

*Дисциплины:*  
«Теория вероятностей»,  
«Математическая статистика»,  
«Теория вероятностей и  
математическая статистика»

Тема: Проверка  
статистических гипотез



Московский  
Технологический  
Институт

## Проверка домашнего задания

5. В результате измерений некоторой физической величины одним прибором (без систематических ошибок) получены следующие результаты: 10, 13, 13. Тогда несмещенная оценка дисперсии измерений равна:

- а) 6;
- б) 2;
- в) 12;
- г) 3.

6. По городской телефонной сети было произведено 100 наблюдений и установлено, что средняя продолжительность телефонного разговора составляет 4 минут при среднеквадратичном отклонении 2 мин. Предельная ошибка выборки с вероятностью 0,954 составляет

- а) 0,2;
- б) 0,3;
- в) 0,4;
- г) 0,5.



# Вопрос 5. $X_i$ : 10, 13, 13

Несмещенной оценкой дисперсии является  
**исправленная выборочная дисперсия.**

Вычислим выборочное среднее:

$$\bar{x} = \frac{10 + 13 + 13}{3} = \frac{36}{3} = 12$$

Вычислим величину:

$$\overline{x^2} = \frac{10^2 + 13^2 + 13^2}{3} = \frac{100 + 169 + 169}{3} = \frac{438}{3} = 146$$

Вычислим выборочную дисперсию:

$$S^2 = \overline{x^2} - \bar{x}^2 = 146 - 12^2 = 146 - 144 = 2$$

Найдем исправленную дисперсию:

$$\overline{S^2} = \frac{3 \cdot S^2}{2} = \frac{3 \cdot 2}{2} = 3$$



Вопрос 6.  $n=100$ ,  $\bar{x} = 4$ ,  $S=2$ ,  $\gamma=0,954$ ,  
 $\Delta$ -?

При определении среднего значения  
величины  $\Delta$  используется формула

$$\Delta = \frac{t \cdot S}{\sqrt{n}}, \text{ где } \Phi(t) = 0,954 \Rightarrow t \cong 2$$

Получим

$$\Delta = \frac{t \cdot S}{\sqrt{n}} = \frac{2 \cdot 2}{\sqrt{100}} = \frac{4}{10} = 0,4$$

Можно найти и соответствующий  
доверительный интервал для среднего  
значения:

$$\bar{x} - \Delta < a < \bar{x} + \Delta$$

$$4 - 0,4 < a < 4 + 0,4$$

$$3,6 < a < 4,4$$



# Таблица значений функции Лапласа (фрагмент)

| Целые и<br>десятые<br>доли $x$ | Сотые доли $x$ |        |        |        |        |        |        |        |        |        |
|--------------------------------|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|                                | 0              | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      |
| 1,3                            | 0,8054         | 0,8098 | 0,8132 | 0,8165 | 0,8198 | 0,8230 | 0,8262 | 0,8293 | 0,8324 | 0,8355 |
| 1,4                            | 0,8385         | 0,8415 | 0,8444 | 0,8473 | 0,8501 | 0,8529 | 0,8557 | 0,8584 | 0,8611 | 0,8638 |
| 1,5                            | 0,8664         | 0,8690 | 0,8715 | 0,8740 | 0,8764 | 0,8789 | 0,8812 | 0,8836 | 0,8859 | 0,8882 |
| 1,6                            | 0,8904         | 0,8926 | 0,8948 | 0,8969 | 0,8990 | 0,9011 | 0,9031 | 0,9051 | 0,9070 | 0,9090 |
| 1,7                            | 0,9109         | 0,9127 | 0,9146 | 0,9164 | 0,9181 | 0,9199 | 0,9216 | 0,9233 | 0,9249 | 0,9265 |
| 1,8                            | 0,9281         | 0,9297 | 0,9312 | 0,9327 | 0,9342 | 0,9357 | 0,9371 | 0,9385 | 0,9392 | 0,9412 |
| 1,9                            | 0,9426         | 0,9439 | 0,9451 | 0,9464 | 0,9476 | 0,9488 | 0,9500 | 0,9512 | 0,9523 | 0,9533 |
| 2,0                            | 0,9545         | 0,9556 | 0,9566 | 0,9576 | 0,9586 | 0,9596 | 0,9606 | 0,9916 | 0,9925 | 0,9634 |
| 2,1                            | 0,9643         | 0,9651 | 0,9660 | 0,9668 | 0,9676 | 0,9684 | 0,9692 | 0,9700 | 0,9707 | 0,9715 |
| 2,2                            | 0,9722         | 0,9729 | 0,9736 | 0,9743 | 0,9749 | 0,9756 | 0,9762 | 0,9768 | 0,9774 | 0,9780 |
| 2,3                            | 0,9786         | 0,9791 | 0,9797 | 0,9802 | 0,9807 | 0,9812 | 0,9817 | 0,9822 | 0,9827 | 0,9832 |
| 2,4                            | 0,9836         | 0,9841 | 0,9845 | 0,9849 | 0,9853 | 0,9857 | 0,9861 | 0,9865 | 0,9869 | 0,9872 |
| 2,5                            | 0,9876         | 0,9879 | 0,9883 | 0,9886 | 0,9889 | 0,9892 | 0,9895 | 0,9898 | 0,9901 | 0,9904 |
| 2,6                            | 0,9907         | 0,9910 | 0,9912 | 0,9915 | 0,9917 | 0,9920 | 0,9922 | 0,9924 | 0,9926 | 0,9928 |
| 2,7                            | 0,9931         | 0,9933 | 0,9935 | 0,9937 | 0,9939 | 0,9940 | 0,9942 | 0,9944 | 0,9946 | 0,9947 |
| 2,8                            | 0,9949         | 0,9951 | 0,9952 | 0,9953 | 0,9955 | 0,9956 | 0,9958 | 0,9959 | 0,9960 | 0,9961 |
| 2,9                            | 0,9963         | 0,9964 | 0,9965 | 0,9966 | 0,9967 | 0,9968 | 0,9969 | 0,9970 | 0,9971 | 0,9972 |



**Статистической гипотезой** называется любое предположение о свойствах распределения вероятностей, лежащего в основе наблюдаемых явлений.

### **Типы гипотез:**

- 1) Гипотезы о виде закона распределения исследуемой случайной величины.
- 2) Гипотезы о числовых значениях параметров случайной величины.
- 3) Гипотезы об общем виде модели, описывающей статистическую зависимость между признаками.
- 4) Гипотезы о принадлежности некоторого признака к тому или иному классу.



- Гипотеза **H** называется **простой**, если она полностью определяет теоретическое распределение случайной величины по имеющейся выборке ее значений.

*Например: H:  $a=5$ ; H:  $p=0,71$ ; H:  $\sigma=2,3$ .*

В противном случае она называется **сложной**.

*Например: H:  $a>5$ ; H:  $p\leq 0,71$ ; H:  $\sigma\neq 2,3$ .*

- Гипотеза **H<sub>0</sub>** является **основной** в том смысле, что желательно убедиться в ее справедливости (проверяемая гипотеза). Ей противопоставляются гипотезы  $H_1, H_2, \dots, H_k$ , называемые **альтернативными (конкурирующими)**. Принятие гипотезы **H<sub>0</sub>** или ее альтернативы **H<sub>1</sub>** основано на исследовании статистических данных, т.е. выборки из генеральной совокупности.
- Правило **K**, по которому принимается или отвергается гипотеза, называется **критерием**.
- Критерии согласия** – критерии, устанавливающие закон распределения случайной величины (например, критерий Пирсона и критерий Колмогорова).



- Вероятность отвергнуть основную гипотезу  $H_0$  если она верна (т.е. принять гипотезу  $H_1$ ), называется **вероятностью ошибки первого рода** или **уровнем значимости**:  $P(H_1 | H_0) = \alpha$ .
- Величина  $\gamma = 1 - \alpha = P(H_0 | H_0)$  называется **уровнем доверия**.
- Вероятность принять основную гипотезу  $H_0$  если она неверна, называется **вероятностью ошибки второго рода**:  $P(H_0 | H_1) = \beta$ .
- Величина  $1 - \beta = P(H_1 | H_1)$  называется **мощностью критерия**.

