

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«УФИМСКИЙ ГОСУДАРСТВЕННЫЙ НЕФТЯНОЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра математики

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Лабораторный практикум
для бакалавров по направлениям «Экономика», «Менеджмент»

Учебное пособие

Уфа
2012

Лабораторный практикум по дисциплине «Теория вероятностей и математическая статистика» предназначен для практического закрепления навыков статистического анализа и содержит прикладные задачи экономики по основным разделам теории вероятностей и математической статистики.

Задачи актуализированы применительно к условиям современной экономики и менеджмента и рекомендуются для получения навыков статистического анализа при решении экономических задач с использованием MS Excel.

Практикум предназначается для бакалавров по направлениям 080100.62 «Экономика», 080200.62 «Менеджмент» и студентов экономических специальностей.

Составитель: Мухаметзянов И.З., профессор, доктор физ.-мат. наук

Рецензент: Сулейманов И.Н., доц., канд. техн. наук

Содержание

	с.
1 Лабораторная работа 1. Моделирование случайных величин	4
2 Лабораторная работа 2. Предельные теоремы	10
3 Лабораторная работа 3. Описательные статистики. Расчет точечных и интервальных оценок генерального математического ожидания и дисперсии	17
4 Лабораторная работа 4. Проверка статистических гипотез	23
5 Лабораторная работа 5. Статистическое исследование зависимостей	27
6 Отчетность, оформление лабораторной работы	49

1 Лабораторная работа №1. Моделирование случайных величин

Цель работы: С использованием Excel провести моделирование случайных величин заданного закона распределения.

Задание

1. Провести моделирование случайных величин заданного закона распределения $R(1;3)$, $E(3)$, $N(10;1)$ (100 величин) используя стандартный метод, что интегральная функция распределения любой непрерывной случайной величины равномерно распределена в интервале $(0, 1)$.
2. Провести моделирование случайных величин заданного закона распределения (Бернулли, Биномиальное, Модельное, Дискретное) используя “Анализ данных–Генерация случайных чисел”
3. Построить гистограмму для иллюстрации закона распределения случайных величин.

Краткая теория и пример выполнения

1. Моделирование непрерывных случайных величин

Пусть нам нужно получить значения случайной величины ξ , распределенной в интервале (a, b) с плотностью вероятности $f(x)$.

Стандартный метод моделирования основан на том, что интегральная функция распределения любой непрерывной случайной величины равномерно распределена в интервале $(0, 1)$, т.е. для любой случайной величины x с плотностью распределения $f(x)$ случайная величина

$F(x) = \int_{-\infty}^x f(t)dt$ равномерно распределена в интервале $(0, 1)$. Тогда

случайную величину ξ с произвольной плотностью распределения $f(x)$ можно найти следующим образом (рис.1).

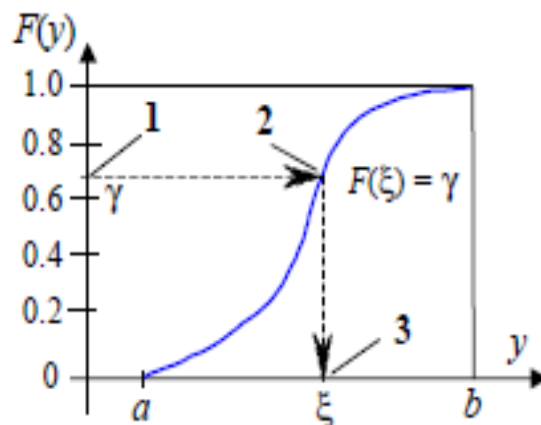


Рис. 1 Моделирование случайной величины ξ

1) получаем случайную величину γ , равномерно распределенную в интервале $(0, 1)$;

2) полагаем $F(\xi) = \int_a^{\xi} f(t)dt = \gamma$;

3) решая уравнение $\xi = F^{-1}(\gamma)$, находим искомое значение ξ .

Такой способ получения случайных величин называется методом обратных функций. На рис.2 приведен фрагмент моделирования в Excel

	A	B	C	D	E	F	G	H	I
1		=1+2*СЛЧИС()							
2									
3									
4	n	*R [0, 1]	R [1, 3]	E (3)	N (0, 4)				
5	1	0,062	1,551	0,009	-6,141				
6	2	0,278	2,244	0,047	-2,361				
7	3	0,607	2,484	0,135	1,083				
8	4	0,752	2,880	0,202	2,718				

Рис. 2 Фрагмент моделирования случайных величин в Excel

2. Моделирование дискретных случайных величин

Рассмотрим дискретную случайную величину ξ , принимающую n значений $X_1, X_2, \dots, X_n$ с вероятностями $P_1, P_2, \dots, P_n$. Эта величина задается таблицей распределения

$$\xi = \begin{pmatrix} X_1 & X_2 & \dots & X_n \\ P_1 & P_2 & \dots & P_n \end{pmatrix}, \quad \sum_i P_i = 1$$

Для моделирования такой дискретной величины разбивают отрезок $[0;1]$ на n последовательных отрезков $\Delta_1, \Delta_2, \dots, \Delta_n$, длины которых соответствуют вероятностям $P_1, P_2, \dots, P_n$. Получаем случайную величину γ , равномерно распределенную в интервале $(0, 1)$ и полагаем $\xi = X_k$, если $\gamma \in \Delta_k$.

Аналогичным образом осуществляется моделирование случайных событий. Пусть необходимо смоделировать реализацию какого либо события S_k из ряда событий $S_1, S_2, \dots, S_n$, образующих полную группу. Свяжем с этой системой событий случайную величину ξ , равную номеру события

$$\xi = \frac{1}{P_1} \frac{2}{P_2} \dots \frac{n}{P_n}, \quad \sum_i P_i = 1, \text{ где } P_k - \text{вероятность наступления } S_k$$

Тогда полагаем, что произошло случайное событие S_k , если случайная величина ξ приняла значение, равное k .

Пример моделирования с использованием MS Excel - функция **Анализ Данных/ Генерация случайных чисел**, приведен на рис.3.

	Xi	Pi	Результат
			XX
	10	0,25	11
	11	0,2	13
	12	0,1	14
	13	0,3	13
	14	0,15	11

Генерация случайных чисел

Число переменных: 1

Число случайных чисел: 20

Распределение: Дискретное

Параметры

Входной интервал значений и вероятностей: \$B\$110:\$C\$114

Случайное рассеивание:

Параметры вывода

☒ Выходной интервал: \$E\$110

☐ Новый рабочий лист:

☐ Новая рабочая книга

Рис. 3 Фрагмент моделирования случайных величин в Excel

Пример для **биномиальной** модели $Bi(n, p)$ - описывает распределение числа “успехов” в n независимых испытаниях с двумя исходами (“успех”, “неуспех”) и неизменной вероятностью успеха p (схема Бернулли). Частный случай этой модели – **бернуллиевская** модель $Bi(1, p)$ – особенно часто встречается в приложениях теории вероятностей и математической статистики, так как эксперимент с двумя возможными исходами является простейшим статистическим экспериментом. При этом, случайная величина с распределением $Bi(n, p)$ может быть реализована как сумма n независимых случайных величин с одинаковым распределением $Bi(1, p)$.

Пример моделирования с использованием MS Excel - функция **Анализ Данных/ Генерация случайных чисел**, приведен на рис.4.

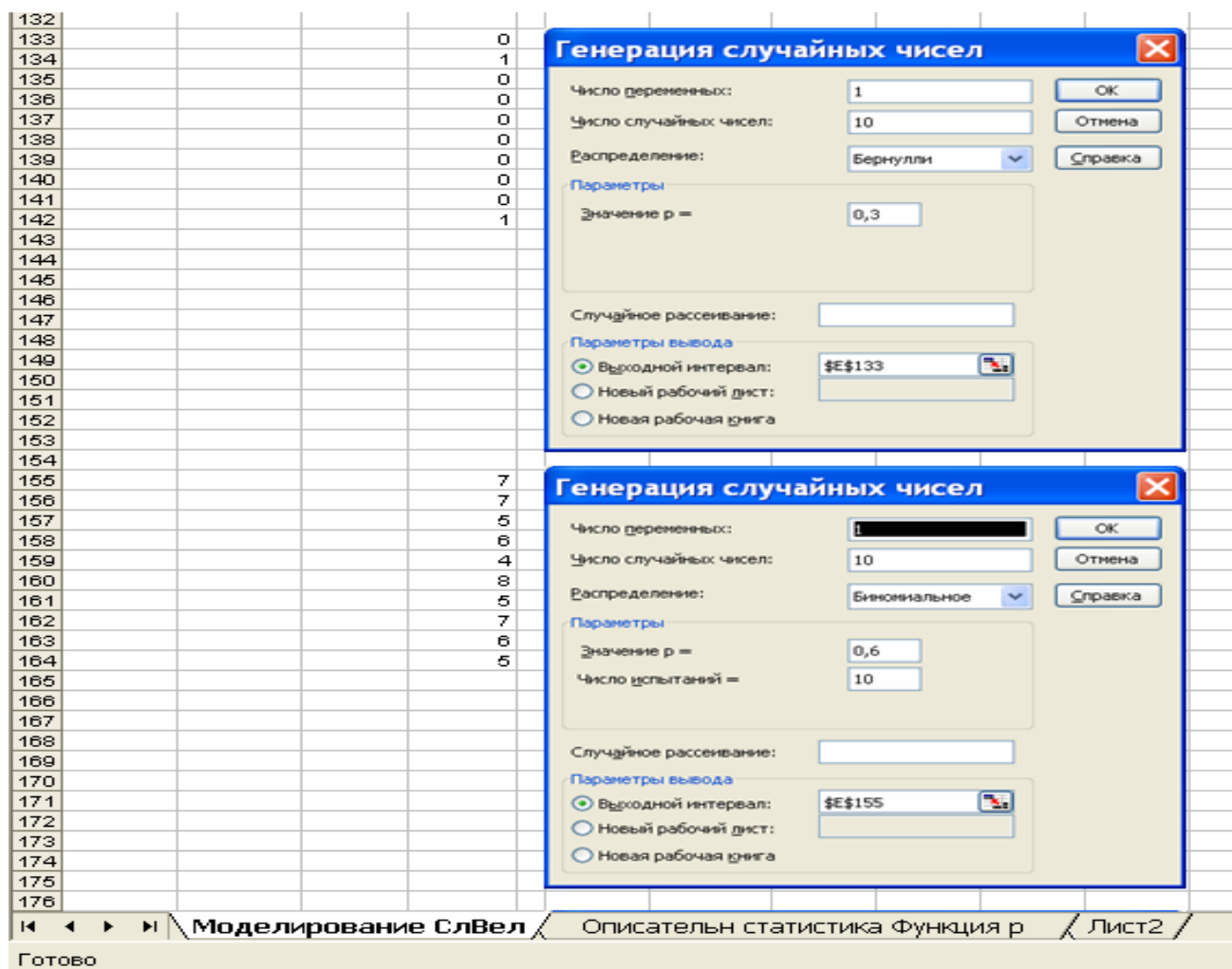


Рис. 4 Фрагмент моделирования случайных величин бернуллиевой и биномиальной модели в Excel

3. Построение гистограммы с использованием функции ЧАСТОТА()

Для этого разобьем диапазон случайных величин на интервалы. Число интервалов группировки $k = k(n) = 1 + [\log_2 n]$. Шаг $h = (maxX - minX)/k$.

Вершины интервалов определим по формуле $a_i = a_0 + i * h$, $i = 1, k$; $a_0 = minX$. Получим массив границ интервалов, который используем как аргумент для функции ЧАСТОТА().

Функции ЧАСТОТА() вычисляет частоту появления значений в интервале значений и возвращает массив чисел. Функцией ЧАСТОТА можно воспользоваться, например, для подсчета количества результатов тестирования, попадающих в интервалы результатов. Поскольку данная функция возвращает массив, ее необходимо вводить как формулу массива.

Синтаксис ЧАСТОТА(массив_данных, массив_интервалов)

Функция ЧАСТОТА имеет аргументы, указанные ниже.

Массив_данных — обязательный аргумент. Массив или ссылка на множество значений, для которых вычисляются частоты. Если аргумент "массив_данных" не содержит значений, функция ЧАСТОТА возвращает массив нулей.

Массив_интервалов — обязательный аргумент. Массив или ссылка на множество интервалов, в которые группируются значения аргумента "массив_данных". Если аргумент "массив_интервалов" не содержит значений, функция ЧАСТОТА возвращает количество элементов в аргументе "массив_данных".

Замечания

Функция ЧАСТОТА вводится как формула массива после выделения диапазона смежных ячеек, в которые требуется вернуть полученный массив распределения.

Количество элементов в возвращаемом массиве на единицу больше числа элементов в массиве "массив_интервалов". Дополнительный элемент в возвращаемом массиве содержит количество значений, превышающих верхнюю границу интервала, содержащего наибольшие значения. Например, при подсчете трех диапазонов значений (интервалов), введенных в три ячейки, убедитесь в том, что функция ЧАСТОТА возвращает значения в четырех ячейках. Дополнительная ячейка возвращает число значений в аргументе "массив_данных", превышающих значение верхней границы третьего интервала.

Функция ЧАСТОТА пропускает пустые ячейки и текст.

Формулы, возвращающие массивы, необходимо вводить как формулы массива.

