu \nu}{n} = \frac{10\cdot(-2)+18\cdot(-1)+44\cdot 0+22\cdot 1+6\cdot 2}{100} = -0,04.$$

Таблица 2.4

$Y \backslash X$	20	25	30	35	40	n_y
16	4	6	–	–	–	10
26	–	8	10	–	–	18
36	–	–	32	3	9	44
46	–	–	4	12	6	22
56	–	–	–	1	5	6
n_x	4	14	46	16	20	$n = 100$

Таблица 2.5

$Y \backslash X$	–2	–1	0	1	2	n_y
–2	4	6	–	–	–	10
–1	–	8	10	–	–	18
0	–	–	32	3	9	44
1	–	–	4	12	6	22
2	–	–	–	1	5	6
n_u	4	14	46	16	20	$n = 100$

Найдем вспомогательные величины $\overline{u}^2$ и $\overline{\nu}^2$:

$$\overline{u}^2 = \frac{\sum n_u u}{n} = \frac{4\cdot 4+14\cdot 1+16\cdot 1+20\cdot 4}{100} = 1,26,$$

$$\overline{\nu}^2 = \frac{\sum n_\nu \nu}{n} = \frac{10\cdot 4+18\cdot 1+22\cdot 1+6\cdot 4}{100} = 1,04.$$

Найдем $\sigma_u \sigma_\nu$:

$$\sigma_u = \sqrt{u^2-(u)^2} = \sqrt{1,26-0,34^2} = 1,07,$$

$$\sigma_\nu = \sqrt{\nu^2-(\nu)^2} = \sqrt{1,04-0,04^2} = 1,02.$$

Найдем $\sum n_{u\nu} u\nu$ для чего составим расчетную таблицу 2.6.

Суммируя числа последнего столбца таблицы 2.6, находим

$$\sum_{\nu} V \cdot u = \sum n_{u\nu} u\nu = 82.$$

Для контроля вычислений находим сумму чисел последней строки:

$$\sum_{u} u \cdot V = \sum n_{u\nu} u\nu = 82.$$

Совпадение сумм свидетельствует о правильности вычислений.

Пояснения к составлению таблицы 2.6:

1. Произведение частоты $n_{u\nu}$ на варианту u, т. е $n_{u\nu}u$, записывают в правом верхнем углу клетки, содержащей значение частоты. Например, в правых верхних углах клеток первой строки записаны произведения

$$4(2) = -8;\ 6(-1) = -6.$$

2. Складывают все числа, помещенные в правых верхних углах клеток одной строки, и их сумму помещают в клетку этой же строки «столбца U».
   Например, для первой строки

$$U = -8 + (-6) = -14$$

Таблица 2.6

ν \ u	-2	-1	0	1	2	$U=\sum n_{u\nu}\cdot u$	$\nu\cdot U$
-2	-8, 4, -8	-6, 6, -12				-14	28
-1	-	-8, 8, -8	0, 10, -10			-8	8
0	-	-	0, 32, 0	3, 3, 0	18, 9, 0	21	24
1	-	-	0, 4, 4	12, 12, 12	12, 6, 6	24	24
2	-	-	-	1, 1, 2	10, 5, 10	11	22
$V=\sum n_{u\nu}\cdot \nu$	-8	20	-6	14	16		$\sum_{\nu} \nu U=82$
$u\cdot V$	16	20	0	14	32	$\sum_{u} uV=82$	← Контроль

3. Наконец, умножают варианту ϑ на U и полученное произведение записывают в соответствующую клетку «столбца $\nu \cdot U$». Например, в первой строке таблицы $\vartheta = -2$, $U = -14$; следовательно $\nu \cdot U = (-2)\cdot(-14) - 28$. Сложив все числа «столбца $\nu \cdot U$», получают сумму $\sum_{\nu} \nu \cdot U$, которая равна иско-

мой сумме $\sum n_{uv}uv$. Например, для таблицы 2.6: $\sum_v V \cdot U = 82$, следовательно, искомая $\sum n_{uv}uv = 82$.

