

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное
учреждение высшего образования
«СЕВАСТОПОЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

С.А. Чернявская, Л.В. Третьякова

**СТАТИСТИЧЕСКАЯ
ОЦЕНКА ДАННЫХ
ДЛЯ НЕКОТОРЫХ ЗАДАЧ
В ЭКОЛОГИИ
(Рабочая тетрадь)**

Учебное пособие

Севастополь
СевГУ
2019

УДК 529.24 (07)
Ч-489

Р е ц е н з е н т:
Ю.Е. Обжерин – доктор техн. наук, профессор

Чернявская С.А., Третьякова Л.В.

Ч-489 Статистическая оценка данных для некоторых задач в экологии (рабочая тетрадь): учеб. пособие / С.А. Чернявская, Л.В. Третьякова. – Севастополь: СевГУ, 2019. – 69 с.

Содержит материалы для практических занятий по математической статистике. Приведены основные теоретические сведения, решения типовых задач и задачи для самостоятельного решения с ответами.

Учебное пособие предназначено для студентов, обучающихся по специальности «Экология и природопользование», но может быть полезно и для других инженерных специальностей.

УДК 529.24 (07)

Рассмотрено и утверждено на заседании кафедры «Высшая математика» СевГУ, протокол № 6 от 25 февраля 2019 г.

© ФГАОУ ВО «Севастопольский
государственный университет», 2019

СОДЕРЖАНИЕ

Введение	4
§ 1. Проверка гипотезы о нормальном законе распределения. Критерий согласия χ^2 («хи-квадрат»).	7
1.1. Проверка гипотез	7
1.2. Проверка гипотезы о нормальном распределении генеральной совокупности	8
1.2.1. Эмпирическое распределение задано в виде последовательности интервалов одинаковой длины и соответствующих им частот	9
1.2.2. Эмпирическое распределение задано в виде последовательности равноотстоящих вариантов и соответствующих им частот	10
1.3. Расчет доверительного интервала для математического ожидания	11
1.4. Задача 1	12
§ 2. Функциональная зависимость и регрессия. Корреляционное отношение, его свойства и оценки	16
2.1. Коэффициенты корреляции, уравнения линейной регрессии	16
2.2. Задача 2	18
2.3. Метод наименьших квадратов	22
2.4. Задача 3	24
Варианты заданий лабораторных работ	26
Задача 1	26
Задача 2	30
Задача 3	41
РАБОЧАЯ ТЕТРАДЬ	42
Лабораторная работа № 1 «Проверка статистических гипотез. Критерий Пирсона»	42
Лабораторная работа № 2 «Корреляционный анализ»	53
Лабораторная работа № 3 «Метод наименьших квадратов»	58
Приложение 1	63
Приложение 2	65
Приложение 3	67
Приложение 4	68
Список рекомендуемой литературы	69

ВВЕДЕНИЕ

К числу наиболее актуальных проблем современной экологии относится проблема адекватного выбора и применения математических методов для решения экологических задач, связанных с исследованиями и эксплуатацией природных популяций и экосистем в целом.

Современное естествознание обладает большим многообразием методов исследований, среди которых эксперимент – наиболее эффективное и действенное средство познания.

В любом естественно-научном эксперименте можно выделить три основных этапа:

- 1) подготовительный (постановка задачи, планирование эксперимента);
- 2) получение экспериментальных данных;
- 3) обработка результатов эксперимента и их анализ.

Анализ результатов эксперимента позволяет оценить тот или иной параметр исследуемого объекта и сопоставить его либо с соответствующим теоретическим значением, либо с экспериментальным значением, полученным другими техническими средствами, что очень важно при определении правильности и степени достоверности полученных результатов.

Обработка данных любого эксперимента проводится с помощью методов математической статистики.

Математическая статистика — наука о математических методах, позволяющих по статистическим данным, например по реализациям случайной величины, построить теоретико-вероятностную модель исследуемого явления. Задачи математической статистики являются, в некотором смысле, обратными к задачам теории вероятностей.

При математической обработке и анализе результатов эксперимента рассматривают вопросы подбора эмпирических формул и оценки их параметров, вопросы оценки истинных значений измеряемых величин и точности измерений, вопросы исследования корреляционных зависимостей, а также вопросы анализа.

При статистическом анализе данных удобно различать следующие четыре общих этапа:

- 1) начальную обработку данных, т.е. представление данных в подходящей для подробного анализа форме;
- 2) предварительный анализ, цель которого выяснить общую форму данных и предложить пути более обстоятельного анализа;
- 3) итоговый анализ, цель которого дать основу для вывода;
- 4) представление выводов в точной, краткой и ясной форме.

Роль математической статистики в экологическом мониторинге особенно велика, так как результаты наблюдения определяют гарантии жизнеобеспечения на настоящий и будущее (при прогнозировании) моменты времени.

Рассматриваемые в данном разделе статистические процедуры широко применяются при обработке данных экологического мониторинга. Они могут представлять самостоятельный интерес при решении конкретных задач и, кроме того, входят в комплекс методов, используемых при статистической обработке результатов многофакторных экспериментов.

Генеральная совокупность и выборка

Математической статистикой называется раздел математики, занимающийся методами обработки опытных данных, полученных в результате наблюдений над случайными явлениями. Изучаемая случайная величина в математической статистике называется признаком (X), а различные значения признака называются вариациями (x).

Вся подлежащая изучению совокупность объектов (наблюдений) называется генеральной совокупностью, а та часть объектов, которая отбирается для непосредственного изучения из генеральной совокупности, называется выборочной совокупностью или выборкой. Числа объектов (наблюдений) в генеральной или выборочной совокупности называются их объемами.

Чтобы по данным выборки иметь возможность судить о генеральной совокупности, она должна быть отобрана случайно и иметь достаточный объем (предста-

вительность выборки). Исследуемая же генеральная совокупность должна быть однородной по составу, чтобы этим свойством обладала и выборка.

Статистическое распределение выборки

Первый шаг к осмыслению имеющегося статистического материала — это его упорядочение, т.е. расположение вариантов в порядке возрастания (ранжирование). В результате получается так называемый вариационный ряд.

Вариационный ряд может быть дискретным, если любые его варианты отличаются на постоянную величину, и непрерывным, если варианты в заданном интервале могут принимать любые значения.

Для более удобного представления выборочных данных варианты разбиваются на отдельные интервалы, т.е. производится группировка данных.

Оптимальное число интервалов, при котором не теряются особенности распределения признака, определяется по формуле

$$k = 1 + 3,22 \cdot \lg n,$$

где n — общее число наблюдений.

Величина интервала или ширина интервала h определяется путем деления величины размаха ($x_{\max} - x_{\min}$) на число принятых интервалов K и полученный результат округляют.

Числа, показывающие, сколько раз встречаются варианты в данном интервале, называются частотами (m), а отношение их к общему числу наблюдений называются относительными частотами W_i

$$W_i = \frac{n_i}{n}$$

Статистическим распределением выборки называют перечень всех возможных вариантов и соответствующих им частот или относительных частот.

Сумма всех частот равна объему выборки; сумма всех относительных частот равна единице.

§ 1. Проверка гипотезы о нормальном законе распределения.

Критерий согласия χ^2 («хи-квадрат»)

1.1. Проверка гипотез

При исследовании генеральной совокупности, всегда рассматривается репрезентативная выборка из этой совокупности. По этой выборке получают точечные и интервальные оценки. Все полученные сведения, как правило, не гарантируют, что данная выборка дает достаточно полное представление о генеральной совокупности. Если мы сумеем показать, что выборка подчиняется тому, или иному закону распределения, то это еще не значит, что генеральная совокупность подчиняется этому закону.

В связи с этим возникает необходимость проверки предполагаемых статистических гипотез.

Сформулируем основные принципы, позволяющие принимать обоснованные выводы о состоятельности выдвигаемых гипотез. Введем несколько основных понятий.

Выдвигаемая гипотеза называется **нулевой (основной)** и обозначается H_0 .

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

При проверке гипотез возможны два вида ошибок:

1. Отвергнута правильная гипотеза (ошибка первого рода).
2. Принята неправильная гипотеза (ошибка второго рода).

Вероятность допустить ошибку 1 или 2 рода обозначается α и β , соответственно. Для того, чтобы принять или отвергнуть нулевую гипотезу необходимо использовать такую характеристику, по которой можно бы было судить стоит ли отвергать гипотезу или нет.

Критерием называют случайную величину k , которая служит для проверки гипотезы.

Значение k , вычисленное по данной выборке, называется наблюдаемым и обозначается $k_{набл}$.

Областью принятия гипотезы называют совокупность значений критерия, при которых нулевую гипотезу принимают.

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

Нулевая гипотеза будет принята, если наблюдаемое значение $k_{набл}$ принадлежит области принятия гипотезы.

Критическими точками называют точки $k_{кр}$, отделяющие критическую область от области принятия гипотезы.

1.2. Проверка гипотезы о нормальном распределении генеральной совокупности

Пусть X – случайная величина генеральной совокупности. Выдвинем гипотезу H_0 , что данная выборка подчинена нормальному закону распределения и H_1 – она подчинена любому другому закону. Эти предположения можно проверить различными способами. Для этого используют специально подобранную случайную величину – критерий согласия. Критерием согласия называют критерий проверки гипотезы о предполагаемом законе неизвестного распределения. Имеется несколько критериев согласия. Рассмотрим один из них – критерий Пирсона.

Для проверки гипотезы H_0 произведем выборку, которую оформим в виде статистического ряда.

x_i	x_1	x_2	...	x_m
n_i	n_1	n_2	...	n_m

Требуется сделать заключение: согласуются ли результаты наблюдений с высказанным предположением. При этом необходимо иметь выборку достаточно большого объема: $n \geq 50$. Диапазон изменения выходной величины в этой выбор-

ке разбивается на s интервалов так, чтобы эти интервалы покрывали всю ось от $-\infty$ до $+\infty$, и в каждый интервал при этом попало не менее пяти значений из исследуемой выборки. Предполагая известным теоретический закон распределения случайной величины X , всегда можно найти p_i – теоретические вероятности. Тогда теоретические частоты значений случайной величины X , попавших в каждый интервал, можно рассчитать по формуле np_i .

