Дисциплина: «Основы научных исследований»

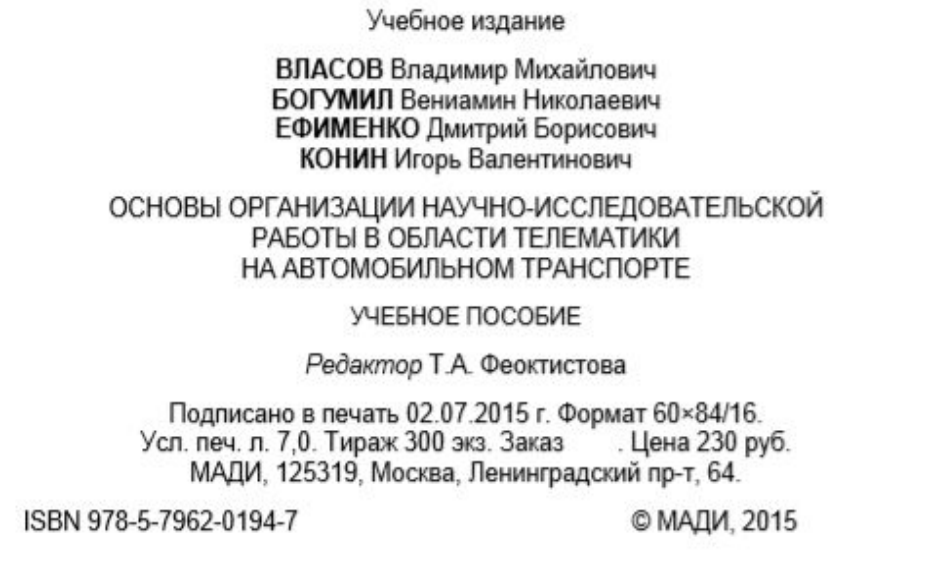
Лабораторная работа № 1

Проверка статистических гипотез.

УЧЕБНЫЕ ВОПРОСЫ:

1. Основные задачи математической статистики.
2. Основные понятия выборочного метода.
3. Вариационные ряды и их характеристики.

Рекомендуемая литература:



Учебное издание

ВЛАСОВ Владимир Михайлович
БОГУМИЛ Вениамин Николаевич
ЕФИМЕНКО Дмитрий Борисович
КОНИН Игорь Валентинович

ОСНОВЫ ОРГАНИЗАЦИИ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ В ОБЛАСТИ ТЕЛЕМАТИКИ НА АВТОМОБИЛЬНОМ ТРАНСПОРТЕ

УЧЕБНОЕ ПОСОБИЕ

Редактор Т.А. Феоктистова

Подписано в печать 02.07.2015 г. Формат 60×84/16.
Усл. печ. л. 7,0. Тираж 300 экз. Заказ . Цена 230 руб.
МАДИ, 125319, Москва, Ленинградский пр-т, 64.

ISBN 978-5-7962-0194-7 © МАДИ, 2015

Учебный вопрос №1 Основные задачи математической статистики.

- указать способ сбора и группирования статистических сведений;
- разработать метод анализа в зависимости от цели исследования.

- Задача статистического моделирования состоит в создании методов сбора и обработки статистической информации для получения научных и практических выводов.

Учебный вопрос №2 Основные понятия выборочного метода.

Сущность выборочного метода состоит в том, чтобы по некоторой части генеральной совокупности вынести суждение о её свойствах в целом.

Генеральная совокупность – это все возможные реализации наблюдаемого явления. Та часть данных, которая отобрана из генеральной совокупности в результате проведенных наблюдений над объектом, называется *выборкой* или выборочной совокупностью.

При изучении некоторой случайной величины X, закон распределения которой неизвестен, и требуется определить этот закон из опыта или проверить экспериментально гипотезу о том, что величина X подчинена тому или иному закону. С этой целью над случайной величиной X производится ряд независимых опытов (наблюдений). В каждом из этих опытов случайная величина X принимает определенное значение. Совокупность данных значений образуют выборку.

Количество наблюдений в выборочной совокупности называется *объём выборки*. Различные выборочные значения изучаемого признака случайной величины называются *вариантами*.

3.2.4. Виды выборок

1. *Собственно-случайная выборка.* Образуется случайным выбором элементов из генеральной совокупности, без их дальнейшего расчленения.

2. *Цензурированная выборка* – это выборка, которая получена из исходной выборочной совокупности путём отбрасывания некоторых выборочных данных.

3. *Выборка с группированием данных.*

4. *Механическая выборка* отбирается путём включения в выборку элементов через определённый интервал (например, каждый 10 – 10% выборка).