Для контроля аналогичные вычисления производят по столбцам:

Произведения $n_{uv} \cdot \vartheta$ записывают в левый нижний угол клетки, содержащей значение частоты; все числа, помещенные в левых нижних углах одного столбца, складывают и их сумму помещают в «строку V»; наконец, умножают каждую варианту на V и результат записывают в клетках последней строки.

Сложив все числа последней строки, получают сумму $\sum_u u \cdot V$, которая также равна искомой сумме Е $\sum n_{uv}uv$. Например, для таблицы 2.7 $\sum_u u \cdot V = 82$, следовательно, $\sum n_{uv}uv = 82$.

Найдем искомый выборочный коэффициент корреляции:

$$r_B = \frac{\sum n_{uv}uv - n\overline{uv}}{n\sigma_u\sigma_v} = \frac{82 - 100 \cdot 0,34(-0,04)}{100 \cdot 1,07 \cdot 1,02} = 0,76.$$

Найдем шаги h_1 и h_2 (разности между любыми двумя соседними вариантами):

$$h_1 = 25 - 20 = 5; h_2 = 26 - 16 = 10.$$

Найдем $\overline{x}$ и $\overline{y}$, учитывая, что $C_1 = 30,\ C_2 = 36$:

$\overline{x} = \overline{u} \cdot h_1 + C_1 = 0,34 \cdot 5 + 30 = 31,70;$

$\overline{y} = \overline{v} \cdot h_2 + C_2 = (-0,04) \cdot 10 + 36 = 35,60.$

Найдем σ_x, σ_y:

$$\sigma_x = h_1 \cdot \sigma_u = 5 \cdot 1,07 = 5,35,$$
$$\sigma_y = h_2 \cdot \sigma_v = 10 \cdot 1,02 = 10,2.$$

Подставив найденные величины в соотношение (20), получим искомое уравнение прямой линии регрессии Y на X

$$\overline{y}_x - 35,60 = 0,76\frac{10,2}{5,35}(x - 31,70).$$

Или окончательно

$$\overline{y}_x = 1,45x - 10,36.$$

2.4 Статистическая проверка статистических гипотез

Статистической называют гипотезу о виде неизвестного распределения или о параметрах известных распределений.

Нулевой (основой) называют выдвинутую гипотезу H_0.

Конкурирующей (альтернативной) называют гипотезу H_1, которая противоречит нулевой.

В итоге проверки гипотезы могут быть допущены ошибки двух родов.

Ошибка первого рода состоит в том, что будет отвергнута правильная гипотеза. Вероятность ошибки первого рода называют *уровнем значимости и* обозначают через ос.

Ошибка второго рода состоит в том, что будет принята неправильная гипотеза. Вероятность ошибки второго рода обозначают через α.

Статистическим критерием (или просто *критерием)* называют случайную величину *К*, которая служит для проверки гипотезы.

Наблюдаемым (эмпирическим) значением называют то значение критерия, которое вычислено по выборкам.

Критической областью называют совокупность значений критерия, при которых нулевую гипотезу отвергают.

Областью принятия гипотезы (областью допустимых значений) называют совокупность значений критерия, при которых нулевую гипотезу принимают.

Основной принцип проверки статистических гипотез: если наблюдаемое значение критерия принадлежит критической области, то нулевую гипотезу отвергают; если наблюдаемое значение критерия принадлежит области принятия гипотезы, то гипотезу принимают.

Критическими точками (границами) называют точки, отделяющие критическую область от области принятия гипотезы.

Правосторонней называют критическую область, определяемую неравенством $K > K_{kp}$, где K_{kp} – положительное число.

Левосторонней называют критическую область, определяемую неравенством $K > K_{kp}$, где K_{kp} – отрицательное число.

Односторонней называют правостороннюю или левостороннюю критическую область.

