Министерство сельского хозяйства Российской Федерации

ФГБОУ «Воронежский государственный аграрный университет имени императора Петра I»

Факультет бухгалтерского учета и финансов

Кафедра прикладной математики и математических методов в экономике

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Методические указания и контрольные задания для студентов заочной формы обучения центра ДОТ по направлению 080100 – «Экономика»

Воронеж
2012

Составители: доцент кафедры прикладной математики и математических методов в экономике А.М. Слиденко, доцент кафедры высшей математики и теоретической механики С.Н. Дементьев

ОГЛАВЛЕНИЕ

Программа курса «Математическая статистика».....	2
Общие методические указания.....	3
Рекомендуемая литература.....……..	4
Задачи контрольной работы ……..	5
Решение типовых задач контрольной работы ……	11
Приложение 1. Значения функции Лапласа $\Phi(x)$	31
Приложение 2. Критические точки распределения χ^2 ……..	33
Приложение 3. Критические точки распределения Стьюдента……	34
Приложение 4. Критические точки распределения Фишера…….	35

ПРОГРАММА КУРСА
«Математическая статистика»

1. Предмет и задачи математической статистики. Генеральная и выборочная совокупности. Выборочной метод. Статистическое распределение выборки. Эмпирическая функция распределения. Полигон и гистограмма.

2. Статистические оценки параметров распределения. Несмещенные, эффективные и состоятельные оценки. Генеральная и выборочная средние. Генеральная и выборочная дисперсии. Исправленная выборочная дисперсия.

3. Точность оценки, доверительная вероятность (надежность). Доверительный интервал. Доверительные интервалы для оценки параметров нормального распределения.

4. Проверка статистических гипотез. Ошибки первого и второго рода. Статистический критерий проверки нулевой гипотезы. Наблюдаемое значение критерия. Критическая область. Сравнение двух дисперсий нормальных генеральных совокупностей.

Проверка гипотезы о нормальном распределении признака с помощью критерия Пирсона.

5. Понятие о дисперсионном анализе. Общая, факторная и остаточная суммы квадратов отклонений. Общая, факторная и остаточная дисперсии. Сравнение нескольких средних методом дисперсионного анализа.

6. Функциональная, статистическая и корреляционная зависимости. Условные средние. Выборочное уравнение регрессии. Отыскание параметров выборочного уравнения прямой регрессии с помощью метода наименьших квадратов. Выборочный коэффициент корреляции и его свойства. Оценка значимости регрессии и коэффициента корреляции. Построение доверительных интервалов для коэффициентов уравнения прямой регрессии.

7. Понятие случайных процессов и их характеристик.

ОБЩИЕ МЕТОДИЧЕСКИЕ УКАЗАНИЯ

В первой части Методических указаний представлены задачи, составляющие содержание контрольной работы, а во второй предлагаются подробные решения задач каждого типа.

Контрольная работа должна быть выполнена в ученической тетради, на внешней обложке которой следует указать название контрольной работы, фамилию и инициалы студента, полный учебный шифр, место учебы.

Перед решением задач необходимо записывать их условия. Решения задач и пояснения к ним должны быть достаточно подробными. Все вычисления необходимо делать полностью. Для замечаний преподавателя нужно оставлять поля.

В контрольной работе студент выполняет все те задачи, номера которых заканчиваются цифрой, совпадающей с последней цифрой его учебного шифра. *Эта цифра является вариантом студента.* Например, если Ваш учебный шифр заканчивается цифрой 4, то у Вас вариант 4 и Вы решаете задачи №№ 4, 14,

24, ... Если Ваш учебный шифр заканчивается цифрой 0, то у Вас вариант 0 и Вы решаете задачи №№ 10, 20, 30, ...

Контрольная работа должна быть представлена на проверку преподавателем не позднее трех недель до начала экзаменационной сессии.

К тестированию студент допускается лишь в случае, если его контрольная работа зачтена.

Р Е К О М Е Н Д У Е М А Я Л И Т Е Р А Т У Р А

1. Основы теории вероятностей и математической статистики с использованием компьютерного пакета Mathcad: [учебное пособие для студентов по экономическим специальностям] /С.Н. Дементьев, А.М. Слиденко, С.О. Стрыгина.–Воронеж: ФГОУ ВПО Воронежский ГАУ, 2010.–206 с.
2. *Кремер Н.Ш.* Теория вероятностей и математическая статистика: Учебник для вузов / *Н.Ш. Кремер*– М.: ЮНИТИ-ДАНА, 2001. – 543 с.
3. *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математической статистике: Учеб. пособие для студентов втузов / *В.Е. Гмурман* –3-е изд.–М.: Высшая школа, 1999.–400 с.

ЗАМЕЧАНИЕ.

Электронная версия учебника [1] имеется на сайте ДОТ.

ЗАДАЧИ КОНТРОЛЬНОЙ РАБОТЫ

В ЗАДАЧАХ 1–10 признак X представлен таблицей, которая является выборкой его значений, полученных в результате 100 независимых наблюдений. Требуется:

1. Составить интервальное выборочное распределение.

2. Построить гистограмму относительных частот.

3. Перейти от составленного интервального к точечному выборочному распределению, взяв при этом за значения признака середины частичных интервалов.

4. Построить полигон относительных частот.

5. Получить аналитический вид эмпирической функции распределения $F^*(x)$ и построить ее график.

6. Вычислить все точечные выборочные оценки числовых характеристик признака: выборочное среднее $\overline{x}$; выборочную дисперсию σ_x^2 и исправленную выборочную дисперсию s^2; выборочное среднее квадратичное отклонение σ_x и исправленное выборочное с.к.о. s.

7. Считая первый столбец таблицы выборкой значений нормально распределенного признака Y, построить доверительные интервалы, покрывающие неизвестные математическое ожидание и дисперсию этого признака с надежностью $\gamma = 0,95$.

8. При уровне значимости $\alpha = 0,05$ проверить с помощью критерия Пирсона гипотезу о нормальном распределении признака X.

9. Считая, что первый и второй столбцы заданной таблицы являются выборками значений нормально распределенных признаков Y и Z соответственно, проверить при уровне значимости $\alpha = 0,05$ гипотезу $H_0: D(Y) = D(Z)$.

1.	51,5	55,3	42,3	43,3	59,5	60,6	86,1	43,3	77,8	59,6
	11,3	22,3	46,3	22,8	47,3	45,3	43,8	56,3	50,3	50,0
	76,3	64,3	16,6	56,3	47,8	54,3	64,1	79,8	68,3	35,8
	51,2	50,1	51,0	70,8	31,3	33,3	23,7	53,3	71,7	58,5
	25,1	51,3	72,5	24,3	49,1	48,7	52,1	79,6	28,3	57,9
	52,6	59,9	29,7	43,7	55,7	53,0	50,1	50,7	58,8	46,7
	34,8	51,3	28,3	41,0	58,8	49,1	19,7	36,9	29,7	38,9
	50,8	28,0	35,3	69,9	30,6	64,0	32,5	45,1	45,3	70,4
	47,6	78,0	38,4	70,5	40,6	31,3	44,3	47,4	91,3	64,3
	31,3	45,1	66,1	23,3	40,1	43,6	66,1	42,3	19,1	31,3

2.	66,7	70,5	57,5	58,5	74,7	75,8	99,9	58,5	93,0	74,8
	26,7	37,5	61,5	38,0	62,5	60,5	59,0	71,5	65,5	65,2
	91,5	79,5	31,8	71,5	63,0	69,5	79,3	95,0	83,5	51,0
	66,4	65,3	66,2	85,5	46,5	48,5	36,9	68,5	86,9	73,7
	40,3	66,5	87,7	39,5	64,3	63,9	67,3	94,8	43,5	73,1
	67,8	75,1	44,9	58,9	70,9	68,2	65,3	65,9	74,0	63,9
	50,0	66,5	43,5	56,2	74,0	64,3	34,9	52,1	44,9	54,1
	66,0	43,2	70,5	85,1	45,8	79,2	47,7	60,3	60,5	85,6
	62,8	93,2	53,6	85,7	55,8	46,5	59,5	62,6	92,8	79,5
	46,5	60,3	81,3	38,5	55,3	58,8	81,3	57,5	34,3	46,5

3.	64,5	68,3	55,3	72,5	73,6	99,1	56,3	90,8	72,6	56,3
	19,2	35,3	59,3	60,2	58,3	56,8	69,3	63,3	63,0	36,8
	89,3	77,3	29,6	69,3	60,8	67,3	77,1	92,8	81,3	48,8
	24,3	63,1	64,0	83,3	44,3	46,3	36,7	66,3	84,7	71,5
	38,1	64,3	85,5	37,3	62,1	61,7	65,1	92,6	41,3	70,9
	65,6	72,9	42,7	56,7	68,7	66,0	63,1	63,7	71,8	61,7
	47,8	64,3	41,3	54,0	71,8	62,1	32,7	49,9	42,7	51,9
	63,8	41,0	68,3	82,9	43,6	77,0	45,5	58,1	58,3	83,4
	60,6	91,0	51,4	83,5	53,6	44,3	57,3	60,4	99,2	77,3
	44,3	58,1	79,1	36,3	53,1	56,6	79,1	55,3	32,1	44,3