**Например:** в условии задачи задан

- уровень значимости  $\alpha = 0,21$ , тогда можем найти уровень доверия  $\gamma = 1 - \alpha = 1 - 0,21 = 0,79$ .
- Уровень доверия  $\gamma = 0,995$ , тогда можем найти уровень значимости  $\alpha = 1 - 0,995 = 0,005$ .

**Замечание:** иногда значения  $\alpha$  или  $\beta$  задаются в процентах:



## Тренировочные вопросы:

1. Если основная гипотеза имеет вид  $H_0: p=0,7$ , то конкурирующими могут быть гипотезы ...

- а)  $H_1: p>0,7$     б)  $H_1: p<0,7$     в)  $H_1: p\neq 0,7$   
г)  $H_1: p\leq 0,7$     д)  $H_1: p\geq 0,7$     е)  $H_1: p<0,8$

2. Гипотеза  $H: a<7$  является    3. Гипотеза  $H: \sigma=12$  является

- а) простой                                  а) простой  
б) сложной                                б) сложной

3. Ошибкой первого рода называется вероятность ...

- А) отвергнуть верную альтернативную гипотезу  
Б) отвергнуть верную основную гипотезу.  
В) принять ложную основную гипотезу  
Г) принять ложную альтернативную гипотезу .

4. Ошибкой второго рода называется вероятность ...

- А) отвергнуть верную альтернативную гипотезу.  
Б) отвергнуть верную основную гипотезу  
В) принять ложную основную гипотезу.  
Г) принять ложную альтернативную гипотезу



## Процедура построения критерия для проверки статистической гипотезы:

1. По имеющимся данным выборки составляется статистика  $\bar{\theta}_n$  так, чтобы в случае, если гипотеза  $H_0$  верна, распределение случайной величины  $\bar{\theta}_n = \bar{\theta}_n(X_1, \dots, X_n)$  было известным.

2. Задается доверительная вероятность  $\gamma$  критерия для принятия гипотезы  $H_0$ .

3. Выбор критерия, в зависимости от вида гипотезы  $H_0$ , заключается в выборе таких значений  $\theta_1$  и  $\theta_2$ , что если  $\bar{\theta}_n < \theta_1$  или  $\bar{\theta}_n \geq \theta_2$ , то  $H_0$  принимается. При этом возможно, что  $\theta_1 = -\infty$  или  $\theta_2 = +\infty$ .

Значения  $\theta_1$  и  $\theta_2$  называются **критическими**.  
 $T_D = \{\theta : \theta_1 < \theta < \theta_2\}$

Область  $\theta_n \in T_D$  называется **областью допустимых значений**.  
 $\theta_n \notin T_D$

4. Если  $\bar{\theta}_n < \theta_1$ , то  $H_0$  принимается. Если  $\bar{\theta}_n \geq \theta_2$ , то  $H_0$  отвергается и принимается альтернативная гипотеза  $H_1$ .  
 $T_{кр} = \{\theta : \theta \leq \theta_1, \theta \geq \theta_2\}$

При этом область  $T_{кр}$  называется **критической областью**.

В зависимости от альтернативной гипотезы  $H_1$  критическая область может быть:

1) **правосторонней**:  $P(\bar{\theta}_n \leq \theta_1) = P(\bar{\theta}_n \geq \theta_2) = \alpha / 2$

2) **левосторонней**:



# характеристик

Пусть случайная величина  $X \in N(a, \sigma)$  и имеется выборка ее значений  $\{x_1, x_2, \dots, x_n\}$ .

## 1. О численной величине среднего значения.

Рассмотрим гипотезу  $H_0: a = a_0$ , где  $a_0$  - некоторое число, и ее альтернативу  $H_1$ .

Проверим гипотезу  $H_0$  на уровне доверия  $\gamma$ . Возможны два случая:

**А) Дисперсия  $\sigma^2$  известна.**

В этом случае случайная величина (статистика)  $\frac{(\bar{x} - a)\sqrt{n}}{\sigma}$  имеет нормальное стандартное распределение  $N(0,1)$ , если верна гипотеза  $H_0: a = a_0$ .

