

Ганичева Антонина Валериановна
Тверская государственная сельскохозяйственная академия
Ганичев Алексей Валерианович
Тверской государственный технический университет

**Проблемная лекция «Проверка гипотез о законе
распределения случайной величины»**

Аннотация. Одной из важных форм интерактивного обучения является интерактивная лекция. С помощью проблемной лекции достигается, например, более качественное усвоение теоретических знаний, формирование исследовательской компетентности будущих специалистов, мотивация студентов к самостоятельному изучению учебного материала, развитие творческого мышления, повышение познавательного интереса к содержанию учебного предмета. В данной статье рассмотрены особенности проблемной лекции по одной из самых сложных тем математической статистики – проверке гипотез о виде закона распределения наблюдаемых экспериментальных данных. Для освоения данной темы требуется не только глубокое понимание основ теории вероятностей и математической статистики, но и применения навыков статистического анализа, а также понимания существа вероятностных суждений. Критерий согласия Пирсона предлагается рассматривать на конкретных примерах проверки гипотез о равномерном, показательном законах распределения, а также распределении Пуассона.

Ключевые слова: гипотеза, закон распределения вероятностей, генеральная совокупность, частота, интервал, гистограмма, критерий

Ganicheva Antonina Valerianovna
Tver State Agricultural Academy
Ganichev Alexey Valerianovich
Tver State Technical University

**Problematic lecture «Checking hypotheses about the law
of distribution of a random variable»**

Abstract. One important form of interactive learning is the interactive lecture. Problem-based lectures can, for example, enhance the acquisition of theoretical knowledge, develop research competence in future specialists, motivate students to independently study course material, develop creative thinking, and enhance cognitive interest in the subject matter. This article examines the specifics of problem-based lectures on one of the most complex topics in mathematical statistics: testing hypotheses about the distribution law of observed experimental data. Mastering this topic requires not only a deep understanding of the fundamentals of probability theory and mathematical statistics but also the application of statistical analysis skills and an understanding of the nature of probability judgments. Pearson's goodness-of-fit test is proposed for examination using specific examples of testing hypotheses about uniform, exponential, and Poisson distributions.

Keywords: hypothesis, probability distribution law, general population, frequency, interval, histogram, criteria

Введение

Проблемные лекции рассмотрены во многих источниках научной литературы. Отмечается их актуальность [4], роль в современном образовательном пространстве [3], особенности подготовки [6], методика проведения [7], результаты [2, 5, 8].

Индивидуальные особенности проведения проблемной лекции имеются во многих учебных дисциплинах, и их следует учитывать в педагогической практике.

Проверка статистических гипотез – одна из основных проблем социально – производственной сферы. Как определить согласованность статистических данных с выдвинутыми предположениями относительно функционального и числового описания той или иной ситуации, того или иного явления, той или иной проблемы. Для решения подобных задач используется метод проверки статистических гипотез. Выделяют проверку параметрических (о параметрах законов распределения случайных величин) и проверку гипотез о виде закона распределения.

Для учащихся тема «Проверка гипотез о законе распределения случайной величины» является одной из самых сложных для усвоения. Поэтому для ее изучения применяются интерактивные методы, которые помогают лучше понять материал и закрепить навыки. Для изложения теоретического материала по данной теме может использоваться проблемная лекция.

Данная статья посвящена рассмотрению особенностей проблемной лекции по оценке статистических гипотез о законе распределения случайной величины (генеральной совокупности, признака).

Материалы и методы

На проблемной лекции рассматривается проверка статистической гипотезы, связанной со статистическим рядом некоторой генеральной совокупности, закон распределения которой неизвестен, и мы можем судить о нем только с некоторой вероятностью, равной, например, 0,9; 0,95; 0,99. Однако абсолютно точно сделать вывод о законе распределения случайной величины (генеральной совокупности) мы не можем.