5. *Стратифицированная выборка* – элементы генеральной совокупности предварительно разбиваются на группы (страты). Далее из каждого слоя (страты) элементы выбираются случайным образом.

6. *Серийная выборка* – отбираются не элементы, а группы (серии) наблюдений. Сама серия подвергается сплошному изучению.

3.2.5. Основные обозначения

N, n – объём генеральной и выборочной совокупности,

x_i – значение признака выборочной совокупности для с.в. X,

N_i, n_i – число значений в генеральной совокупности и в выборочной совокупности с признаком x_i.

Пример 3.1. Произведена статистическая выборка значений пробегов до отказа бортовых навигационно-связных терминалов, установленных на автомобили КамАЗ объемом n = 30. Полученные выборочные значения сведены в табл. 3.1.

Таблица 3.1

Пробеги до отказа бортовых навигационно-связных терминалов, установленных на автомобили КамАЗ (тыс. км)

№ п/п	x_i	№ п/п	x_i	№ п/п	x_i	№ п/п	x_i	№ п/п	x_i	№ п/п	x_i
1	**81**	6	**72**	11	**82**	16	**112**	21	**134**	26	**84**
2	**99**	7	**52**	12	**96**	17	**118**	22	**90**	27	**66**
3	**88**	8	**25**	13	**90**	18	**74**	23	**46**	28	**88**
4	**63**	9	**126**	14	**58**	19	**62**	24	**94**	29	**116**
5	**106**	10	**104**	15	**98**	20	**98**	25	**82**	30	**78**

Учебный вопрос №3 - Вариационные ряды и их характеристики.

3.3.1. Одним из способов обработки простого статистического ряда является построение вариационного ряда

Определение. *Вариационный ряд* – ранжированные в порядке возрастания (или в порядке убывания) выборочные значения.

Вариационный ряд, построенный по данным пробегов до отказа бортовых навигационно-связных терминалов, установленных на автомобили КамАЗ (тыс. км)

№ п/п	x_i	№ п/п	x_i	№ п/п	x_i	№ п/п	x_i	№ п/п	x_i	№ п/п	x_i
1	**25**	6	**63**	11	**81**	16	**88**	21	**98**	26	**112**
2	**46**	7	**66**	12	**82**	17	**90**	22	**98**	27	**116**
3	**52**	8	**72**	13	**82**	18	**90**	23	**99**	28	**118**
4	**58**	9	**74**	14	**84**	19	**94**	24	**104**	29	**126**
5	**62**	10	**78**	15	**88**	20	**96**	25	**106**	30	**134**

Пример построения гистограммы (ступенчатого графика) случайной величины

Если выборка состоит из достаточно большого числа единиц (несколько десятков и более) и распределение данных выборки близко к нормальному, то можно использовать формулу Стерджесса для выбора количества равных интервалов разбиения:

$$N = 1 + 3{,}322 \lg N, \qquad (3.2)$$

где n – число интервалов разбиения, N – объем выборки.

Величина интервала определяется по формуле:

$$\ln t = \frac{x_{max} - x_{min}}{n = 1 + 3{,}322 \times \lg N}, \qquad (3.3)$$

где x_{max} – максимальное значение признака в изучаемой совокупности; x_{min} – минимальное значение признака в изучаемой совокупности; n – количество групп; N – объем выборки.

Тогда левая граница первого интервала ($x_{1л}$) совпадает с величиной x_{min} ($x_{1л} = x_{min}$) Правая граница первого интервала определится по формуле:

$$x_{1п} = x_{1л} + \ln t. \qquad (3.4)$$

Левая граница второго интервала совпадает с правой границей первого интервала ($x_{2л} = x_{1п}$). Правая граница первого интервала определится по формуле:

$$x_{1п} = x_{1л} + \ln t. \qquad (3.5)$$

Гистограмма строится следующим образом. По оси абсцисс откладываются интервалы, и на каждом интервале как на основании строится прямоугольник. Высота i-го прямоугольника (i = 1, 2, …, n) равна количеству данных выборки (m_i), попавших в i-й интервал.

Пример 3.3. В качестве примера построим гистограмму по данным статистического ряда, рассмотренного в примере 3.1 (табл. 3.1, 3.2). По формуле Стерджесса определяем число групп:

$$N = 1 + 3{,}322 \log N = 1 + 3{,}322 \log 30 = 5{,}9. \qquad (3.6)$$

Принимаем $n = 6$.

1. Определяем величину интервала по формуле:

$$\text{Int}\,t = (x_{max} - x_{min})/n = (134 - 25)/6 = 109/6 \approx 18. \qquad (3.7)$$

2. Найдем границы интервалов. Произведем группировку выборочных данных в соответствии с построенными интервалами. Сведем полученные значения в табл. 3.3.