В качестве меры расхождения между n_i и np_i Пирсон предложил величину

$$\chi^2 = \sum_{i=1}^m \frac{(n_i - np_i)^2}{np_i}$$

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

Распределение χ^2 зависит от параметра $k = s - r - 1$, называемого числом «степеней свободы» распределения. Где s – число интервалов выборки; r – число параметров предполагаемого распределения. Если предполагаемое распределение нормальное, то оценивают два параметра (μ и σ), поэтому число степеней свободы $k = s - 3$.

1.2.1. Эмпирическое распределение задано в виде последовательности интервалов одинаковой длины и соответствующих им частот

Если эмпирическое распределение значений X задано в виде последовательности интервалов (x_i, x_{i+1}) и соответствующих им частот n_i (n_i – сумма частот, которые попали в i -й интервал)

(x_i, x_{i+1})	(x_1, x_2)	(x_2, x_3)	...	(x_m, x_{m+1})
n_i	n_1	n_2	...	n_m

Для того чтобы при заданном уровне значимости α проверить гипотезу о нормальном законе распределения, используя критерий Пирсона, необходимо:

1. Для каждого интервала выбрать $x_i^* = \frac{x_i + x_{i+1}}{2}$.

2. Найти $\bar{x}_g^* = \frac{\sum_{i=1}^m x_i^* n_i}{\sum_{i=1}^m n_i}$ и $\sigma_g = \sqrt{D_g}$, где $D_g = \sum_{i=1}^m (x_i - \bar{x}_g^*)^2 \frac{n_i}{n}$.

3. Найти p_i , пользуясь формулой $p_i = P(x_i < x < x_{i+1}) = \Phi\left(\frac{x_{i+1} - \bar{x}_g^*}{\sigma_g}\right) - \Phi\left(\frac{x_i - \bar{x}_g^*}{\sigma_g}\right)$,

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ – функция Лапласа.

4. Вычислить теоретические частоты $n_T = n \cdot p_i$.

5. Вычислить

$$\chi_{набл}^2 = \sum_{i=1}^m \frac{(n_i - np_i)^2}{np_i}.$$

6. По таблице критических точек распределения χ^2 по заданному числу α и числу $k = s - 3$ находим $\chi_{кр}^2$.

7. Если $\chi_{набл}^2 < \chi_{кр}^2$, то нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности. Если $\chi_{набл}^2 > \chi_{кр}^2$, то гипотеза отвергается.

Результаты заносим в таблицу:

№	x_i	x_{i+1}	$z_i = \frac{x_i - \bar{x}_g^*}{\sigma_g}$	$z_{i+1} = \frac{x_{i+1} - \bar{x}_g^*}{\sigma_g}$	$\Phi(z_i)$	$\Phi(z_{i+1})$	$np_i = n_i (\Phi(z_{i+1}) - \Phi(z_i))$	$\frac{(n_i - np_i)^2}{np_i}$
---	-------	-----------	--	--	-------------	-----------------	--	-------------------------------

1.2.2. Эмпирическое распределение задано в виде последовательности равноотстоящих вариантов и соответствующих им частот

Пусть эмпирическое распределение задано в виде последовательности равностоящих вариантов x и частот n_i .

x_i	x_1	x_2	...	x_m
n_i	n_1	n_2	...	n_m

Для того чтобы при заданном уровне значимости α проверить гипотезу о нормальном законе распределения, используя критерий Пирсона, необходимо:

1. Найти $\bar{x}_s$ и σ_s .
2. Вычислить теоретические частоты

$$n_i^T = \frac{n \cdot h}{\sigma_s} \varphi(u_i),$$

где $h = x_{i+1} - x_i$, ($i = \overline{1 \dots m}$) – шаг, $u_i = \frac{x_i - \bar{x}_s}{\sigma_s}$; $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$.

3. Вычислить

$$\chi_{набл}^2 = \sum_{i=1}^m \frac{(n_i - n_i^T)^2}{n_i^T}.$$

4. Остальное, так же как и в предыдущем.

Все результаты целесообразно занести в таблицу:

i	x_i	u_i	$\varphi(u_i)$	n_i	n_i^T	$\frac{(n_i - n_i^T)^2}{n_i^T}$
-----	-------	-------	----------------	-------	---------	---------------------------------

1.3. Расчет доверительного интервала для математического ожидания

Доверительным интервалом называют такой интервал, который с заданной вероятностью γ покрывает оцениваемый параметр, т.е.

$$t_b - \varepsilon < T \leq t_b + \varepsilon,$$

Оценка t_b , ε – точность оценки, T – неизвестная оценка генеральной совокупности.

Пусть случайные ошибки измерения подчинены нормальному закону распределения вероятностей. Тогда доверительный интервал для математического

ожидания в предположении, что дисперсия измеряемой величины заранее не известна, имеет вид

$$\bar{x}_B - t \cdot \frac{s}{\sqrt{n}} < a < \bar{x}_B + t \cdot \frac{s}{\sqrt{n}},$$

где $s^2 = \frac{n}{n-1} D_B$.

Так как выборка распределена нормально, то t находим по таблице значений функции Лапласа $\Phi(t) = \frac{\gamma}{2}$.

1.4. Задача 1

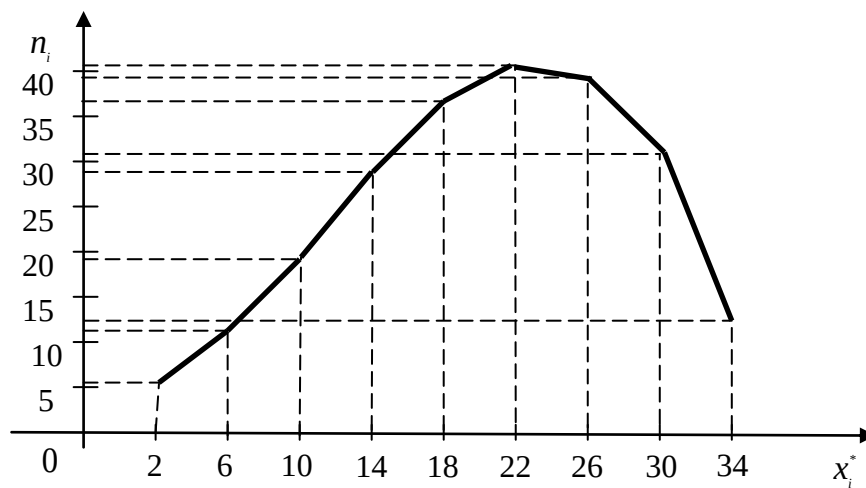
Для химического процесса очень важно, чтобы раствор, используемый в реакции, имел pH , равный 8,30. Метод определения pH дает значения, которые распределены по нормальному закону около истинного pH раствора со стандартным отклонением 0,01. Девять определений pH раствора, полученные в результате проведения процесса, представляют следующие величины, приведенные в таблице. Используя критерий Пирсона при уровне значимости $\alpha = 0,01$, проверить, согласуется ли гипотеза о нормальном распределении значений pH раствора около истинного pH . Если гипотеза согласуется, то найти доверительный интервал с надежностью $\gamma = 0,99$ неизвестного математического ожидания, по исправленной выборочной дисперсии S^2 и выборочной средней $\bar{x}_s$. Построить полигон частот и гистограмму частот значений pH раствора.

$(x_i; x_{i+1})$	(0;4)	(4;8)	(8;12)	(12;16)	(16;20)	(20;24)	(24;28)	(28;32)	(32;36)
n_i	6	12	19	28	37	41	39	31	13

Решение.

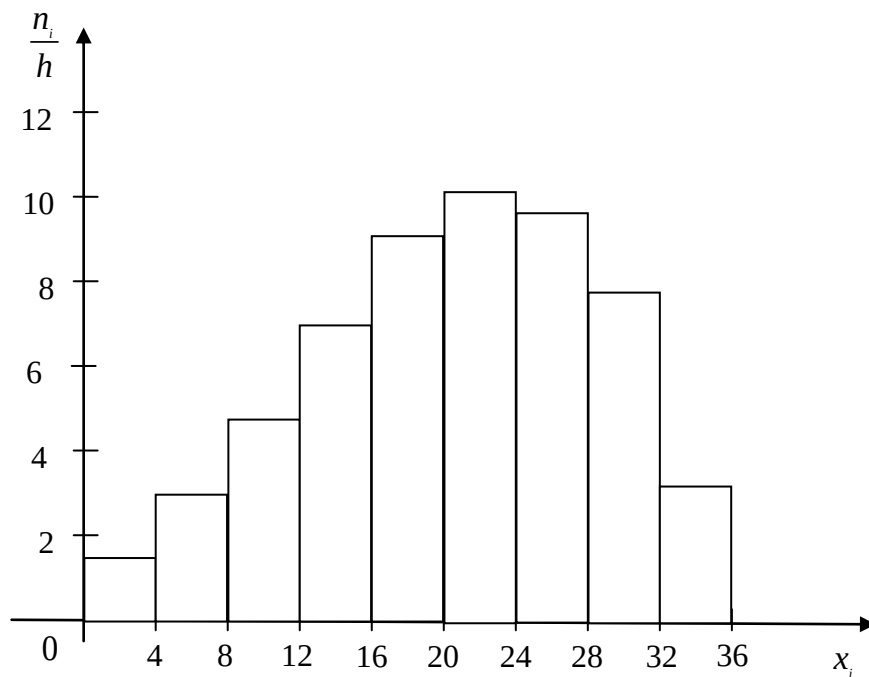
а) Построим полигон частот – ломаную с вершинами в точках $(x_i^*; n_i)$

x_i^*	2	6	10	14	18	22	26	30	34
n_i	6	12	19	28	37	41	39	31	13



б) Построим гистограмму частот - ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиной $h = 4$, а высоты равны отношению $\frac{n_i}{h}$.

$(x_i; x_{i+1})$	(0; 4)	(4; 8)	(8; 12)	(12; 16)	(16; 20)	(20; 24)	(24; 28)	(28; 32)	(32; 36)
n_i	6	12	19	28	37	41	39	31	13
$\frac{n_i}{h}$	1,5	3	4,75	7	9,25	10,75	9,75	7,75	3,25



Случайную величину – отклонение – обозначим через X .