Возникает проблема проверки закона распределения. Учащиеся подводят к мысли, что для этого нужно сравнить данное статистическое распределение с теоретическим распределением, выдвинутым согласно гипотетическому предположению. Они вспоминают, какие законы распределения знают. Преподаватель предлагает рассмотреть конкретный пример. По статистическому ряду обучаемые строят гистограмму, на основе которой выдвигают гипотезу о виде закона распределения случайной величины. Преподаватель ставит вопрос, насколько они уверены в своем предположении? Возникает сложная проблема оценки данной уверенности. Здесь уместно поговорить об ошибках первого и второго рода. Вместе с обучаемыми делается вывод о возможной формуле для вычисления погрешности, т. е. о формуле Пирсона. Далее на конкретных примерах рассматривается эта формула для различных законов распределения непрерывных и дискретных случайных величин. Обучаемые должны очень хорошо почувствовать разницу в определении погрешности для разных законов распределения вероятностей.

В заключение предложить учащимся самим придумать аналогичные примеры и выдвинуть гипотезы о законе распределения. Затем построить гистограмму и на ее основе откорректировать свое предположение, после чего воспользоваться формулой Пирсона и сделать вывод о согласовании экспериментальных наблюдений с выдвинутой гипотезой.

Рассматриваемый на проблемной лекции критерий согласия Пирсона (χ^2 - хи-квадрат) является одним из наиболее распространённых методов для проверки гипотезы о законе распределения генеральной совокупности. Он основан на сравнении эмпирических (наблюдаемых) и теоретических частот.

Распределение случайной величины задано интервальной таблицей частот (табл. 1).

Таблица 1 – Интервалы частот

$x_1 \div x_2$	$x_2 \div x_3$		$x_m \div x_{m+1}$
n_1	n_2		n_m

Обозначим: m – количество интервалов, x_i, x_{i+1} – границы интервалов, $(x_{i+1}-x_i)$ – величина интервала, частоты $i = \overline{1, m}$.

Проверить гипотезу о том, что наблюдения имеют данное теоретическое распределение.

Критерий Пирсона имеет вид:

$$\chi^2 = \sum_{i=1}^m \frac{(n_i - n_i^T)^2}{n_i^T} < \chi_{m-r-1; \alpha}^2, \quad (1)$$

где $n_i^T = n \cdot p_i$ – теоретические частоты попадания в i -ый интервал; p_i – вероятность попадания в i -ый интервал; $\chi_{m-r-1; \alpha}^2$ – табличное значение распределения χ^2 для числа степеней свободы $k = m - r - 1$ и уровня значимости α ; r – число параметров гипотетического закона распределения.

Следует обратить внимание учащихся, что при применении данного критерия должно выполняться условие $np_i \geq 5$. В противном случае соответствующий интервал присоединяется к соседнему и число степеней свободы вычисляется для нового числа интервалов. В некоторых источниках используется условие $n_i \geq 5$. В литературе указывается на достаточное выполнение неравенств: $np_i \geq 1$ и $n \geq 50$.

Количество параметров некоторых законов распределения указано в следующей таблице (табл. 2).

Таблица 2 – Количество параметров законов распределения

Закон распределения		Количество параметров
1.	Равномерный	2
2.	Показательный	1
3.	Нормальный	2
4.	Биноминальный	1
5.	Пуассоновский	1

Результаты и их обсуждение

Рассмотрим пример на проверку гипотезы о равномерном законе распределения.

По условию задачи студенты Иванов и Сидоров каждые вторник и пятницу занимаются в спортивной секции. Встречаются в одном и том же месте и идут в спортивный зал. Из-за проблемы с транспортом Сидоров всегда приходит позже Иванова. В табл. 3 приведено время его задержки.

Таблица 3 – Время задержек Сидорова

Время задержки X , в минутах	2-4	4-6	6-8	8-10	10-12	12-14	14-16
n_i (количество)	6	7	6	8	7	8	8

При уровне значимости $\alpha = 0,05$ проверим гипотезу о равномерном распределении времени X . В случае принятия этой гипотезы оценить вероятность того, что в данный вторник Сидоров:

- придет на 2 минуты раньше Иванова?
- придет позднее Иванова более, чем на 10 минут?

в) среднее время опоздания.

Для наглядного представления исходного материала построим гистограмму (рис.

1).

