

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«Московский авиационный институт  
(национальный исследовательский университет)»  
Кафедра «Моделирование систем и информационные технологии»

**ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТОВ**  
**Первичная статистическая обработка**  
**(малые выборки)**

Методические указания к практическим занятиям  
по дисциплине «Математическая статистика»

**Составители:** Егорова Ю.Б.  
Мамонов И.М.

МОСКВА 2019

## ВВЕДЕНИЕ

Цель работы - ознакомиться с порядком проведения первичной статистической обработки результатов измерений в случае сверхмалых выборок (менее 10 измерений).

### 1. СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ МАЛЫХ ВЫБОРОК

В процессе испытаний измерения не могут быть выполнены абсолютно точно. Это связано с появлением различных погрешностей (неоднородность структуры реальных металлов и сплавов, наличие дефектов, инструментальные погрешности и т.п.). Поэтому для получения достоверных данных численное значение изучаемого свойства определяют по результатам нескольких измерений. Совокупность этих значений обозначается  $n$  и называется статистической выборкой или выборочной совокупностью.

Основная цель первичной статистической обработки в случае малых выборок – оценка истинного значения изучаемого свойства и ошибки его определения по выборке.

Порядок и этапы проведения первичной статистической обработки зависят от объема выборки. Объем выборки при испытаниях может меняться в широких пределах: от 2-5 до нескольких сотен измерений.

Рассмотрим, как проводится первичная статистическая обработка, если объем выборки составляет менее 10 измерений. В этом случае определяют выборочные числовые характеристики и строят доверительный интервал для истинного значения (математического ожидания) измеряемого свойства.

По данным выборки определяют следующие числовые характеристики:

- 1) **выборочное среднее**  $\bar{x}$  характеризует среднее значение исследуемой величины и является статистической оценкой ее истинного значения:

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

- 2) **выборочная дисперсия** характеризует средний квадрат отклонения значений случайной величины от выборочного среднего; в случае малых выборок определяют «исправленную» дисперсию  $S^2$ :

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1},$$

- 3) **выборочное среднее квадратическое отклонение** (стандартное отклонение)  $S$  характеризуют отклонение значений случайной величины от среднего:

$$S = \sqrt{S^2}.$$

- 4) **выборочная медиана  $Me^*$**  – значение, соответствующее середине вариационного ряда.  
5) **выборочная мода  $Mo^*$**  - наиболее часто встречающееся значение.  
6) **выборочный коэффициент асимметрии** характеризует симметричность эмпирического распределения:

$$A^* = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)S^3};$$

- 6) **выборочный эксцесс  $E^*$  и коэффициент эксцесса  $\varepsilon^*$ :**

$$E^* = \varepsilon^* - 3 = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(n-1)S^4} - 3.$$

**Доверительный интервал для истинной величины** измеряемого свойства (математического ожидания) можно

построить только в том случае, если результаты измерений имеют нормальный закон распределения (или близкий к нему).

**Приблизительная проверка выборочного распределения на нормальность** в случае малых выборок может быть проведена несколькими способами, которые дополняют друг друга:

- 1) визуальный метод предназначен для предварительной субъективной оценки и может быть осуществлен по виду полигона (применяется только в том случае, если есть повторяющиеся значения);
- 2) приблизительная проверка, состоящая в том, что для нормального теоретического распределения  $M(X)=M_0=Me$ . Если по выборке  $\bar{x} \approx M_0^* \approx Me^*$ , то распределение можно считать приблизительно нормальным;
- 3) построение трехсигмового интервала:  $m \pm 3\sigma \approx \bar{x} \pm 3S$ . Если все значения случайной величины попали в трехсигмовый интервал, то распределение можно считать приблизительно нормальным.

**Доверительный интервал** для истинной величины измеряемого свойства  $m$  при уровне надежности  $\gamma=0,95$  можно найти по следующей формуле:

$$\bar{x} - \delta < m < \bar{x} + \delta ,$$

$$\delta = t_{\gamma,n} \frac{S}{\sqrt{n}},$$

где  $\delta$  – точность оценки (погрешность измерения):

$t_{\gamma,n}$  – квантиль распределения Стьюдента, который определяется по специальным таблицам в зависимости от  $\gamma$  и  $k=n-1$ ;  $\gamma$  – доверительная вероятность (надежность оценки);  $n$  – объем выборки. Для  $\gamma=0,95$  и  $n=5$  квантиль  $t_{\gamma,n}=2,78$ .