Пример:

	А	В
1	Баллы	Интервалы
2	79	70
3		
4	85	79
5	78	89
6	85	
7	50	
8		
9	81	
10	95	
11	88	
12	97	

13	Формула	Описание (результат)
14	=ЧАСТОТА(A2:A10;B2:B4)	Количество оценок, меньших или равных 70 (1)
15		Количество оценок в интервале 71-79 (2)
		Количество оценок в интервале 80-89 (4)
		Количество оценок, больших или равных 90 (2)

ПРИМЕЧАНИЕ. Формулу в этом примере необходимо ввести как формулу массива. После копирования примера на пустой лист выделите диапазон A13:A16, нажмите клавишу F2, а затем нажмите клавиши CTRL + SHIFT + ВВОД. Если формула не будет введена как формула массива, отобразится только один результат в ячейке A12 (1).

Для построения гистограммы используем мастер диаграмм по диапазону частот (рис.5).

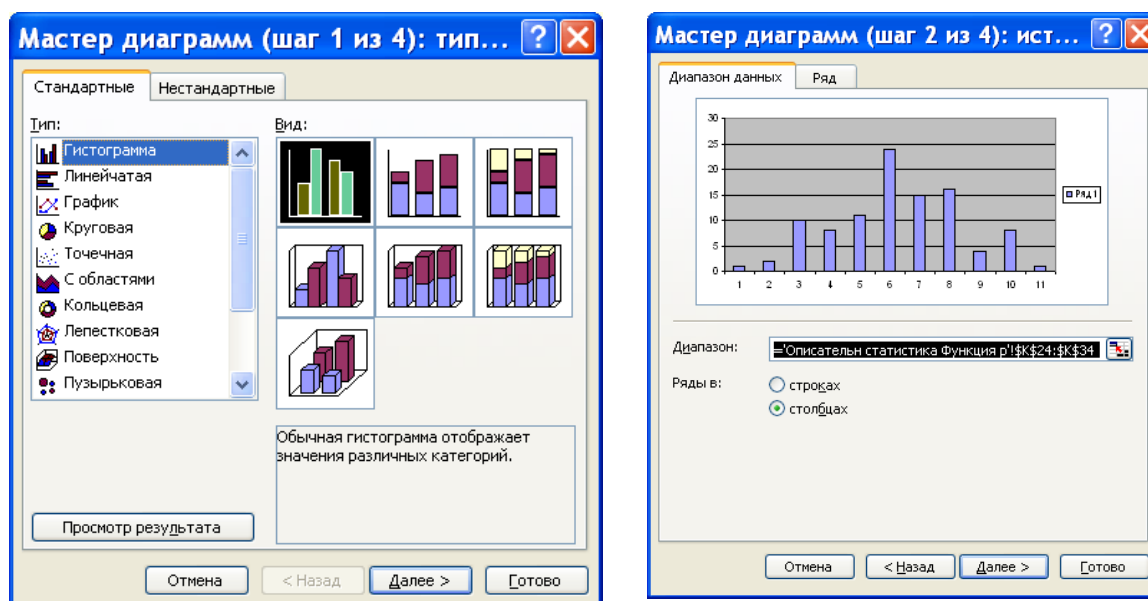


Рис. 5 Построения гистограммы с использованием мастера диаграмм в Excel

Выводы по работе

С использованием MS Excel, провели моделирование случайных величин заданных законов распределения. Результаты моделирования представили в виде гистограммы.

Лабораторная работа №2. Предельные теоремы

Цель работы: статистически пронаблюдать существо основных предельных теорем.

Задание

1. Пронаблюдать содержание теоремы Бернулли на примере моделирования бросания монеты.
2. Пронаблюдать содержание закон больших чисел в форме Чебышева для равномерно распределенной случайной величины в интервале (0,1).
3. Пронаблюдать свойство сжатия распределения с ростом объема выборки на 20 выборках из $R(0,1)$.
4. Пронаблюдать содержание центральной предельной теоремы на 12 одинаково распределенных слагаемых из $R(0,1)$ при различных объемах выборки n . Построить гистограмму и сравнить ее визуалью с нормальной плотностью.

Краткая теория и пример выполнения

1. Теорема Бернулли

Если проводится n независимых испытаний случайного события A , вероятность которого $P(A) = p$, то относительная частота μ/n появления события A (μ – число появлений A) при большом n приближенно равна вероятности p :

$$\frac{\mu}{n} \approx p.$$

Уточнение: будем писать

$$\frac{\mu}{n} \rightarrow p \quad \text{при } n \rightarrow \infty,$$

если для любого $\varepsilon > 0$ и для достаточно больших n соотношение

$$\left| \frac{\mu}{n} - p \right| < \varepsilon \tag{1}$$

выполняется с вероятностью, стремящейся к 1 с ростом n ; запишем это так:

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} \rightarrow 1 \text{ при } n \rightarrow \infty.$$

В этом состоит теорема Бернулли. Заметим, что теорема не утверждает, что соотношение (1) достоверно, однако, если n достаточно велико, то вероятность его выполнения близка к 1 (например, 0.98 или 0.999), что **практически достоверно**. Если собираемся провести эксперимент, состоящий из этого достаточно большого числа n испытаний, то можем быть уверены, что соотношение (1) будет выполнено. Проверим это не абсолютно достоверное утверждение.

Пример: Бросание симметричной монеты

Вероятность появления герба $p=0.5$. Можно показать (с помощью центральной предельной теоремы), что, например, если $n \geq (1.5/\varepsilon)^2$, то соотношение (1) выполняется с вероятностью 0.997, а если $n \geq (1.3/\varepsilon)^2$, то – с вероятностью 0.99; последняя в данном случае нас вполне устраивает как практическая достоверность. Положим $\varepsilon = 0.1$; тогда соотношение

$$|\mu/n - 0.5| < 0.1 \quad (a)$$

выполняется с вероятностью 0.99 при $n \geq 170$. Если $\varepsilon=0.03$, то соотношение

$$|\mu/n - 0.5| < 0.03 \quad (б)$$

выполняется с вероятностью 0.99 при $n \geq 1850$. Мы уверены, что, проведя 170 бросаний монеты, получим (а), а, проведя 1850 бросаний, получим (б).

Бросание монеты моделируем генерацией случайной величины α , принимающей значения 1 ("герб") и 0 ("цифра") с вероятностями 1/2. Число появлений "герба" в n испытаниях

$$\mu = \sum_{k=1}^n \alpha_k,$$

где α_k – результат k -го испытания. Результаты моделирования приведены на рис. 6

	A	B	C	D	E	F
1	n/n	Сл число+0,5	Целое	1- герб, 0- решка		
2	1	1,430	1			
3	2	0,938	0			
4	3	0,831	0			
5	4	0,841	0			
6	5		0			
7	6		0			
8	7	0,551	0			
9	8	0,871	0			
10	9	0,922	0			
11	10	0,736	0			
161	160	0,951	0			
162	161	1,404	1			
163	162	1,154	1			
164	163	1,433	1			
165	164	1,278	1			
166	165	0,847	0			
167	166	0,842	0			
168	167	1,308	1			
169	168	1,121	1			
170	169	0,732	0			
171	170	0,697	0			
172			91	Относительная частота появления герба		
173				$ f_n - 0.5 < 0.1.$		
174			0,5353	0,0353		
175						
176						
177						
178						
179						
180						

Рис. 6 Реализация выполнения теоремы Бернулли на примере бросания монеты в MS Excel

2. Закон больших чисел в форме Чебышева

Основное утверждение Одно из основных утверждений закона

больших чисел состоит в том, что значение среднеарифметического $\frac{1}{n} \sum_{i=1}^n \xi_i$

случайных величин с равными математическими ожиданиями $M\xi_i = a$ при большом n (при некоторых широких условиях) оказывается приближенно равным a :

$$\frac{1}{n} \sum_{i=1}^n \xi_i \approx a$$

Уточним: будем писать

$$\frac{1}{n} \sum_{i=1}^n \xi_i \rightarrow a \text{ при } n \rightarrow \infty,$$

если для любого $\varepsilon > 0$ и достаточно больших n соотношение

$$\left| \frac{1}{n} \sum_{i=1}^n \xi_i - a \right| < \varepsilon \quad (2)$$

выполняется с вероятностью, стремящейся к 1 с ростом n ; запишем это так:

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n \xi_i - a\right| < \varepsilon\right\} \rightarrow 1 \text{ при } n \rightarrow \infty.$$

Это одно из утверждений закона больших чисел. Заметим, что, как и теорема Бернулли, оно не означает, что соотношение (2) достоверно; однако, если n достаточно велико, то вероятность его выполнения близка к 1, например, 0.98 или 0.999, что означает **практически достоверно**. Приведем полную формулировку одной из теорем закона больших чисел в форме Чебышева,

Теоремы Чебышева. Если $\xi_1, \dots, \xi_n, \dots$ - последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной:

$$D\xi_1 < c, D\xi_2 < c, \dots, D\xi_n < c, \dots,$$

то для любого $\varepsilon > 0$

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n M\xi_i\right| < \varepsilon\right\} \rightarrow 1 \text{ при } n \rightarrow \infty.$$

Испытание практически достоверного события

Убедимся в выполнении (2) статистически на примере.

Пример. Случайные величины распределены равномерно на отрезке $[0,1]$. Если значение ε задавать произвольно, а число испытаний выбирать из условия $n \geq (9D\xi/\varepsilon^2)$, то (как нетрудно показать) соотношение (2) выполняется с вероятностью $P=0.997$, а если $n \geq (5.4D\xi/\varepsilon^2)$ - то с $P=0.98$. Последняя нас устраивает, как практическая достоверность.

Положим $\varepsilon_1 = 0.1$ и $\varepsilon_2 = 0.02$, определим два соответствующих значения

$$n_1 = 45 \text{ и } n_2 = 1125$$

Далее, генерируем случайное число ($n=45$) (аналогично для $n=1125$) и убеждаемся в выполнении неравенства (рис.7).

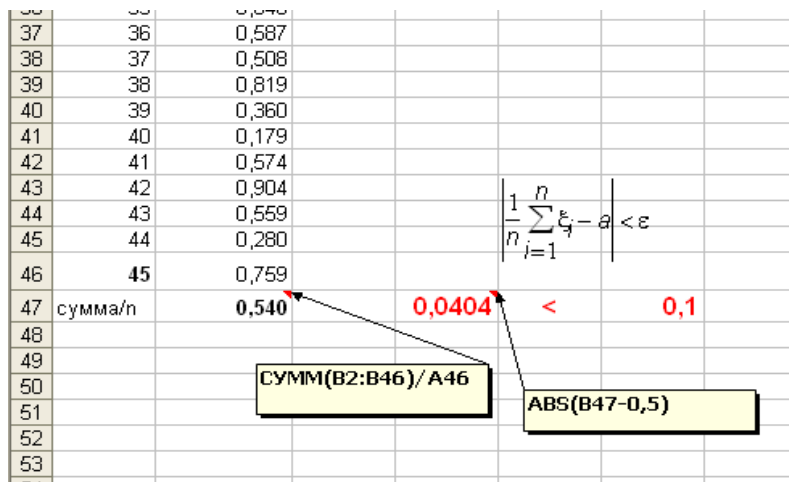


Рис. 7 Реализация выполнения теоремы Чебышева в MS Excel

3. Сжатие распределения с ростом числа слагаемых

Закон больших чисел в форме Чебышева означает, что распределение случайной величины

$$\bar{\xi}_n = \frac{1}{n} \sum_{i=1}^n \xi_i$$

сжимается с ростом n . Если математические ожидания одинаковы, т.е. $M\xi_i = a$, то сжатие происходит в окрестности точки a .

Аналитически иллюстрировать сжатие можно, если распределение для $\bar{\xi}_n$ легко выписывается. Например, если ξ_i распределены нормально $N(a, \sigma^2)$, то случайная величина $\bar{\xi}_n$ распределена по $N(a, \sigma^2/n)$. Построим графики плотностей для $n = 1, 4, 25, 100$ и $\sigma = 1, a = 1$ (сделаем это в целях освоения пакета).

Статистически убедиться в сжатии можно, наблюдая гистограммы при различных значениях n (например, для $n = 10, 40, 160, 640$). Сгенерируем k раз (например, хотя бы $k = 20$) случайную величину $\bar{\xi}_n \equiv \bar{\xi} : \bar{x}_1, \dots, \bar{x}_k$ (рис.8) и

построим для этой выборки средних гистограмму H_n . Сравнивая гистограммы для различных n , мы заметим сжатие (сделать самостоятельно).

	A	B	C	D	E	F	G
1							
2			СлЧисло X_i				
3		n	X1	X2	X3	X4	X5
4		1	0,827	0,194	0,858	0,899	0,519
5		2	0,913	0,285	0,704	0,168	0,958
6		3	0,789	0,483	0,146	0,073	0,257
7		4	0,222	0,629	0,569	0,579	0,019
297		294	0,216	0,002	0,121	0,743	0,873
298		295	0,263	0,496	0,147	0,075	0,557
299		296	0,272	0,997	0,751	0,517	0,412
300		297	0,213	0,122	0,122	0,122	0,656
301		298	0,101	0,122	0,122	0,122	0,122
302		299	0,378	0,018	0,028	0,403	0,381
303		300	0,474	0,516	0,693	0,188	0,700
304							
305	Средние по	50	0,501	0,584	0,506	0,419	0,491
306	n=	100	0,502	0,524	0,471	0,482	0,522
307	выборкам	200	0,509	0,489	0,500	0,473	0,537
308		300	0,474	0,495	0,528	0,484	0,521
309							
310		Y	50	50	50	50	50
311		для построения	100	100	100	100	100
312			200	200	200	200	200
313			300	300	300	300	300

Рис. 8 Генерация выборок в MS Excel

Сжатие можно увидеть на рис. 9 определением для каждого n по $\bar{x}_1, \dots, \bar{x}_k$ минимального $\bar{x}_{\min}$, максимального $\bar{x}_{\max}$ значений и размаха $w = \bar{x}_{\max} - \bar{x}_{\min}$.