Двусторонней называют критическую область, определяемую неравенствами $K < k_{k1}, K > k_2$, где $k_2 > k_1$. В частности, если критические точки симметричны относительно нуля, то двусторонняя критическая область определяется неравенствами (в предположении, что $K_{kp} > 0$)

$$K < -k_{kp}, K > k_{kp},$$

или равносильным неравенством

$$|K| > -k_{kp}.$$

Для отыскания критической области задаются уровнем значимости а и ищут критические точки, исходя из следующих соотношений:

а) ддя правосторонней критической области

$$P(K > k_{kp}) = \alpha \ (k_{kp} > 0);$$

б) для левосторонней критической области

$$P(K < k_{kp}) - \alpha \ (k_{kp} < 0);$$

в) для двусторонней симметричной области

$$P(K >^{k} k_{kp}) = \frac{\alpha}{2} \ (k_{kp} > 0),$$

$$P(K < -k_{kp}) = \frac{\alpha}{2} \ .$$

Обозначим через n и m объемы больших ($n > 30$, $m > 30$) независимых выборок, по которым найдены соответствующие выборочные средние $\bar{x}$ и $\bar{y}$.

Генеральные дисперсии $D(X)$ и $D(Y)$ известны.

Для того, чтобы при заданном уровне значимости а проверить нулевую гипотезу $H_0: M(X) = M(Y)$ *о равенстве математических ожиданий (генеральных средних) двух нормальных генеральных совокупностей с известными дисперсиями (в случае больших выборок) при конкурирующей гипотезе* $H_1: M(X) \neq M(Y)$, *надо вычислить наблюдаемое значение критерия*

$$Z_{набл} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{D(X)}{n} + \frac{D(Y)}{m}}} \tag{27}$$

И по таблице функции Лапласа найти критическую точку $Z_{кр}$ *из равенства*

Если $|Z_{набл}| < Z_{кр}$ *– нет оснований отвергнуть нулевую гипотезу.*

Если $|Z_{набл}| > Z_{кр}$ *– нулевую гипотезу отвергают.*

При конкурирующей гипотезе $H_1: M(X) > M(Y)$ *находят критическую мочку* $Z_{кр}$ *по таблице функции Лапласа из равенства*

$$\Phi(Z_{кр}) = \frac{1 - 2\alpha}{2}.$$

Если $Z_{набл} < Z_{кр}$ *– нет оснований отвергнуть нулевую гипотезу.*

Если $Z_{набл} > Z_{кр}$ *– нулевую гипотезу отвергают.*

При конкурирующей гипотезе $H_1 : M(X) < M(Y)$ *находят вспомогательную точку* $Z_{кр}$.

Если $Z_{набл} > Z_{кр}$ *– нет основания отвергнуть нулевую гипотезу.*

Если $Z_{набл} < Z_{кр}$ *– нулевую гипотезу отвергают.*

ЗАДАЧА № 9

По двум независимым выборкам, объем которых *n* = 40 и *m* = 50, извлеченных из нормальных генеральных совокупностей найдены выборочные средние *x* = 130 и *y* = 140. Генеральные дисперсии известны $D(X) = 80$, $D(Y) = 100$. Требуется при уровне значимости 0,01 проверить нулевую гипотезу $H_0 : M(X) = M(Y)$, при конкурирующей гипотезе $H_1 : M(X) \neq M(Y)$

Решение. Найдем наблюдаемое значение критерия

$$Z_{набл} = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{D(X)}{n} + \frac{D(Y)}{m}}} = \frac{130 - 120}{\sqrt{\frac{80}{40} + \frac{100}{50}}} = -5.$$

По условию конкурирующая гипотеза имеет вид $M(X) \neq M(Y)$, поэтому критическая область – двусторонняя. Найдем правую критическую точку из равенства

$$\Phi(Z_{кр}) = \frac{1-a}{2} = \frac{1-1,01}{2} = 0,495.$$

По таблице функции Лапласа (приложение А) находим $Z_{кр} = 2,58$.