Здесь $n = 50$, $\Delta = 2$. Найдем высоты:

$$h_1 = \frac{6}{2 \cdot 50} = 0,06; \quad h_2 = \frac{7}{2 \cdot 50} = 0,07; \quad h_3 = \frac{6}{2 \cdot 50} = 0,06;$$

$$h_4 = \frac{8}{2 \cdot 50} = 0,08; \quad h_5 = \frac{7}{2 \cdot 50} = 0,07; \quad h_6 = h_7 = \frac{8}{2 \cdot 50} = 0,08.$$

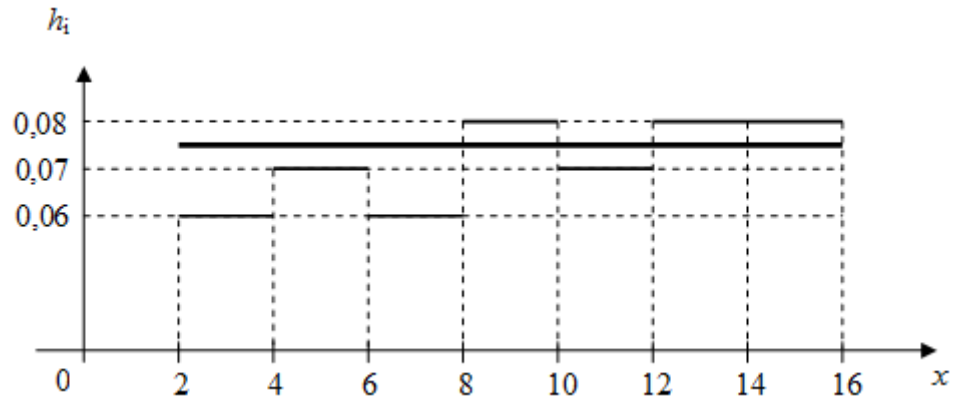


Рис. 1 – Гистограмма задержек Сидорова

Из гистограммы видно, что ее можно аппроксимировать прямой горизонтальной линией (жирная линия на рисунке). Для равномерного распределения плотность распределения вероятностей равна:

$$f(x) = \begin{cases} \frac{1}{(b-a)}, & a < x < b; \\ 0, & x \notin (a,b). \end{cases}$$

Находим среднее значение, дисперсию и среднее квадратическое отклонение:

$$\bar{X} = \frac{1}{50}(3 \cdot 6 + 5 \cdot 7 + 7 \cdot 6 + 9 \cdot 8 + 11 \cdot 7 + 13 \cdot 8 + 15 \cdot 8) = 9,36;$$

$$S^2 = \frac{1}{50}(3^2 \cdot 6 + 5^2 \cdot 7 + 7^2 \cdot 6 + 9^2 \cdot 8 + 11^2 \cdot 7 + 13^2 \cdot 8 + 15^2 \cdot 8) - 9,36^2 = 15,79;$$

$$S = \sqrt{15,79} = 3,97.$$

Для равномерного распределения:

$$a = \bar{X} - S\sqrt{3}, \quad b = \bar{X} + S\sqrt{3};$$

$$a = 9,36 - 3,97 \cdot \sqrt{3} = 2,6; \quad b = 9,36 + 3,97 \cdot \sqrt{3} = 16,11.$$

Полученные значения a и b существенно отличаются от значений в табл. 2. На основе табл. 2 имеем: $a = 2$ и $b = 16$.

Вероятность p_i попадания в i -ый интервал длины 2 будет одинаковой для всех интервалов; $n_i^T = np_i$ - теоретические частоты. Расчеты представлены в табл. 4.

Таблица 4 – Расчетная таблица

Интервалы	(2,4)	(4,6)	(6,8)	(8,10)	(10,12)	(12,14)	(14,16)	Σ
p_i	0,143	0,143	0,143	0,143	0,143	0,143	0,143	1,001
$n_i^T = np_i$	7,15	7,15	7,15	7,15	7,15	7,15	7,15	50,05
n_i	6	7	6	8	7	8	8	50

$n_i - n_i^T$	-1,15	-0,15	-1,15	0,85	-0,15	0,85	0,85	-
$(n_i - n_i^T)^2 / np_i$	0,18	0,003	0,18	0,1	0,003	0,1	0,1	0,67

Получили: $\chi^2 = \sum_{i=1}^7 (n_i - n_i^T)^2 / np_i = 0,67$. Из приложение 4 [1] при $\alpha = 0,05$

и числе степеней свободы $k = 7 - 2 = 5$ находим $\chi_{5;0,05}^2 = 11,1$, т.е. $\chi_{5;0,05}^2 > \chi^2$.