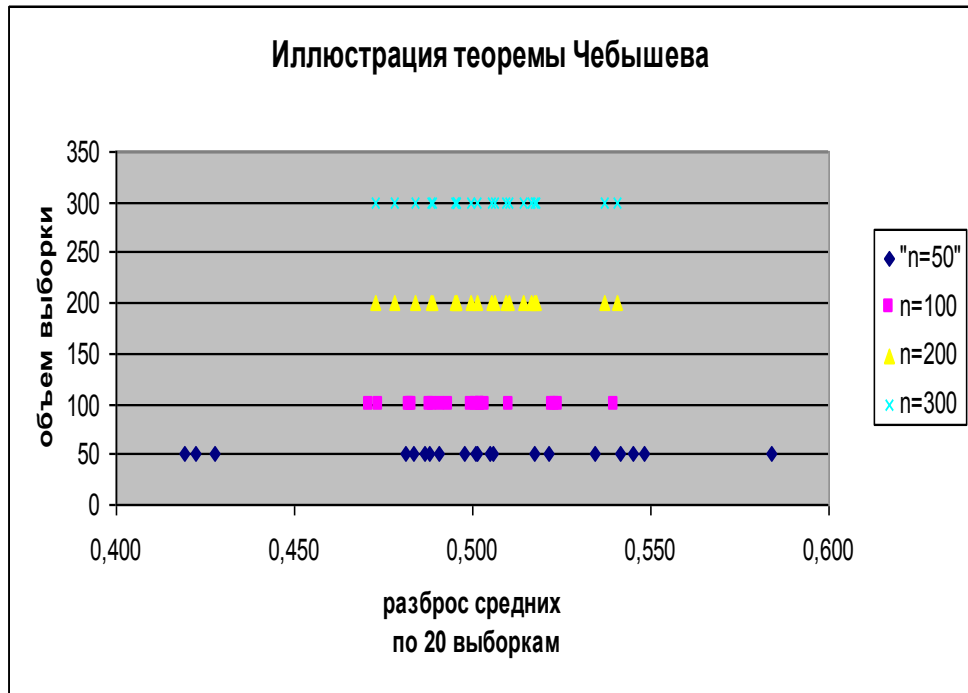


Рис. 9 Реализация сжатия распределения с ростом числа слагаемых в MS Excel

4. Центральная предельная теорема (ЦПТ)

Содержание теоремы Закон больших чисел утверждает, что при $n \rightarrow \infty$

$$\frac{1}{n} \sum_{i=1}^n \xi_i \rightarrow a,$$

где $a = M\xi_i$. Центральная предельная теорема утверждает нечто большее, а, именно, что при этом стремлении происходит нормализация:

$$\frac{1}{n} \sum_{i=1}^n \xi_i \sim N\left(c, \frac{\sigma^2}{n}\right), \quad (10)$$

где $\sigma^2 = D\xi_i$, т.е. среднеарифметическое при больших n распределено приближенно по нормальному закону с дисперсией σ^2/n ; этот факт записывают иначе, нормируя сумму:

$$\frac{\sum_{i=1}^n \xi_i - na}{\sigma\sqrt{n}} \sim N(0,1).$$

Сгенерируем **12** раз случайную величину $\bar{\xi}_n \equiv \bar{\xi} : \bar{x}_1, \dots, \bar{x}_k$ (рис. 10) и построим для выборки средних гистограмму (рис. 10, рис. 11).

Сл.Число X_i													СУММ(B4:M4)/12	
n	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Сумма X_i/n	
1	0,432	0,918	0,270	0,426	0,762	0,314	0,787	0,557	0,551	0,007	0,930	0,705	0,555	
2	0,299	0,471	0,129	0,422	0,891	0,026	0,677	0,543	0,518	0,799	0,639	0,777	0,516	
3	0,236	0,033	0,035	0,063	0,365	0,639	0,849	0,057	0,047	0,291	0,097	0,625	0,278	
4	0,688	0,937	0,879	0,719	0,043	0,785	0,575	0,215	0,786	0,045	0,556	0,079	0,526	Центра
5	0,259	0,409	0,557	0,109	0,412	0,883	0,535	0,368	0,198	0,147	0,260	0,042	0,348	Солонж

Рис. 10 Генерация выборок для иллюстрации ЦПТ в MS Excel

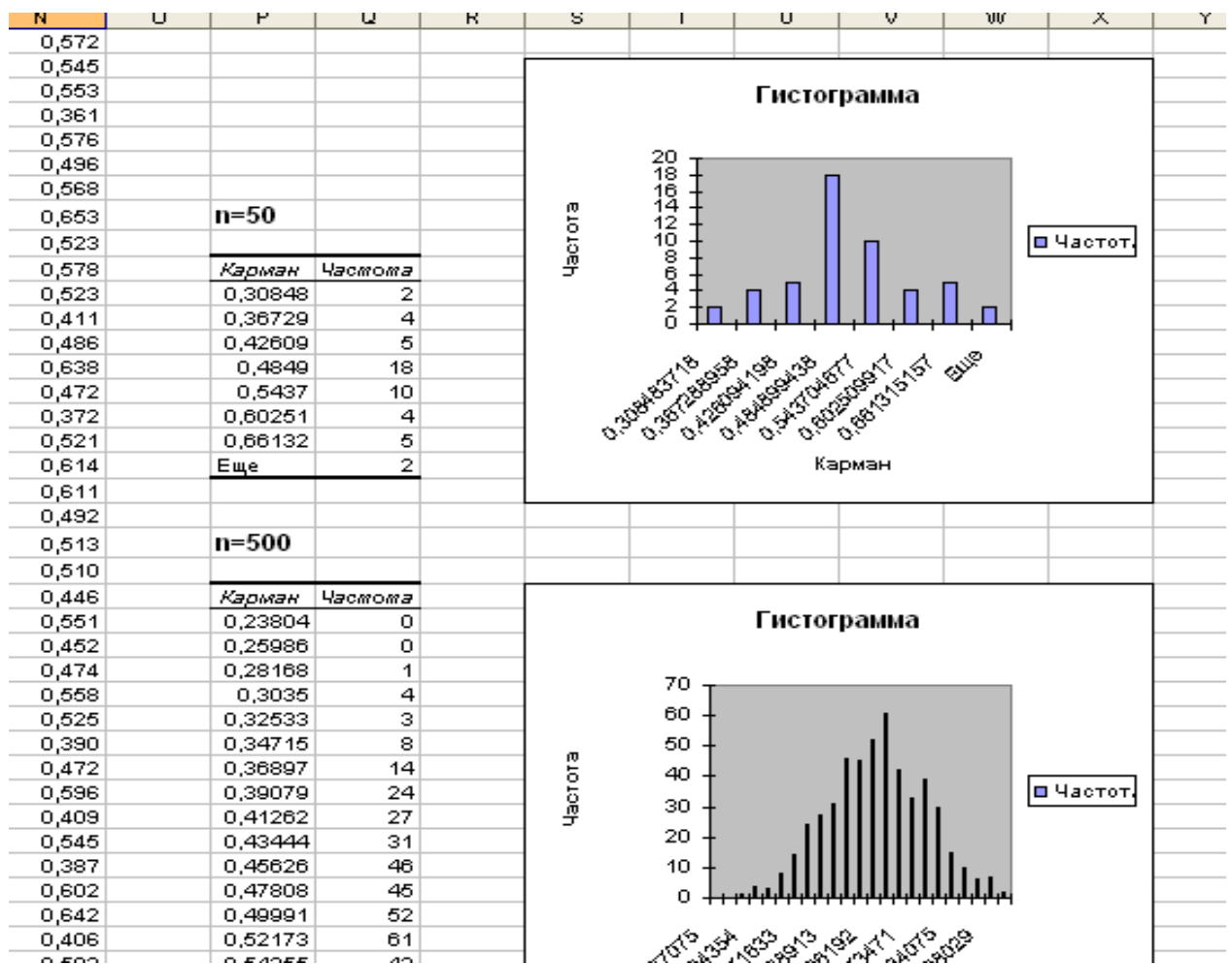


Рис. 11 Сравнение гистограммы с нормальной плотностью.

Лабораторная работа № 3. Описательные статистики. Расчет точечных и интервальных оценок генерального математического ожидания и дисперсии

Цель С использованием MS Excel провести расчет точечных и интервальных оценок генерального математического ожидания и дисперсии.

Задание

1. Сгенерировать выборку объема $n=100$ из заданного закона распределения (согласно вариантам заданий) используя стандартный метод, что интегральная функция распределения любой непрерывной случайной величины равномерно распределена в интервале $(0, 1)$. Использовать статистические функции Excel: СЛЧИС();НОРМОБР() или выражение для обратной функции.
2. Используя **Анализ данных - Описательная статистика** получить числовые характеристики выборки
3. Построить эмпирическую функцию распределения и гистограмму.
 - а) использовать статистические функции СЧЁТЕСЛИ(), ЧАСТОТА() для расчета частоты попадания случайной величины в частичные интервалы.
Наилучшим числом интервалов группировки (формула Стерджесса) является $k = k(n) = 1 + [\log_2 n]$.
 - б) использовать **Анализ данных - Гистограмма** для построения гистограммы
4. Получить точечные оценки математического ожидания и дисперсии
5. Сгенерировать выборку объема $N=100$ биномиального распределения используя "**Анализ данных – Генерация случайных чисел**" для заданного значения вероятности успеха в одном испытании и числа испытаний n (по вариантам заданий).
Рассчитать оценку генеральной доли через частотность события.
6. Для выборки из задания 1 получить интервальные оценки математического ожидания и дисперсии в случаях с уровнем надежности 0,95.
 - 1) Оцениваемый параметр μ
 - а) σ известно
 - б) σ неизвестно
 - 2) Оцениваемый параметр σ
 - а) $n \leq 30$
 - 1) μ известно
 - 2) μ неизвестно
 - б) $n > 30$

Пример выполнения (пошагово)

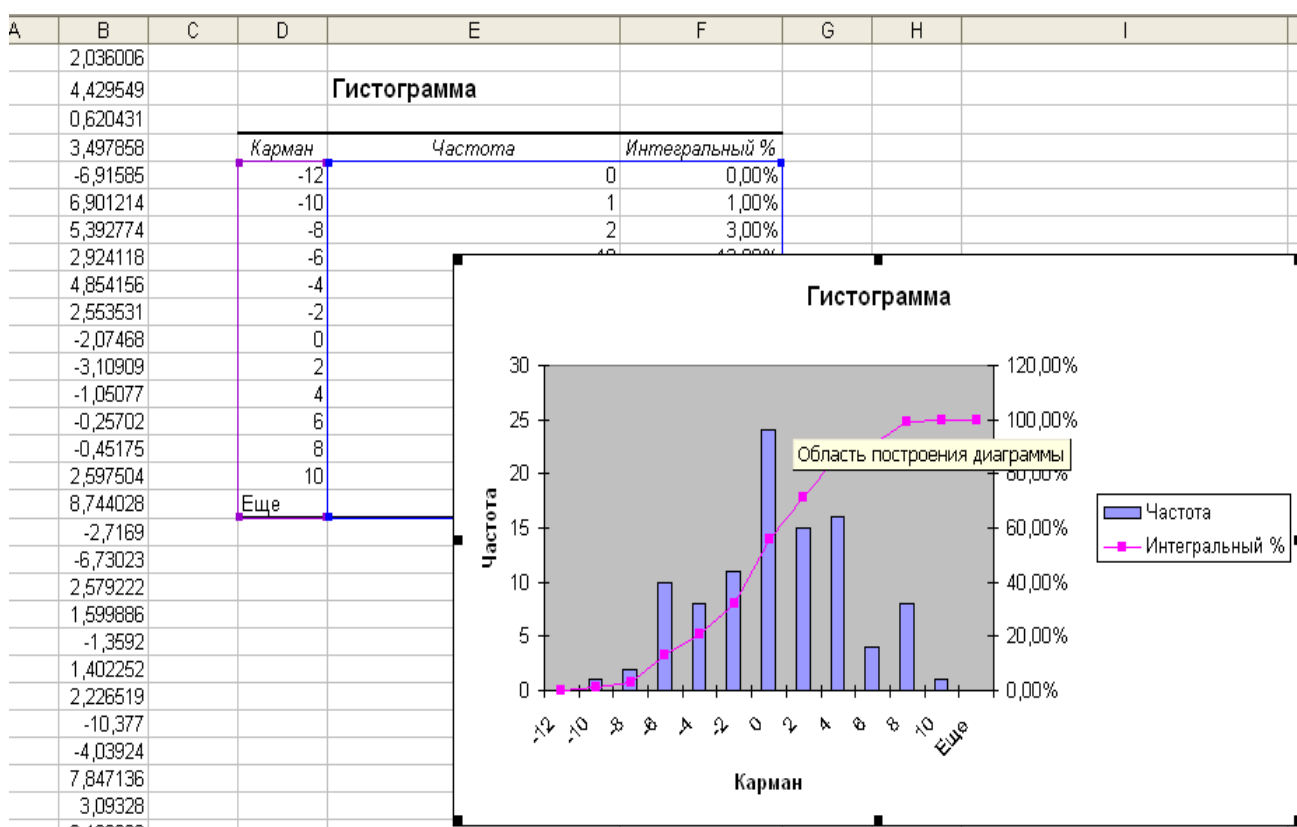
Шаг 1. Генерация выборки из $N(0,4)$. Получение описательных статистик (генерацию выполнить по вариантам)