Для вычисления вероятности p_i необходимо вычислить параметры определяющие нормальный закон распределения (a, σ) . Для чего в таблице перейдём от интервальных значений случайной величины x_i к вариантам, для чего в качестве вариант выберем середины интервалов:

$x_i^* = \frac{x_i + x_{i+1}}{2}$	2	6	10	14	18	22	26	30	34
n_i	6	12	19	28	37	41	39	31	13

$$m=9, n=6+12+19+28+37+41+39+31+13=226.$$

Найдём выборочную среднюю $\overline{x_g^*}$:

$$\overline{x_g^*} = \frac{\sum_{i=1}^6 x_i^* n_i}{\sum_{i=1}^6 n_i} = \frac{2 \cdot 6 + 6 \cdot 12 + 10 \cdot 19 + 14 \cdot 28 + 18 \cdot 37 + 22 \cdot 41 + 26 \cdot 39 + 30 \cdot 31 + 34 \cdot 13}{226} \approx 20,44.$$

Найдём выборочную дисперсию D_g :

$$D_6^* = \frac{\sum_{i=1}^6 (x_i^* - \bar{x}_6^*)^2 n_i}{\sum_{i=1}^6 n_i} = \frac{1}{226} [(2-20,44)^2 \cdot 6 + (6-20,44)^2 \cdot 12 + (10-20,44)^2 \cdot 19 + (14-20,44)^2 \cdot 28 + (18-20,44)^2 \cdot 37 + (22-20,44)^2 \cdot 41 + (26-20,44)^2 \cdot 39 + (30-20,44)^2 \cdot 31 + (34-20,44)^2 \cdot 13] \approx 64,26$$

$$\sigma^* = \sqrt{D_6^*} = \sqrt{64,26} \approx 8,02$$

Заполним таблицу

№	n_i	x_{i+1}	x_i	$z_{i+1} = \frac{x_{i+1} - \bar{x}_6^*}{\sigma_6^*}$	$z_i = \frac{x_i - \bar{x}_6^*}{\sigma_6^*}$	$\Phi(z_{i+1})$	$\Phi(z_i)$	$np_i = n_i (\Phi(z_{i+1}) - \Phi(z_i))$	$\frac{(n_i - np_i)^2}{np_i}$
1	6	4	0	-2,05	-2,55	-0,479	-0,495	0,096	1,316
2	12	8	4	-1,55	-2,05	-0,439	-0,479	0,48	0,0813
3	19	12	8	-1,05	-1,55	-0,353	-0,439	1,634	0,004
4	28	16	12	-0,55	-1,05	-0,209	-0,353	4,032	0,747
5	37	20	16	-0,055	-0,55	-0,024	-0,209	6,845	0,956
6	41	24	20	0,44	-0,055	0,17	-0,024	7,954	0,241
7	39	28	24	0,94	0,44	0,326	0,17	19,344	0,336
8	31	32	28	1,44	0,94	0,425	0,326	23,281	2,553
9	13	36	32	1,94	1,44	0,474	0,425	11,687	0,384

Суммируя результаты последнего столбца, находим $\chi_{набл}^2 = \sum_{i=1}^m \frac{(n_i - np_i)^2}{np_i} = 1,07$

Находим число степеней свободы. По выборке рассчитано два параметра, интервалов шесть, следовательно $k = 6 - 2 - 1 = 3$.

По таблице находим $\chi_{3;0,01}^2 = 11,3$.

Так как $\chi_{набл}^2 \leq \chi_{3;0,01}^2$ ($1,07 < 11,3$), то нет оснований отвергнуть проверяемую гипотезу.

Найдём $S^2 = \frac{n}{n-1} D_B = \frac{226}{225} \cdot 64,26 = 64,5456 \Rightarrow S = \sqrt{64,5456} \approx 8,034$

Искомый интервал для неизвестной дисперсии генеральной совокупности определяется из неравенства $\overline{x_B} - t_\gamma \frac{S}{\sqrt{n}} < a < \overline{x_B} + t_\gamma \frac{S}{\sqrt{n}}$, где S – исправленное среднеквадратическое отклонение, определяемое из формулы $S^2 = \frac{n}{n-1} D_B$. Так как выборка распределена нормально, то t находим по таблице значений функции Лапласа, при этом $\Phi(t) = \frac{\gamma}{2} = 0,495$, отсюда $t = 2,58$. Отсюда имеем

$$\alpha = \overline{x_e^*} - t_\gamma \frac{S}{\sqrt{n}} = 19,61; \quad \beta = \overline{x_e^*} + t_\gamma \frac{S}{\sqrt{n}} = 21,27.$$

Тогда доверительный интервал: $(19,61; 21,27)$.

Вывод. Подтверждается утверждение о том, что метод определения pH дает значения, которые распределены по нормальному закону около истинного pH раствора с отклонением $\alpha = 0,01$.

§ 2. Функциональная зависимость и регрессия. Корреляционное отношение, его свойства и оценки

2.1. Коэффициенты корреляции, уравнения линейной регрессии

Во многих случаях целью экспериментальных исследований является установление и изучение зависимости между некоторыми величинами. Если каждая из этих величин является случайной, то при этом используют методы корреляционного анализа. Так, методами корреляционного анализа можно оценить степень взаимосвязи между пределом влажности воздуха и количеством дисперсных загрязняющих частиц и т.д.

Будем говорить, что между двумя случайными величинами имеется статистическая связь, если при изменении одной из них меняется распределение другой. Для оценки статистической связи по данным эксперимента широко используется выборочный коэффициент корреляции.

Корреляционная связь величин заключается в том, что при задании одной из них, устанавливается зависимость между каждой из них и соответствующим математическим ожиданием другой.

Корреляционная таблица вида:

y	y_1	y_2	y_3	$\dots$	y_m	n_x
x						
x_1	n_{11}	n_{12}	n_{13}	$\dots$	n_{1m}	n_{x1}
x_2	n_{21}	n_{22}	n_{23}	$\dots$	n_{2m}	n_{x2}
x_3	n_{31}	n_{32}	n_{33}	$\dots$	n_{3m}	n_{x3}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_k	n_{k1}	n_{k2}	n_{k3}	$\dots$	n_{km}	n_{xk}
n_y	n_{y1}	n_{y2}	n_{y3}	$\dots$	n_{ym}	n

где

$$n_{x1} = n_{11} + n_{12} + n_{13} + \dots + n_{1m}$$

$$n_{y1} = n_{11} + n_{21} + n_{31} + \dots + n_{k1}$$

$$n = \sum n_x = \sum n_y$$

Силу связи между X и Y характеризует корреляционный момент K_{XY} , определяемый равенством

$$K_{XY} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

При вычислении корреляционного момента удобно пользоваться формулой

$$K_{XY} = \overline{xy} - \bar{x} \cdot \bar{y}, \text{ где } \overline{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i.$$

Выборочным коэффициентом корреляции r_e называется отношение выборочного корреляционного момента K_{XY} к произведению выборочных среднеквадратических отношений

$$r_e = \frac{K_{XY}}{\sigma_x \sigma_y}$$

или

$$r_{\epsilon} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Удобно пользоваться формулой

$$r_{\epsilon} = \frac{\overline{xy}}{\sigma_x \sigma_y}$$

Уравнение прямой регрессии Y по X имеет вид

$$y - \bar{y} = r_{\epsilon} \frac{\sigma_y}{\sigma_x} (x - \bar{x}).$$

Если $|r_{\epsilon}|$ близок к единице, то связь между X и Y близка к функциональной.

Для того чтобы при уровне значимости α проверить нулевую гипотезу о равенстве нулю генерального коэффициента корреляции нормальной двумерной случайной величины при конкурирующей гипотезе $H_1: r \neq 0$ надо вычислить

$$T_{набл} = \frac{r_{\epsilon} \sqrt{n-2}}{\sqrt{1-r_{\epsilon}^2}}$$

и по таблице критических точек Стьюдента по заданному α и числу степеней свободы $k = n - 2$ найти $t_{кр}(\alpha, k)$. Если $T_{набл} > t_k$ нулевую гипотезу отвергают.

2.2. Задача 2

При 100 пробах руды Каданского рудника были получены следующие данные о процентном содержании в руде свинца и серебра. Результаты занесены в корреляционную таблицу (X – содержание свинца в руде (в %), Y – содержание серебра в руде (в %)).

По данной корреляционной таблице

- а) найти коэффициент корреляции;
- б) найти выборочное уравнение прямой $y_{нах}$;

в) при уровне значимости $\alpha = 0,05$ определить, значимо или незначимо отличается генеральный коэффициент корреляции от нуля;

г) сделать вывод о тесноте корреляции у на х;

д) построить графики эмпирической корреляционной зависимости и полученной теоретической.

х у	7-11	11-15	15-19	19-23	23-27	n_y
12			5	7	9	21
15		12	10	8		30
18	7	9	4			20
21	14	15				29
n_x	21	36	19	15	9	100

Решение.