Для двусторонней альтернативной гипотезы  $H_1: a \neq a_0$  при заданном уровне доверия  $\gamma$  находим по таблицам функции Лапласа такое  $t_{кр}$ , чтобы  $|\frac{(\bar{x} - a)\sqrt{n}}{\sigma}| < t_{кр}$ .

Если статистика  $|\frac{(\bar{x} - a_0)\sqrt{n}}{\sigma}| < t_{кр}$ , то гипотеза  $H_0$  принимается, иначе принимается  $H_1$ .

Для односторонних альтернативных гипотез ( $H_1: a > a_0$  или  $H_1: a < a_0$ ) для определения критического значения вместо  $\gamma = 1 - \alpha$  используется величина  $1 - 2\alpha$ .

**Б) Дисперсия  $\sigma^2$  неизвестна.**

В этом случае случайная величина (статистика)  $\frac{(\bar{x} - a)\sqrt{n}}{S}$  где  $\bar{S}^2 = \frac{1}{n-1} \sum_{i=1}^m (x_i - \bar{x})^2$

исправленная выборочная дисперсия,  $t$  имеет распределение Стьюдента с  $(n-1)$  степенью свободы, если гипотеза  $H_0$  верна.

Для двусторонней альтернативной гипотезы  $H_1: a \neq a_0$  при заданном уровне доверия  $\gamma$  по таблицам распределения Стьюдента определяется критическое значение при условии  $|\frac{(\bar{x} - a_0)\sqrt{n}}{S}| < t_{кр}$ .



# Пример 1.

По результатам 9 замеров установлено, что выборочное среднее время изготовления детали равно 48 секунд. Предполагая, что время изготовления – нормально распределенная случайная величина с дисперсией 9, рассмотреть на уровне 0,95 гипотезу  $H_0: a=49$  против гипотезы  $H_1: a \neq 49$ .

*Решение.* По условию задачи  $n=9$ ,  $\bar{x}=48$ ,  $\sigma^2=9$ ,  $\gamma=0,95$ . Вычислим статистику

$$t = \frac{(\bar{x} - a)\sqrt{n}}{\sigma} = \frac{(48 - 49)\sqrt{9}}{\sqrt{9}} = -1$$

По таблице распределения Лапласа из условия  $\Phi(t)=0,95$  найдем  $t_{кр}$  1,96. Так как  $|t| = 1 < t_{кр} = 1,96$

то гипотеза  $H_0$  принимается.

| Целые и<br>десятые<br>доли x | Сотые доли x |        |        |        |        |        |        |        |        |        |
|------------------------------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|                              | 0            | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      |
| 1,3                          | 0,8054       | 0,8098 | 0,8132 | 0,8165 | 0,8198 | 0,8230 | 0,8262 | 0,8293 | 0,8324 | 0,8355 |
| 1,4                          | 0,8385       | 0,8415 | 0,8444 | 0,8473 | 0,8501 | 0,8529 | 0,8557 | 0,8584 | 0,8611 | 0,8638 |
| 1,5                          | 0,8664       | 0,8690 | 0,8715 | 0,8740 | 0,8764 | 0,8789 | 0,8812 | 0,8836 | 0,8859 | 0,8882 |
| 1,6                          | 0,8904       | 0,8926 | 0,8948 | 0,8969 | 0,8990 | 0,9011 | 0,9031 | 0,9051 | 0,9070 | 0,9090 |
| 1,7                          | 0,9109       | 0,9127 | 0,9146 | 0,9164 | 0,9181 | 0,9199 | 0,9216 | 0,9233 | 0,9249 | 0,9265 |
| 1,8                          | 0,9281       | 0,9297 | 0,9312 | 0,9327 | 0,9342 | 0,9357 | 0,9371 | 0,9385 | 0,9392 | 0,9412 |
| 1,9                          | 0,9426       | 0,9439 | 0,9451 | 0,9464 | 0,9476 | 0,9488 | 0,9500 | 0,9512 | 0,9523 | 0,9533 |
| 2,0                          | 0,9545       | 0,9556 | 0,9566 | 0,9576 | 0,9586 | 0,9596 | 0,9606 | 0,9616 | 0,9625 | 0,9634 |



# 1. Гипотезы о значениях числовых характеристик

## 2. О числовом значении дисперсии.

Рассмотрим гипотезу  $H_0: \sigma^2 = \sigma_0^2$ , где  $\sigma_0$  - некоторое число, и ее альтернативу  $H_1$ . Проверим гипотезу  $H_0$  на уровне доверия  $\gamma$ .