4.	54,2	58,0	45,0	46,0	62,2	63,3	88,8	46,0	80,5	62,3
	14,0	25,0	49,0	25,5	50,0	48,0	46,5	59,0	53,0	52,7
	79,0	67,0	19,3	59,0	50,5	57,0	66,8	82,5	71,0	38,5
	53,9	52,8	53,7	73,0	34,0	36,0	26,4	56,0	74,4	61,2
	27,8	54,0	75,2	27,0	51,8	51,4	54,8	82,3	31,0	60,6
	55,3	62,6	32,4	46,4	58,4	55,7	52,8	53,4	61,5	51,4
	37,5	54,0	31,0	43,7	61,5	51,8	22,4	39,6	32,4	41,6
	53,5	30,7	58,0	72,6	33,3	66,7	35,2	47,8	48,0	73,1
	50,3	80,7	41,1	73,2	43,3	34,0	47,0	50,1	94,0	67,0
	34,0	47,8	68,8	26,0	42,8	46,3	68,8	45,0	21,8	34,7

5.	54,3	58,1	45,1	46,1	62,3	63,4	88,9	46,1	60,6	62,4
	14,1	25,1	49,1	25,6	50,1	48,1	46,6	59,1	53,1	52,8
	79,1	67,1	19,4	59,1	50,6	57,1	66,9	82,6	71,1	38,6
	54,0	52,9	53,8	73,1	34,1	36,1	26,5	56,1	74,5	63,1
	27,9	54,1	75,3	27,1	51,9	51,5	54,9	82,4	31,1	60,7
	55,4	62,7	32,5	46,5	58,5	55,8	52,9	53,5	61,6	51,7
	37,6	54,1	31,1	43,8	61,6	51,9	22,5	39,7	32,5	41,7
	53,6	30,8	58,1	72,7	33,4	66,8	35,3	47,9	48,1	73,2
	50,4	80,8	41,2	73,3	43,4	34,1	47,1	50,2	94,1	67,1
	34,2	47,9	68,9	26,1	42,9	46,4	68,9	45,1	21,9	34,1

6.	51,4	55,2	42,2	43,2	59,4	60,5	86,0	43,2	77,7	59,5
	11,2	22,2	46,2	47,2	45,2	43,7	56,2	50,2	49,9	22,7
	76,2	64,2	16,5	56,2	47,7	54,2	64,0	79,7	68,2	35,7
	51,1	50,0	50,9	7,2	31,2	33,2	23,6	53,2	71,6	58,4
	25,0	51,2	72,4	24,2	49,0	56,6	52,0	79,5	28,2	57,8
	52,5	59,8	29,6	43,6	55,6	52,9	50,0	50,6	58,7	48,6
	34,7	51,2	28,2	40,9	58,7	49,0	19,6	36,8	29,6	38,8
	50,7	27,9	55,2	69,8	30,5	63,9	32,4	45,0	45,2	70,3
	47,5	77,9	38,3	70,4	40,5	31,2	44,2	47,3	91,2	64,2
	31,3	45,0	66,0	23,2	40,0	43,5	66,0	42,2	19,0	31,2

7.	54,5	58,3	45,3	46,3	62,5	63,6	89,1	46,4	80,8	62,6
	14,3	25,3	49,3	25,8	61,8	48,3	46,8	59,3	53,3	53,0
	79,3	67,3	19,6	59,3	50,3	57,3	61,7	82,8	71,3	38,8
	54,2	53,1	54,0	73,8	50,8	36,3	25,7	56,3	74,7	61,5
	28,1	54,3	75,5	27,3	34,3	51,7	55,1	82,6	31,3	60,9
	55,6	62,9	32,7	46,7	52,1	56,0	53,1	53,7	61,8	51,7
	37,8	54,3	31,3	44,0	58,7	52,1	22,7	39,9	32,7	41,9
	53,8	31,0	58,3	72,9	33,6	67,0	35,5	48,1	48,3	73,4
	50,6	81,0	41,4	73,5	43,6	34,3	47,3	50,4	94,3	67,3
	34,3	48,1	69,1	26,3	43,1	46,6	69,1	45,3	22,1	34,3

8.	42,8	46,6	33,6	34,6	50,8	51,9	77,4	34,6	69,1	50,9
	2,6	13,6	37,6	14,1	38,6	36,6	35,1	47,6	41,6	41,3
	67,6	55,6	7,9	47,6	39,1	45,6	55,4	71,1	59,6	27,1
	42,5	41,4	42,3	61,6	22,6	24,6	15,0	44,6	63,0	49,8
	16,4	42,6	63,8	15,6	40,4	40,0	43,4	70,9	19,6	49,2
	43,9	51,2	21,0	35,0	47,0	44,3	41,4	42,0	50,1	40,2
	26,1	42,6	19,6	32,3	50,1	40,4	11,0	28,2	21,0	30,2
	42,1	19,3	46,6	61,2	21,9	55,3	23,8	36,4	36,6	61,7
	38,9	69,3	29,7	61,8	81,9	22,6	35,6	38,7	82,6	55,6
	22,6	36,4	57,4	14,6	31,4	34,9	57,4	33,6	10,4	22,6

9.	56,7	60,5	47,5	48,5	64,7	65,8	91,3	48,5	83,0	64,8
	16,5	27,5	51,5	28,0	52,5	50,5	49,0	61,5	55,5	55,2
	81,5	69,5	21,8	61,5	53,0	59,5	69,3	85,0	73,5	41,0
	56,4	55,3	56,2	75,5	36,5	38,5	26,9	58,5	76,9	63,7
	30,3	56,5	77,7	29,5	54,3	53,9	57,3	84,8	33,5	63,1
	57,8	65,1	34,9	60,9	58,2	55,3	55,9	64,0	53,9	48,9
	40,0	56,5	33,5	46,2	64,0	54,3	24,9	42,1	44,9	44,1
	56,0	33,2	60,5	75,1	35,8	69,2	37,7	50,3	50,5	75,6
	52,8	83,2	43,6	75,7	45,8	36,5	49,5	52,6	96,5	69,5
	36,5	50,3	71,3	28,5	45,3	48,8	71,3	47,5	24,3	36,5

10.	54,1	57,9	44,9	45,9	62,1	62,2	88,7	45,8	80,4	63,2
	13,9	24,9	48,9	47,9	46,4	58,9	52,9	52,6	25,4	49,9
	78,9	65,9	19,2	58,9	50,4	56,9	66,7	82,4	70,9	38,4
	53,8	52,7	53,6	72,6	33,9	35,9	26,3	55,9	74,3	61,1
	27,7	53,9	75,1	26,9	51,8	51,3	54,7	82,2	30,9	60,5
	55,2	62,5	32,3	46,3	58,3	55,6	52,7	53,1	61,4	51,3
	37,4	53,9	30,9	43,6	61,4	51,7	22,3	39,5	32,3	41,5
	53,4	30,6	57,9	75,2	33,2	66,6	35,1	47,7	47,9	73,0
	50,2	80,6	41,0	73,1	43,2	33,9	46,9	50,0	93,9	66,9
	33,9	47,7	68,7	25,9	42,7	46,2	68,7	44,9	21,7	33,9

В ЗАДАЧАХ 11–20 даны таблицы с выборками пар значений признаков X и Y. **Требуется:**

1. Вычислить выборочный коэффициент корреляции r и сделать выводы о тесноте и направлении линейной корреляционной зависимости между признаками X и Y.

2. При уровне значимости $\alpha = 0,05$ проверить гипотезу о значимости коэффициента корреляции.

3. Составить выборочное уравнение прямой регрессии Y на X, построить полученную прямую в системе координат вместе с исходными данными.

4. Вычислить коэффициент детерминации R^2 и оценить качество регрессии.

5. При уровне значимости $\alpha = 0,05$ оценить значимость регрессии с помощью критерия Фишера.

6. При уровне значимости $\alpha = 0,05$ получить доверительные интервалы для оценки генеральных параметров регрессии.