Заменяем интервальные значения переменных вариантами, выбрав середины интервалов за их значения. Получим корреляционную таблицу вида:

х у	9	13	17	21	25	n_y
12			5	7	9	21
15		12	10	8		30
18	7	9	4			20
21	14	15				29
n_x	21	36	19	15	9	100

а) Найдём коэффициент корреляции по формуле $r = \frac{K_{xy}}{\sigma_x \sigma_y}$, для чего вычислим:

1) выборочные средние признаков Х и У

$$\bar{x} = \frac{1}{n} \sum_{i=1}^5 x_i n_{xi} = \frac{1}{100} (9 \cdot 21 + 13 \cdot 36 + 17 \cdot 19 + 21 \cdot 15 + 25 \cdot 9) = 15,2$$

$$\bar{y} = \frac{1}{n} \sum_{j=1}^4 y_j n_{yj} = \frac{1}{100} (12 \cdot 21 + 15 \cdot 30 + 18 \cdot 20 + 21 \cdot 29) = 16,71$$

2) выборочные среднеквадратические отклонения, предварительно вычислив соответствующие дисперсии,

$$D_x = \frac{1}{n} \sum_{i=1}^5 (x_i - \bar{x})^2 n_{xi} = \frac{1}{100} [(9 - 15,2)^2 \cdot 21 + (13 - 15,2)^2 \cdot 36 + (17 - 15,2)^2 \cdot 19 + (21 - 15,2)^2 \cdot 15 + (25 - 15,2)^2 \cdot 9] = 24,12$$

$$\sigma_x = \sqrt{24,12} = 4,91$$

$$D_y = \frac{1}{n} \sum_{j=1}^4 (y_j - \bar{y})^2 n_{yj} = \frac{1}{100} [(12 - 16,71)^2 \cdot 21 + (15 - 16,71)^2 \cdot 30 + (18 - 16,71)^2 \cdot 20 + (21 - 16,71)^2 \cdot 29] = 11,21$$

$$\sigma_y = \sqrt{11,206} = 3,35$$

3) выборочный корреляционный момент

$$K_{xy} = \frac{1}{n} \sum_{i=1}^5 \sum_{j=1}^4 x_i y_j n_{ij} - \bar{x} \bar{y} = \frac{1}{100} (12 \cdot 5 \cdot 17 + 12 \cdot 7 \cdot 21 + 12 \cdot 9 \cdot 25 + 15 \cdot 12 \cdot 13 + 15 \cdot 10 \cdot 17 + 15 \cdot 8 \cdot 21 + 18 \cdot 7 \cdot 9 + 18 \cdot 9 \cdot 13 + 18 \cdot 4 \cdot 17 + 21 \cdot 14 \cdot 9 + 21 \cdot 15 \cdot 13) - 16,71 \cdot 15,2 = -13,00$$

Тогда коэффициент корреляции равен $r = \frac{K_{xy}}{\sigma_x \sigma_y} = \frac{-13,00}{4,911 \cdot 3,35} = -0,79$

б) Составим выборочное уравнение прямой У на Х вида $y - \bar{y} = \rho_{yx}(x - \bar{x})$, где коэф-

фициент регрессии $\rho_{yx} = r \frac{\sigma_y}{\sigma_x} = -0,79 \frac{3,35}{4,91} = -0,54$,

тогда уравнение линии регрессии У на Х будет иметь вид:

$$y - 16,71 = -0,54(x - 15,2)$$

$$y = -0,54x + 24,9$$

в) Выясним, значимо или незначимо отличается генеральный коэффициент кор-

реляции от нуля. Для этого найдём $T_{набл.} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{-0,79\sqrt{98}}{\sqrt{1-(-0,79)^2}} = -12,79$.

По таблице критических точек распределения Стьюдента имеем

$$t_{кр.}(\alpha; n-2) = t_{кр.}(0,05; 98) = 1,99$$

Так как $T_{набл.} > t_{кр.}$, то гипотезу о равенстве нулю генерального коэффициента корреляции отвергаем.

г) Так как $|r|$ близок к 1, то можно утверждать, что X и Y хорошо коррелируемы. Для того чтобы построить график эмпирического распределения, необходимо

найти условные средние значения $\bar{y}_i = \frac{\sum_{j=1}^5 y_{ij} n_{ij}}{n_{xi}}$.

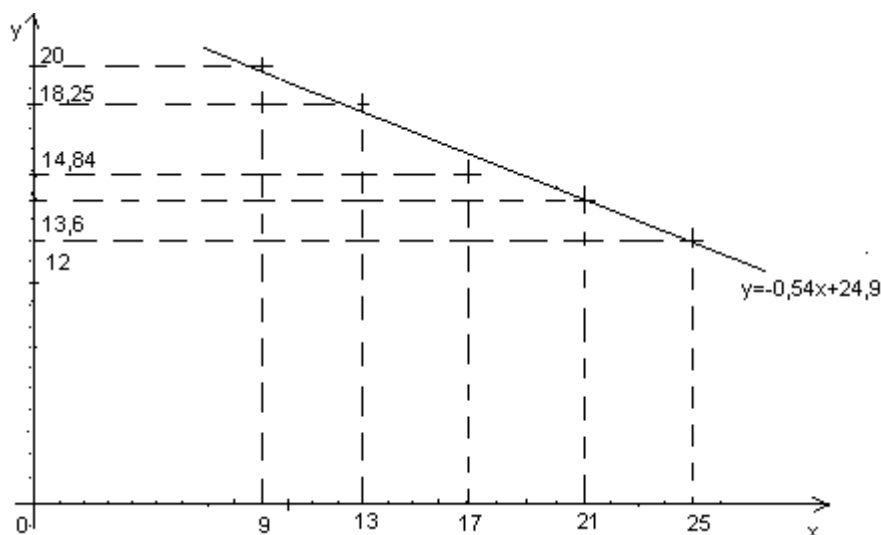
$$\bar{y}_1 = \frac{17 \cdot 7 + 21 \cdot 14}{21} = 20; \quad \bar{y}_2 = \frac{15 \cdot 12 + 18 \cdot 9 + 21 \cdot 15}{36} = 18,25;$$

$$\bar{y}_3 = \frac{12 \cdot 5 + 10 \cdot 15 + 18 \cdot 4}{19} = 14,84; \quad \bar{y}_4 = \frac{12 \cdot 7 + 15 \cdot 8}{15} = 13,6; \quad \bar{y}_5 = \frac{12 \cdot 9}{9} = 12.$$

Получим следующую таблицу, определяющую соответствие между значениями x_i и условными средними $\bar{y}_i$:

x_i	9	13	17	21	25
$\bar{y}_i$	20	18,25	14,84	13,6	12

д) Построим в декартовой системе координат точки с координатами $(x_i; \bar{y}_i)$. В той же системе координат построим график теоретического распределения $y = -0,54x + 24,9$.



Из чертежа видно, что теоретическая прямая хорошо коррелирует эмпирические данные.

Ответ: $\bar{y} = -0,54x + 24,9$.

2.3. Метод наименьших квадратов

При эмпирическом (экспериментальном) изучении функциональной зависимости одной величины y от другой величины x производят ряд измерений величины y при различных значениях величины x .

Задача заключается в аналитическом представлении искомой функциональной зависимости, т.е. подборе формулы, описывающей результаты эксперимента.

Пусть в результате опыта получены " n " - значений функции соответствующие " n "-значениям аргумента, записанные в таблицу

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

По этим данным надо составить уравнение функции в виде:

$$f(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \dots + a_k\varphi_k(x),$$

где $a_1, a_2, \dots, a_k$ некоторые коэффициенты, которые необходимо найти, $\varphi_1(x), \varphi_2(x), \dots, \varphi_k(x)$ неизвестные функции.

Одним из методов нахождения коэффициентов $a_1, a_2, \dots, a_k$ - является метод наименьших квадратов. Суть его состоит в том, что коэффициенты $a_1, a_2, \dots, a_k$ подбираются так, чтобы квадрат разности

$$S = \sum_{i=1}^n (y_i - f(x_i))^2$$

принимал наименьшее значение.

[illegible]
$$\begin{cases} a_1 \sum \varphi_1^2(x_i) + a_2 \sum \varphi_1(x_i)\varphi_2(x_i) + \dots + a_k \sum \varphi_1(x_i)\varphi_k(x_i) = \sum y_i\varphi_1(x_i) \\ a_1 \sum \varphi_1(x_i)\varphi_2(x_i) + a_2 \sum \varphi_2^2(x_i) + \dots + a_k \sum \varphi_2(x_i)\varphi_k(x_i) = \sum y_i\varphi_2(x_i) \\ \vdots \\ a_1 \sum \varphi_1(x_i)\varphi_k(x_i) + a_k \sum \varphi_2(x_i)\varphi_k(x_i) + \dots + a_k \sum \varphi_k^2(x_i) = \sum y_i\varphi_k(x_i) \end{cases}$$

Решая эту систему, находят значения $a_1, a_2, \dots, a_k$.

$$\text{системы уравнений} \left\{ \begin{array}{l} \alpha \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i; \\ \alpha \sum_{i=1}^n x_i + b n = \sum_{i=1}^n y_i. \end{array} \right.$$

Если зависимость между величинами X и Y искать в форме

то параметры α, b и c находятся из системы

$$\left\{ \begin{array}{l} \alpha \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 y_i; \\ \alpha \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i; \\ \alpha \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + c \cdot n = \sum_{i=1}^n y_i. \end{array} \right.$$

2.4. Задача 3

Проведено экспериментальное исследование зависимости площади нефтяного пятна на поверхности воды от изменения температуры $X = t, ^\circ\text{C}$.

Пусть задана таблица экспериментальных данных

x	-2	-1	0	1	2
y	3,8	1,1	0	0,3	4,1

Построим корреляционную зависимость с помощью функции $f(x) = a + bx + cx^2$.

Решение.

Если зависимость между величинами X и Y искать в форме $\bar{y}_x = \alpha + bx + cx^2$, то параметры α, b и c находятся из системы

$$\begin{cases} \alpha \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 y_i; \\ \alpha \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i; \\ \alpha \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + c \cdot n = \sum_{i=1}^n y_i. \end{cases}$$

Для нахождения коэффициентов составим предварительно таблицу:

	1	2	3	4	5	Σ
x_i	-2	-1	0	1	2	0
x^2	4	1	0	1	4	10
x^3	-3	-1	0	1	8	0
x^4	16	1	0	1	16	34
y	3,3	1,1	0	0,8	4,1	9,8
yx	-7,6	-1,1	0	0,8	8,2	0,3
yx^2	15,2	1,1	0	0,8	16,4	33,5

Отсюда система имеет вид:
$$\begin{cases} 34c + 10a = 33,5 \\ 10b = 0,3 \\ 10c + 5a = 9,8 \end{cases}.$$

Решая эту систему, получим $a = 0,02; b = 0,03; c = 0,99$,

т.е. $f(x) = 0,99x^2 + 0,03x - 0,02$

Вычислим теоретические значения ординат при $x = -2, -1, 0, 1, 2$. Имеем

$$\begin{aligned} y(-2) &= 3,88; & y(0) &= -0,02; & y(2) &= 4,0 \\ y(-1) &= 0,94; & y(1) &= 1,0; \end{aligned}$$

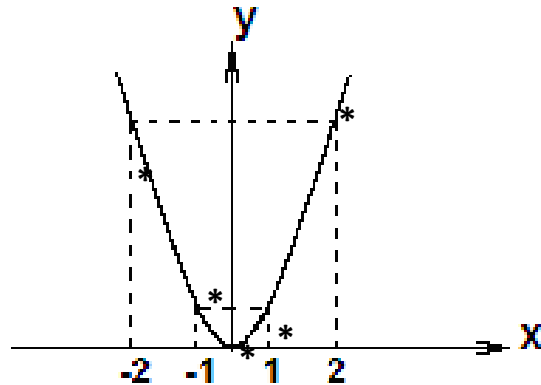
Ошибки в этом случае:

$$\begin{aligned} \delta_1 &= 3,8 - 3,88 = -0,08; & \delta_3 &= 0 + 0,02 = 0,02; & \delta_5 &= 4,1 - 4 = 0,1 \\ \delta_2 &= 1,1 - 0,94 = 0,16; & \delta_4 &= 0,8 - 1 = -0,2; \end{aligned}$$

Среднеквадратичная ошибка:

$$\sigma = \sqrt{(0,08^2 + 0,16^2 + 0,02^2 + 0,2^2) : 5} \approx 0,12.$$

Полученная ошибка является минимально возможной. Строим график функции и наносим на плоскость заданные экспериментальные значения.