Рассмотрим статистику  $\chi^2 = \frac{(n-1)\bar{S}^2}{\sigma_0^2}$ , которая имеет распределение  $\chi_{n-1}^2$ , если гипотеза  $H_0$  верна.

Для заданного уровня доверия  $\gamma$  по таблицам Хи-квадрат распределения выбираем такие  $U$  и  $V$ , чтобы  $P(U < t < V) = \gamma$ . Если статистика  $\chi^2 < U$  или  $\chi^2 > V$ , то гипотеза  $H_0$  принимается.

- 1) Для *двусторонней* альтернативной гипотезы  $U$  определяется при  $1-\alpha/2$ , а  $V$  при  $\alpha/2$ .
- 2) Для *правосторонней* альтернативной гипотезы определяется только величина  $V$  при  $\alpha$ ;
- 3) Для *левосторонней* альтернативной гипотезы определяется только величина  $U$  при  $1-\alpha$ .



## Пример 2.

Точность работы станка-автомата проверяется по дисперсии контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена оценка дисперсии  $\bar{s}^2 = 0,25$ . При уровне значимости 0,1 выяснить, обеспечивает ли станок требуемую точность.

**Решение.** По условию задачи  $n=25$ ,  $\bar{s}^2 = 0,25$ ,  $\sigma_0^2 = 0,15$ ,  $\alpha=0,1$ . Проверим гипотезу

**$H_0: \sigma^2=0,15$**  против гипотезы  **$H_1: \sigma^2>0,15$** . В этом случае гипотеза  **$H_0$**  принимается, если  $t < V$ . Найдем  $V$  из распределения Хи-квадрат для  $\alpha=0,1$  и степени свободы  $n-1=24$ , получим  $V=33,2$ . Вычислим статистику  $t$  по формуле:

$$t = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(25-1) \cdot 0,25}{0,15} = 40$$

Так как  $t > V$ , то гипотеза  **$H_0$**  отвергается. Следовательно станок **не обеспечивает** требуемую точность.

| Число<br>степеней<br>свободы | Вероятность $\alpha$ |      |      |      |      |      |      |      |      |      |      |      |      |
|------------------------------|----------------------|------|------|------|------|------|------|------|------|------|------|------|------|
|                              | 0,99                 | 0,98 | 0,95 | 0,90 | 0,80 | 0,70 | 0,50 | 0,30 | 0,20 | 0,10 | 0,05 | 0,02 | 0,01 |
| 22                           | 9,54                 | 10,6 | 12,3 | 14,0 | 16,3 | 18,1 | 21,3 | 24,9 | 27,3 | 30,8 | 33,9 | 37,7 | 40,3 |
| 23                           | 10,2                 | 11,3 | 13,1 | 14,8 | 17,2 | 19,0 | 22,3 | 26,0 | 28,4 | 32,0 | 35,2 | 39,0 | 41,6 |
| 24                           | 10,9                 | 12,0 | 13,8 | 15,7 | 18,1 | 19,9 | 23,3 | 27,1 | 29,6 | 33,2 | 36,4 | 40,3 | 43,0 |
| 25                           | 11,5                 | 12,7 | 14,6 | 16,5 | 18,9 | 20,9 | 24,3 | 28,2 | 30,7 | 34,4 | 37,7 | 41,7 | 44,3 |

# I. Гипотезы о значениях числовых характеристик

## 3. О числовом значении доли признака.

Рассмотрим гипотезу  $H_0: p=p_0$ , где  $p_0$  - некоторое число, и ее альтернативу  $H_1$ . Проверим гипотезу  $H_0$  на уровне доверия  $\gamma$ .