Так как $|Z_{набл}| > Z_{кр}$, то, в соответствии с правилом, нулевую гипотезу отвергаем. Другими словами, выборочные средние различаются значимо.

ЗАДАЧА № 10

По двум независимым выборкам, объемы которых *n* = 45 и *m* = 40, извлеченныч из нормальных генеральных совокупностей, найдены выборочные средние: $\bar{x}$ = 120; $\bar{y}$ = 123,8.

Генеральные дисперсии известны $D(X) = 90$, $D(Y) = 80$. Требуется при уровне значимости 0,01 проверить нулевую гипотезу $H_0 : M(X) = M(Y)$ при конкурирующей гипотезе $H_1 : M(X) \neq M(Y)$.

Решение. Объемы выборок большие (*m* > 30), дисперсии генеральные известны, совокупности распределены нормально. Для проверки нулевой гипотезы используем критерий *Z*, имеющий нормальное распределение.

Найдем наблюдаемое значение критерия

$$Z_н = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{D(X)}{n} + \frac{D(Y)}{m}}} = \frac{120 - 123,8}{\sqrt{\frac{50}{25} + \frac{80}{40}}} = -1,9.$$

По условию задания конкурирующая гипотеза имеет вид

$H_1 : M(X) \neq M(Y).$

Поэтому критическая область – двусторонняя.

Найдем правую критическую точку из равенства

$$\Phi(Z_{кр}) = \frac{1-\alpha}{2} = \frac{1-1,01}{2} = 0,495.$$

По таблице функции Лапласа (приложение А) находим критическое значение критерия $Z_{кр} = 2,58.$

Так как $|Z_H| < Z_K$, то нет основания отвергать нулевую гипотезу. Другими словами, выборочные средние различаются незначимо.

2.5 Проверка гипотезы о нормальном законе распределения генеральной совокупности по критерию Пирсона

По данным выборки проверить гипотезу о том, что генеральная совокупность распределена по нормальному закону, т. е. выдвигаем Но: генеральная совокупность распределена по нормальному закону.

Проверку Но проводим в следующем порядке:

1. Весь интервал наблюдаемых значений признака разбиваем на S интервалов равной длины точками $x_i, i = 0,1,2,...,S$.

2. За варианты берем середины интервалов $x_i^* = \frac{x_i + x_{i+1}}{2}$.

По ним вычисляем среднее выборочное $\overline{x}_B$ и выборочное среднеквадратичное отклонение σ_B.

3. Вычисляем вероятность попадания вариант в каждый интервал по формуле

$$P_i = P(x_i < X < x_{i+1}) = \Phi(\frac{\overline{x}_{i+1} - \overline{x}_B}{\sigma_B}) - \Phi(\frac{\overline{x}_i - \overline{x}_B}{\sigma_B}).$$

4. частоты $n_i = nP_i$, где n – объем выборки; $n = \sum n_i$.

5. Вычисляем наблюдаемое значение критерия Пирсона

$$x_H^2 = \sum \frac{(n_i - n_i')^2}{n_i'} \tag{28}$$

6. По таблице В.1 (приложение В) критических точек распределения x^2 по уровню значимости α и числу степеней свободы $K = S - 3$ находим критическую точку $K = S - 3$ правосторонней критической области x_K^2.

7. При $x_H^2 < x_K^2$ нет оснований отвергать нулевую гипотезу о нормальном законе распределения генеральной совокупности, т. е. эмпирические и теоретические частоты различаются незначимо.

Если $x_H^2 < x_K^2$, отвергаем гипотезу о нормальном распределении. Это означает, что данные наблюдений не согласуются с теоретическими расчетами, т. е. различие эмпирических и теоретических частот значимо.