Варианты заданий лабораторных работ

Задача 1.

Из генеральной совокупности извлечена выборка объёма n

- а) Построить полигон частот и гистограмму частот.
- б) Используя критерий Пирсона при уровне значимости $\alpha = 0,01$ проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности с эмпирическим распределением.
- в) Если гипотеза согласуется, то найти доверительный интервал с надежностью $\gamma = 0,99$ неизвестного математического ожидания a по исправленной выборочной дисперсии S^2 и выборочной средней $\bar{x}$.

Вариант 1

x_i	3,5	4,1	4,7	5,3	5,9	6,5	7,1	7,7	8,3	8,9	
n_i	12	20	30	33	39	37	31	26	15	5	$n = 248$

Вариант 2

x_i	3-7	7-11	11-15	15-19	19-23	23-27	27-31	31-35	35-39	
n_i	4	7	12	14	13	8	6	4	2	$n = 70$

Вариант 3

x_i	0-4	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	
n_i	6	12	19	28	37	41	39	31	13	$n = 226$

Вариант 4

x_i	3-5	5-7	7-9	9-11	11-13	13-15	15-17	17-19	19-21	
n_i	8	10	13	18	22	27	19	13	7	$n = 137$

Вариант 5

x_i	2,7	3,1	3,5	3,9	4,3	4,7	5,1	5,5	5,9	6,3	
n_i	4	7	9	13	17	22	19	14	11	8	$n = 124$

Вариант 6

x_i	-2,1	-1,1	-0,1	0,9	1,9	2,9	3,9	4,9	5,9	
n_i	3	7	12	20	27	35	28	22	14	$n = 168$

Вариант 7

x_i	0,7	1,2	1,7	2,2	2,7	3,2	3,7	4,2	4,7	5,2	
n_i	7	16	35	41	49	56	51	44	23	15	$n = 337$

Вариант 8

x_i	2,3-2,9	2,9-3,5	3,5-4,1	4,1-4,7	4,7-5,3	5,3-5,9	5,9-6,5	6,5-7,1	7,1-7,7	
n_i	2	11	17	25	31	22	18	12	5	$n = 143$

Вариант 9

x_i	-0,3	-0,1	0,1	0,3	0,5	0,7	0,9	1,1	1,3	1,5	$n=196$
n_i	4	9	15	22	30	35	29	23	17	12	

Вариант 10

x_i	3,5	4,1	4,7	5,3	5,9	6,5	7,1	7,7	8,3	8,9	
n_i	12	20	30	33	39	37	31	26	15	5	$n = 248$

Вариант 11

x_i	3-7	7-11	11-15	15-19	19-23	23-27	27-31	31-35	35-39	
n_i	4	7	12	14	13	8	6	4	2	$n = 70$

Вариант 12

x_i	0-4	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	
n_i	6	12	19	28	37	41	39	31	13	$n = 226$

Вариант 13

x_i	3-5	5-7	7-9	9-11	11-13	13-15	15-17	17-19	19-21	
n_i	8	10	13	18	22	27	19	13	7	$n = 137$

Вариант 14

x_i	2,7	3,1	3,5	3,9	4,3	4,7	5,1	5,5	5,9	6,3	
n_i	4	7	9	13	17	22	19	14	11	8	$n = 124$

Вариант 15

x_i	-2,1	-1,1	-0,1	0,9	1,9	2,9	3,9	4,9	5,9	
n_i	3	7	12	20	27	35	28	22	14	$n = 168$

Вариант 16

x_i	0,7	1,2	1,7	2,2	2,7	3,2	3,7	4,2	4,7	5,2	
n_i	7	16	35	41	49	56	51	44	23	15	$n = 337$

Вариант 17

x_i	2,3-2,9	2,9-3,5	3,5-4,1	4,1-4,7	4,7-5,3	5,3-5,9	5,9-6,5	6,5-7,1	7,1-7,7	
n_i	2	11	17	25	31	22	18	12	5	$n=143$

Вариант 18

x_i	-0,3	-0,1	0,1	0,3	0,5	0,7	0,9	1,1	1,3	1,5	$n=196$
n_i	4	9	15	22	30	35	29	23	17	12	

Вариант 19

x_i	3,5	4,1	4,7	5,3	5,9	6,5	7,1	7,7	8,3	8,9	
n_i	12	20	30	33	39	37	31	26	15	5	$n=248$

Вариант 20

x_i	3-7	7-11	11-15	15-19	19-23	23-27	27-31	31-35	35-39	
n_i	4	7	12	14	13	8	6	4	2	$n=70$

Вариант 21

x_i	0-4	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	
n_i	6	12	19	28	37	41	39	31	13	$n=226$

Вариант 22

x_i	3-5	5-7	7-9	9-11	11-13	13-15	15-17	17-19	19-21	
n_i	8	10	13	18	22	27	19	13	7	$n = 137$

Вариант 23

x_i	2,7	3,1	3,5	3,9	4,3	4,7	5,1	5,5	5,9	6,3	
n_i	4	7	9	13	17	22	19	14	11	8	$n = 124$

Вариант 24

x_i	-2,1	-1,1	-0,1	0,9	1,9	2,9	3,9	4,9	5,9	
n_i	3	7	12	20	27	35	28	22	14	$n = 168$

Вариант 25

x_i	0,7	1,2	1,7	2,2	2,7	3,2	3,7	4,2	4,7	5,2	
n_i	7	16	35	41	49	56	51	44	23	15	$n = 337$

Вариант 26

x_i	2,3-2,9	2,9-3,5	3,5-4,1	4,1-4,7	4,7-5,3	5,3-5,9	5,9-6,5	6,5-7,1	7,1-7,7	
n_i	2	11	17	25	31	22	18	12	5	$n = 143$

Вариант 27

x_i	-0,3	-0,1	0,1	0,3	0,5	0,7	0,9	1,1	1,3	1,5	n=196
n_i	4	9	15	22	30	35	29	23	17	12	

Вариант 28

x_i	3,5	4,1	4,7	5,3	5,9	6,5	7,1	7,7	8,3	8,9	
n_i	12	20	30	33	39	37	31	26	15	5	n = 248

Вариант 29

x_i	3-7	7-11	11-15	15-19	19-23	23-27	27-31	31-35	35-39	
n_i	4	7	12	14	13	8	6	4	2	n = 70

Вариант 30

x_i	0-4	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	
n_i	6	12	19	28	37	41	39	31	13	n = 226

Задача 2.

По данной корреляционной таблице

- найти коэффициент корреляции;
- найти выборочное уравнение прямой y/x ;
- при уровне значимости $\alpha = 0,05$ определить, значимо или незначимо отличается генеральный коэффициент корреляции от нуля;
- сделать вывод о тесноте корреляции y на x ;
- построить графики эмпирической корреляционной зависимости и полученной теоретической.

Вариант 1.

$y \backslash x$	12-16	16-20	20-24	24-28	28-32	n_y
12	12	11				23
15		9	14	5		28
18			9	9	8	26
21				13	10	23
n_x	12	20	23	27	18	100

Ответ: $\overline{y_x} = 0,54x + 4,28$

Вариант 2.

$y \backslash x$	7-11	11-15	15-19	19-23	23-27	n_y
12			5	7	9	21
15		12	10	8		30
18	7	9	4			20
21	14	15				29
n_x	21	36	19	15	9	100

Ответ: $\overline{y_x} = -0,539x + 25,44$

Вариант 3.

$y \backslash x$	12-15	15-18	18-21	21-24	24-27	n_y
10				12	17	29
15		10	9	11		30
20	8	10	7			25
25	7	9				16
n_x	15	29	16	23	17	100

Ответ: $\overline{y_x} = -1,01x + 35,78$

Вариант 4.

$y \backslash x$	6-10	10-14	14-18	18-22	22-26	n_y
7	5	7	9			21
11		8	10	12		30
15			4	9	7	20
19				15	14	29
n_x	5	15	23	36	21	100

Ответ: $\overline{y_x} = 0,755x - 0,397$

Вариант 5.

$y \backslash x$	17-21	21-25	25-29	29-33	33-37	n_y
12	13	11				24
14	8	10	7			25
16		11	9	5		25
18			7	9	10	26
n_x	21	32	23	14	10	100

Ответ: $\overline{y_x} = 0,358x + 5,97$

Вариант 6.

$y \backslash x$	6	8	10	12	14	n_y
11			5	7	9	21
13		12	10	8		30
15	7	9	4			20
17	14	15				29
n_x	21	36	19	15	9	100

Ответ: $\overline{y_x} = -0,719x + 21,681$

Вариант 7.

$y \backslash x$	43	49	55	61	67	n_y
3	17	12				29
10		11	9	10		30
17			7	10	8	25
24				9	7	16
n_x	17	23	16	29	15	100

Ответ: $\overline{y_x} = 0,761x + 30$

Вариант 8.

$y \backslash x$	6-10	10-14	14-18	18-22	22-26	n_y
7	12	11				23
11		9	14	5		28
15			9	9	8	26
19				13	10	23
n_x	12	20	23	27	18	100

Ответ: $\overline{y_x} = 0,714x + 0,99$

Вариант 9.

$y \backslash x$	13,5	16,5	19,5	22,5	25,5	n_y
10				12	17	29
15		10	9	11		30
20	8	10	7			25
25	7	9				16
n_x	15	29	16	23	17	100

Ответ: $\overline{y_x} = -1,088x + 37,54$

Вариант 10.