Рассмотрим статистику  $u = \frac{(w - p_0)\sqrt{n}}{\sqrt{p_0(1 - p_0)}}$ , которая при больших  $n$  ( $n > 30$ ) имеет распределение  $N(0,1)$ , если  $H_0$  верна.

Для двусторонней альтернативной гипотезы  $H_1: p \neq p_0$  при заданном уровне доверия  $\gamma$  по таблице значений функции Лапласа  $\Phi(t)$  находим критическое значение так, чтобы  $\Phi(t_{кр}) = \gamma$ .

Если статистика  $|t| < t_{кр}$ , то гипотеза  $H_0$  принимается, иначе принимается  $H_1$ .

Для односторонних альтернативных гипотез ( $H_1: p > p_0$  или  $H_1: p < p_0$ ) для определения критического значения вместо  $\gamma = 1 - \alpha$  используется величина  $1 - 2\alpha$ .



### Пример 3.

Партия изделий принимается в том случае, если вероятность того, что изделие окажется соответствующим стандарту, будет не менее 0,95. Среди случайно отобранных 100 изделий оказалось 98 соответствующих стандарту. Можно ли на уровне значимости 0,1 принять партию?

**Решение.** По условию задачи  $m=98$ ,  $n=100$ ,  $\alpha=0,1$ ,  $p_0=0,95$ . Необходимо проверить гипотезу  **$H_0: p=0,95$**  против гипотезы  **$H_1: p<0,95$** . Вычислим статистику  $t$ :

$$t = \frac{(w - p_0)\sqrt{n}}{\sqrt{p_0(1 - p_0)}} = \frac{\left(\frac{m}{n} - p_0\right)\sqrt{n}}{\sqrt{p_0(1 - p_0)}} = \frac{(0,98 - 0,95)\sqrt{100}}{\sqrt{0,95 \cdot 0,05}} = \frac{0,03 \cdot 10}{\sqrt{0,0475}} = \frac{0,3}{0,218} = 1,377$$

Она имеет распределение  $N(0,1)$  при  $n>30$ , если  **$H_0$**  верна. Критическое значение найдем по таблицам функции Лапласа для вероятности  $1-2\alpha=0,8$   $t_{кр}$

Так как  $|t| > t_{кр}$ , то гипотеза  **$H_0$  отвергается**. Следовательно партия **не принимается**

| Целые и<br>десятые<br>доли x | Сотые доли x |        |        |        |        |        |        |        |        |        |
|------------------------------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|                              | 0            | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      |
| 1,1                          | 0,7287       | 0,7330 | 0,7373 | 0,7415 | 0,7457 | 0,7499 | 0,7540 | 0,7580 | 0,7620 | 0,7660 |
| 1,2                          | 0,7699       | 0,7737 | 0,7775 | 0,7813 | 0,7850 | 0,7887 | 0,7923 | 0,7959 | 0,7994 | 0,8029 |
| 1,3                          | 0,8054       | 0,8098 | 0,8132 | 0,8165 | 0,8198 | 0,8230 | 0,8262 | 0,8293 | 0,8324 | 0,8355 |
| 1,4                          | 0,8385       | 0,8415 | 0,8444 | 0,8473 | 0,8501 | 0,8529 | 0,8557 | 0,8584 | 0,8611 | 0,8638 |
| 1,5                          | 0,8664       | 0,8690 | 0,8715 | 0,8740 | 0,8764 | 0,8789 | 0,8812 | 0,8836 | 0,8859 | 0,8882 |

## II. Гипотезы о равенствах значений числовых характеристик

Пусть случайные величины  $X_1 \in N(a_1, \sigma_1)$ ,  $X_2 \in N(a_2, \sigma_2)$ .

Имеются две выборки объемов  $n_1$  и  $n_2$  этих величин  $\{x_1, x_2, \dots, x_{n_1}\}$   $\{y_1, y_2, \dots, y_{n_2}\}$ .

1. О равенстве средних значений:  $H_0: a_1 = a_2$ . Проверяется на уровне доверия  $\gamma$ .