	A	B	C	D	E	F	G
1		N (0, 4)			Анализ данных- Описательная статистика		
2		-6,14149			Столбец1		
3		-2,36086					
4		1,083499			Среднее	-0,466	
5		2,718447			Стандартная ошибка	0,419	
6		-3,70821			Медиана	-0,694	
7		-0,71392			Мода	#N/D	
8		-1,78405			Стандартное отклонение	4,191	
9		2,085715			Дисперсия выборки	17,565	
10		-2,74788			Эксцесс	-0,309	
11		-1,35351			Асимметричность	-0,023	
12		2,165786			Интервал	19,121	
13		1,676333			Минимум	-10,377	
14		2,472435			Максимум	8,744	
15		-9,11737			Сумма	-46,634	
16		-2,10643			Счет	100,000	
17		-1,8283			Наибольший(1)	8,744	
18		-4,63462			Наименьший(1)	-10,377	
19		-0,67376			Уровень надежности(95,0%)	0,832	
20		3,0835					
21		0,652264					

Варианты заданий к п. 1. Нормальное распределение $N(\mu, \sigma)$

вариант	μ	σ	вариант	μ	σ
1	2	0,5	16	720	50
2	12	2	17	842	20
3	50	5	18	840	50
4	7	1	19	758	10
5	20	4	20	776	40
6	46	6	21	893	60
7	16	3	22	855	30
8	28	7	23	756	70
9	24	5	24	784	50
10	37	4	25	801	40
11	24	6	26	846	60
12	31	5	27	755	50
13	18	3	28	720	30
14	28	4	29	600	40
15	31	5	30	560	50

Шаг 2. Построение гистограммы



Шаг 3. Получение оценок для математического ожидания и дисперсии

	A	B	C	D	E	F	G	H	I	J
1		N (1, 4)	(Xi-μ)^2	n=	100	μ=1	σ ² =16			
2	6,699	2,280103	1,63866356							
3	-3,445	4,045841	9,277147397							
4	0,039	3,01454	4,058370272							
5	-4,204	0,659279	0,116090885	1) Оценка математического ожидания						
6	-3,871	0,207759	0,627645408	СРЗНАЧ(B2:B101)						
7	6,774	5,509669	20,33711317	Хср=	0,7206	2)	DXср=σ ² /n	0,16		
8	-0,337	3,50153	6,257654784	3)	Смещенная оценка дисперсии - функция			ДИСП(B2:B101)		
9	-1,371	1,453056	0,205259542	S2=	15,99		S=	3,999		
10	1,213	-2,61895	13,09678454							
11	2,963	0,244275	0,571120837	Несмещенная оценка дисперсии функция					ДИСП(B2:B101)	
12	-1,585	1,16444	0,027040593	S2=	16,456		S=	4,057		
13	-3,692	17,73722	280,1344785							
14	-5,197	-2,75337	14,08775127	4)	Для известного M(X) μ=1			по формуле 2.11 используя столбец (		
15	3,544	-0,43705	2,065122542	S2=	16,068		S*=	4,009		
16	1,155	1,511891	0,262032665							
17	0,025	1,997146	0,994300366							
18	5,551	-1,49167	6,20840346							

Шаг 4. Генерация случайных чисел биномиального закона

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Выборка X путем генерации случайных чисел из "Анализ данных"										N=	100		p=0,65	n=6
2															
3		событие A произошло в 6 независимых испытаниях													
4		ровно m раз			частотность =m/N										
5	6	m=	0	0	0										
6	1		1	3	0,03										
7	5		2	8	0,08										
8	2		3	24	0,24										
9	5		4	28	0,28										
10	4		5	27	0,27										
11	3		6	10	0,1										
12	5	Сумма		100	1										
13	6														
14	4														
15	2														
16	5														
17	5														
18	3														
19	5														
20	6														
21	5														
22	3														
23	5														
24	4														
25	5														
26	3														
27	4														
28	5														
29	3														
30	5														
31	5														
32	4														

1.2.5. Биномиалы

Биномиальное распределение

Число переменных: 1

Число случайных чисел: 100

Распределение: Биномиальное

Параметры

Значение p = 0,65

Число испытаний = 6

Случайное рассеивание:

Параметры вывода

☒ Выходной интервал: \$A\$3

☐ Новый рабочий лист:

☐ Новая рабочая книга

Готово

Описательная статистика / Точечные оценки / Биномиальное расп / Интервальные оценки

Готово

$$P_{m,n} = \frac{n!}{m!(n-m)!}$$

$$X = \begin{cases} m \\ c_n^m \end{cases}$$

Шаг 5. Получение интервальных оценок

Рисунок 5										
	A	B	C	D	E	F	G	H	I	J
1	Для И(1,4) используя данные лист "Точечные оценки"									
2	n=	100	σ=	4	х _{ср} =	0,721	s=	4,057		
3	Оцениваемый параметр μ							S ² =	16,46	
4			σ известно					НОРМОБР(E6;0;1)		
5										
6	задали	γ=	0,95	(1+γ)/2=	0,975		найдем	t _γ =	1,960	
7	Δ=	0,784	16*D2/B2^0,5							
8	или используя функцию									
9	Δ=	0,784	ДОВЕРИТ(0,05;D2;100)							
10			σ неизвестно					СТЫЮДРАСПОБР(C12;D11)		
11		степеней свободы =n-1		99						
12		α=1-γ=	0,05		найдем	t _α =	1,984			
13	Δ=	0,809	величина доверительного интервала получилась шире предыдущего случая							
14										
15										
16										
17	Оцениваемый параметр σ		Δ ₁ < σ ² < Δ ₂							
18										
19	для примера пусть n=	25						ХИ2ОБР(B22;D22)		
20			μ известно=1		S _* =	4,009				
21										
22	1-α/2=	0,975	n-1=	24	X ² =	12,4	Δ ₂ =	32,39257	C19*F20^2/F22	
23	α/2=	0,025			X ² =	39,36	Δ ₁ =	10,20487		
24			μ неизвестно			s=	4,057	Δ ₂ =	33,17334	
25								Δ ₁ =	10,45084	
26										
27	n>30	n=	100	Δ ₁ < σ < Δ ₂	σ=	4		НОРМОБР(E29;0;1)		
28										
29	задали	γ=	0,95	(1+γ)/2=	0,975		найдем	t _γ =	1,960	
30										
31								Δ ₂ =	4,750705	
32			Δ ₁ ² < σ ² < Δ ₂ ²					Δ ₁ =	3,586486	
33										
34		12,863	σ ² <	22,5692						
35										
36								(2*C27)^0,5/(((2*C27-3)^0,5-129)*F24		
37										
38										

Выводы: проведен расчет точечных и интервальных оценок генерального математического ожидания и дисперсии с использованием MS Excel.

Справочно

Основные формулы, используемые при интервальном оценивании параметров распределений

Оцениваемый параметр	Условия оценки	Используемое распределение	Основные формулы	Доверительный интервал
μ	σ известно	$\Phi(t)$	$\gamma \leftrightarrow t_\gamma;$ $\Delta = t_\gamma \frac{\sigma}{\sqrt{n}}$	$\bar{x} \pm \Delta$
	σ не известно	St	$\left. \begin{array}{l} \alpha \\ v = n - 1 \end{array} \right\} \leftrightarrow t_\alpha;$ $\Delta = t_\alpha \frac{S}{\sqrt{n-1}}$	
σ^2	$n \leq 30$ μ известно	χ^2	$\left. \begin{array}{l} 1 - \frac{\alpha}{2} \\ v = n - 1 \end{array} \right\} \leftrightarrow \chi_1^2;$ $\left. \begin{array}{l} \frac{\alpha}{2} \\ v = n - 1 \end{array} \right\} \leftrightarrow \chi_2^2;$ $\gamma = P(\chi_1^2) - P(\chi_2^2)$	$\frac{nS_0^2}{\chi_2^2} \leq \sigma^2 \leq \frac{nS_0^2}{\chi_1^2}$
	$n \leq 30$ μ не известно	χ^2	$\left. \begin{array}{l} 1 - \frac{\alpha}{2} \\ v = n - 1 \end{array} \right\} \leftrightarrow \chi_1^2;$ $\left. \begin{array}{l} \frac{\alpha}{2} \\ v = n - 1 \end{array} \right\} \leftrightarrow \chi_2^2;$ $\gamma = P(\chi_1^2) - P(\chi_2^2)$	$\frac{nS^2}{\chi_2^2} \leq \sigma^2 \leq \frac{nS^2}{\chi_1^2}$
	$n > 30$	$\Phi(t)$	$\gamma \rightarrow t_\gamma$	$\frac{\sqrt{2n}}{\sqrt{2n-3} + t_\gamma} \cdot S \leq \sigma \leq S \cdot \frac{\sqrt{2n}}{\sqrt{2n-3} - t_\gamma}$
p	$n \rightarrow \infty$	$\Phi(t)$	$\gamma \leftrightarrow t_\gamma$ $\Delta = t_\gamma \sqrt{\frac{\frac{m}{n} \left(1 - \frac{m}{n}\right)}{n}}$	$\frac{m}{n} \pm \Delta$

Лабораторная работа № 4. Проверка статистических гипотез

Цель С использованием MS Excel проверить гипотезы о значении генеральной средней. Провести проверку гипотезы нормальности генеральной совокупности.

Задание

Для задач 1)-3):

- выполнить t-тест для проверки гипотезы о значении генеральной средней для задач анализ запасов, анализ продаж, анализ клиентов с использованием формул.
- выполнить парный двухвыборочный t-тест для средних для зависимых и независимых выборок с использованием функций Excel «Анализ данных». Провести сравнение результатов.
- выполнить сравнение двух средних с использованием функций Excel «Анализ данных»/ «Однофакторный дисперсионный анализ». Провести сравнение результатов.

Для задачи 4) провести проверку гипотезы нормальности генеральной совокупности.

Задача 1. Уровень материальных запасов на вашей фабрике в прошлом году определяли 12 раз в случайно выбранное время. Результаты таковы: (таблица варианты по столбцам). Значимо ли отличается среднее проведенных измерений уровня материальных запасов от значения $\bar{X}_{ср}$, которое руководство использует для управления бюджетом? Обоснуйте свой ответ.

Варианты заданий:

	Вариант									
n/n	1	2	3	4	5	6	7	8	9	10
1	528	657	794	563	1578	4344	3156	6344	5256	3178
2	498	595	692	441	1168	3321	2337	5321	4437	2768
3	471	543	605	336	820	2450	1640	4450	3740	2420
4	487	574	657	399	1030	2974	2059	4974	4159	2630
5	482	565	642	380	966	2815	1932	4815	4032	2566
6	506	612	720	474	1279	3598	2558	5598	4658	2879
7	490	580	667	411	1070	3074	2139	5074	4239	2670
8	506	612	721	475	1282	3606	2565	5606	4665	2882
9	487	574	657	399	1028	2971	2057	4971	4157	2628
10	494	587	678	424	1114	3184	2227	5184	4327	2714
11	521	642	771	535	1482	4106	2965	6106	5065	3082
12	518	637	761	523	1445	4012	2890	6012	4990	3045
$\bar{X}_{ср}$	500	600	700	450	1200	3400	2400	5400	4500	2800
	Вариант									
n/n	11	12	13	14	15	16	17	18	19	20
1	813	1538	1732	1477	2221	2440	3467	2554	4081	3054

2	771	1420	1631	1393	2154	2314	3257	2386	3829	2886
3	852	1645	1824	1554	2283	2556	3659	2707	4311	3207
4	721	1279	1511	1292	2074	2164	3006	2185	3527	2685
5	804	1512	1710	1459	2207	2413	3422	2517	4026	3017
6	889	1749	1913	1628	2342	2667	3845	2856	4534	3356
7	804	1511	1710	1458	2206	2412	3420	2516	4024	3016
8	873	1704	1875	1595	2316	2618	3764	2791	4436	3291
9	796	1488	1690	1442	2193	2387	3379	2483	3975	2983
10	826	1574	1763	1503	2242	2479	3532	2605	4158	3105
11	740	1333	1557	1331	2105	2221	3102	2262	3642	2762
12	792	1478	1681	1434	2187	2376	3361	2468	3953	2968
Хср	800	1500	1700	1450	2200	2400	3400	2500	4000	3000
Вариант										
n/n	21	22	23	24	25	26	27	28	29	30
1	844	1622	1805	1537	2270	2531	3618	2674	4261	3174
2	868	1691	1864	1587	2309	2605	3742	2773	4410	3273
3	774	1428	1638	1398	2159	2322	3271	2396	3845	2896
4	765	1402	1616	1380	2144	2295	3224	2359	3789	2859
5	729	1301	1529	1308	2086	2186	3044	2215	3573	2715
6	809	1525	1721	1468	2214	2426	3444	2535	4053	3035
7	849	1637	1817	1548	2278	2547	3645	2696	4293	3196
8	740	1332	1556	1330	2104	2219	3099	2259	3639	2759
9	784	1455	1662	1418	2175	2352	3320	2436	3905	2936
10	847	1631	1812	1544	2275	2540	3634	2687	4281	3187
11	858	1662	1839	1566	2293	2574	3690	2732	4348	3232
12	890	1752	1916	1630	2344	2670	3850	2860	4540	3360
Хср	800	1500	1700	1450	2200	2400	3400	2500	4000	3000

Задача 2. Туристическая фирма в среднем реализует X путевку в день. Количество проданных путевок за последнюю неделю составило (варианты по столбцам). Полагая, что уровень значимости равен 5%, определить:

- выполнит ли фирма план по продаже путевок
- увеличилось ли среднее число продаж путевок за неделю.