11.

x_i	5	10	15	18	23	19	22	18
y_i	7	9	19	17	24	22	18	21

12.

x_i	3	7	12	13	21	19	16	17
y_i	5	14	16	10	23	22	20	18

13.

x_i	1	3	4	5	6	8	9	10
y_i	6	8	11	7	14	17	15	14

14.

x_i	2	5	8	4	3	13	9	5
y_i	4	9	12	6	8	16	16	6

15.

x_i	3	7	8	11	13	8	10	12
y_i	21	9	9	6	4	14	10	5

16.

x_i	1	5	7	13	18	19	21	17
y_i	2	14	13	14	21	20	26	19

17.

x_i	2	7	3	12	8	9	10	4
y_i	1	4	1	7	7	5	8	3

18.

x_i	2	7	3	12	8	9	10	4
y_i	9	9	10	9	8	8	5	10

19.

x_i	3	5	2	6	7	10	12	4
y_i	11	12	8	9	12	7	6	8

20.

x_i	3	5	2	6	7	10	12	4
y_i	6	5	3	5	4	7	8	4

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ КОНТРОЛЬНОЙ РАБОТЫ

РЕШЕНИЕ ЗАДАЧИ ТИПА 1–10

Задача. Пусть признак X представлен таблицей 1, которая является выборкой его значений, полученных в результате 100 независимых наблюдений:

Таблица 1

50,2	54,0	41,0	42,0	58,2	59,3	84,8	45,0	76,5	58,3
21,0	55,0	45,0	21,5	46,0	44,0	42,5	49,0	48,7	75,0
15,3	55,0	23,8	46,5	53,0	62,8	78,5	67,0	34,5	49,9
49,7	63,0	30,0	32,0	42,4	22,4	52,0	70,4	57,2	50,0
23,0	47,8	47,4	50,8	78,3	27,0	56,6	51,3	58,6	28,4
51,7	50,0	48,8	49,4	57,5	47,4	33,5	27,0	39,7	57,5
18,4	35,6	28,4	37,6	49,5	26,7	54,0	68,6	29,3	62,7
43,8	44,0	69,1	46,3	76,7	37,1	69,2	39,3	30,0	43,0
85,0	63,0	30,0	43,8	64,8	22,0	38,8	42,3	64,8	41,0
30,0	**10,0**	63,0	48,8	71,2	54,4	47,8	31,2	46,1	17,8

Требуется:

1. Составить интервальное выборочное распределение.

2. Построить гистограмму относительных частот.

3. Перейти от составленного интервального к точечному выборочному распределению, взяв при этом за значения признака середины частичных интервалов.

4. Построить полигон относительных частот.

5. Получить аналитический вид эмпирической функции распределения $F^*(x)$ и построить ее график.

6. Вычислить все точечные выборочные оценки числовых характеристик признака: выборочное среднее $\overline{x}$; выборочную дисперсию σ_x^2 и исправленную выборочную дисперсию s^2; вы-

борочное среднее квадратичное отклонение σ_x и исправленное выборочное с.к.о. s.

7. Считая первый столбец таблицы выборкой значений нормально распределенного признака Y, построить доверительные интервалы, покрывающие неизвестные математическое ожидание и дисперсию этого признака с надежностью $\gamma = 0,95$.

8. При уровне значимости $\alpha = 0,05$ проверить с помощью критерия Пирсона гипотезу о нормальном распределении признака X.

9. Считая, что первый и второй столбцы заданной таблицы являются выборками значений нормально распределенных признаков Y и Z соответственно, проверить при уровне значимости $\alpha = 0,05$ гипотезу H_0: $D(Y) = D(Z)$.

Решение.

1. Составим интервальное выборочное распределение (интервальный вариационный ряд). Для этого, прежде всего, отметим, что у нас $x_{\min} = 10$, $x_{\max} = 85$, а размах выборочных значений $R = x_{\max} - x_{\min} = 85 - 10 = 75$.

Теперь определим длину каждого частичного интервала (иногда их называют классовыми интервалами), воспользовавшись формулой Стерджеса

$$l = \frac{R}{1 + 3,322 \lg n},$$

где n – объем выборки. В нашем случае

$$l = \frac{75}{1 + 3,322 \lg 100} \approx 9,81 \approx 10.$$

Далее устанавливаем границы частичных интервалов: левую границу первого интервала принимаем равной $x_0 = x_{\min} - \frac{l}{2} = 10 - 5 = 5$, далее полагаем $x_1 = x_0 + l = 5 + 10 = 15$, $x_2 = 25, \ldots,$ $x_9 = 85$. На этом указанная процедура заканчивается, т.к. последующие частичные интервалы не будут содержать выборочных значений признака.

Приступаем к распределению по частичным интервалам выборочных значений признака, ставя в соответствие интервалу с номером i частоту n_i как число выборочных значений признака, попавших в интервал. При этом договоримся, *что если некоторое из выборочных значений совпадет с границей двух соседних интервалов, то будем относить его к предыдущему из них* (так, например, число 55 дважды будет отнесено к интервалу (45;55)).

В итоге реализации данных рекомендаций получим таблицу 2, в первых двух столбцах которой разместим искомое интервальное распределение выборки, в третьем – относительные частоты $w_i = n_i / n$, а в последнем четвертом – плотности распределения относительных частот на частичных интервалах: $\tilde{p}_i = w_i / l$ (величины w_i и $\tilde{p}_i$ нам потребуются в дальнейшем).

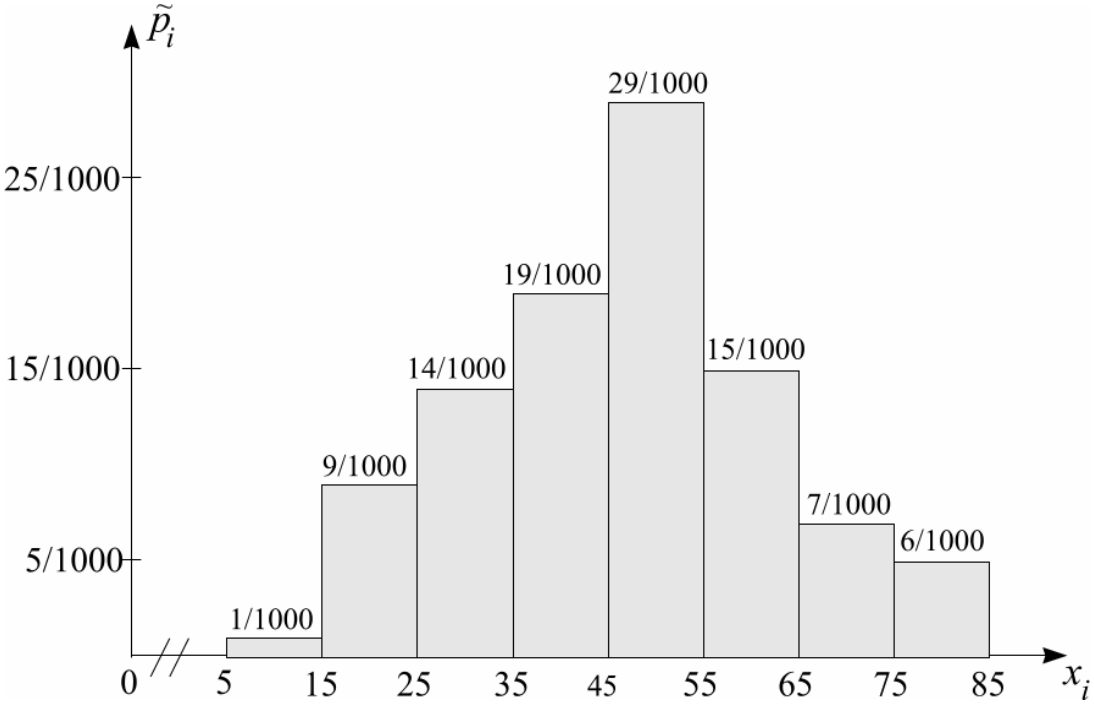
Таблица 2

$(x_{i-1};\ x_i)$	n_i	w_i	$\tilde{p}_i$
(5; 15)	1	$1/100$	$1/1000$
(15; 25)	9	$9/100$	$9/1000$
(25; 35)	14	$14/100$	$14/1000$
(35; 45)	19	$19/100$	$19/1000$
(45; 55)	29	$29/100$	$29/1000$
(55; 65)	15	$15/100$	$15/1000$
(65; 75)	7	$7/100$	$7/1000$
(75; 85)	6	$6/100$	$6/1000$
$\sum$	**100**	**1**	

2. Распределение непрерывной случайной величины принято графически представлять кривой распределения, которая является графиком ее плотности вероятностей (дифференциальной функции распределения). В статистике одной из оценок кривой распределения является ***гистограмма относительных частот***.

Это ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы, а высотами являются плотности относительных частот $\tilde{p}_i$ на частичных интервалах.