А) Дисперсии  $\sigma_1^2$  и  $\sigma_2^2$  известны. Если гипотеза  $H_0$  выполняется, то статистика  $t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$  имеет распределение  $N(0,1)$ .

$$t_{кр} \quad \Phi(t_{кр}) = \gamma$$

Для двусторонней альтернативной гипотезы  $H_1: a_1 \neq a_2$  при заданном уровне доверия  $\gamma$  находим по таблицам функции Лапласа такое  $t_{кр}$ , чтобы  $|t| < t_{кр}$ .

Если статистика  $t < t_{кр}$ , то гипотеза  $H_0$  принимается, иначе принимается  $H_1$ .

Для односторонних альтернативных гипотез ( $H_1: a_1 > a_2$  или  $H_1: a_1 < a_2$ ) для определения критического значения вместо  $\gamma = 1 - \alpha$  используется величина  $1 - 2\alpha$ .

Б) Дисперсии  $\sigma_1^2$  и  $\sigma_2^2$  неизвестны, но равны. Если гипотеза  $H_0$  выполняется, то статистика  $t = \frac{\bar{x} - \bar{y}}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$  имеет распределение Стьюдента.



## Пример 4

Было произведено 12 измерений диаметра вала. При этом оказалось, что выборочное среднее равно 10,2 мм, а стандартное отклонение – 0,05 мм. Затем вал поместили в условия с высокой температурой и произвели еще 8 замеров. На этот раз среднее значение оказалось равным 10,25 мм, а стандартное отклонение – 0,06 мм. Можно ли сделать вывод при 5%-ном уровне значимости, что диаметр вала существенно увеличивается при нагревании?

**Решение.** По условию задачи  $n_1=12, \bar{x}=10,2$ ,  $\sigma_1=0,05$ ,  $n_2=8, \bar{y}=10,25$ ,  $\sigma_2=0,06$ ,  $\alpha=0,05$ . Проверяется гипотеза  $H_0: a_1=a_2$  против гипотезы  $H_1: a_1 \neq a_2$ . Вычислим статистику

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{10,2 - 10,25}{\sqrt{\frac{0,05^2}{12} + \frac{0,06^2}{8}}} = \frac{-0,05}{0,02566} = -1,95$$

$$t_{кр} = 1,96$$

По таблице распределения Лапласа из условия  $\Phi(t)=1-0,05=0,95$  находим  $t=1,95 < t_{кр}=1,96$

. Так как  $t < t_{кр}$ , то  $H_0$  принимается. Диаметр вала существенно не увеличивается.

| Целые и<br>десятые<br>доли x | Сотые доли x |        |        |        |        |        |        |        |        |        |
|------------------------------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|                              | 0            | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      |
| 1,9                          | 0,9426       | 0,9439 | 0,9451 | 0,9464 | 0,9476 | 0,9488 | 0,9500 | 0,9512 | 0,9523 | 0,9533 |
| 2,0                          | 0,9545       | 0,9556 | 0,9566 | 0,9576 | 0,9586 | 0,9596 | 0,9606 | 0,9616 | 0,9625 | 0,9634 |



## II. Гипотезы о равенствах значений числовых характеристик

Имеются выборочные доли  $w_1 = \frac{m_1}{n_1}$   $w_2 = \frac{m_2}{n_2}$ , где  $m_1$  и  $m_2$  соответственно число элементов первой и второй выборки, обладающих данным признаком.

**2. О равенстве долей признака:  $H_0: p_1 = p_2$ .** Проверяется на уровне доверия  $\gamma$ .

При  $n_1 > 30$ ,  $n_2 > 30$ , если гипотеза  $H_0$ , то статистика

$$t = \frac{w_1 - w_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

имеет распределение  $N(0,1)$ . При вычислении статистики вместо величины  $p$  используют общую выборочную долю признака:  $\bar{p} = \frac{m_1 + m_2}{n_1 + n_2}$ .