Вывод: выдвинутую гипотезу можно принять на уровне значимости $\alpha = 0,05$. Вероятность того, что Сидоров придет на 2 минуты раньше, равна нулю; вероятность того, что он опоздает более, чем на 10 мину, будет равна $\frac{16-10}{16-2} = 0,43$. Среднее время

задержки равно $\frac{16+2}{2} = 9$ (мин).

Теперь проверим гипотезу о показательном законе распределения. В табл. 5 приведены выборочные данные о времени разговора по телефону.

Таблица 5 - Время разговора по телефону

Время разговора (в мин.)	1-6	6-10	10-15	15-20	20-25	25-30
n_i (частота)	30	15	11	8	4	2

На уровне значимости $\alpha = 0,05$ проверить гипотезу о показательном законе распределения и оценить среднее время продолжительности разговора, разброс этого времени, а также вероятность того, что время разговора будет более 15-ти минут.

Для наглядности сначала построим гистограмму (рис. 2), объединив два последних интервала, чтобы выполнялось условие $n_i \geq 5$.

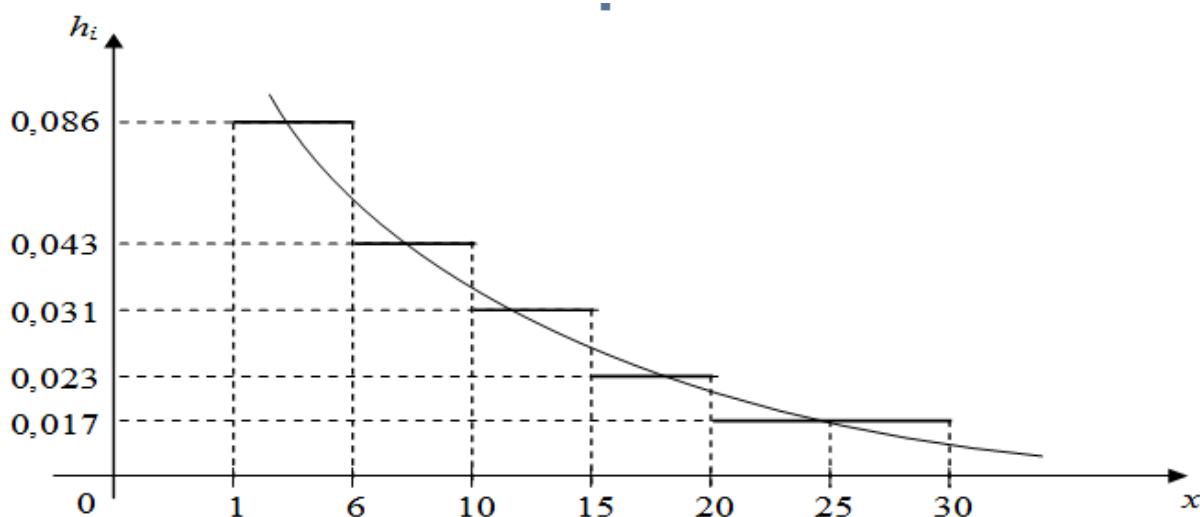


Рис. 2 – Гистограмма времени разговора по телефону

Для данного примера интервал $\Delta = 5$ (мин.), $n = 70$, а высоты равны:

$$h_1 = \frac{30}{70 \cdot 5} = 0,086; h_2 = \frac{15}{70 \cdot 5} = 0,042; h_3 = \frac{11}{70 \cdot 5} = 0,031;$$

$$h_4 = \frac{8}{70 \cdot 5} = 0,023; h_5 = \frac{6}{70 \cdot 5} = 0,017.$$

Из рис. 1 можно предположить показательный закон распределения указанного времени:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Поскольку $\lambda = 1/\bar{X}$, а $\bar{X} = \frac{1}{70}(3,5 \cdot 30 + 8 \cdot 15 + 12,5 \cdot 10 + 17,5 \cdot 8 + 22,5 \cdot 6) = 8,61$, находим $\lambda = 1/8,61 = 0,12$. Тогда