При построении гистограммы относительных частот в нашей задаче используем первый и последний столбцы таблицы 2:



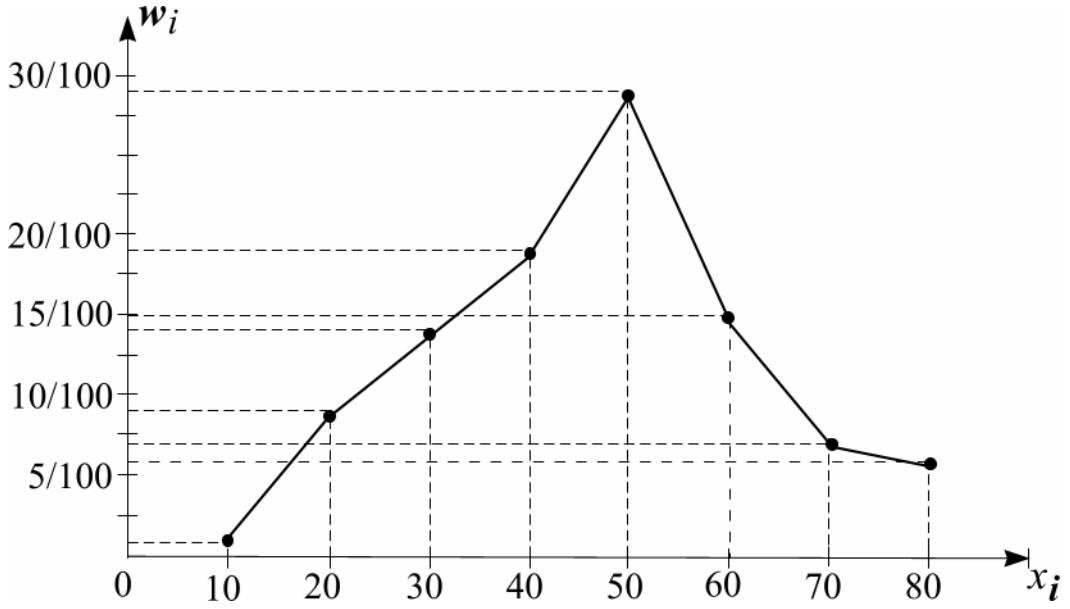
Гистограмма относительных частот

3. Нередко от интервального распределения выборки бывает удобно перейти к точечному (дискретному) распределению, взяв за новые выборочные значения признака середины частичных интервалов. В рассматриваемой задаче такое распределение, очевидно, имеет вид следующей таблицы 3:

Таблица 3

x_i	10	20	30	40	50	60	70	80
n_i	1	9	14	19	29	15	7	6

4. По полученной таблице 3 может быть построен ***полигон относительных частот***, который является, как и гистограмма относительных частот, статистической оценкой кривой распределения признака. *Это ломаная линия, вершины которой находятся в точках* $(x_i; w_i)$. В рассматриваемом случае в соответствии с первой строкой таблицы 3 и третьим столбцом таблицы 2 полигон относительных частот имеет следующий вид:



Полигон относительных частот

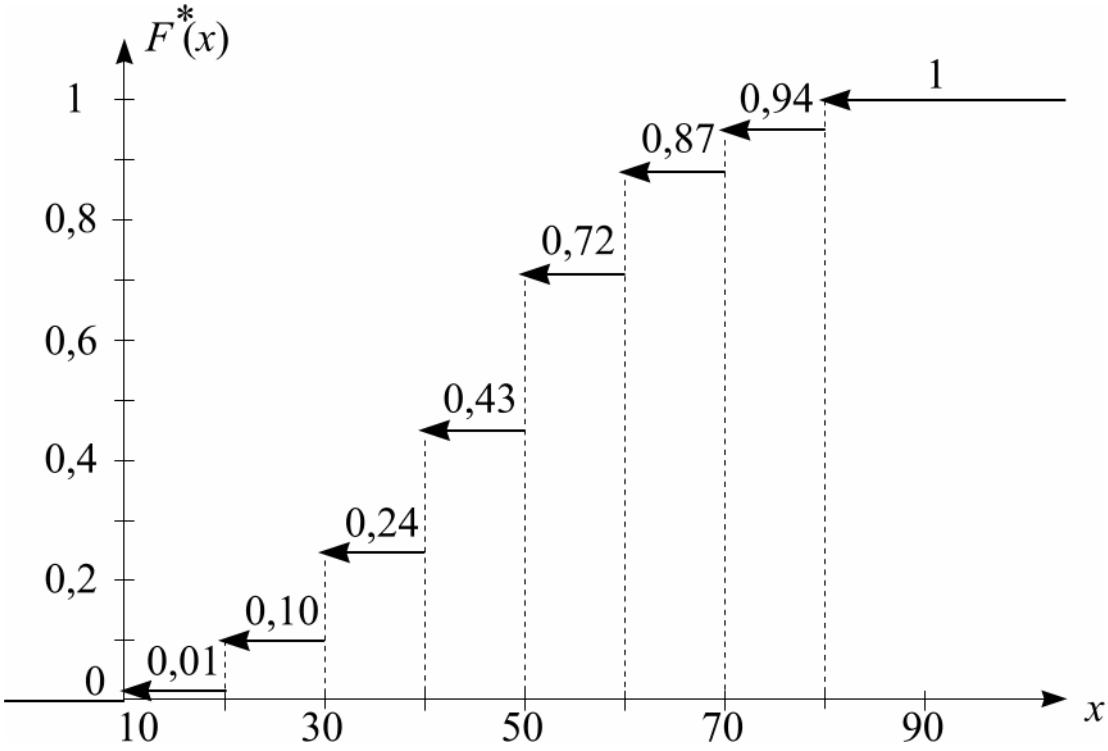
5. Для точечного распределения выборки может быть получена ***эмпирическая функция распределения*** $F^*(x)$**,** которая является статистической оценкой функции распределения вероятностей признака X (интегрального закона распределения) и строится по формуле

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

где n – объем выборки, а n_x – сумма частот выборочных значений признака, которые меньше x**.**

В нашей задаче, очевидно,

$$F^*(x)=\begin{cases} 0 & при \quad x \le 10, \\ 0,01 & при \quad 10 < x \le 20, \\ 0,10 & при \quad 20 < x \le 30, \\ 0,24 & при \quad 30 < x \le 40, \\ 0,43 & при \quad 40 < x \le 50, \\ 0,72 & при \quad 50 < x \le 60, \\ 0,87 & при \quad 60 < x \le 70, \\ 0,94 & при \quad 70 < x \le 80, \\ 1 & при \quad x > 80. \end{cases}$$



6. Важнейшими числовыми характеристиками признака X являются, как известно, математическое ожидание, дисперсия и среднее квадратичное отклонение (с.к.о.). Точечными статистическими оценками этих параметров служат соответственно выборочное среднее $\bar{x}$, выборочная дисперсия $\sigma_x^2 = \sigma_n^2$ и исправленная выборочная дисперсия $\sigma_{n-1}^2 = s^2$, выборочное с.к.о. $\sigma_x = \sigma_n$ и исправленное выборочное с.к.о. $\sigma_{n-1} = s$, которые вычисляются по формулам

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i ,$$

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

или

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2 \text{ (универсальная формула, где } \overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i \text{),}$$

$$s^2 = \frac{n}{n-1}\sigma_x^2 ,$$

$$\sigma_x = \sqrt{\sigma_x^2} , \quad s = \sqrt{s^2} ,$$

где x_i – выборочные значения (варианты) признака X, n_i – частоты этих значений, n – объем выборки.

По приведенным выше формулам вычислим точечные статистические оценки генеральных параметров распределения признака X, используя при этом данные из таблицы 3:

1) Находим выборочное среднее:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i = \frac{1}{100}(10\cdot 1 + 20\cdot 9 + 30\cdot 14 + 40\cdot 19 + 50\cdot 29 +$$

$$+60\cdot 15 + 70\cdot 7 + 80\cdot 6) = 46,9.$$

2) Находим выборочную дисперсию, используя универсальную формулу ее вычисления $\sigma_x^2 = \overline{x^2} - (\bar{x})^2$. Имеем:

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i = \frac{1}{100}(10^2\cdot 1 + 20^2\cdot 9 + 30^2\cdot 14 + 40^2\cdot 19 + 50^2\cdot 29 +$$

$$+60^2\cdot 15 + +70^2\cdot 7 + 80^2\cdot 6) = 2459;$$

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2 = 2459 - (46,9)^2 = 259,39.$$

3) Находим исправленную выборочную дисперсию:

$$s^2 = \frac{n}{n-1}\sigma_x^2 = \frac{100}{100-1} \cdot 259,39 \approx 262,01.$$

4) Находим выборочное с.к.о. и исправленное выборочное с.к.о.:

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{259,39} \approx 16,11; \quad s = \sigma_{n-1} = \sqrt{262,01} \approx 16,19.$$

7. *Числовые характеристики признака принято оценивать с помощью доверительных интервалов, покрывающих эти характеристики с заданной надежностью (доверительной вероятностью).*

Для нормально распределенного признака X, представленного выборкой объема n, доверительные интервалы, покрывающие с надежностью γ его неизвестные математическое ожидание a и дисперсию σ^2, имеют соответственно вид

$$(\bar{x} - t_\gamma \cdot \frac{s}{\sqrt{n}}; \bar{x} + t_\gamma \cdot \frac{s}{\sqrt{n}}) \qquad (1)$$

и

$$(s^2 \cdot \frac{n-1}{r_2}; \; s^2 \cdot \frac{n-1}{r_1}). \qquad (2)$$

Здесь значения t_γ являются критическими точками распределения t_{n-1}. Они ищутся в зависимости от заданного уровня значимости $\alpha = 1-\gamma$ и числа степеней свободы $\nu = n-1$ по таблицам приложения 3.