ЗАДАЧА № 11

Используя критерий Пирсона при уровне значимости 0,01, установить, случайно или значимо расхождение между эмпирическими частотами n_i и теоретическими частотами n_i', которые вычислены, исходя из гипотезы о нормальном распределении генеральной совокупности *X*:

n_i	8	16	40	72	36	18	10
n_i	6	18	36	76	39	18	7

Решение. Найдем наблюдаемое значение критерия Пирсона

$$x_H^2 = \sum \frac{(n_i - n_i')^2}{n_i'}.$$

Составим расчетную таблицу 2.7.

Из таблицы 2.7 находим наблюдаемое значение критерия:

$$x_H^2 = 3,068.$$

По таблице критических точек распределения x^2 (приложение В) по уровню значимости 0,01 и числу степеней свободы $k = s - 3 = 7 - 3 = 4$ находим критическую точку правосторонней критической области

$$x_{kp}^2(0,01;4) = 13,3.$$

Так как $x_H^2 < x_{kp}^2$, то нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности. Другими словами, расхождение между эмпирическими и теоретическими частотами незначимо (случайно).

Таблица 2.7.

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)}{n'_i}$
1	8	6	2	4	0.667
2	16	18	–2	4	0,224
3	40	36	4	16	0,448
4	72	76	–4	16	0,208
5	36	39	–3	9	0,234
6	18	18	–	–	–
7	10	7	3	9	1,287
e	n = 200				$x_H^2 = 3,068.$

3 Задания к самостоятельной работе

1–10. Найти методом произведений: а) выборочную среднюю; б) выборочную дисперсию; в) выборочное среднее квадратическое отклонение по данному статистическому распределению выборки (в первой строке указаны выборочные варианты x_i, а во второй – соответственные частоты n_i, количественного признака X)

№ 1

x_i,	105	110	115	120	125	130	135
n_i	4	6	10	10	20	12	8

№ 2

x_i,	12,5	13,0	13,5	14,0	14,5	15,0	15,5
n_i	5	14	16	25	21	12	7

№ 3

x_i,	10,2	10,9	11,6	12,3	13,0	13,7	14,6
n_i	8	10	60	12	5	3	2

№ 4

x_i,	45	50	55	60	65	1270	75
n_i	4	6	10	40	20	12	7

№ 5

x_i,	110	115	120	125	130	135	140
n_i	5	10	30	25	15	10	5

№ 6

x_i,	12,4	16,4	20,4	24,4	28,4	32,4	36,4
n_i	5	14	16	25	21	12	7

№ 7

x_i,	26	32	38	44	50	56	62
n_i	5	15	40	25	8	4	3

№ 8

x_i,	10,6	15,6	20,6	25,6	30,6	35,6	40,6
n_i	8	10	60	12	5	3	2

№ 9

x_i,	100	110	120	130	140	150	160
n_i	4	6	10	40	20	12	8

№ 10

x_i,	130	140	150	160	170	180	190
n_i	5	10	30	25	15	10	5

11–20. Найти доверительные интервалы для оценки математического ожидания a нормального распределения с надежностью 0,95, зная выборочную среднюю X, объем выборки n и среднее квадратическое отклонение σ.

№ 11 $\overline{x} = 75{,}17$ $\sigma = 6$ $n = 36$
№ 12 $\overline{x} = 75{,}16$ $\sigma = 7$ $n = 49$
№ 13 $\overline{x} = 75{,}15$ $\sigma = 8$ $n = 64$
№ 14 $\overline{x} = 75{,}14$ $\sigma = 9$ $n = 81$
№ 15 $\overline{x} = 75{,}13$ $\sigma = 10$ $n = 100$
№ 16 $\overline{x} = 75{,}12$ $\sigma = 11$ $n = 121$
№ 17 $\overline{x} = 75{,}11$ $\sigma = 12$ $n = 144$
№ 18 $\overline{x} = 75{,}10$ $\sigma = 13$ $n = 169$
№ 19 $\overline{x} = 75{,}09$ $\sigma = 14$ $n = 196$
№ 20 $\overline{x} = 75{,}08$ $\sigma = 15$ $n = 225$

21–30. Найти выборочное уравнение прямой $\overline{y}_x - \overline{y} = r_B \dfrac{\sigma_y}{\sigma_x}(x - \overline{x})$ регрессии Y на X по данной корреляционной таблице.