**ПРИМЕР.** Провести первичную статистическую обработку результатов измерений твердости образцов из легированной стали, приведенных в протоколе испытаний:

Протокол испытаний (простой статистический ряд)

№ опыта	1	2	3	4	5
<i>HV</i>	47	55	53	52	48

**Решение.**

1. Определим выборочные числовые характеристики:

1.1) Выборочное среднее (среднее арифметическое):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{47 + 55 + 53 + 52 + 48}{5} = 51 \text{ HV};$$

1.2) «Исправленная» выборочная дисперсия:

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} = \frac{(47 - 51)^2 + (55 - 51)^2 + (53 - 51)^2 + (52 - 51)^2 + (48 - 51)^2}{5 - 1} =$$

$$= 11,5 \text{ (HV)}^2;$$

1.3) Стандартное отклонение:

$$S = \sqrt{S^2} = \sqrt{11,5} = 3,39 \text{ HV}.$$

1.4) Выборочную моду определить нельзя, так как нет повторяющихся значений.

1.5) Выборочная медиана:  $Me=52 \text{ HV}$  в соответствии с вариационным статистическим рядом:

№ пп	1	2	3	4	5
<i>HV</i>	47	48	52	53	55

2. Проведем приблизительную проверку нормальности распределения:

2.1) по полигону проверить нельзя, так как нет повторяющихся значений;

2.2) статистическое распределение можно считать приблизительно нормальным, так как:

$$\bar{x} \approx Me^*$$

2.3) найдем «трехсигмовый интервал»:

$$\bar{x} \pm 3S = 51 \pm 3 \cdot 3,39 = 40,83; 61,17.$$

Все измеренные значения находятся в этом интервале, поэтому распределение можно считать приблизительно нормальным.

3. Найдем доверительный интервал для истинного значения твердости  $m$  по формуле:

$$\bar{x} - \delta < m < \bar{x} + \delta.$$

По таблице распределения Стьюдента при доверительной вероятности  $\gamma=0,95$  найдем:

$$t_{\gamma,n} = t_{0,95;5} = 2,78.$$

Определим предельную ошибку выборки:

$$\delta = t_{\gamma,n} \frac{S}{\sqrt{n}} = 2,78 \cdot \frac{3,39}{\sqrt{5}} = 4,2 \text{ HB}.$$

Найдем доверительный интервал:

$$\begin{aligned} 51-4,2 < m < 51+4,2; \\ 46,8 < m < 55,2. \end{aligned}$$

**Выводы:**

- 1) Выборочное среднее значение твердости стали равно 51 HB.
- 2) Погрешность измерения твердости составляет 4,2 HB.
- 3) С доверительной вероятностью 0,95 можно ожидать, что истинное значение твердости стали находится в интервале от 46,8 до 55,2 HB.

## 2. ЗАДАНИЯ

Провести первичную статистическую обработку результатов измерений температуры в зоне резания при точении титановых сплавов (табл. 1).

Таблица 1

№ вар-та	Сплав	$x_{\text{в}}, ^\circ\text{C}$				
		1	2	3	4	5
1	BT1-0	445	450	448	450	490
2	BT5Л	774	753	795	742	754
3	BT5-1	698	647	732	658	698
4	BT20	690	640	735	650	696
5	BT18	736	723	725	725	725
6	BT6	750	755	758	745	758
7	BT3-1	750	755	760	745	765
8	BT9	755	760	755	745	770
9	BT8	775	765	780	770	768
10	BT23	825	815	810	825	830
11	BT25	820	815	810	823	830
12	BT35	850	852	860	855	856
13	BT14	755	755	765	745	765
14	BT16	765	760	750	745	770
15	BT15	775	765	785	770	768
16	BT1-0	448	452	448	450	493
17	BT5Л	775	753	795	742	754
18	BT5-1	690	647	732	658	698
19	BT20	690	640	735	650	696
20	BT18	735	723	725	725	725
21	BT6	755	755	758	745	758
22	BT3-1	755	755	760	745	765
23	BT9	755	760	755	745	770
24	BT8	778	765	780	770	768
25	BT23	820	815	810	825	830
26	BT25	825	815	810	823	830
27	BT35	855	852	860	855	856
28	BT14	750	755	765	745	765
29	BT16	760	760	750	745	770
30	BT25	765	765	785	770	768

## **КОНТРОЛЬНЫЕ ВОПРОСЫ**

1. Что называется генеральной совокупностью и выборкой?
2. Цель и порядок проведения первичной статистической обработки результатов измерений в случае малых выборок.
3. Назовите выборочные числовые характеристики, дайте их определения.
4. Дайте определения точечной и интервальной оценок.
5. Какие требования предъявляются к точечным оценкам?
6. Назовите наилучшие точечные оценки математического ожидания и дисперсии генеральной совокупности в зависимости от объема выборки.
7. Что называется доверительным интервалом и доверительной вероятностью?
8. Как можно провести приблизительную проверку гипотезы о нормальном распределении генеральной совокупности в случае малых выборок?