Величины r_1 и r_2 являются критическими точками распределения χ^2_{n-1}. Их находят по таблицам приложения 2 в зависимости от числа степеней свободы $\nu = n-1$, а также уровней значимости $\alpha_1 = 0,5 \cdot (1+\gamma)$ и $\alpha_2 = 0,5 \cdot (1-\gamma)$ соответственно.

Пусть нормально распределенный признак Y представлен выборкой значений из первой колонки таблицы 1:

y_i	50,2	21,0	15,3	49,7	23,0	51,7	18,4	43,8	85,0	30,0

Построим доверительные интервалы, покрывающие с надежностью $\gamma = 0{,}95$ математическое ожидание a и дисперсию σ^2 этого признака. Для этого по заданной выборке и таблицам приложений 2, 3 находим:

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{10}(50,2 + 21,0 + 15,3 + 49,7 + 23,0 + 51,7 + 18,4 +$$
$$+43,8 + 85,0 + 30,0) = 38,81;$$

$$\overline{y^2} = \frac{1}{n}\sum_{i=1}^{n} y_i^2 = \frac{1}{10}[(50,2)^2 + (21,0)^2 + (15,3)^2 + (49,7)^2 + (23,0)^2 +$$
$$+(51,7)^2 + (18,4)^2 + (43,8)^2 + (85,0)^2 + (30,0)^2] = 1924,911;$$

$$\sigma_y^2 = \overline{y^2} - (\overline{y})^2 = 1924,911 - (38,81)^2 = 418,695;$$

$$s^2 = \frac{n}{n-1}\cdot\sigma_y^2 = \frac{10}{10-1}\cdot 418,695 = 465,217;$$

$$s = \sqrt{s^2} = \sqrt{465,217} = 21,57;$$

$$t_\gamma = t_\gamma(\alpha;\ \nu) = t_\gamma(0,05;\ 9) = 2,26;$$

$$r_1 = r_1(\alpha_1;\ \nu) = r_1(0,975;\ 9) = 2,70;$$

$$r_2 = r_2(\alpha_2;\ \nu) = r_2(0,025;\ 9) = 19,0.$$

Теперь по формулам (1), (2) после несложных вычислений получаем искомые доверительные интервалы $(23,39;\ 54,23)$, $(220,37;\ 1550,72)$, покрывающие с надежностью $\gamma = 0,95$ параметры нормального распределения a и σ^2 соответственно.

8. Пусть непрерывная случайная величина (признак) X представлена выборкой значений в виде интервального распределения, причем известны выборочное среднее $\overline{x}$ и исправленное выборочное с.к.о. s.

Пусть имеются основания предполагать, что случайная величина X подчинена нормальному закону распределения (например, из визуального соответствия гистограммы и нормальной кривой).

Проверка этой гипотезы при уровне значимости α с помощью критерия Пирсона осуществляется по следующей схеме.

1) Нужно проанализировать интервальное распределение выборки, объем которой должен быть не менее 50, и в случае, если какому-нибудь частичному интервалу выборочных значений соответствует эмпирическая частота n_i, которая меньше, чем 5, этот интервал следует объединить с соседним (соседними), поставив в соответствие новому интервалу сумму эмпирических частот объединенных интервалов. Так как нормальное распределение определено для всех действительных значений x, то *принято левую границу первого частичного интервала расширить до* $-\infty$, *а правую границу последнего – до* $+\infty$. По окончании описанной процедуры будем обозначать число частичных интервалов через m.

2) В предположении, что исследуемая случайная величина X действительно распределена нормально с параметрами $\bar{x}$ и s ($X \sim N(\bar{x}, s)$), нужно вычислить вероятности P_i попадания ее значений в каждый из m частичных интервалов по формуле

$$P_i = P(x_{i-1} < X < x_i) = \Phi\left(\frac{x_i - \bar{x}}{s}\right) - \Phi\left(\frac{x_{i-1} - \bar{x}}{s}\right);\ i = 1, \ldots, m, \tag{I}$$

где x_0 и x_m заменены соответственно на $-\infty$ и $+\infty$, а значения функции Лапласа можно найти в таблицах приложения 1. При безошибочном счете должно выполняться условие

$$P_1 + \ldots + P_m = 1.$$

3) Нужно вычислить теоретические частоты по формуле

$$n_i^T = nP_i, \tag{II}$$

где n – объем выборки. Отметим, что при этом должно выполняться условие $\sum_{i=1}^{m} n_i^T = n$.

4) Теперь требуется вычислить опытное значение критерия χ^2 :

$$\chi_{on}^2 = \sum_{i=1}^{m} \frac{(n_i - n_i^T)^2}{n_i^T} \quad . \qquad \text{(III)}$$

Кроме того, нужно найти критическое значение критерия χ^2 ($\chi_{кр}^2$) в зависимости от выбранного уровня значимости α и числа степеней свободы $\nu = m - 3$. Это осуществляется с помощью таблиц приложения 2.

5) Наконец, необходимо сравнить полученные значения χ_{on}^2 и $\chi_{кр}^2$:

если $\chi_{on}^2 > \chi_{кр}^2$, то гипотеза о нормальном распределении случайной величины X при уровне значимости α отвергается;

если $\chi_{on}^2 < \chi_{кр}^2$, то считают, что при заданном уровне значимости α нет оснований отвергнуть гипотезу о нормальном распределении рассматриваемой случайной величины X .

Приступаем к решению задания из п.8.

Имеем интервальное распределение выборки значений признака X, содержащееся в таблице 2:

$(x_{i-1}; x_i)$	(5;15)	(15;25)	(25;35)	(35;45)	(45;55)	(55;65)	(65;75)	(75;85)
n_i	1	9	14	19	29	15	7	6

В п.6 получены значения $\bar{x} = 46,9$; $s = 16,19$.

Согласно рекомендациям, данным выше, объединим первые два интервала последней таблицы. В результате получим уже 7 (m = 7) частичных интервалов (вместо первоначальных восьми). Теперь заполним следующую таблицу:

Таблица 4

Интервалы значений признака	P_i	n_i	$n_i^T = nP_i$	$\frac{\left(n_i - n_i^T\right)^2}{n_i^T}$
(-∞; 25)	0,0885	10 (1+9)	8,85	0,15
(25; 35)	0,1442	14	14,42	0,01
(35; 45)	0,2195	19	21,95	0,40
(45; 55)	0,2393	29	23,93	1,07
(55; 65)	0,1771	15	17,71	0,41
(65; 75)	0,0905	7	9,05	0,46
(75; +∞)	0,0409	6	4,09	0,89
Сумма	**1**	**100**	**100**	**3,39**

Здесь

$$P_1 = p(-\infty < X < 25) = \Phi\left(\frac{25-46,9}{16,19}\right) - \Phi\left(-\infty\right) = \Phi\left(-1,35\right) + \Phi(+\infty) =$$

$$= -\Phi(1,35) + 0,5 = -0,4115 + 0,5 = 0,0885;$$

$$P_2 = p(25 < X < 35) = \Phi\left(\frac{35-46,9}{16,19}\right) - \Phi\left(\frac{25-46,9}{16,19}\right) =$$

$$= \Phi\left(-0,73\right) - \Phi(-1,35) = -0,2673 + 0,4115 = 0,1442;$$

$$P_3 = p(35 < X < 45) = \Phi\left(\frac{45-46,9}{16,19}\right) - \Phi\left(\frac{35-46,9}{16,19}\right) =$$

$$= \Phi\left(-0,12\right) - \Phi(-0,73) = -0,0478 + 0,2673 = 0,2195;$$

$$P_4 = p(45 < X < 55) = \Phi\left(\frac{55-46,9}{16,19}\right) - \Phi\left(\frac{45-46,9}{16,19}\right) =$$

$$= \Phi\left(0,50\right) - \Phi(-0,12) = 0,1915 + 0,0478 = 0,2393;$$

$$P_5 = p(55 < X < 65) = \Phi\left(\frac{65-46,9}{16,19}\right) - \Phi\left(\frac{55-46,9}{16,19}\right) =$$

$$= \Phi\left(1,12\right) - \Phi(0,50) = 0,3686 - 0,1915 = 0,1771;$$

$$P_6 = p(65 < X < 75) = \Phi\left(\frac{75-46,9}{16,19}\right) - \Phi\left(\frac{65-46,9}{16,19}\right) =$$

$$= \Phi\left(1,74\right) - \Phi(1,12) = 0,4591 - 0,3686 = 0,0905;$$

$$P_7 = p(75 < X < +\infty) = \Phi\left(+\infty\right) - \Phi\left(\frac{75-46,9}{16,19}\right) =$$

$$= \Phi\left(+\infty\right) - \Phi(1,74) = 0,5 - 0,4591 = 0,0409.$$

Поскольку

$$\chi^2_{on} = \sum_{i=1}^{m} \frac{(n_i - n_i^T)^2}{n_i^T} = 3,39,$$

$$\chi^2_{кр}(0,05;\ 7–3) = \chi^2_{кр}\ (0,05;\ 4)=9,5,$$

то

$$\chi^2_{on} < \chi^2_{кр}.$$

В ы в о д. В нашем случае при уровне значимости $\alpha = 0,05$ нет оснований отвергнуть гипотезу о нормальном распределении признака *X*.