$y \backslash x$	12-16	16-20	20-24	24-28	28-32	n_y
12				11	12	23
15		5	14	9		28
18	8	9	9			26
21	10	13				23
n_x	18	27	23	20	12	100

Ответ: $\overline{y_x} = -0,536x + 27,847$

Вариант 11.

$y \backslash x$	12-16	16-20	20-24	24-28	28-32	n_y
12	12	11				23
15		9	14	5		28
18			9	9	8	26
21				13	10	23
n_x	12	20	23	27	18	100

Ответ: $\overline{y_x} = 0,54x + 4,28$

Вариант 12.

$y \backslash x$	7-11	11-15	15-19	19-23	23-27	n_y
12			5	7	9	21
15		12	10	8		30
18	7	9	4			20
21	14	15				29
n_x	21	36	19	15	9	100

Ответ: $\overline{y_x} = -0,539x + 25,44$

Вариант 13.

$y \backslash x$	12-15	15-18	18-21	21-24	24-27	n_y
10				12	17	29
15		10	9	11		30
20	8	10	7			25
25	7	9				16
n_x	15	29	16	23	17	100

Ответ: $\overline{y_x} = -1,01x + 35,78$

Вариант 14.

$y \backslash x$	6-10	10-14	14-18	18-22	22-26	n_y
7	5	7	9			21
11		8	10	12		30
15			4	9	7	20
19				15	14	29
n_x	5	15	23	36	21	100

Ответ: $\overline{y_x} = 0,755x - 0,397$

Вариант 15.

$y \backslash x$	17-21	21-25	25-29	29-33	33-37	n_y
12	13	11				24
14	8	10	7			25
16		11	9	5		25
18			7	9	10	26
n_x	21	32	23	14	10	100

Ответ: $\overline{y_x} = 0,358x + 5,97$

Вариант 16.

$y \backslash x$	6	8	10	12	14	n_y
11			5	7	9	21
13		12	10	8		30
15	7	9	4			20
17	14	15				29
n_x	21	36	19	15	9	100

Ответ: $\overline{y_x} = -0,719x + 21,681$

Вариант 17.

$y \backslash x$	43	49	55	61	67	n_y
3	17	12				29
10		11	9	10		30
17			7	10	8	25
24				9	7	16
n_x	17	23	16	29	15	100

Ответ: $\overline{y_x} = 0,761x + 30$

Вариант 18.

$y \backslash x$	6-10	10-14	14-18	18-22	22-26	n_y
7	12	11				23
11		9	14	5		28
15			9	9	8	26
19				13	10	23
n_x	12	20	23	27	18	100

Ответ: $\overline{y_x} = 0,714x + 0,99$

Вариант 19.

$y \backslash x$	13,5	16,5	19,5	22,5	25,5	n_y
10				12	17	29
15		10	9	11		30
20	8	10	7			25
25	7	9				16
n_x	15	29	16	23	17	100

Ответ: $\overline{y_x} = -1,088x + 37,54$

Вариант 20.

$y \backslash x$	12-16	16-20	20-24	24-28	28-32	n_y
12				11	12	23
15		5	14	9		28
18	8	9	9			26
21	10	13				23
n_x	18	27	23	20	12	100

Ответ: $\overline{y_x} = -0,536x + 27,847$

Вариант 21.

$y \backslash x$	12-16	16-20	20-24	24-28	28-32	n_y
12	12	11				23
15		9	14	5		28
18			9	9	8	26
21				13	10	23
n_x	12	20	23	27	18	100

Ответ: $\overline{y_x} = 0,54x + 4,28$

Вариант 22.

$y \backslash x$	7-11	11-15	15-19	19-23	23-27	n_y
12			5	7	9	21
15		12	10	8		30
18	7	9	4			20
21	14	15				29
n_x	21	36	19	15	9	100

Ответ: $\overline{y_x} = -0,539x + 25,44$

Вариант 23.

$y \backslash x$	12-15	15-18	18-21	21-24	24-27	n_y
10				12	17	29
15		10	9	11		30
20	8	10	7			25
25	7	9				16
n_x	15	29	16	23	17	100

Ответ: $\overline{y_x} = -1,01x + 35,78$

Вариант 24.

$y \backslash x$	6-10	10-14	14-18	18-22	22-26	n_y
7	5	7	9			21
11		8	10	12		30
15			4	9	7	20
19				15	14	29
n_x	5	15	23	36	21	100

Ответ: $\overline{y_x} = 0,755x - 0,397$

Вариант 25.

$y \backslash x$	17-21	21-25	25-29	29-33	33-37	n_y
12	13	11				24
14	8	10	7			25
16		11	9	5		25
18			7	9	10	26
n_x	21	32	23	14	10	100

Ответ: $\overline{y_x} = 0,358x + 5,97$

Вариант 26.

$y \backslash x$	6	8	10	12	14	n_y
11			5	7	9	21
13		12	10	8		30
15	7	9	4			20
17	14	15				29
n_x	21	36	19	15	9	100

Ответ: $\overline{y_x} = -0,719x + 21,681$

Вариант 27.

$y \backslash x$	43	49	55	61	67	n_y
3	17	12				29
10		11	9	10		30
17			7	10	8	25
24				9	7	16
n_x	17	23	16	29	15	100

Ответ: $\overline{y_x} = 0,761x + 30$

Вариант 28.

$y \backslash x$	6-10	10-14	14-18	18-22	22-26	n_y
7	12	11				23
11		9	14	5		28
15			9	9	8	26
19				13	10	23
n_x	12	20	23	27	18	100

Ответ: $\overline{y_x} = 0,714x + 0,99$

Вариант 29.

$y \backslash x$	13,5	16,5	19,5	22,5	25,5	n_y
10				12	17	29
15		10	9	11		30
20	8	10	7			25
25	7	9				16
n_x	15	29	16	23	17	100

Ответ: $\overline{y_x} = -1,088x + 37,54$

Вариант 30.

$y \backslash x$	12-16	16-20	20-24	24-28	28-32	n_y
12				11	12	23
15		5	14	9		28
18	8	9	9			26
21	10	13				23
n_x	18	27	23	20	12	100

Ответ: $\overline{y_x} = -0,536x + 27,847$

Задача 3

По данным исходной таблицы значения y_i , соответствующих аргументам

$x_1 = 1, \quad x_2 = 2, \quad x_3 = 3, \quad x_4 = 4, \quad x_5 = 5,$ аппроксимировать функцией

$$f(x) = ax + b.$$

№	Y01	Y02	Y03	Y04	Y05
1	1,9	3,0	3,3	4,0	4,9
2	1,4	3,0	4,4	4,8	5,4
3	1,6	2,7	4,2	5,7	6,4
4	2,1	2,9	4,0	5,3	7,0
5	3,2	3,6	4,3	5,3	6,5
6	2,5	4,2	5,1	6,2	7,6
7	2,2	4,3	6,3	7,2	8,2
8	2,6	4,2	6,1	8,2	9,5
9	3,3	4,6	6,1	7,9	10,1
10	4,8	5,5	6,6	8,1	9,8
11	4,4	7,7	10,1	12,7	15,6
12	4,7	8,2	11,7	14,3	16,8
13	5,6	8,6	12,0	15,7	18,5
14	6,9	9,6	12,6	15,8	19,5
15	8,7	11,0	13,7	16,6	19,7
16	-0,5	-0,3	-0,2	-1,2	-2,1
17	-0,4	-0,7	-0,7	-0,5	-1,1
18	0,0	-0,6	-1,0	-1,1	-0,8
19	1,0	-0,0	-0,8	-1,2	-1,4
20	1,0	0,9	-0,1	-0,9	-1,4
21	-2,3	-4,9	-7,3	-10,6	-14,2
22	-1,4	-4,4	-7,1	-9,4	-12,4
23	-0,0	-3,4	-6,4	-9,2	-11,6
24	1,7	-1,9	-5,3	-8,4	-11,3
25	2,2	-0,2	-3,7	-7,2	-10,4
26	-0,4	-1,9	-4,0	-5,9	-7,4
27	-0,7	-1,6	-2,8	-4,8	-6,8
28	-0,4	-1,8	-2,8	-3,8	-5,6
29	0,3	-1,4	-2,9	-4,0	-4,9
30	1,6	-0,6	-2,4	-3,9	-5,2

РАБОЧАЯ ТЕТРАДЬ

Студента _____

ФИ

класс

Вариант № _____

Лабораторная работа №1

Проверка статистических гипотез. Критерий Пирсона.

Из генеральной совокупности извлечена выборка объёма n

- а) Построить полигон частот и гистограмму частот.
- б) Используя критерий Пирсона при уровне значимости $\alpha = 0,01$ проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности с эмпирическим распределением.
- в) Если гипотеза согласуется, то найти доверительный интервал с надежностью $\gamma = 0,99$ неизвестного математического ожидания a по исправленной выборочной дисперсии S^2 и выборочной средней $\bar{x}$.