Таблица 3.3

Распределение данных выборки пробегов до отказа бортовых навигационно-связных терминалов, установленных на автомобили КамАЗ

Интервал	Диапазон	Частота, *f*
1	25–43	1
2	43–61	3
3	61–79	6
4	79–97	10
5	97–115	7
6	115–134	4

3. С использованием средств «Microsoft Graph» строим ступенчатый график гистограммы выборки (рис. 3.1).

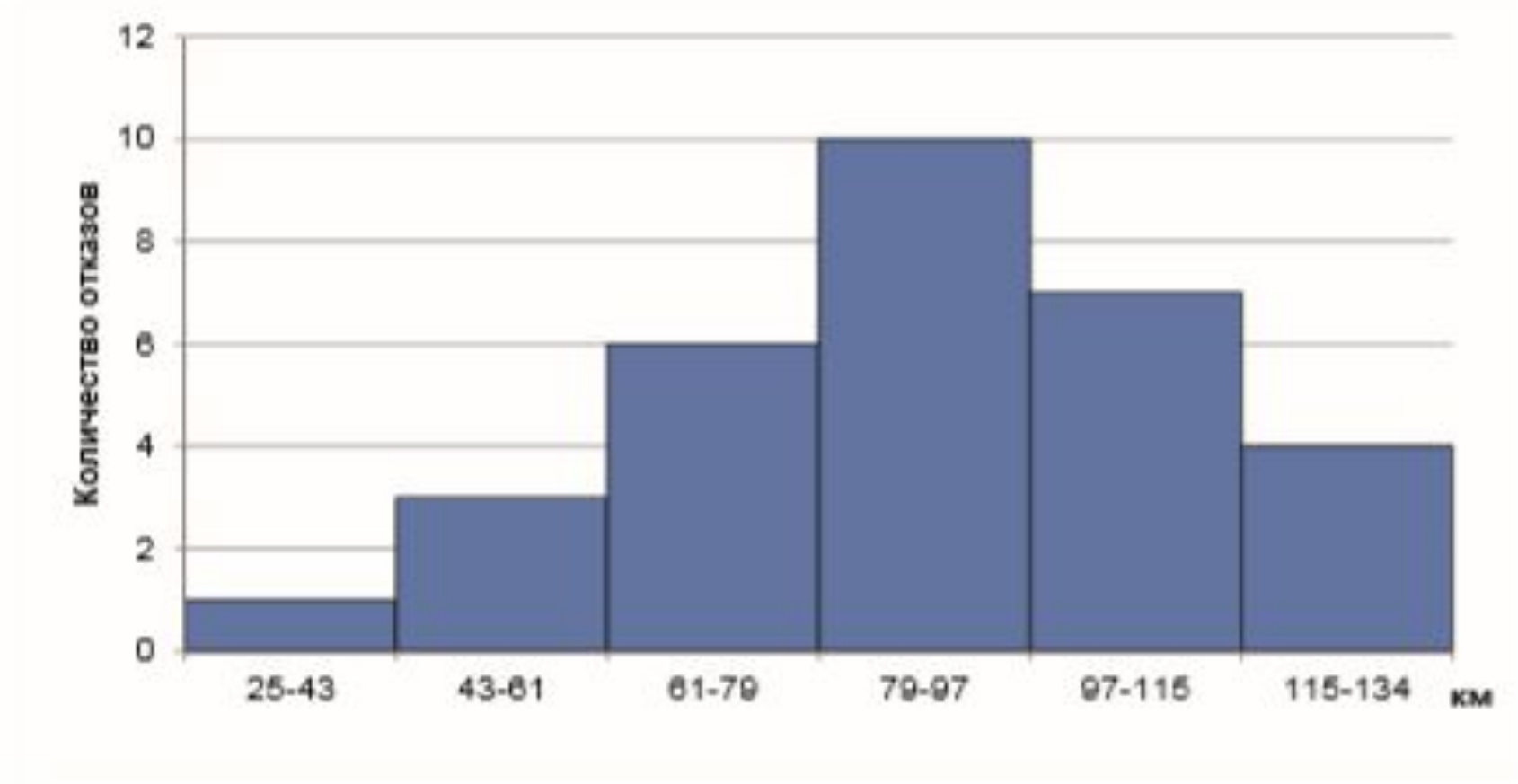


Рис. 3.1. Гистограмма выборочных значений пробегов до отказа бортовых навигационно-связных терминалов, установленных на автомобили КамАЗ