9. Пусть имеются выборки значений нормально распределенных признаков *Y* и *Z* объемов n_y и n_z соответственно. Пусть при уровне значимости α требуется проверить гипотезу $H_0:\ D(Y) = D(Z)$.

Поступаем следующим образом:

1) Вычисляем s_y^2 и s_z^2 (п**усть оказалось, что** $s_y^2 > s_z^2$**).**

2) Находим опытное значение критерия Фишера $F_{on} = \dfrac{s_y^2}{s_z^2}$ и его критическое значение $F_{кр}(\alpha;\ n_y - 1;\ n_z - 1)$ по таблицам приложения 4 в зависимости от уровня значимости α и чисел степеней свободы $\nu_1 = n_y - 1$, $\nu_2 = n_z - 1$.

3) Сравниваем полученные значения критерия:

если $F_{on} < F_{кр}$, то при заданном уровне значимости α нет оснований отвергнуть гипотезу $H_0:\ D(Y) = D(Z)$;

если $F_{on} > F_{кр}$, то при заданном уровне значимости α гипотеза H_0: $D(Y) = D(Z)$ должна быть отвергнута.

Замечание. Если оказалось, что $s_z^2 > s_y^2$, то следует брать

$$F_{on} = \frac{s_z^2}{s_y^2},\ F_{кр}(\alpha;\ n_z - 1;\ n_y - 1).$$

По условию нашей задачи имеем выборки

Y	50,2	21,0	15,3	49,7	23,0	51,7	18,4	43,8	85,0	30,0
Z	54,0	55,0	55,0	63,0	47,8	50,0	35,6	44,0	63,0	10,0

Выше в п.7 вычислено $s_y^2 = 465,217$. Находим s_z^2:

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{10}(54,0 + 55,0 + 55,0 + 63,0 + 47,8 + 50,0 + 35,6 +$$
$$+44,0 + 63,0 + 10,0) = 47,74;$$

$$\overline{z^2} = \frac{1}{n}\sum_{i=1}^{n} z_i^2 = \frac{1}{10}[(54)^2 + (55)^2 + (55)^2 + (63)^2 + (47,8)^2 +$$
$$+(50)^2 + (35,6)^2 + (44)^2 + (63)^2 + (10)^2] = 2499,22;$$

$$\sigma_z^2 = \overline{z^2} - (\bar{z})^2 = 2499,22 - (47,74)^2 = 220,112;$$

$$s_z^2 = \frac{n}{n-1}\cdot\sigma_z^2 = \frac{10}{10-1}\cdot 220,112 = 244,569;$$

Получили s_y^2(465,217)> s_z^2(244,112), поэтому

$$F_{on} = \frac{s_y^2}{s_z^2} = \frac{465,217}{244,569} = 1,90;$$

$$F_{кр}(\alpha;\ n_y - 1;\ n_z - 1) = F_{кр}(0,05;\ 9;\ 9) = 3,18.$$

Так как $F_{on} < F_{кр}$, то при заданном уровне значимости $\alpha = 0,05$ нет оснований отвергнуть гипотезу H_0: $D(Y) = D(Z)$.

РЕШЕНИЕ ЗАДАЧИ ТИПА 11–20

Задача. Дана выборка пар значений признаков X и Y:

x_i	3	4	5	6	7	8	9	10
y_i	60	53	50	40	37	33	27	20

Требуется:

1. Вычислить выборочный коэффициент корреляции r и сделать выводы о тесноте и направлении линейной корреляционной зависимости между признаками X и Y.

2. При уровне значимости $\alpha = 0,05$ проверить гипотезу о значимости коэффициента корреляции.

3. Составить выборочное уравнение прямой регрессии Y на X, построить полученную прямую в системе координат вместе с исходными данными.

4. Вычислить коэффициент детерминации R^2 и оценить качество регрессии.

5. При уровне значимости $\alpha = 0,05$ проверить значимость регрессии с помощью критерия Фишера.

6. При уровне значимости $\alpha = 0,05$ получить доверительные интервалы для оценки генеральных параметров регрессии.

Решение.

1. Проводим вычисление выборочного коэффициента корреляции по формуле $r = \dfrac{\overline{x \cdot y} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y}$:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{8}\left(3+4+5+6+7+8+9+10\right) = 6,5;$$

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{8}\left(3^2+4^2+5^2+6^2+7^2+8^2+9^2+10^2\right) = 47,5;$$

$$\sigma_x^2 = \overline{x^2} - (\overline{x})^2 = 47,5 - (6,5)^2 = 5,25;$$

$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{5,25} \approx 2,29.$

Аналогично, $\overline{y} = 40$; $\overline{y^2} = 1762$; $\sigma_y^2 = 162$; $\sigma_y \approx 12,73$.

$$\overline{x \cdot y} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i = \frac{1}{8} \cdot (3 \cdot 60 + 4 \cdot 53 + 5 \cdot 50 + 6 \cdot 40 +$$

$$+7 \cdot 37 + 8 \cdot 33 + 9 \cdot 27 + 10 \cdot 20) = 231;$$

$$r = \frac{\overline{x \cdot y} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y} = \frac{231 - 6,5 \cdot 40}{2,29 \cdot 12,73} = -0,99.$$

В ы в о д. Линейная корреляционная зависимость сильная, т.к. $|r|$ близок к 1, и отрицательная, т.к. $r < 0$, т.е. с ростом значений одного признака значения другого признака имеют тенденцию к убыванию.

2. Проверяем значимость коэффициента корреляции.

Вычисляем ошибку репрезентативности:

$$S_r = \sqrt{\frac{1 - r^2}{n - 2}} = \sqrt{\frac{1 - (-0,99)^2}{8 - 2}} = 0,06.$$

Находим $t_{набл}$ и $t_{кр}$ (по таблицам приложения 3):

$$t_{набл} = \frac{r}{S_r} = \frac{-0,99}{0,06} = -16,5; \quad t_{кр}(\alpha,\ n-2) = t_{кр}(0,05,\ \ 6) = 2,45.$$

В ы в о д. Так как $|t_{набл}| > t_{кр}$, то при уровне значимости $\alpha = 0,05$ можно утверждать достоверность коэффициента корреляции (значимость отличия от нуля), т.е. линейная корреляционная зависимость между рассматриваемыми признаками существует не только в выборочной, но и в генеральной совокупности.

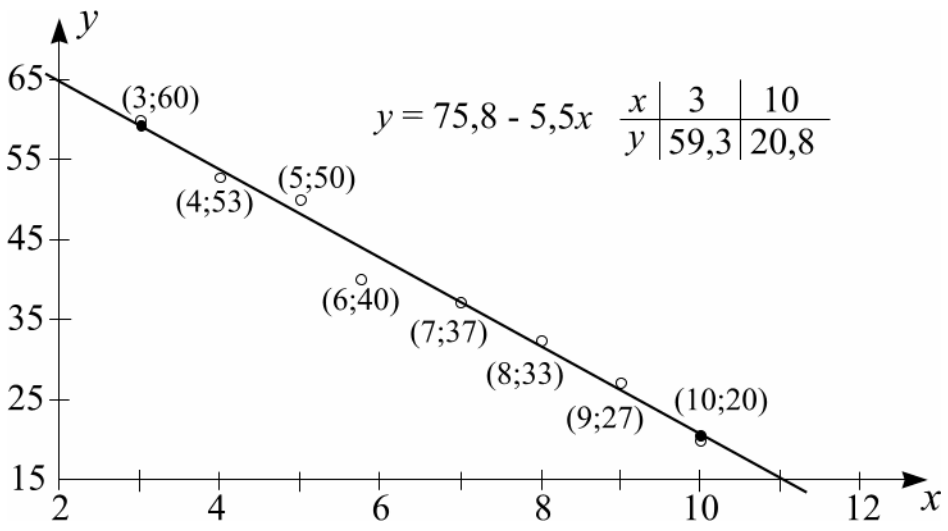
3. Получаем выборочное уравнение прямой регрессии Y на X по формуле

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

В нашем случае искомое уравнение имеет вид

$$y - 40 = -0,99 \cdot \frac{12,73}{2,29} \cdot (x - 6,5) \Rightarrow \underline{y = 75,75 - 5,50x}.$$

Строим полученную прямую регрессии и исходные данные на координатной плоскости. Для построения прямой находим координаты двух точек, через которые она проходит. В качестве ординат этих точек берем минимальное и максимальное значения переменной x из исходной таблицы, а соответствующие абсциссы точек получаем из уравнения прямой регрессии, параметры которого записываем с точностью 0,1 ($y = 75,8 - 5,5x$):



4. Для нахождения коэффициента детерминации R^2 используем формулу

$$R^2 = \frac{SS_{мод}}{SS_{мод} + SS_{ост}},$$

где

$SS_{ост} = \sum_{i=1}^{n} (y_i^{мод} - y_i)^2$ – остаточная сумма квадратов (значения $y_i^{мод}$ получают из выборочного уравнения прямой регрессии при подстановке в него значений $x = x_i$);

$SS_{мод} = \sum_{i=1}^{n} (y_i^{мод} - \overline{y})^2$ – модельная (факторная) сумма квадратов.