Решение - 1

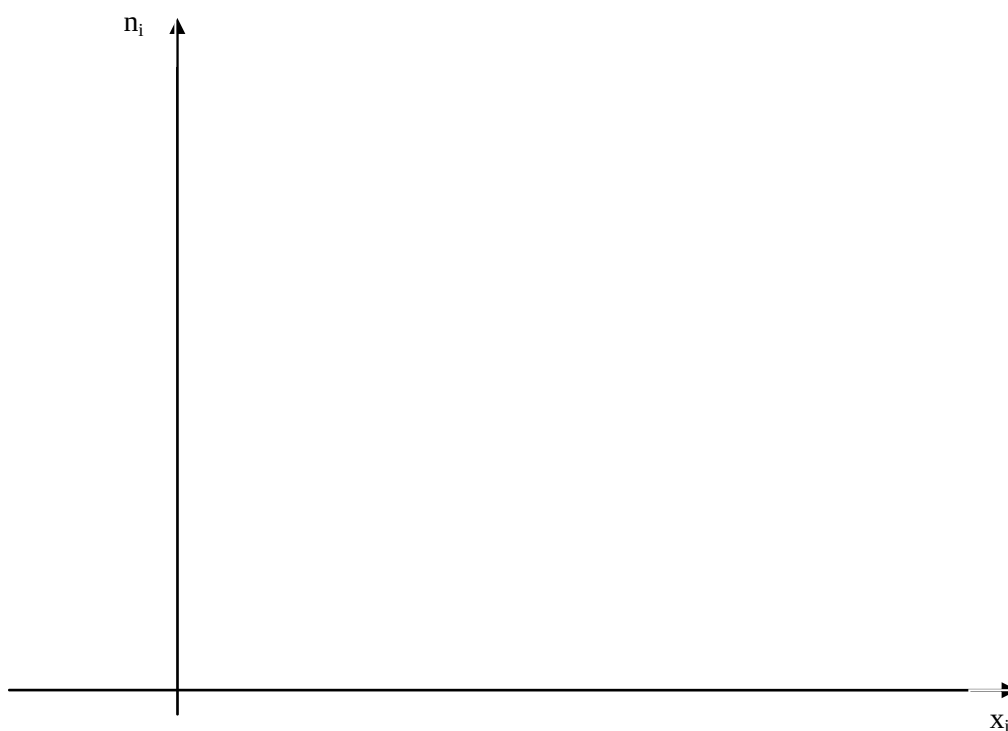
1. Записываем условие задачи.
2. Если случайная величина x_i задана интервальными значениями, то переходим к вариантам x_i^* . Для этого выбираем вариант середины интервалов

$$x_i^* = \frac{x_i + x_{i+1}}{2}$$

x_i	
n_i	

3. Построим полигон частот – ломаную с вершинами в точках $(x_i^*; n_i)$.

x_i^*	
n_i	



4. Построим гистограмму частот – ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиной $h(x_i^*; x_{i+1}^*)$, а высоты равны отношению $\frac{n_i}{h}$.

x_i	
n_i	
$\frac{n_i}{h}$	



5. Записываем условие задачи. Если случайная величина x_i задана интервальными значениями, то переходим к вариантам. Для этого выбираем вариант середины интервалов.

x_i^*	
n_i	

6. Находим параметры выборки:

- объем выборки:

$$n = \sum_{i=1}^m n_i = \underline{\hspace{2cm}};$$

- интервал между значениями вариантов:

$$h = x_i - x_{i+1} = \underline{\hspace{2cm}};$$

- число групп выборки:

$$m = \underline{\hspace{2cm}};$$

- значение постоянной:

$$\frac{nh}{\sigma_{\epsilon}} = \underline{\hspace{2cm}};$$

7. Находим

a . Выборочную среднюю $\bar{x}_{\epsilon}$

$$\bar{x}_b = \frac{\sum_{i=1}^m x_i \cdot n_i}{n} = \underline{\hspace{10cm}}$$

$$\underline{\hspace{10cm}} = \underline{\hspace{10cm}}.$$

b. выборочную D_ϵ

$$D_\epsilon = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x}_\epsilon)^2 \cdot n_i = \underline{\hspace{10cm}}$$

$$\underline{\hspace{10cm}} = \underline{\hspace{10cm}}.$$

c. среднее квадратическое отклонение

$$\sigma_\epsilon = \sqrt{D_\epsilon} = \underline{\hspace{10cm}}.$$

8. Составим таблицу, в которую внесем значения $x_i, u_i, \varphi(u_i), n_i^T$, а затем найдем

$$\chi_{набл}^2.$$

№	x_i	$u_i = \frac{x_i - \bar{x}_\epsilon}{\sigma_\epsilon}$	$\varphi(u_i)$	n_i	$n_i^T = \frac{nh}{\sigma_\epsilon} \varphi(u_i)$	$\frac{(n_i - n_i^T)^2}{n_i^T}$

						$\chi^2_{набл} =$

9. Просуммировав значения последнего столбца, найдем $\chi^2_{набл} =$ _____.
10. По таблице «Критические точки распределения χ^2 » при значимости $\alpha = 0,01$, и числу степеней свободы $k = m - 3 =$ _____, находим
- $\chi^2_{кр} =$ _____.
11. Если $\chi^2_{набл} < \chi^2_{кр}$, то нет оснований отвергать гипотезу о нормальном распределении генеральной совокупности.
12. Искомый интервал для неизвестной дисперсии генеральной совокупности определяется из неравенства

$$\bar{x}_B - t \cdot \frac{S}{\sqrt{n}} < a < \bar{x}_B + t \cdot \frac{S}{\sqrt{n}}, \text{ где } S^2 = \frac{n}{n-1} D_B.$$

Где t находим из функции Лапласа $\Phi(t) = \frac{\gamma}{2} = \frac{0,99}{2} = 0,495, \Rightarrow t = 2,58$.

$$\bar{x}_B - t \cdot \frac{S}{\sqrt{n}} = \underline{\hspace{10cm}}$$

$$\bar{x}_B + t \cdot \frac{S}{\sqrt{n}} = \underline{\hspace{10cm}}$$

Т.о., математическое ожидание генеральной совокупности с надежностью $\gamma = 0,99$ находится внутри интервала

(_____; _____).

Вывод:

Решение - 2

Записываем условие задачи.

x_i	
n_i	

Если случайная величина x_i задана интервальными значениями, то переходим к вариантам x_i^* . Для этого выбираем вариант середины интервалов $x_i^* = \frac{x_i + x_{i+1}}{2}$

13. Построим полигон частот – ломаную с вершинами в точках $(x_i^*; n_i)$.

x_i^*	
n_i	



14. Построим гистограмму частот – ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиной h $(x_i^*; x_{i+1}^*)$, а высоты равны отношению $\frac{n_i}{h}$.

x_i	
n_i	
$\frac{n_i}{h}$	



15. Записываем условие задачи. Если случайная величина x_i задана интервальными значениями, то переходим к вариантам. Для этого выбираем вариант середины интервалов.

x_i^*	
n_i	

16. Находим объем выборки:

$$n = \sum_{i=1}^m n_i = \underline{\hspace{2cm}};$$

5. Находим

а) Выборочную среднюю $\bar{x}_g$

$$\bar{x}_g^* = \frac{\sum_{i=1}^m x_i^* \cdot n_i}{n} = \underline{\hspace{10cm}}$$

$$\underline{\hspace{10cm}} = \underline{\hspace{10cm}}$$

b) выборочную D_g

$$D_g^* = \frac{1}{n} \sum_{i=1}^m \left(x_i^* - \bar{x}_g^* \right)^2 \cdot n_i = \underline{\hspace{10cm}}$$

$$\underline{\hspace{10cm}} = \underline{\hspace{10cm}}.$$

c) среднее квадратическое отклонение

$$\sigma_g^* = \sqrt{D_g^*} = \underline{\hspace{10cm}}.$$

6. Составим таблицу, в которую внесем значения $x_i, x_{i+1}, z_i, z_{i+1}, \Phi(z_i), \Phi(z_{i+1}), p_i, n_i^T$, а затем найдем $\chi_{набл}^2$, просуммировав значения последнего столбца.

№	x_i	x_{i+1}	$z_i =$ $= \frac{x_i - \bar{x}_g^*}{\sigma_g^*}$	$z_{i+1} =$ $= \frac{x_{i+1} - \bar{x}_g^*}{\sigma_g^*}$	$\Phi(z_i)$	$\Phi(z_{i+1})$	$p_i =$ $= \Phi(z_{i+1}) - \Phi(z_i)$	$n_i^T = np_i$	$\frac{(n_i - n_i^T)^2}{n_i^T}$

--	--	--	--	--	--	--	--	--	--

									$\chi^2_{набл} =$
--	--	--	--	--	--	--	--	--	-------------------

7. Просуммировав значения последнего столбца, найдем $\chi^2_{набл} =$ _____.

8. По таблице «Критические точки распределения χ^2 » при значимости $\alpha = 0,01$, и числу степеней свободы $k = m - 3 =$ _____, находим $\chi^2_{кр} =$ _____.

9. Если $\chi^2_{набл} < \chi^2_{кр}$, то нет оснований отвергать гипотезу о нормальном распределении генеральной совокупности.

10. Искомый интервал для неизвестной дисперсии генеральной совокупности определяется из неравенства

$$\bar{x}_B^* - t \cdot \frac{S}{\sqrt{n}} < a < \bar{x}_B^* + t \cdot \frac{S}{\sqrt{n}}, \text{ где } S^2 = \frac{n}{n-1} D_B,$$

где t находим из функции Лапласа $\Phi(t) = \frac{\gamma}{2} = \frac{0,99}{2} = 0,495, \Rightarrow t = 2,58$.

$$\bar{x}_B^* - t \cdot \frac{S}{\sqrt{n}} = \underline{\hspace{10cm}}$$

$$\bar{x}_B^* + t \cdot \frac{S}{\sqrt{n}} = \underline{\hspace{10cm}}$$

Т.о., математическое ожидание генеральной совокупности с надежностью $\gamma = 0,99$ находится внутри интервала

(_____ ; _____).

Вывод:

Лабораторная работа № 2

Корреляционный анализ

По данной корреляционной таблице

- а) найти коэффициент корреляции;
- б) найти выборочное уравнение прямой y/x ;
- в) при уровне значимости $\alpha = 0,05$ определить, значимо или незначимо отличается генеральный коэффициент корреляции от нуля;
- г) сделать вывод о тесноте корреляции y на x ;

д) построить графики эмпирической корреляционной зависимости и полученной теоретической.

Решение.