№ 21

Y \ X	5	10	15	20	25	30	n_y
45	2	4	–	–	–	–	6
55	–	3	5	–	–	–	8
65	–	–	5	35	5	–	45
75	–	–	2	8	17	–	27
85	–	–	–	4	7	3	14
n_x	2	7	12	47	29	3	$n = 100$

№ 22

Y \ X	10	15	20	25	30	35	n_y
40	2	4	–	–	–	–	6
50	–	3	7	–	–	–	10
60	–	–	5	30	10	–	45
70	–	–	7	10	8	–	25
80	–	–	–	5	6	3	14
n_x	2	7	19	45	24	3	$n = 100$

№ 23

Y \ X	15	20	25	30	35	40	n_y
15	4	1	–	–	–	–	5
25	–	3	4	–	–	–	10
35	–	–	2	50	2	–	54
45	–	–	1	9	7	–	17
55	–	–	–	4	3	7	14
n_x	4	7	7	63	12	7	$n = 100$

№ 24

Y \ X	10	7	12	17	22	35	n_y
110	1	5	–	–	–	–	6
120	–	5	3	–	–	–	8
130	–	–	3	40	12	–	55
140	–	–	2	10	5	–	17
150	–	–	–	3	4	7	14
n_x	1	10	8	45	21	7	$n = 100$

№ 25

Y \ X	5	10	15	20	25	30	n_y
10	3	5	–	–	–	–	8
20	–	4	4	–	–	–	8
30	–	–	7	35	8	–	50
40	–	–	2	10	8	–	20
50	–	–	–	5	6	3	14
n_x	3	9	13	50	22	3	$n = 100$

№ 26

Y \ X	12	17	22	27	32	37	n_y
25	2	4	–	–	–	–	6
35	–	6	3	–	–	–	9
45	–	–	6	35	4	–	45
55	–	–	2	8	6	–	16
65	–	–	–	14	7	3	24
n_x	2	10	11	57	17	3	$n = 100$

№ 27

Y \ X	15	20	25	30	35	40	n_y
25	3	4	–	–	–	–	7
35	–	6	3	–	–	–	9
45	–	–	6	35	5	–	43
55	–	–	12	8	6	–	26
65	–	–	–	4	7	4	15
n_x	3	10	21	47	15	4	$n = 100$

№ 28

Y \ X	4	9	14	19	24	29	n_y
30	3	3	–	–	–	–	6
40	–	5	4	–	–	–	9
50	–	–	40	2	8	–	50
60	–	–	5	10	6	–	21
70	–	–	–	4	7	3	14
n_x	3	8	49	16	21	3	$n = 100$

№ 29

Y \ X	5	10	15	20	25	30	n_y
30	2	6	–	–	–	–	8
40	–	5	3	–	–	–	8
50	–	–	7	40	2	–	49
60	–	–	4	9	6	–	19
70	–	–	–	4	7	5	16
n_x	2	11	14	53	15	35	n = 100

№ 30

Y \ X	10	15	20	25	30	35	n_y
20	5	1	–	–	–	–	6
30	–	6	2	–	–	–	8
40	–	–	5	40	5	–	50
50	–	–	2	8	7	–	17
60	–	–	–	4	7	8	19
n_x	5	7	9	52	19	8	n = 100

31–40. Заданы выборочные средние x и y, найденные по выборкам объемов n = 60 m = 50, извлеченных из нормальных генеральных совокупностей X и Y с известными дисперсиями $D(X)$ и $D(Y)$. Требуется при уровне значимости a = 0,05 проверить нулевую гипотезу $H_0 : M(X) \neq M(Y)$ при конкурирующей гипотезе $H_1 : M(X)M(Y)$, т. е. требуется установить, значимо или незначимо различаются выборочные средние.