Находим модельные значения $y_i^{мод}$ отклика Y по формулам $y_i^{мод} = 75,75 - 5,50x_i$ и присоединяем их к таблице исходных данных:

x_i	3	4	5	6	7	8	9	10
y_i	60	53	50	40	37	33	27	20
$y_i^{мод}$	59,25	53,75	48,25	42,75	37,25	31,75	26,25	20,75

(здесь, например,

$$y_1^{мод} = 75,75 - 5,50 \cdot 3 = 59,25,$$

$$y_3^{мод} = 75,75 - 5,50 \cdot 5 = 48,25,$$

$$y_8^{мод} = 75,75 - 5,50 \cdot 10 = 20,75).$$

Теперь вычисляем

$$SS_{мод} = \sum_{i=1}^{n} (y_i^{мод} - \overline{y})^2 = (59,25 - 40)^2 + ... + (20,75 - 40)^2 = 1270,5;$$

$$SS_{ост} = \sum_{i=1}^{n} (y_i^{мод} - y_i)^2 = (59,25 - 60)^2 + ... + (20,75 - 20)^2 = 14,5.$$

$$R^2 = \frac{SS_{мод}}{SS_{мод} + SS_{ост}} = \frac{1270,5}{1270,5 + 14,5} = 0,99.$$

В ы в о д. Полученное значение коэффициента детерминации R^2 близко к 1, поэтому в нашем случае выборочное уравнение прямой регрессии хорошо объясняет вариации значений отклика Y.

5. Проверка значимости регрессии по критерию Фишера проводится следующим образом:

1) Вычисляют $\nu_{ост} = n - 2$; $\nu_{мод} = 1$ – числа степеней свободы остаточной и модельной сумм квадратов.

2) Вычисляют

$$s^2_{ост} = \frac{SS_{ост}}{\nu_{ост}} = \frac{SS_{ост}}{n-2}$$ – остаточную дисперсию;

$$s^2_{мод} = \frac{SS_{мод}}{\nu_{мод}} = \frac{SS_{мод}}{1}$$ – модельную дисперсию.

3) Находят $F_{оп} = \frac{s^2_{мод}}{s^2_{ост}}$ и по таблицам приложения 4 $F_{кр}(\alpha,\ \nu_{мод},\ \nu_{ост}) = F_{кр}(\alpha,\ 1,\ n-2)$.

4) Если $F_{оп} > F_{кр}$, регрессия признается значимой, если же $F_{оп} < F_{кр}$, то регрессия считается незначимой.

Проверяем с помощью критерия Фишера значимость регрессии в нашей задаче при уровне значимости $\alpha = 0,05$.

1) Вычисляем:

$$s^2_{мод} = \frac{SS_{мод}}{\nu_{мод}} = \frac{SS_{мод}}{1} = 1270,5;$$

$$s^2_{ост} = \frac{SS_{ост}}{\nu_{ост}} = \frac{SS_{ост}}{n-2} = \frac{14,5}{8-2} = 2,42;$$

$$F_{оп} = \frac{s^2_{мод}}{s^2_{ост}} = \frac{1270,5}{2,42} \approx 525.$$

2) Находим по таблицам приложения 4

$$F_{кр}(\alpha,\ \nu_{мод},\ \nu_{ост}) = F_{кр}(\alpha,\ 1,\ n-2) = F_{кр}(0,05;\ 1;\ 6) = 5,99.$$

3) Так как $F_{оп} > F_{кр}$, регрессия признается значимой.

6. Точечными выборочными оценками генеральных параметров β_0, β_1 уравнения регрессии $y = \beta_0 + \beta_1 x$ являются соот-

ветственно величины b_0, b_1 из выборочного уравнения регрессии $y = b_0 + b_1 x$. Доверительные интервалы, покрывающие указанные генеральные параметры с заданной надежностью γ, имеют соответственно вид

$$[b_0 - t_{кр} \cdot s(b_0); \quad b_0 + t_{кр} \cdot s(b_0)], \qquad (*)$$

$$[b_1 - t_{кр} \cdot s(b_1); \quad b_1 + t_{кр} \cdot s(b_1)], \qquad (**)$$

где значение $t_{кр}(1-\gamma,\ n-2) = t_{кр}(\alpha,\ n-2)$ ищется по таблицам приложения 3, а ошибки $s(b_0)$, $s(b_1)$ вычисляются по формулам

$$s(b_0) = \frac{s_{ост}}{\sqrt{n \cdot \sigma_x^2}} \cdot \sqrt{\overline{x^2}}\ ; \quad s(b_1) = \frac{s_{ост}}{\sqrt{n \cdot \sigma_x^2}}.$$

Строим для нашей задачи доверительные интервалы, покрывающие генеральные параметры регрессии с надежностью $\gamma = 0,95$.

Вычисляем:

$$s_{ост} = \sqrt{s_{ост}^2} = \sqrt{2,42} = 1,56;$$

$$s(b_0) = \frac{s_{ост}}{\sqrt{n \cdot \sigma_x^2}} \cdot \sqrt{\overline{x^2}} = \frac{1,56}{\sqrt{8 \cdot 5,25}} \cdot \sqrt{47,5} = 1,66;$$

$$s(b_1) = \frac{s_{ост}}{\sqrt{n \cdot \sigma_x^2}} = \frac{1,56}{\sqrt{8 \cdot 5,25}} = 0,24;$$

$$t_{кр}(1-\gamma,\ n-2) = t_{кр}(\alpha,\ n-2) = t_{кр}(0,05;\ 6) = 2,45.$$

Подставляем полученные результаты в формулы (*), (**):

$$[75,75 - 2,45 \cdot 1,66;\ 75,75 + 2,45 \cdot 1,66] \Rightarrow [71,68;\ 79,82];$$
$$[-5,50 - 2,45 \cdot 0,24;\ -5,50 - 2,45 \cdot 0,24] \Rightarrow [-6,09;\ -4,91].$$

В ы в о д. Доверительные интервалы, покрывающие генеральные параметры β_0, β_1 регрессии с надежностью $\gamma = 0,95$ имеют соответственно вид [71,68; 79,82] и [–6,09; –4,91].