Варианты заданий:

	Вариант									
n/n	1	2	3	4	5	6	7	8	9	10
1	14	16	23	20	30	42	24	20	162	91
2	8	10	15	12	16	28	13	12	104	62
3	11	13	19	16	24	36	19	16	134	77
4	10	12	18	15	21	33	17	15	124	72
5	10	12	17	14	19	31	16	14	118	69
6	12	14	20	17	26	38	21	17	143	81
7	15	17	25	22	33	45	27	22	173	96
X	10	12	17	14	20	32	16	14	120	70
Вариант										
n/n	11	12	13	14	15	16	17	18	19	20
1	45	58	55	65	40	29	90	65	33	67

2	16	32	26	36	14	-22	-19	-15	15	30
3	48	60	58	68	42	34	101	72	35	70
4	24	39	34	44	21	-8	11	7	20	40
5	57	68	67	77	50	49	133	96	40	81
6	47	59	57	67	41	32	96	69	34	69
7	48	60	58	68	42	35	102	73	35	71
X	40	53	50	60	35	20	70	50	30	60
	Вариант									
n/n	21	22	23	24	25	26	27	28	29	30
1	23	32	43	36	21	26	13	27	23	23
2	32	40	52	59	26	31	16	34	28	35
3	26	35	46	45	23	28	14	30	25	28
4	23	32	43	36	21	26	13	27	23	23
5	10	22	30	6	15	20	8	18	15	8
6	10	21	30	5	15	20	8	17	15	7
7	31	39	51	57	25	30	16	33	28	33
X	20	30	40	30	20	25	12	25	21	20

Задача 3. Имеются данные о количестве выданных пластиковых кредитных карт за 30 дней. Начальник отдела кредитования хочет проверить предположение о том, что отделением в среднем выдается 350 карт в день. Полагая, что уровень значимости $\alpha=0.05$, определите: какой вывод сделает отдел маркетинга на основе статистической информации. Данные представлены в таблице.

день	Номера отделений						
	1	2	3	4	5	6	7
1	300	280	340	350	293	345	298
2	290	270	342	370	311	367	330
3	295	265	335	360	290	372	321
4	280	256	338	352	295	354	309
5	275	279	341	347	280	354	298
6	305	254	330	350	325	347	295
7	310	302	327	320	343	348	278
8	315	307	325	324	354	370	290
9	290	304	343	329	360	380	327
10	295	300	354	326	339	381	325
11	293	298	360	332	337	352	343
12	311	330	339	343	360	324	354
13	314	321	337	345	352	356	360
14	322	309	312	360	347	345	356
15	320	298	330	356	314	367	368
16	317	295	325	368	322	387	371
17	305	278	333	371	320	369	358
18	345	290	340	358	317	345	363

19	312	295	314	363	305	352	325
20	320	293	322	345	345	376	343
21	332	311	320	354	298	342	354
22	330	290	317	356	330	346	360
23	310	295	339	327	321	368	339
24	325	280	335	325	309	367	315
25	333	330	327	343	298	365	290
26	340	321	345	354	322	387	295
27	342	309	348	360	335	354	293
28	345	310	346	339	337	387	311
29	330	312	342	337	348	387	314
30	335	333	335	339	350	383	323

Варианты заданий

Вариант задания	Номера отделений	Вариант задания	Номера отделений	Вариант задания	Номера отделений
1	1,2	7	3,4	13	1,7
2	1,3	8	5,6	14	2,4
3	1,4	9	5,7	15	3,7
4	1,5	10	2,5	16	2,5
5	2,3	11	3,5	17	4,5
6	2,4	12	1,6	18	1,6

Задача 4. Сгенерировать нормально распределенный признак с параметрами **Mean** (среднее) и **StDev** (стандартное отклонение) (по вариантам) с наложением “шума” $\text{StDev} \cdot (1 - \text{rnd}(1))$ с использованием функции

НОРМОБР(СЛЧИС(); Mean; StDev) - StDev * (СЛЧИС() - 0,5).

Провести проверку гипотезы нормальности генеральной совокупности.

Варианты заданий

вариант	1	2	3	4	5	6	7	8	9	10
Mean	100	30	120	400	300	250	120	550	300	750
St.Dev	20	4	15	25	5	12	7	30	20	50

вариант	11	12	13	14	15	16	17	18	19	20
Mean	4	6	7	12	13	3	13	14	52	62
St.Dev	2	1	3	2	5	1	3	3	12	5

вариант	21	22	23	24	25	26	27	28	29	30
Mean	10	30	12	40	30	25	12	55	30	75
St.Dev	2	4	3	5	3	4	3	13	7	5

Лабораторная работа № 5. Статистическое исследование зависимостей

Цель Проведение с применением пакета MS Excel статистического исследования выборочной совокупности предприятий по признакам *Среднегодовая стоимость основных производственных фондов* и *Выпуск продукции*.

Задание

1) изучить структуру исследуемой совокупности предприятий, на основе обобщающих показателей выявляются типические черты и закономерности, присущие наблюдаемой совокупности в целом; провести анализ распределения единиц совокупности по признаку *Среднегодовая стоимость основных производственных фондов* (признак X) и *Выпуск продукции* (признак Y).

- Выявление и исключение из выборки аномальных единиц наблюдения.
- Оценка описательных статистических параметров совокупности.
- Построение гистограмм по признаку X и Y.

2) изучить взаимосвязь между факторным признаком *Среднегодовая стоимость основных производственных фондов* (признак X) и результативным признаком *Выпуск продукции* (признак Y).

3) считая, что исходные данные есть некоторые экспертные оценки (при отсутствии информации) *стоимости основных производственных фондов* и *выпуску продукции*, построить ранговую таблицу данных и посчитать коэффициент ранговой корреляции Спирмена.

4) построить однофакторную линейную регрессионную модель связи признаков X и Y, используя инструмент Регрессия надстройки Пакет анализа, и рассчитать доверительные интервалы коэффициентов уравнения линейной регрессии.

Построить теоретическую кривую регрессии. Дать экономическую интерпретацию коэффициента регрессии. Рассчитать коэффициент эластичности и дать его экономическую интерпретацию.

5) с использованием корреляционного отношения дать оценку нелинейности связи. В случае нелинейности найти наиболее адекватное уравнение регрессии с помощью средств инструмента Мастер диаграмм.

6) введем еще две переменные, влияющие на выпуск продукции – Производительность труда и Коэффициент использования основных фондов (загруженность оборудования). Построить множественную регрессию

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3.$$

Провести регрессионный анализ совокупности по признакам X1, X2, X3. с использованием функции **Регрессия в Анализ данных**.

Дать его экономическую интерпретацию полученных результатов.

Пример:

Задание 1

Шаг 1. В пакете **Анализ данных** используем функцию **Описательная статистика** последовательно, сначала для ячеек фактора X (входной интервал), затем то же самое делаем для признака Y.

Окно функции:

Получим следующие результаты:

X		Y	
Среднее	341,3	Среднее	329,9
Стандартная ошибка	10,4	Стандартная ошибка	12,6
Медиана	348,5	Медиана	325,0
Мода	352,0	Мода	325,0
Стандартное отклонение	58,8	Стандартное отклонение	71,5
Дисперсия выборки	3457,2	Дисперсия выборки	5117,8
Эксцесс	-0,2	Эксцесс	-0,3
Асимметричность	-0,2	Асимметричность	-0,1
Интервал	250,0	Интервал	300,0
Минимум	215,0	Минимум	175,0
Максимум	465,0	Максимум	475,0
Сумма	10921,0	Сумма	10556,0
Счет	32,0	Счет	32,0
Наибольший(1)	465,0	Наибольший(1)	475,0
Наименьший(1)	215,0	Наименьший(1)	175,0
Уровень надежности(95,0%)	21,2	Уровень надежности(95,0%)	25,8

В процессе исследования совокупности необходимо решить ряд статистических задач для выборочной и генеральной совокупностей.

Статистический анализ выборочной совокупности:

1. Выявить наличие среди исходных данных резко выделяющихся значений признаков («выбросов» данных) с целью исключения из выборки аномальных единиц наблюдения.
2. Рассчитать обобщающие статистические показатели совокупности по изучаемым признакам: среднюю арифметическую, моду, медиану, размах вариации, дисперсию, средние отклонения – линейное и квадратическое, коэффициент вариации, структурный коэффициент асимметрии К.Пирсона.
3. На основе рассчитанных показателей в предположении, что распределения единиц по обоим признакам близки к нормальному, оценить:
 - а) степень колебаний значений признаков в совокупности;
 - б) степень однородности совокупности по изучаемым признакам;
 - в) устойчивость индивидуальных значений признаков;
 - г) количество попаданий индивидуальных значений признаков в диапазоны $(x \pm \sigma)$, $(x \pm 2\sigma)$, $(x \pm 3\sigma)$
4. Дать сравнительную характеристику распределений единиц совокупности по двум изучаемым признакам на основе анализа:
 - а) вариации признаков;
 - б) количественной однородности единиц;
 - в) надежности (типичности) средних значений признаков;
 - г) симметричности распределений в центральной части ряда.

Аномальных наблюдений не наблюдается. Можно проверить на аномальность минимальное и максимальное значения в выборке.

Шаг 2. Построим интервальный вариационный ряд и гистограмму распределения единиц совокупности по признакам X и Y с использованием функции **Гистограмма** в **Анализ данных**

Окно функции:

Гистограмма

Входные данные

Входной интервал: \$B\$3:\$B\$34

Интервал карманов:

☐ Метки

Параметры вывода

☒ Выходной интервал: \$E\$28

☐ Новый рабочий лист:

☐ Новая рабочая книга

☐ Парето (отсортированная гистограмма)

☐ Интегральный процент

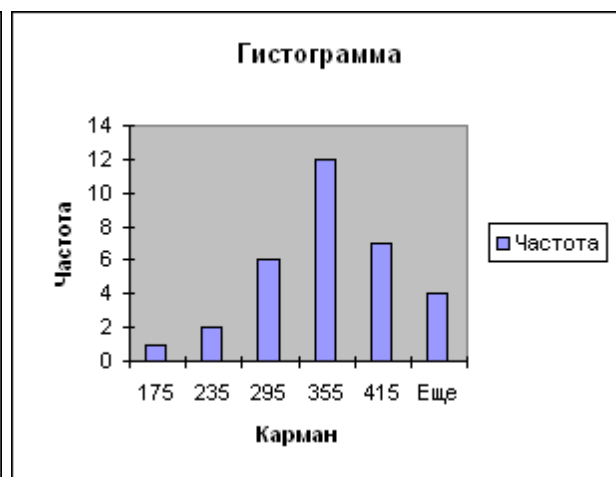
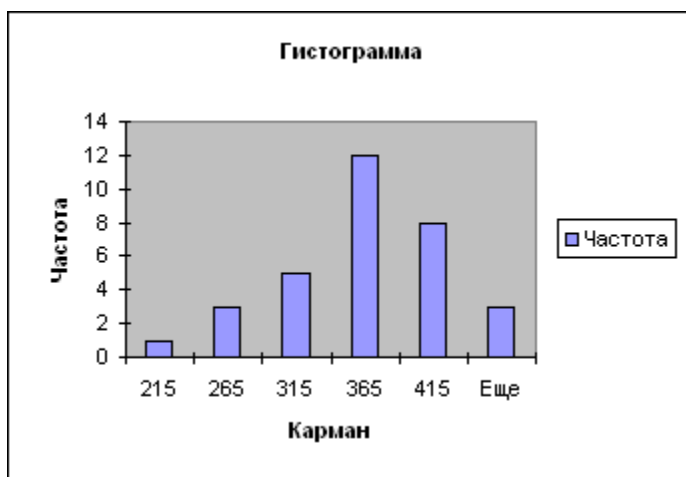
☐ Вывод графика

OK Отмена Справка

Получим следующие результаты:

Карман	Частота
215	1
265	3
315	5
365	12
415	8
Еще	3

Карман	Частота
175	1
235	2
295	6
355	12
415	7
Еще	4



Визуальный анализ гистограмм, а также основных описательных статистик позволяет выдвинуть гипотезу о нормальном законе распределении генеральных совокупностей X и Y .