1. Заменяем, если это нужно, интервальные значения переменных вариантами, выбрав середины интервалов за их значения. Получим по условию варианта корреляционную таблицу:

$x \backslash y$						n_y
n_x						100

2. Найдем коэффициент корреляции по формуле $r = \frac{\frac{1}{n} [\sum X_i Y_i] - \bar{X} \bar{Y}}{\sigma_x \sigma_y}$, для чего

вычислим:

а. выборочные средние признаков X и Y :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^5 x_i n_{xi} = \underline{\hspace{15cm}}$$

$$\underline{\hspace{15cm}} = \underline{\hspace{15cm}}$$

$$\bar{y} = \frac{1}{n} \sum_{j=1}^4 y_j n_{yj} = \underline{\hspace{15cm}}$$

_____ = _____.

Выборочные среднеквадратические отклонения, предварительно вычислив соответствующие дисперсии:

$$D_x = \frac{1}{n} \sum_{i=1}^5 (x_i - \bar{x})^2 n_{xi} = \underline{\hspace{10cm}}$$

_____ = _____.

$$\sigma_x = \underline{\hspace{10cm}}.$$

$$D_y = \frac{1}{n} \sum_{j=1}^4 (x_j - \bar{x})^2 n_{yj} = \underline{\hspace{10cm}}$$

_____ = _____.

$$\sigma_y = \underline{\hspace{10cm}}.$$

б. выборочный корреляционный момент:

$$K_{xy} = \frac{1}{n} \sum_{i=1}^5 \sum_{j=1}^4 x_i y_j n_{ij} - \bar{xy} = \underline{\hspace{10cm}}$$

$$=$$

с. коэффициент корреляции:

$$r = \frac{K_{xy}}{\sigma_x \sigma_y} =$$

3. Составим выборочное уравнение прямой Y на X вида $y - \bar{y} = \rho_{yx}(x - \bar{x})$, где коэффициент регрессии

$$\rho_{yx} = r \frac{\sigma_y}{\sigma_x} =$$

уравнение регрессии Y на X будет иметь вид:

4. Выясним, значимо или незначимо отличается генеральный коэффициент корреляции от нуля. Для этого найдем

$$T_{набл} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} =$$

По таблице критических точек распределения Стьюдента имеем

$$t_{кр.} = (\alpha; n-2) =$$

Если $T_{набл.} > t_{кр.}$, то гипотезу о равенстве нулю генерального коэффициента корреляции отвергаем.

5. Так как $|r|$ близок к 1, то можно утверждать, X и Y хорошо коррелируемы.

Для того, чтобы построить график эмпирического распределения, необхо-

димо найти условные средние значения $\bar{y}_i = \frac{\sum_{j=1}^5 y_j n_{ij}}{n_{xi}}$.

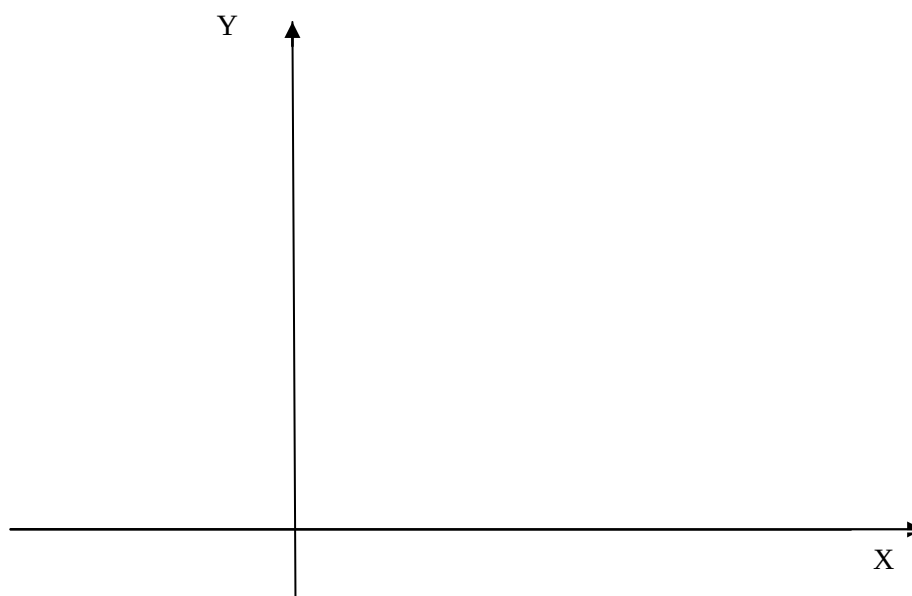
$\bar{y}_1 =$ _____; $\bar{y}_2 =$ _____; $\bar{y}_3 =$ _____;

$\bar{y}_4 =$ _____; $\bar{y}_5 =$ _____.

Получим таблицу, определяющую соответствие между значениями x_i и условными средними $\bar{y}_i$:

x_i					
$\bar{y}_i$					

6. Построим в декартовой системе координат точки с координатами $(x_i; \bar{y}_i)$. В той же системе координат построим график теоретического распределения.



Вывод:

Лабораторная работа № 3 Метод наименьших квадратов

По данным исходной таблицы значения y_i , соответствующих аргументам $x_1 = 1, \quad x_2 = 2, \quad x_3 = 3, \quad x_4 = 4, \quad x_5 = 5$, аппроксимировать функцией $f(x) = ax + b$.

Решение.

1. Заполнить таблицу исходных данных:

x	1	2	3	4	5
y					

2. Составить систему для нахождения a и b (должна получиться система двух уравнений с двумя неизвестными). Для нахождения коэффициентов системы предварительно удобно заполнить следующую таблицу:

№пп	1	2	3	4	5	Σ
x	1	2	3	4	5	15

x^2						
y						
xy						

В последнем столбце стоят суммы элементов строк.

3. Составить систему:

$$\begin{cases} a \sum_{i=1}^5 x_i^2 + b \sum_{i=1}^5 x_i = \sum_{i=1}^5 x_i y_i \\ a \sum_{i=1}^5 x_i + 5b = \sum_{i=1}^5 y_i \end{cases}.$$

4. Решить полученную систему.

$$a = \underline{\hspace{10cm}}; b = \underline{\hspace{10cm}}.$$

Т.о. уравнение имеет вид:

$$f(x) = ax + b \underline{\hspace{10cm}}.$$

5. Найти теоретические ординаты $f(x_i) = \bar{y}_i = ax_i + b$, ($i = \overline{1,5}$), при $x = 1, 2, 3, 4, 5$.

$$\bar{y}_1 = \underline{\hspace{10cm}}; \bar{y}_2 = \underline{\hspace{10cm}}; \bar{y}_3 = \underline{\hspace{10cm}};$$

$$\bar{y}_4 = \underline{\hspace{10cm}}; \bar{y}_5 = \underline{\hspace{10cm}}.$$

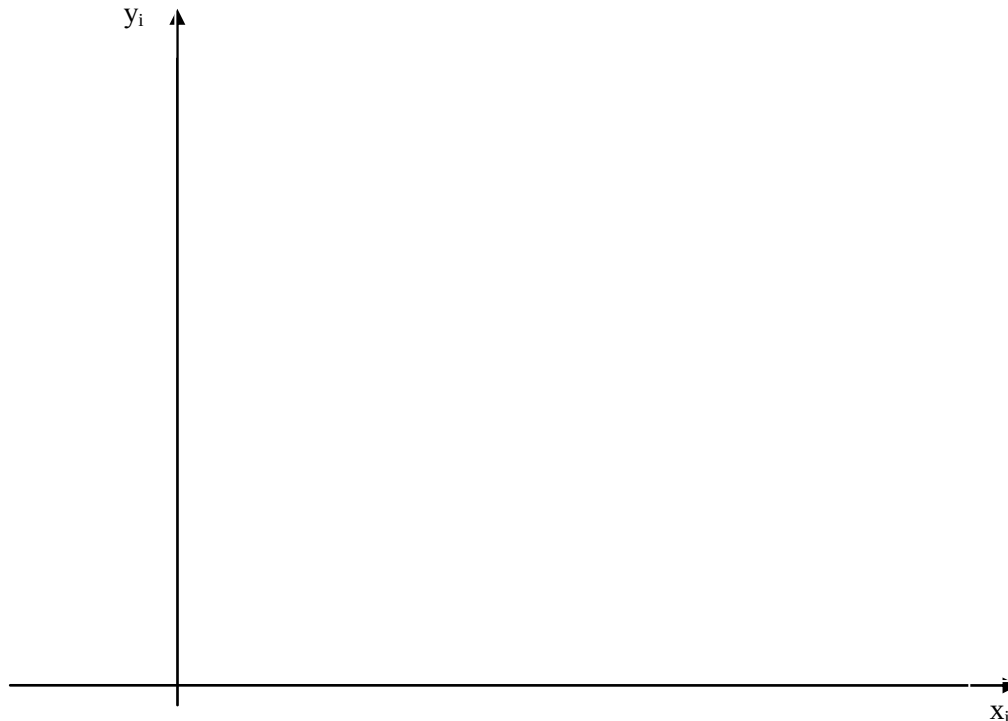
6. Найти отклонения экспериментальных данных от полученных теоретических:

$$\delta_i = y_i - \bar{y}_i = y_i - (a \cdot x_i + b), \quad (i = \overline{1,5}).$$

7. Найти среднеквадратичную ошибку

$$\sigma = \sqrt{\frac{\sum_{i=1}^5 \delta_i^2}{5}} = \underline{\hspace{15cm}}.$$

8. Построить полученную теоретическую кривую $f(x) = ax + b$ и заданные точки.



Вывод:

ПРИЛОЖЕНИЕ 1

Таблица значений функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127

Продолжение прилож. 1

[illegible]

ПРИЛОЖЕНИЕ 2

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

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,4738	2,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997
1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499997

ПРИЛОЖЕНИЕ 3

Таблица значений $t_{\gamma} = t(\gamma, n)$

$\gamma \backslash n$	0,95	0,99	0,999	$\gamma \backslash n$	0,95	0,99	0,099
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

ПРИЛОЖЕНИЕ 4

Критические точки распределения χ^2

Число степеней свободы k	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Гмурман В.Е. Теория вероятностей и математическая статистика: учеб. пособие для вузов. – 10-е изд., стер. - М.: Высш.шк., 2003. - 479 с.
2. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учеб. пособие для вузов. - 9-е изд., стер. – М.: Высш. шк., 2004. - 404 с.
3. Теория вероятностей и математическая статистика в задачах: учеб. пособие для вузов / В.А.Ватутин, Г.А.Ивченко, Ю.И.Медведев и др. – 2-е изд., испр. – М.: Дрофа, 2003. – 328 с.: ил.

**Чернявская Светлана Алексеевна
Третьякова Лилия Владимировна**

**СТАТИСТИЧЕСКАЯ
ОЦЕНКА ДАННЫХ
ДЛЯ НЕКОТОРЫХ ЗАДАЧ
В ЭКОЛОГИИ
(Рабочая тетрадь)**

Учебное пособие

В авторской редакции

Изд. № 137/19. Объём 4,5 п.л.
РИИЦМ ФГАОУ ВО "Севастопольский государственный университет"