ПРИЛОЖЕНИЯ

Приложение 1

Таблица значений функции Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.0000	0.35	0.1368	0.70	0.2580	1.05	0.3531
0.01	0.0040	0.36	0.1406	0.71	0.2611	1.06	0.3554
0.02	0.0080	0.37	0.1443	0.72	0.2642	1.07	0.3577
0.03	0.0120	0.38	0.1480	0.73	0.2673	1.08	0.3599
0.04	0.0160	0.39	0.1517	0.74	0.2703	1.09	0.3621
0.05	0.0199	0.40	0.1554	0.75	0.2734	1.10	0.3643
0.06	0.0239	0.41	0.1591	0.76	0.2764	1.11	0.3665
0.07	0.0279	0.42	0.1628	0.77	0.2794	1.12	0.3686
0.08	0.0319	0.43	0.1664	0.78	0.2823	1.13	0.3708
0.09	0.0359	0.44	0.1700	0.79	0.2852	1.14	0.3729
0.10	0.0398	0.45	0.1736	0.80	0.2881	1.15	0.3749
0.11	0.0438	0.46	0.1772	0.81	0.2910	1.16	0.3770
0.12	0.0478	0.47	0.1808	0.82	0.2939	1.17	0.3790
0.13	0.0517	0.48	0.1844	0.83	0.2967	1.18	0.3810
0.14	0.0557	0.49	0.1879	0.84	0.2995	1.19	0.3830
0.15	0.0596	0.50	0.1915	0.85	0.3023	1.20	0.3849
0.16	0.0636	0.51	0.1950	0.86	0.3051	1.21	0.3869
0.17	0.0675	0.52	0.1985	0.87	0.3078	1.22	0.3883
0.18	0.0714	0.53	0.2019	0.88	0.3106	1.23	0.3907
0.19	0.0753	0.54	0.2054	0.89	0.3133	1.24	0.3925
0.20	0.0793	0.55	0.2088	0.90	0.3159	1.25	0.3944
0.21	0.0832	0.56	0.2123	0.91	0.3186	1.26	0.3962
0.22	0.0871	0.57	0.2157	0.92	0.3212	1.27	0.3980
0.23	0.0910	0.58	0.2190	0.93	0.3238	1.28	0.3997
0.24	0.0948	0.59	0.2224	0.94	0.3264	1.29	0.4015
0.25	0.0987	0.60	0.2257	0.95	0.3289	1.30	0.4032
0.26	0.1026	0.61	0.2291	0.96	0.3315	1.31	0.4049
0.27	0.1064	0.62	0.2324	0.97	0.3340	1.32	0.4066
0.28	0.1103	0.63	0.2357	0.98	0.3365	1.33	0.4082
0.29	0.1141	0.64	0.2389	0.99	0.3389	1.34	0.4099
0.30	0.1179	0.65	0.2422	1.00	0.3413	1.35	0.4115
0.31	0.1217	0.66	0.2454	1.01	0.3438	1.36	0.4131
0.32	0.1255	0.67	0.2486	1.02	0.3461	1.37	0.4147
0.33	0.1293	0.68	0.2517	1.03	0.3485	1.38	0.4162
0.34	0.1331	0.69	0.2549	1.04	0.3508	1.39	0.4177

Продолжение приложения 1

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1.40	0.4192	1.79	0.4633	2.36	0.4909	5.00	0.5000
1.41	0.4207	1.80	0.4641	2.38	0.4913		
1.42	0.4222	1.81	0.4649	2.40	0.4918		
1.43	0.4236	1.82	0.4656	2.42	0.4922		
1.44	0.4251	1.83	0.4664	2.44	0.4927		
1.45	0.4265	1.84	0.4671	2.46	0.4931		
1.46	0.4279	1.85	0.4678	2.48	0.4934		
1.17	0.4292	1.86	0.4686	2.50	0.4938		
1.48	0.4306	1.87	0.4693	2.52	0.4941		
1.49	0.4319	1.88	0.4699	2.54	0.4945		
1.50	0.4332	1.89	0.4706	2.56	0.4948		
1.51	0.4345	1.90	0.4713	2.58	0.4951		
1.52	0.4357	1.91	0.4719	2.60	0.4953		
1.53	0.4370	1.92	0.4726	2.62	0.4956		
1.54	0.4382	1.93	0.4732	2.64	0.4959		
1.55	0.4394	1.94	0.4738	2.66	0.4961		
1.56	0.4406	1.95	0.4744	2.68	0.4963		
1.57	0.4418	1.96	0.4750	2.70	0.4965		
1.58	0.4429	1.97	0.4756	2.72	0.4967		
1.59	0.4441	1.98	0.4761	2.74	0.4969		
1.60	0.4452	1.99	0.4767	2.76	0.4971		
1.61	0.4463	2.00	0.4772	2.78	0.4973		
1.62	0.4474	2.02	0.4783	2.80	0.4974		
1.63	0.4484	2.04	0.4793	2.82	0.4976		
1.64	0.4495	2.06	0.4803	2.84	0.4977		
1.65	0.4505	2.08	0.4812	2.86	0.4979		
1.66	0.4515	2.10	0.4821	2.88	0.4980		
1.67	0.4525	2.12	0.4830	2.90	0.4981		
1.68	0.4535	2.14	0.4838	2.92	0.4982		
1.69	0.4545	2.16	0.4846	2.94	0.4984		
1.70	0.4554	2.18	0.4854	2.96	0.4985		
1.71	0.4564	2.20	0.4861	2.98	0.4986		
1.72	0.4573	2.22	0.4868	3.00	0.4986		
1.73	0.4582	2.24	0.4875	3.20	0.4993		
1.74	0.4591	2.26	0.4881	3.40	0.4997		
1.75	0.4599	2.28	0.4887	3.60	0.4998		
1.76	0.4608	2.30	0.4893	3.80	0.4999		
1.77	0.4616	2.32	0.4898	4.00	0.5000		
1.78	0.4625	2.34	0.4904	4.50	0.5000		

Приложение 2

Критические точки распределения χ^2

Число степеней свободы ν	**Уровень значимости α**			
$\downarrow$	**0.01**	**0.025**	**0.05**	**0.975**
1	6.6	5.0	3.8	0.00098
2	9.2	7.4	6.0	0.051
3	11.3	9.4	7.8	0.216
4	13.3	11.1	9.5	0.484
5	15.1	12.8	11.1	0.831
6	16.8	14.4	12.6	1.24
7	18.5	16.0	14.1	1.69
8	20.1	17.5	15.5	2.18
9	21.7	19.0	16.9	2.70
10	23.2	20.5	18.3	3.25
11	24.7	21.9	19.7	3.82
12	26.2	23.3	21.0	4.40
13	27.7	24.7	22.4	5.01
14	29.1	26.1	23.7	5.63
15	30.6	27.5	25.0	6.26
16	32.0	28.8	26.3	6.91
17	33.4	30.2	27.6	7.56
18	34.8	31.5	28.9	8.23
19	36.2	32.9	30.1	8.91
20	37.6	34.2	31.4	9.59
21	38.9	35.5	32.7	10.3
22	40.3	36.8	33.9	11.0
23	41.6	38.1	35.2	11.7
24	43.0	39.4	36.4	12.4
25	44.3	40.6	37.7	13.1
26	45.6	41.9	38.9	13.8
27	47.0	43.2	40.1	14.6
28	48.3	44.5	41.3	15.3
29	49.6	45.7	42.6	16.0
30	50.9	47.0	43.8	16.8

Приложение 3

Критические точки распределения Стьюдента

ν	**Уровень значимости α (двусторонняя критическая область)**					
↓	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
1	6.31	12.7	31.82	63.7	318.3	637.0
2	2.92	4.30	6.97	9.92	22.33	31.6
3	2.35	3.18	4.54	5.84	10.22	12.9
4	2.13	2.78	3.75	4.60	7.17	8.61
5	2.01	2.57	3.37	4.03	5.89	6.86
6	1.94	2.45	3.14	3.71	5.21	5.96
7	1.89	2.36	3.00	3.50	4.79	5.40
8	1.86	2.31	2.90	3.36	4.50	5.04
9	1.83	2.26	2.82	3.25	4.30	4.78
10	1.81	2.23	2.76	3.17	4.14	4.59
11	1.80	2.20	2.72	3.11	4.03	4.44
12	1.78	2.18	2.68	3.05	3.93	4.32
13	1.77	2.16	2.65	3.01	3.85	4.22
14	1.76	2.14	2.62	2.98	3.79	4.14
15	1.75	2.13	2.60	2.95	3.73	4.07
16	1.75	2.12	2.58	2.92	3.69	4.01
17	1.74	2.11	2.57	2.90	3.65	3.96
18	1.73	2.10	2.55	2.88	3.61	3.92
19	1.73	2.09	2.54	2.86	3.58	3.88
20	1.73	2.09	2.53	2.85	3.55	3.85
21	1.72	2.08	2.52	2.83	3.53	3.82
22	1.72	2.07	2.51	2.82	3.51	3.79
23	1.71	2.07	2.50	2.81	3.49	3.77
24	1.71	2.06	2.49	2.80	3.47	3.74
25	1.71	2.06	2.49	2.79	3.45	3.72
26	1.71	2.06	2.48	2.78	3.44	3.71
27	1.71	2.05	2.47	2.77	3.42	3.69
28	1.70	2.05	2.46	2.76	3.40	3.66
29	1.70	2.05	2.46	2.76	3.40	3.66
30	1.70	2.04	2.46	2.75	3.39	3.65
40	1.68	2.02	2.42	2.70	3.31	3.55
60	1.67	2.00	2.39	2.66	3.23	3.46
120	1.66	1.98	2.36	2.62	3.17	3.37
∞	1.64	1.96	2.33	2.58	3.09	3.29
	0.05	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
	Уровень значимости α (односторонняя критическая область)					

Приложение 4

Критические точки распределения Фишера
(уровень значимости $\alpha = 0{,}05$;
ν_1 – число степеней свободы большей дисперсии,
ν_2 – число степеней свободы меньшей дисперсии)

ν_2	ν_1												
↓	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**24**	∞
1	161	200	216	225	230	234	237	239	241	242	244	249	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.69	2.51	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.01	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	1.98	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	1.96	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.79	1.51
∞	3.84	3.00	3.60	2.37	2.21	2.10	2.01	1.94	1.83	1.83	1.75	1.52	1.00