Для расчета теоретических частот используем группированные по интервалам (карманам) данные, полученные на предыдущем шаге. Теоретическая частота рассчитывается по нормальному закону распределения так:

$$=n \cdot h \cdot \text{НОРМРАСП}(x; \mu; \sigma; \text{ЛОЖЬ}),$$

где n - объем выборки, h - величина интервала, x - значение «кармана»; μ - выборочное среднее; σ - стандартное отклонение; ЛОЖЬ - для функции НОРМРАСП возвращает плотность вероятности. Пример из расчета $1,08 = 32 \cdot 50 \cdot \text{НОРМРАСП}(G48; 341,3; 58,8; \text{ЛОЖЬ})$

Для X

ср инт	по частота	теорет частота		
215	1	1,08	0,00655	
265	3	4,68	0,93804	
315	5	9,82	4,6508	
365	12	10,01	0,33046	
415	8	4,95	1,16363	
455	3	1,67	0,58622	
		Хи2_Набл	7,68	
		Хи2_Крит	7,81473	
		Хи2_Набл<=Хи2_Крит		
		нет оснований отвергать		
		гипотезу нормальности X		

Для Y

ср инт	по частота	теорет частота		
175	1	0,20	0,6417	
235	2	2,12	0,00699	
295	6	7,96	0,64151	
355	12	10,56	0,17163	
415	7	4,95	0,60099	
465	4	1,19	1,9776	
		Хи2_Набл	4,04	
		Хи2_Крит	7,81473	
		Хи2_Набл<=Хи2_Крит		
		нет оснований отвергать		
		гипотезу нормальности Y		

Выводы: X и Y – две нормальные генеральные совокупности, к которым можно обоснованно применять методы корреляционно-регрессионного анализа.

Задание 2

Результаты с применением статистических функций:

коэф.корреляции	
Пирсона=	0,91
Значение Т критерия	
T=	11,97
Tkr (Стьюдента)	2,04
T>Tkr, => нулевую гипотезу о равенстве нулю выборочного коэфф. корреляции Пирсона отвергаем	

коэф.корреляции Пирсона
 =ПИРСОН(B3:B34;C3:C34)

Значение Т критерия
 =B41*КОРЕНЬ(A34-2)/КОРЕНЬ(1-B41*B41)

Tkr (Стьюдента)
 =СТЮДРАСПОБР(0,05;A34-2)

Вывод: Связь между X и Y сильная

Задание 3

Считая, что исходные данные есть некоторые экспертные оценки (при отсутствии информации) *стоимости основных производственных фондов и выпуску продукции*, построить ранговую таблицу данных и посчитать коэффициент ранговой корреляции Спирмена по формуле:

$$r_s = 1 - 6 \cdot \frac{\sum_{i=1}^n (R_{x_i} - R_{y_i})^2}{n^3 - n}$$

Пересчитаем ранги X и Y используя статистическую функцию РАНГ(x_i ; $x_1:x_n$; 1). Затем по формуле найдем r_s .

Реализация:

Ранг X	Ранг Y	($x_i - y_i$) ²	
4	6	4	
9	8	1	
10	13	9	
14	21	49	
1	1	0	
17	10	49	
21	28	49	
5	7	4	
13	15	4	
25	27	4	
22	25	9	
30	29	1	
11	19	64	
17	22	25	
29	31	4	
32	32	0	
16	14	4	
24	24	0	
8	4	16	
26	16	100	
31	30	1	
7	5	4	
2	3	1	
27	23	16	
17	16	1	
12	11	1	
3	2	1	
15	12	9	
28	20	64	
20	25	25	
23	16	49	
6	9	9	32
		577	
0,893692			
Спирмен= 0,894245			

Задание 4

Проведем регрессионный анализ совокупности по признакам X и Y с использованием функции **Регрессия** в **Анализ данных**. Итоги выведем на новый рабочий лист.

Окно функции:

Вывод результатов

<i>Регрессионная статистика</i>	
Множественный R	0,91
R-квадрат	0,83
Нормированный R-квадрат	0,82
Стандартная ошибка	30,26
Наблюдения	32,00

<i>Дисперсионный анализ</i>					<i>Значимость</i>
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>F</i>
Регрессия	1	131181,88	131181,88	143,27	5,94802E-13
Остаток	30	27469,62	915,65		
Итого	31	158651,50			

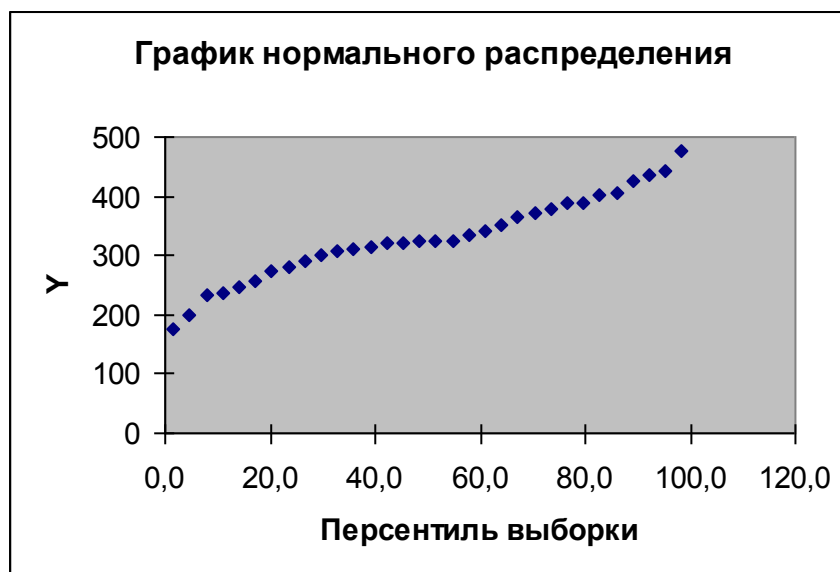
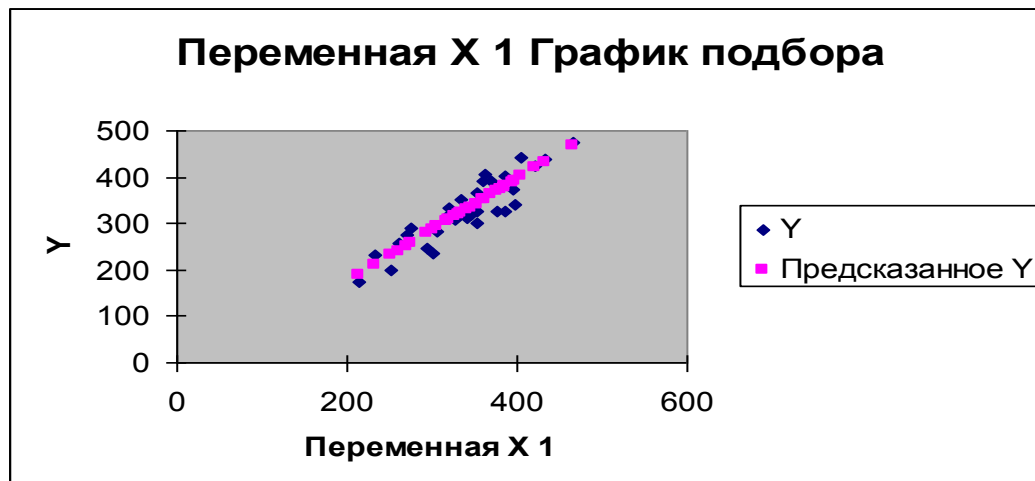
Выполним строгую оценку статистической надежности моделирования с помощью **Ф-критерия Фишера**. Для этого проверим нулевую гипотезу H_0 о статистической незначимости полученного уравнения регрессии. Т.к. расчетное значение F-критерия $F = 143,27$ больше его критического значения ($F_{\text{крит}}$) при очень малом значении уровня значимости $\alpha = 5,94802\text{E-}13$, то

нулевая гипотеза отвергается и полученное уравнение регрессии принимается значимым.

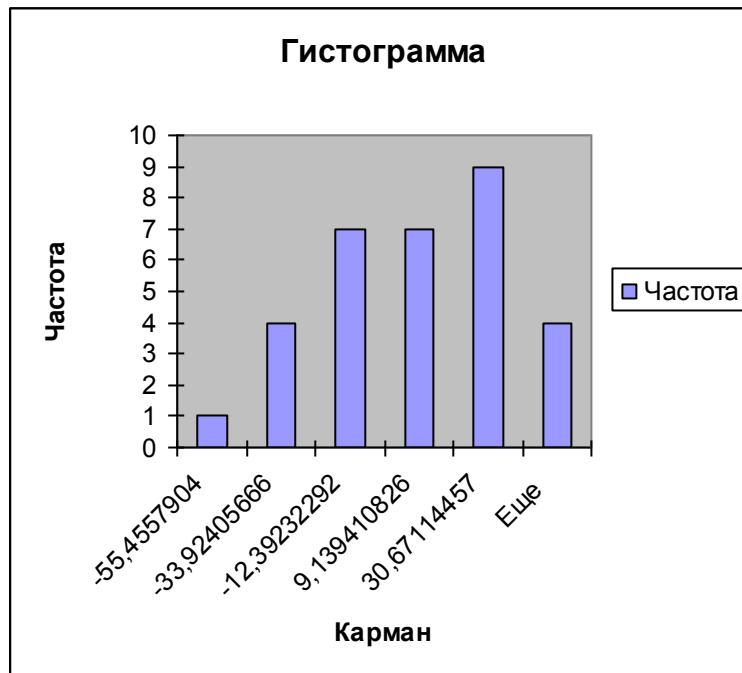
	<i>Козфф. модели</i>	<i>Станд. ошибка</i>	<i>t- статистика</i>	<i>P- Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>	<i>Нижние 95,0%</i>	<i>Верхние 95,0%</i>
Y- пересечение	-47,70	32,00	-1,49	0,15	-113,04	17,64	-113,04	17,64
Переменная X 1	1,11	0,09	11,97	0,00	0,92	1,30	0,92	1,30

Остатки

<i>Y тыс.руб</i>	<i>Предсказанное Y</i>	<i>Остатки</i>
257	239,9	17,1
282	291,9	-9,9
315	303,0	12,0
350	322,9	27,1
175	190,2	-15,2
300	341,7	-41,7
405	352,8	52,2
275	251,0	24,0
322	319,6	2,4
402	378,2	23,8
390	361,6	28,4
425	419,2	5,8
335	306,3	28,7
365	341,7	23,3
442	400,4	41,6
475	466,8	8,2
320	334,0	-14,0
380	374,9	5,1
237	286,4	-49,4
325	380,5	-55,5
437	430,2	6,8
247	278,7	-31,7
232	209,0	23,0
372	389,3	-17,3
325	341,7	-16,7
307	314,1	-7,1
200	231,1	-31,1
312	330,7	-18,7
342	391,5	-49,5
390	350,6	39,4
325	369,4	-44,4
290	256,5	33,5



Гистограмма остатков



Нормальность остатков

	Карман	Частота		
h=	-55,46	1	9,2274	67,69
	-33,92	4	9,1497	6,6297
	-12,39	7	8,9809	0,5606
	9,14	7	8,9809	0,5606
	30,67	9	8,8202	0,0036
	52,20	4	9,1497	6,6297
			Хи2_Набл	82,07
			Хи2_Крит(0,05; 3)	7,8147
			Хи2_Набл > Хи2_Крит	

Вывод: гипотезу нормальности остатков отвергаем для уровня значимости 0,05.

Задание 5

С использованием корреляционного отношения проведем оценку нелинейности связи.

Вычислить корреляционное отношение выборочных данных по формуле

$$\eta_{yx} = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}} = \sqrt{1 - \frac{S_{ост}^2}{S_y^2}}$$

С использованием результатов **Регрессии**, полученных на предыдущем шаге

Sост=SSостаток=27469,6

Sy =SSитого=158651,5

S2ост= 27469,62

$\eta =$

0,90932

S2y= 158651,50

В качестве одного из простых критериев оценки нелинейной связи можно использовать следующий: если $K_H > 2,5$, то корреляционную связь можно считать нелинейной, где

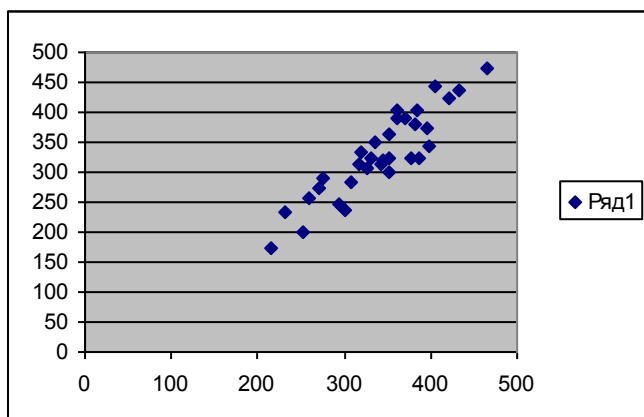
$$K_H = \frac{\sqrt{n}}{0,67499} \cdot \frac{1}{2} \sqrt{\eta_{yx}^2 - r_{yx}^2}.$$

Оценка нелинейности связи

K=

6,244E-08 < 2,5

Вывод: Корреляционную связь можно считать линейной



Задание 6

Исходные данные для новых факторов сгенерировать в Excel по нормальному закону распределения с параметрами Mean и StDev с наложением шума $StDev \cdot (1 - rnd(1))$ следующим образом:

Задать Среднее (Mean) и Стандартное отклонение (St.Dev) для переменных X1 и X2 в ячейки B и C

		B	C
2	Mean	100	0,5
3	St.Dev	20	0,3

С помощью генератора случайных чисел $rnd(1)$ – функция EXCEL: СЛЧИС() генерирует число в интервале (0;1) – задать случайные значения в столбце D (в задаче 32 значения).