$$f(x) = \begin{cases} 0,12 \cdot e^{-0,12x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Вероятности $p_i (i = \overline{1,5})$ вычисляются по формуле:

$$p_i = e^{-0,12 \cdot x_i} - e^{-0,12 \cdot x_{i+1}},$$

где x_i - левая, x_{i+1} - правая граница i -ого интервала, причем x_1 полагаем равной нулю, а последнюю границу x_6 полагаем равной $+\infty$. Отсюда находим (см. приложение 7 [1]):

$$\begin{aligned} P_1 &= e^{-0} - e^{-0,12 \cdot 6} = 0,51; & P_2 &= e^{-0,12 \cdot 6} - e^{-0,12 \cdot 10} = 0,19; \\ P_3 &= e^{-0,12 \cdot 10} - e^{-0,12 \cdot 15} = 0,14; \\ P_4 &= e^{-0,12 \cdot 15} - e^{-0,12 \cdot 20} = 0,08; & P_5 &= e^{-0,12 \cdot 20} - e^{-\infty} = 0,08. \end{aligned}$$

Дальнейшие вычисления приведены в табл. 6:

Таблица 6 - Расчетная таблица

Интервалы	(1,6)	(6,10)	(10,15)	(15,20)	(20,30)	Σ
p_i	0,51	0,19	0,14	0,08	0,08	1
$n_i^T = np_i$	35,7	13,3	9,8	5,6	4,41	61,11
n_i	30	17	10	8	5	70
$n_i - n_i^T$	-5,7	3,7	0,2	2,4	0,59	-
$(n_i - n_i^T)^2 / np_i$	0,91	1,03	0,004	1,03	0,079	3,053

Таким образом, $\chi^2 = \sum_{i=1}^5 (n_i - n_i^T)^2 / np_i = 3,053$. При $\alpha = 0,05$ и $k = 5 - 1 = 4$ находим $\chi_{4;0,05}^2 = 9,49$. Имеем: $\chi_{4;0,05}^2 > \chi^2$.

Вывод: выдвинутую гипотезу можно принять на уровне значимости $\alpha = 0,05$. Среднее время равно 8,61 минуты, разброс также равен 8,61 минуты.

Рассмотрим теперь гипотезу о распределении дискретной случайной величины с гипотетическим распределением Пуассона.

Дискретная случайная величина задается таблицей частот (табл. 7).

Таблица 7 – Таблица частот

x_i	x_1	x_2		x_m
n_i	n_1	n_2		n_m

Распределение Пуассона описывается формулой:

$$p(i) = \frac{a^i \cdot e^{-a}}{i!}$$

(см. приложение 8 [1]). Для дискретной случайной величины при подсчете числа степеней свободы k вместо числа интервалов берут число разрядов с учетом их объединения. Поскольку распределение Пуассона содержит только один параметр a , при нахождении табличного значения распределения χ^2 (приложение 4) следует положить $r = 1$.

Теоретические частоты в этом случае вычисляются по формуле:

$$n_i^T = n \cdot p(i) = n \cdot \frac{a^i \cdot e^{-a}}{i!}.$$

Рассмотрим следующий пример. Пусть 200 рабочих изготавливают однотипные изделия. Для контроля качества работы выборочно проверено по 10 изготовленных каждым из рабочих изделий. Результаты приведены в табл. 8.

Таблица 8 – Количество бракованных изделий

Количество бракованных изделий x_i	0	1	2	3	4	5	6	7	8	9	10
Число рабочих n_i	109	65	22	3	1	0	0	0	0	0	0

На уровне значимости 0,05 проверим гипотезу о распределении Пуассона числа бракованных изделий.

Шесть последних разрядов не рассматриваем ввиду их нулевой численности рабочих.