№ 31 $\bar{x}$ =638, $\bar{y}$ =625, $D(X)$=30, $D(Y)$=25.
№ 32 $\bar{x}$ =130, $\bar{y}$=125, $D(X)$=60, $D(Y)$=150.
№ 33 $\bar{x}$ =260, $\bar{y}$ =250, $D(X)$=90, $D(Y)$=125.
№ 34 $\bar{x}$ =390, $\bar{y}$ =375, $D(X)$=180, $D(Y)$=50.
№ 35 $\bar{x}$ =520, $\bar{y}$=500, $D(X)$=72, $D(Y)$=140.
№ 36 $\bar{x}$ =650, $\bar{y}$ =640, $D(X)$=78, $D(Y)$=135.
№ 37 $\bar{x}$ =780, $\bar{y}$=785, $D(X)$=120 $D(Y)$=100.
№ 38 $\bar{x}$ =910, $\bar{y}$ =885, $D(X)$=180 $D(Y)$=50.
№ 39 $\bar{x}$=1000, $\bar{y}$ =990, $D(X)$=84, $D(Y)$=130.
№ 40 $\bar{x}$ =13.8, $\bar{y}$ =17.1, $D(X)$=96, $D(Y)$=120.

41–50. Требуется при уровне значимости $\alpha = 0{,}05$ проверить по критерию согласия Пирсона гипотезу о нормальном распределении генеральной совокупности, если известны эмпирические частоты n_i и теоретические частоты n_i'

№ 41

n_i	6	10	20	27	19	12	6
n'_i	5	14	16	25	21	12	7

№ 42

n_i	5	12	19	27	20	1210	7
n'_i	8	13	14	27	22	12	6

№ 43

n_i	8	10	18	27	17	11	9
n'_i	4	15	16	25	21	12	7

№ 44

n_i	5	11	22	25	21	11	5
n'_i	5	13	17	25	21	12	7

№ 45

n_i	7	12	16	25	21	11	7
n'_i	5	12	18	19	20	10	6

№ 46

n_i	8	12	16	27	19	12	6
n'_i	5	17	13	25	21	12	7

№ 47

n_i	6	15	16	26	19	12	6
n'_i	5	17	13	25	21	12	7

№ 48

n_i	6	13	21	23	19	12	6
n'_i	5	14	16	25	21	13	76

№ 49

n_i	6	11	20	27	19	12	6
n'_i	5	13	16	25	20	13	7

№ 50

n_i	6	9	21	27	20	11	6
n'_i	5	13	17	25	21	12	7

ПРИЛОЖЕНИЕ А

Таблица А.1 – Значения функции $\varPhi(x)\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$

x	$\varPhi(x)$	x	$\varPhi(x)$	x	$\varPhi(x)$	x	$\varPhi(x)$
1	2	3	4	5	6	7	8
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,9Т	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517.	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734.	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12,	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,8710	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869

Продолжение таблицы А.1

1	2	3	4	5	6	7	8
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264	1,26	0,3962
0,31	0,1217	0,63	0,2357	0,95	0,3289	1,27	0,3980
1,28	0,3997	1,61	0,4463	1,94	0,4738	2,52	0,4941
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,54	0,4945
'1,30	0,4032	1,63	0,4484	1,96	0,4750	2,56	0,4948
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,58	0,4951
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,60	0,4953
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,62	0,4956
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,64	0,4959
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,66	0,4961
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,68	0,4963
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,70	0,4965
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,72	0,4967
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,74	0,4969
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,76	0,4971
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,78	0,4973
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,80	0,4974
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,82	0,4976
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,84	0,4977
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,86	0,4979
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,88	0,4980
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,90	0,4981
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,92	0,4982
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,94	0,4984
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,96	0,4985