Затем, используя функцию EXCEL НОРМОБР(вероятность; среднее; стандартное отклонение) получим требуемые случайные числа X

	D	F	E
	СлЧисло	X1	X2
4			
5	0,081564	79,4	
формула	= СЛЧИС()	=НОРМОБР(\$D5;B\$2;B\$3)- B\$3*(СЛЧИС()-0,5)	=НОРМОБР(\$D5;C\$2;C\$3)- C\$3*(СЛЧИС()-0,5)

Замечание – при выполнении работы у Вас будут ссылки на другие ячейки табличного процессора Excel. В примере ниже - X2 сгенерировано с другими значениями Mean и StDev.

Исходные данные

X1	X2	X3	Y
141,6	110,5	260	257
124,2	108,7	307	282
123,5	115,6	317	315
124,4	110,1	335	350
134,8	115,9	215	175
116	92	352	300
92,5	67,2	362	405
119,5	101,5	270	275
131,2	109,8	332	322
103,8	81	385	402
100,2	72,3	370	390
115,2	106,7	422	425
106	85,4	320	335
149,9	129,8	352	365
114,5	82,1	405	442
109,4	68,7	465	475
118,7	110,9	345	320
131,2	107,9	382	380

112,2	77,3	302	237
127,6	98,5	387	325
128,7	110,8	432	437
123	101,4	295	247
136,1	129,5	232	232
156,3	147,1	395	372
124,5	95	352	325
130,3	109,9	327	307
119,3	114,1	252	200
116,4	87,3	342	312
101,1	75,3	397	342
129,4	107,4	360	390
141,8	113,3	377	325
111,3	77	275	290

Множественная регрессионная линейная модель имеет вид

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3.$$

Проведем регрессионный анализ совокупности по признакам X1, X2, X3 и Y с использованием функции **Регрессия** в **Анализ данных**. Итоги выведем на новый рабочий лист. Окно функции выглядит, как и в случае двух переменных. Необходимо входной интервал X задать как область данных переменных X1, X2, X3.

Соответственно результаты регрессии аналогичные, с отличием - определяются коэффициент множественной регрессии и оценки коэффициентов линейной модели для всех переменных. В данном случае b_0 , b_1 , b_2 , b_3 .

Результаты множественной линейной регрессии

<i>Регрессионная статистика</i>	
Множественный R	0,912786133
R-квадрат	0,833178524
Нормированный R-квадрат	0,815304795
Стандартная ошибка	30,74461473
Наблюдения	32

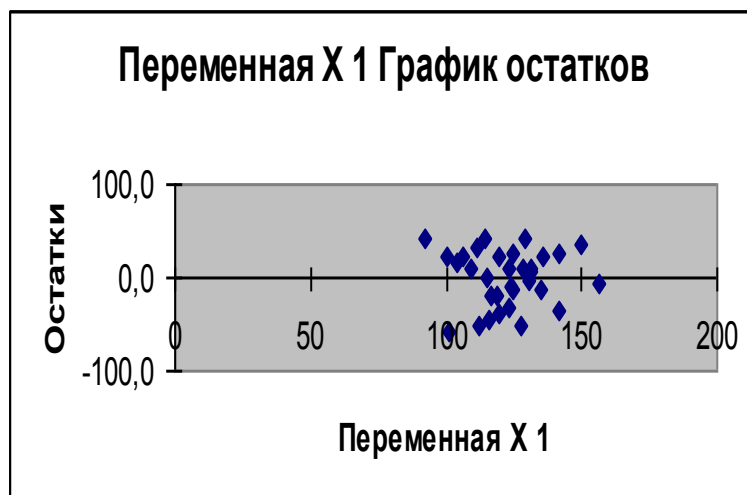
Дисперсионный анализ

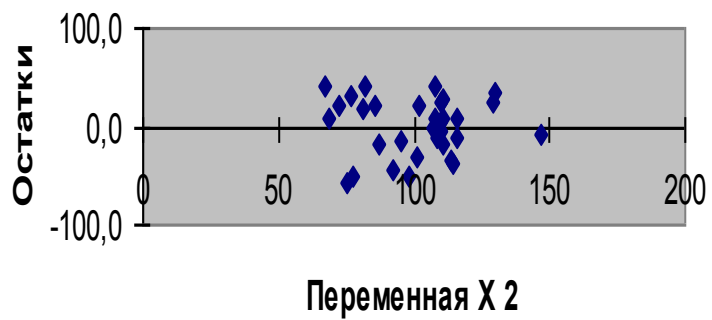
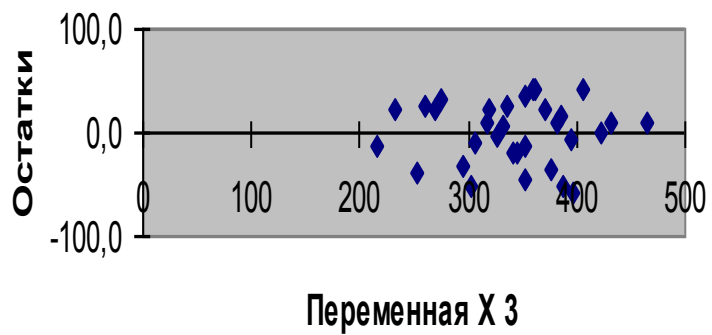
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	3	132185,02	44061,67	46,6147	5,1481E-11
Остаток	28	26466,477	945,2313		
Итого	31	158651,5			

	<i>Коэф. b_i</i>	<i>Станд. ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>	<i>Нижние 95,0%</i>	<i>Верхние 95,0%</i>
Y-перес	11,058	65,873	0,168	0,868	-123,876	145,993	-	-
X 1	-0,699	0,856	-0,817	0,421	-2,452	1,054	-2,452	1,054
X 2	0,289	0,656	0,440	0,663	-1,055	1,632	-1,055	1,632
X 3	1,100	0,098	11,200	0,000	0,899	1,301	0,899	1,301

ВЫВОД ОСТАТКА

Y	<i>Предсказанное Y</i>	<i>Остатки</i>
257	229,9	27,1
282	293,2	-11,2
315	306,7	8,3
350	324,3	25,7
175	186,7	-11,7
300	343,6	-43,6
405	363,9	41,1
275	253,7	21,3
322	316,1	5,9
402	385,2	16,8
390	368,7	21,3
425	425,4	-0,4
335	313,5	21,5
365	330,8	34,2
442	400,1	41,9
475	465,7	9,3
320	339,5	-19,5
380	370,6	9,4
237	287,0	-50,0
325	375,9	-50,9
437	428,1	8,9
247	278,7	-31,7
232	208,4	23,6
372	378,6	-6,6
325	338,5	-13,5
307	311,3	-4,3
200	237,7	-37,7
312	331,0	-19,0
342	398,7	-56,7
390	347,5	42,5
325	359,2	-34,2
290	257,9	32,1



Переменная X 2 График остатков**Переменная X 3 График остатков**

Исходные данные к лабораторной работе 5

При проведении статистического наблюдения за деятельностью предприятий корпорации получены выборочные данные по 32-м предприятиям, выпускающим однородную продукцию, о среднегодовой стоимости основных производственных фондов и о выпуске продукции за год.

Среднегодовая стоимость основных производственных фондов (млн.руб.) (признак X) и Выпуск продукции (признак Y) (млн.руб.).

Варианты заданий

Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5	
X	Y	X	Y	X	Y	X	Y	X	Y
110,5	32,9	141,6	417,9	305,2	256,6	129,7	592,2	323,7	808,3
108,7	30,2	124,2	395,2	301,4	253,2	121,5	567,6	311,4	748,7
115,6	31,3	123,5	419,7	305,1	252,2	126,1	579,2	303,9	764,1
110,1	31,0	124,4	406,9	299,4	247,7	121,6	565,7	309,1	749,9
115,9	33,6	134,8	413,9	302,0	257,3	123,3	580,4	325,9	810,3
92,0	25,8	116,0	374,8	295,5	246,3	116,4	535,5	288,7	718,6
67,2	23,1	92,5	361,4	292,2	226,9	109,7	508,8	257,7	648,0
101,5	28,2	119,5	409,9	297,8	251,1	117,7	553,0	306,1	729,7
109,8	33,9	131,2	409,8	300,1	254,4	121,2	575,1	312,1	798,6
81,0	28,4	103,8	384,0	298,2	246,1	114,0	545,3	287,2	703,1
72,3	26,3	100,2	359,5	295,6	236,7	111,8	513,6	278,1	710,8
106,7	28,5	115,2	392,1	298,4	255,7	122,6	561,1	304,1	740,5
85,4	27,1	106,0	381,0	295,6	233,1	112,1	520,4	285,6	692,7
129,8	37,5	149,9	442,9	309,1	270,1	129,2	586,0	329,4	827,8
82,1	29,7	114,5	391,7	297,6	245,8	114,1	526,2	297,3	738,8
68,7	24,8	109,4	378,9	294,0	235,3	114,2	521,6	271,3	685,5
110,9	30,1	118,7	404,7	302,5	253,1	122,6	551,4	301,4	736,4
107,9	33,2	131,2	418,8	304,1	254,6	124,3	576,2	314,6	785,2
77,3	25,2	112,2	374,4	296,1	237,4	110,3	527,5	280,6	689,2
98,5	29,5	127,6	414,5	299,3	254,0	118,7	557,5	297,6	770,2
110,8	32,8	128,7	408,0	300,8	260,8	122,1	555,5	310,4	763,3
101,4	30,6	123,0	409,4	300,4	248,0	120,3	558,9	288,8	745,3
129,5	34,6	136,1	433,2	304,8	262,7	125,9	585,2	335,7	801,5
147,1	38,3	156,3	453,7	310,3	276,9	140,5	623,2	358,0	873,4
95,0	31,4	124,5	411,8	300,1	254,3	119,2	554,7	312,8	775,3
109,9	34,5	130,3	425,0	302,7	254,4	128,9	585,8	309,8	789,3
114,1	29,6	119,3	409,7	302,7	255,4	121,8	557,4	308,7	784,9
87,3	26,9	116,4	377,1	297,1	246,1	118,7	533,4	297,7	724,3
75,3	23,4	101,1	368,8	294,8	233,2	112,7	498,2	269,6	673,1
107,4	30,3	129,4	414,8	299,5	250,5	120,9	559,7	297,3	782,2
113,3	36,0	141,8	430,2	303,9	258,9	128,1	566,3	321,0	780,9
77,0	26,2	111,3	379,0	293,3	240,1	112,0	507,8	279,4	686,1

Вариант 6		Вариант 7		Вариант 8		Вариант 9		Вариант 10	
X	Y	X	Y	X	Y	X	Y	X	Y
71,0	25,1	100,7	366,9	292,5	240,2	115,8	506,6	268,7	697,3
76,3	28,2	105,3	370,8	294,5	241,2	114,5	526,7	294,4	735,7
103,3	29,5	114,4	396,8	299,9	245,9	118,1	547,3	290,4	762,5
115,1	32,3	130,8	405,1	301,3	261,9	127,3	554,1	301,4	782,0
109,5	32,1	130,4	414,3	305,9	258,8	126,1	585,8	308,9	802,4
68,2	25,8	100,1	364,2	294,5	232,6	108,5	502,8	263,6	679,9
100,6	30,1	122,5	401,5	299,4	251,2	118,3	552,5	302,7	740,4
109,2	29,0	119,3	397,0	299,3	250,3	120,4	541,5	290,4	753,6
72,2	21,9	102,5	373,6	289,9	226,8	108,9	489,8	268,9	667,0
98,4	26,9	115,2	392,7	297,6	246,8	118,0	536,9	295,8	722,7
108,2	31,4	123,1	395,0	302,1	244,6	123,1	556,6	295,0	740,5
102,2	31,5	119,9	413,6	300,6	251,6	124,6	567,0	299,1	788,7
111,2	34,3	133,9	413,2	306,1	254,8	127,3	569,5	322,4	770,1
132,3	34,2	141,1	441,3	306,8	259,8	128,1	581,2	331,5	837,9
96,6	30,3	127,7	406,1	301,5	256,1	120,7	554,4	293,6	736,5
102,3	30,9	128,7	423,1	304,7	261,4	122,4	552,2	311,7	789,2
136,7	40,0	157,0	439,1	311,0	275,8	133,9	612,2	338,5	857,1
74,2	26,6	114,1	367,6	296,9	238,0	116,9	516,2	284,1	727,7
93,7	28,1	106,4	378,9	299,4	239,5	117,4	544,1	289,3	718,4
72,1	22,3	87,2	365,1	289,7	231,3	110,5	508,5	273,8	676,2
124,1	32,8	131,8	426,4	305,8	266,6	127,1	580,7	314,2	779,3
138,5	36,3	150,7	446,3	307,0	268,8	129,4	600,8	328,4	829,8
115,5	31,7	120,7	417,2	301,7	250,7	121,1	550,0	303,6	759,5
80,3	26,5	103,7	379,5	296,3	238,8	112,9	529,4	276,2	694,9
92,6	27,3	110,9	385,1	296,5	246,2	114,6	541,8	286,3	718,8
64,5	25,5	100,4	358,8	290,7	237,5	112,6	498,1	270,1	667,5
91,9	30,5	122,5	408,7	300,1	243,8	117,6	547,0	304,6	725,4
108,0	29,8	122,0	401,4	302,9	249,1	120,9	545,7	313,8	791,1
107,2	32,7	133,5	415,3	301,3	251,8	123,9	561,3	313,9	760,3
90,6	25,1	103,9	383,2	296,5	236,2	117,4	510,7	276,5	705,8
83,6	27,8	114,2	393,4	297,6	243,3	114,7	526,7	281,0	714,5
122,4	32,6	132,2	428,3	304,5	264,4	125,6	570,0	310,0	802,9