Предполагаем, что интересующая нас случайная величина X – число бракованных изделий – имеет закон распределения Пуассона ($n=200$ – число рабочих). Оценкой математического ожидания a случайной величины X является среднее арифметическое числа бракованных изделий.

$$\hat{a} = \frac{1}{n} \sum_{i=0}^4 x_i \cdot n_i = \frac{122}{200} = 0,61.$$

Тогда вероятность принятия случайной величиной X значения i равна:

$$p_i = p(i) = \frac{0,61^i \cdot e^{-0,61}}{i!},$$

а именно: $p_0 = 0,55$; $p_1 = 0,33$; $p_2 = 0,1$; $p_3 = 0,02$; $p_4 = 0,003$.

Теоретические частоты $n_i^T = n \cdot p_i$ приведены в табл. 9.

Таблица 9 – Теоретические частоты

i	0	1	2	3	4
n_i^T	110	66	24,6	4	0,6

Для выполнения условия $np_i \geq 5$ три последних разряда объединяются в один, т. е. $n_2 = 22 + 3 + 1 = 26$; $n_2^T = 20 + 4 + 0,6 = 24,6$.

Расчетное значение критерия будет:

$$\chi^2 = \sum_{i=0}^2 \frac{(n_i - n_i^T)^2}{n_i^T} = \frac{(109 - 110)^2}{110} + \frac{(65 - 66)^2}{66} + \frac{(26 - 24,6)^2}{24,6} = 0,104.$$

Табличное значение критерия χ^2 для числа степеней свободы $3-2=1$ и уровня значимости 0,05:

$$\chi_{1; 0,05}^2 = 3,84.$$

Поскольку $0,104 = \chi^2 < \chi_{1; 0,05}^2 = 3,84$, гипотеза о распределении числа бракованных изделий по закону Пуассона принимается.

Заключение

На проблемной лекции учащиеся познакомились с основными понятиями проверки статистических гипотез, рассмотрели примеры по применению рекомендованных методик для генеральной совокупности непрерывной и дискретной случайной величины, тем самым имели возможность «прочувствовать» и соучаствовать решению проблемы оценки неизвестного закона распределения случайной величины (признака).

Список источников

1. Ганичева А.В. Прикладная статистика. СПб: «Лань», 2017. – 172 с.
2. Жолудева В. В. Проблемная лекция как средство повышения эффективности учебного процесса при изучении дисциплин математического цикла // Научно-методический электронный журнал «Концепт», 2016. – Т. 37. – С. 51–55. – URL: <http://e-koncept.ru/2016/56785.htm>.
3. Кириллова С.С. Роль проблемной лекции в современном образовательном пространстве // Наука и Образование. 2022. Т. 5. № 1.
4. Михайлина С.А. Проблемная лекция как актуальная форма интерактивного обучения // Экономические и социально-гуманитарные исследования, 2017. - № 1 (13). – С. 101-106.
5. Молодцова Т.Д., Шалова С.Ю., Кобышева Л.И. Проблемная лекция как средство формирования исследовательской компетентности будущих педагогов // Современные проблемы науки и образования, 2017. - № 4; URL: <https://science-education.ru/ru/article/view?id=26554> (дата обращения: 08.05.2026).
6. Феоктистова О.В. Особенности подготовки современной проблемной лекции // Педагогический вестник, 2025. - № 41. - С. 119-121.
7. Шукаева Е.М. Проблемная лекция как форма организации проблемного обучения // Вестник научных конференций, 2024. - № 12-2 (112). - С. 104-105.
8. Шущков В.И., Поляков В.Н. Проблемная лекция как метод мотивации студента к самостоятельному изучению некоторых разделов высшей математики // Известия Волгоградского государственного технического университета. Серия: Новые образовательные системы и технологии обучения в вузе, 2014. - Т. 12. - № 15 (142). - С. 77-79.

Сведения об авторах

Ганичева Антонина Валериановна, профессор, Тверская государственная сельскохозяйственная академия, Тверь, Россия

Ганичев Алексей Валерианович, старший преподаватель, Тверской государственный технический университет, Тверь, Россия

Information about the authors

Ganishcheva Antonina Valerianovna, Professor, Tver State Agricultural Academy, Tver, Russia

Ganishchev Alexey Valerianovich, Senior Lecturer, Tver State Technical University, Tver, Russia