Окончание таблицы А.1

1	2	3	4	5	6	7	8
1,51	0,4345	1,84	0,4671	2,34	0,4904	2,98	0,4986
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,00	0,49865
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,20	0,49931
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,40	0,49966
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,60	0,499841
1,56	0,4406	1,89	0,4706	2,44	0,4927	3,80	0,499928
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,00	0,499968
1,58	0,4429	1,91	0,4719	2,48	0,4934	4,50	0,499997
1,59	0,4441	1,92	0,4726	2,50	0,4938	5,00	0,499997
1,60	0,4452	1,93	0,4732				

ПРИЛОЖЕНИЕ Б

Таблица Б.1 – значений $t_\gamma = t(\gamma, n)$

$n \backslash \gamma$	0,95	0,99	0,999	$n \backslash \gamma$	0,95	0,99	0,999
5	2,28	4,06	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	2,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,697	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,001	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

ПРИЛОЖЕНИЕ В

Таблица В.1 – Критические точки распределения x^2

Число степеней свободы k	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,1	0,020
3	11,3	9,3	7,8	0,352	0,2	0,115
4	13,3	11,1	9,5	0,711	0,5	0,297
5	15,1	12,8	11,1	1,15	0,8	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26.1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Список литературы

1. Гмурман В. Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. – М.: Высшая школа, 1977. – 477 с.
2. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. – М.: Высшая школа, 1979. – 400 с.
3. Галушко В. Г. Вероятностно-статистические методы на автотранспорте / В. Г. Галушко. – К.: Вища школа, 1976. – 230 с.
4. Гурский Е. И. Теория вероятностей с элементами математической статистики / Е. И. Гурски. – М.: Высшая школа, 1971. – 327 с.
5. Джонсон Н. Статистика и планирование эксперимента в технике и науке / Н. Джонсон, Ф. Лион. – М.: Мир, 1980. – 516 с.

УЧЕБНО-МЕТОДИЧЕСКОЕ ИЗДАНИЕ

Вовк Леонид Петрович
Королев Евгений Алескандрович

УЧЕБНО–МЕТОДИЧЕСКОЕ ПОСОБИЕ К ВЫПОЛНЕНИЮ КОНТРОЛЬНЫХ РАБОТ ПО ДИСЦИПЛИНЕ «МАТЕМАТИЧЕСКАЯ СТАТИСТИКА» ДЛЯ СТУДЕНТОВ НАПРАВЛЕНИЙ ПОДГОТОВКИ: 23.03.01 «ТЕХНОЛОГИЯ ТРАНСПОРТНЫХ ПРОЦЕССОВ», 27.03.04 «УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ», 23.05.01 «НАЗЕМНЫЕ ТРАНСПОРТНО-ТЕХНОЛОГИЧЕСКИЕ СРЕДСТВА», 23.03.03 «ЭКСПЛУАТАЦИЯ ТРАНСПОРТНО-ТЕХНОЛОГИЧЕСКИХ МАШИН И КОМПЛЕКСОВ», 08.03.01 «СТРОИТЕЛЬСТВО», 20.03.01 «ТЕХНОСФЕРНАЯ БЕЗОПАСНОСТЬ», 08.05.02 «СТРОИТЕЛЬСТВО, ЭКСПЛУАТАЦИЯ, ВОССТАНОВЛЕНИЕ И ТЕХНИЧЕСКОЕ ПРИКРЫТИЕ АВТОМОБИЛЬНЫХ ДОРОГ, МОСТОВ И ТОННЕЛЕЙ», 38.03.02 «МЕНЕДЖМЕНТ» (ВСЕХ ФОРМ ОБУЧЕНИЯ)

Подписано к выпуску 21.07.17 г. Гарнитура Times New.
Усл. печ. л. 2,31. Зак. № 260 .

Государственное образовательное учреждение
Высшего профессионального образования
«Донецкий национальный технический университет»
Автомобильно-дорожный институт
84646, ДНР г. Горловка, ул. Кирова, 51
e-mail: print-adi@adidonntu.ru

Редакционно-издательский отдел