Вариант 11		Вариант 12		Вариант 13		Вариант 14		Вариант 15	
X	Y	X	Y	X	Y	X	Y	X	Y
73,8	24,5	107,5	362,7	293,3	230,6	113,2	511,9	269,3	667,0
108,9	33,3	134,1	401,6	301,6	253,8	125,9	579,2	317,3	768,7
85,4	29,8	106,5	389,1	295,3	248,8	114,4	535,8	298,7	716,9
121,6	34,0	124,0	423,5	303,1	258,2	121,8	572,7	317,2	778,0
91,1	28,6	112,5	393,0	298,7	251,1	117,7	533,3	295,1	734,3
110,8	34,9	137,6	412,7	303,1	257,2	126,9	576,0	317,6	772,5
126,8	34,7	144,4	428,4	304,5	262,1	129,0	581,3	331,4	820,6
117,4	34,7	127,0	410,9	301,6	257,8	127,4	558,7	321,6	779,8
96,3	27,7	117,1	381,0	297,2	247,5	118,2	531,2	288,6	739,8
117,4	31,7	130,2	402,5	304,8	256,1	126,8	562,6	309,1	776,7
78,1	25,2	103,8	390,1	295,3	237,7	114,6	514,8	287,5	735,6
92,8	29,0	120,8	386,7	296,9	246,2	118,5	556,9	293,4	746,5
134,2	36,6	142,1	432,0	308,7	271,0	127,5	604,5	328,1	806,6
85,0	29,0	116,9	385,7	298,4	242,5	118,6	538,1	279,9	710,4
90,1	31,1	112,5	387,1	300,4	244,7	120,0	530,0	290,0	715,3
127,5	37,3	141,1	445,3	307,2	262,6	126,7	591,8	335,7	824,8
102,4	27,6	117,1	388,7	300,6	242,7	120,2	535,3	302,1	723,8
32,5	15,7	66,4	322,2	286,4	207,8	97,3	448,5	235,3	567,4
116,2	35,2	135,5	423,8	306,0	263,7	127,3	576,4	315,4	779,9
65,4	22,8	101,8	359,2	290,3	230,7	108,9	494,1	272,8	673,9
107,7	31,3	131,2	413,9	302,9	262,6	122,8	564,3	312,5	761,3
84,6	27,6	115,0	390,8	299,5	248,5	117,9	530,0	287,8	750,2
106,5	31,8	117,7	397,1	300,3	258,6	119,7	564,4	303,7	738,4
83,9	27,8	115,5	377,6	296,9	241,4	114,9	530,0	298,6	746,0
110,5	34,5	136,5	428,9	304,1	257,7	122,8	578,1	320,2	814,3
94,9	26,8	104,5	385,9	295,8	241,9	115,2	544,4	286,5	740,1
118,8	30,4	129,5	424,9	301,7	261,5	125,3	564,1	317,0	760,6
102,2	29,6	119,8	395,4	298,7	252,3	117,9	542,8	290,4	722,5
93,7	29,4	118,2	408,8	302,3	255,8	120,9	556,8	312,3	739,3
104,3	31,3	117,6	406,9	301,9	253,4	123,0	568,2	301,3	759,1
102,4	27,9	111,3	381,9	298,1	241,8	115,5	542,3	294,4	737,5
101,0	30,5	125,9	407,5	303,2	254,0	123,5	566,3	301,1	786,3

Вариант 16		Вариант 17		Вариант 18		Вариант 19		Вариант 20	
X	Y	X	Y	X	Y	X	Y	X	Y
116,4	32,5	119,8	415,8	300,1	254,5	122,7	569,0	318,8	762,8
93,0	31,2	114,1	392,2	299,5	251,1	120,0	541,7	294,3	762,7
84,7	26,5	103,0	380,4	296,0	242,5	117,1	528,7	280,8	714,5
132,3	34,0	132,2	440,0	306,6	263,0	129,0	584,4	322,6	793,1
115,3	31,5	122,6	397,1	304,0	259,1	122,1	550,0	305,3	787,6
108,6	30,7	125,9	403,5	299,4	255,4	122,9	546,2	313,7	756,2
86,2	29,2	110,1	394,4	298,5	248,8	115,6	523,0	282,7	717,5
90,5	27,2	111,3	369,6	296,2	238,7	115,3	517,3	289,4	707,9
71,9	28,1	103,5	377,2	294,6	243,5	116,6	510,9	283,7	693,8
136,9	38,9	154,3	453,3	312,1	267,8	135,7	618,3	328,8	831,3
82,0	26,8	96,2	375,5	294,8	236,0	113,8	523,5	271,6	705,2
120,4	34,5	135,5	430,1	304,5	264,3	126,2	596,0	328,0	797,6
72,1	23,4	91,2	349,7	292,3	227,4	110,7	508,3	274,1	681,7
97,4	30,3	126,9	394,9	302,4	252,3	117,4	541,4	309,6	752,3
101,0	31,4	133,1	417,0	303,6	256,3	126,6	575,3	317,3	787,2
137,2	34,3	137,3	435,0	304,7	261,9	128,5	595,2	326,7	828,6
81,6	26,7	105,0	375,0	294,1	235,6	117,1	523,6	279,5	711,7
93,7	29,7	108,3	393,1	298,9	251,4	115,6	551,1	288,7	725,6
76,9	26,9	105,3	383,3	293,6	233,6	116,0	520,8	283,0	704,5
104,1	30,2	118,9	386,5	300,8	248,7	120,2	535,1	299,8	745,0
111,4	32,5	132,7	420,9	304,3	263,0	124,7	573,6	320,9	786,7
91,8	29,7	114,4	402,7	299,6	249,0	121,2	557,2	297,1	752,9
97,2	28,6	112,9	388,2	297,2	249,1	118,7	544,4	284,9	734,7
95,0	26,3	114,8	394,2	294,1	236,3	114,6	514,7	279,9	707,0
82,4	24,5	102,0	374,8	292,8	229,9	114,6	526,1	277,1	686,2
73,5	27,9	106,4	377,3	295,8	243,3	116,8	528,7	288,7	709,6
115,8	30,9	122,4	421,2	300,3	255,3	126,0	572,6	317,8	780,4
104,3	29,1	125,6	402,7	300,7	249,8	123,6	546,6	295,2	750,3
48,1	22,2	85,6	330,4	286,6	219,1	101,6	482,7	248,5	614,6
97,9	26,8	118,3	381,9	299,5	239,9	119,4	547,3	292,3	739,5
142,6	38,5	154,0	458,8	312,3	271,4	134,3	604,1	350,5	856,6
150,6	37,9	158,3	461,3	311,8	274,6	137,9	637,4	355,7	891,3

Вариант 21		Вариант 22		Вариант 23		Вариант 24		Вариант 25	
X	Y	X	Y	X	Y	X	Y	X	Y
107,7	35,0	135,5	431,3	301,5	257,4	125,5	585,6	323,3	800,0
93,4	28,8	106,5	388,0	297,6	247,4	112,7	532,2	286,9	741,1
122,4	33,2	139,3	431,9	302,3	254,9	123,5	562,4	317,5	805,1
95,5	30,1	118,0	395,9	300,4	255,8	118,4	559,3	305,4	744,1
130,6	36,7	144,0	447,1	308,7	274,0	134,0	606,9	327,6	815,3
101,5	31,1	116,7	418,6	299,2	255,0	119,6	557,6	299,6	758,8
114,8	33,3	130,7	401,2	301,0	258,8	120,9	573,3	306,7	793,8
92,4	30,0	122,3	393,2	297,9	244,7	116,7	526,7	291,7	752,7
78,8	26,0	105,2	392,8	296,3	245,5	117,5	522,7	290,2	730,2
84,0	30,3	109,1	381,3	299,5	242,4	118,0	550,9	291,0	718,3
99,7	33,1	126,9	408,3	303,1	257,6	125,8	565,7	308,8	779,6
87,5	27,2	107,2	388,7	298,2	240,6	117,3	534,2	285,8	711,8
82,9	24,6	105,6	370,0	296,3	234,3	114,8	509,3	278,6	693,9
75,4	26,9	101,5	381,1	293,6	235,2	114,5	515,1	279,9	711,3
116,5	31,6	132,7	421,8	302,3	257,6	124,2	585,4	322,2	769,7
117,6	31,6	126,2	429,0	303,0	253,5	127,2	573,1	316,6	775,0
116,7	31,8	139,0	416,6	302,1	261,3	126,2	558,8	306,7	779,5
86,9	26,6	103,5	382,3	297,4	247,1	112,6	535,3	282,4	722,9
96,8	29,6	109,8	389,2	299,9	243,2	116,8	549,0	291,6	730,7
76,7	26,4	110,0	380,9	296,3	241,6	113,5	515,3	287,1	715,9
106,1	33,6	129,1	413,0	300,3	256,8	123,7	568,2	309,2	770,4
109,8	32,2	125,3	396,8	300,8	252,3	121,5	557,2	306,0	766,3
111,4	34,2	135,4	434,1	304,3	256,0	123,8	563,9	323,0	816,5
125,8	32,6	133,5	432,2	306,6	264,4	130,2	596,5	331,4	796,9
107,3	31,2	128,0	403,1	299,5	248,8	119,2	556,7	297,0	733,3
115,0	32,2	124,7	408,8	304,0	262,0	121,8	558,5	309,2	789,9
116,9	35,5	142,2	445,1	306,0	263,0	130,5	593,2	326,5	793,1
99,0	33,1	125,2	402,0	304,6	252,3	126,0	569,4	301,5	767,5
109,6	30,4	122,8	405,7	300,8	247,5	122,8	554,7	297,6	745,3
116,4	35,4	137,6	421,3	305,6	264,3	129,7	597,5	330,7	830,5
94,8	29,6	115,9	395,6	301,5	246,5	117,4	538,3	292,1	745,9
93,2	30,4	127,6	400,6	300,1	250,1	119,5	559,9	307,9	771,8

Для решения задачи 6) для всех вариантов принять дополнительные факторы Х2 - инвестиции (млн.руб); Х3 – среднесписочная численность работников (чел.).

Х2	Х3
110,5	260
108,7	307
115,6	317
110,1	335
115,9	215
92	352
67,2	362
101,5	270
109,8	332
81	385
72,3	370
106,7	422
85,4	320

129,8	352
82,1	405
68,7	465
110,9	345
107,9	382
77,3	302
98,5	387
110,8	432
101,4	295
129,5	232
147,1	395
95	352
109,9	327
114,1	252
87,3	342
75,3	397
107,4	360
113,3	377
77	275

Отчетность, оформление лабораторной работы

По результатам выполнения ЛР студент подготавливает отчет, который должен содержать **Текстовую часть** и **Приложение**.

Текстовая часть включает:

- **выводы** о статистических свойствах изучаемой совокупности, сделанные на основе анализа таблиц и графиков, полученных в результате компьютерных расчетов;
- **экономическую интерпретацию** полученных статистических характеристик, раскрывающую их экономический смысл применительно к изучаемой совокупности предприятий.
- **выводы** о наличии статистической и корреляционной связи результативного и факторного признаков;
- **оценку** тесноты взаимосвязи признаков, сделанную на основе анализа показателей аналитической таблицы;
- **математическую модель** взаимосвязи исследуемых признаков в виде однофакторного **линейного** уравнения регрессии (построенного с помощью инструмента **Регрессия** надстройки **Пакет анализа**), а также **результаты анализа адекватности** этой модели;
- **математические модели** зависимости между исследуемыми признаками в виде однофакторных **нелинейных** уравнений регрессии (построенных с помощью инструмента **Мастер диаграмм**) и **обоснование выбора наиболее адекватной** из них.

Приложение к отчету должно включать:

- **таблицу исходных данных** и **результативные таблицы** с рассчитанными обобщающими показателями;
- **рисунки** статистических графиков.
- **графики** нелинейных уравнений регрессии.

Подготовка отчета производится студентом **вне рамок времени, отведенного на выполнение ЛР**.

Отчет сдается для проверки преподавателю, проводившему ЛР. Защита проверенного отчета состоит в обосновании студентом выводов, сделанных

им на основе результатов выполненных компьютерных статистических расчетов.

Результаты защиты отчета учитываются преподавателем при получении студентом зачета. Учебным планом предусмотрен один общий зачет по всем выполненным лабораторным работам.

Государственное бюджетное образовательное учреждение
высшего профессионального образования
Уфимский государственный нефтяной технический университет
Кафедра «Математика»

ЛАБОРАТОРНАЯ РАБОТА № 1

по дисциплине «Теория вероятностей и математическая статистика»

Тема «Моделирование случайных величин»

Вариант №3

Выполнил: студент группы БУТ-12-01

Иванова Д.В.

Проверил: д. ф.-м. н., профессор

Мухаметзянов И. З.