Для *двусторонней* альтернативной гипотезы  $H_1: p_1 \neq p_2$  при заданном уровне доверия  $\gamma$  по таблице значений функции Лапласа  $\Phi(t)$  находим критическое значение так, чтобы  $\Phi(t_{кр}) = \gamma$ .

Если статистика  $|t| < t_{кр}$ , то гипотеза  $H_0$  принимается, иначе принимается  $H_1$ .

Для *односторонних* альтернативных гипотез ( $H_1: p_1 > p_2$  или  $H_1: p_1 < p_2$ ) для определения критического значения вместо  $\gamma = 1 - \alpha$  используется величина  $1 - 2\alpha$ .



## Пример 5

Контрольную работу выполняли студенты двух групп. В первой группе было предложено 100 задач, из которых были правильно решены 58. Во второй группе из 120 задач верно решены 65. На уровне значимости 0,02 проверить гипотезу о том, что материал одинаково усвоен студентами обеих групп.

**Решение.** По условию задачи  $n_1=100$ ,  $m_1=58$ ,  $n_2=120$ ,  $m_2=65$ ,  $\alpha=0,02$ . Проверяется гипотеза  $H_0: p_1=p_2$  против гипотезы  $H_1: p_1 \neq p_2$ . Вычислим статистику

$$t = \frac{w_1 - w_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}, \text{ где } w_1 = \frac{m_1}{n_1}, \quad w_2 = \frac{m_2}{n_2}, \text{ а вместо } p \text{ используем } \bar{p} = \frac{m_1 + m_2}{n_1 + n_2}$$

Тогда:

$$w_1 = \frac{m_1}{n_1} = \frac{58}{100} = 0,58, \quad w_2 = \frac{m_2}{n_2} = \frac{65}{120} = 0,542, \quad \bar{p} = \frac{m_1 + m_2}{n_1 + n_2} = \frac{58 + 65}{100 + 120} = \frac{123}{220} = 0,559$$

$$t = \frac{0,58 - 0,542}{\sqrt{0,559(1 - 0,559)\left(\frac{1}{100} + \frac{1}{120}\right)}} = \frac{0,038}{\sqrt{0,246519 \frac{220}{12000}}} = \frac{0,038}{0,06723} = 0,565$$

Для  $\gamma=1-0,02=0,98$  по таблице функции Лапласа найдем  $t_{кр} = 2,33$ . Так как  $|t| < t_{кр}$ , то

Гипотеза  $H_0$  принимается. Материал усвоен одинаково.

| Целые и<br>десятые<br>доли x | Сотые доли x |        |        |        |        |        |        |        |        |        |
|------------------------------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|                              | 0            | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      |
| 2,3                          | 0,9786       | 0,9791 | 0,9797 | 0,9802 | 0,9807 | 0,9812 | 0,9817 | 0,9822 | 0,9827 | 0,9832 |

## II. Гипотезы о равенствах значений числовых характеристик

3. О равенстве дисперсий:  $H_0: \sigma_1^2 = \sigma_2^2$  проверяется на уровне доверия  $\gamma$ .

Статистика  $t = \frac{\bar{S}_1^2}{\bar{S}_2^2}$  имеет распределение Фишера  $F_{n_1-1, n_2-1}$ .

По таблицам распределения Фишера находим такие  $U$  и  $V$ , чтобы при заданном уровне доверия  $\gamma = 1 - \alpha$ :  $P(F_{n_1-1, n_2-1} \leq U) = P(F_{n_1-1, n_2-1} \geq V) = \alpha / 2$ . Если  $U < t < V$ , то  $H_0$  принимается.

)Для двусторонней альтернативной гипотезы  $U$  определяется при  $1 - \alpha/2$ , а  $V$  при  $\alpha/2$ .

)Для правосторонней альтернативной гипотезы определяется величина  $V$  при  $\alpha$ .

)Для левосторонней альтернативной гипотезы определяется величина  $U$  при  $1 - \alpha$ .

### Пример 6.

Расходы сырья  $X$  и  $Y$  на единицу продукции по старой и новой технологиям приведены в таблице. Предполагается, что генеральные совокупности  $X$  и  $Y$  имеют нормальное

|                                                  |                      |     |     |     |                     |     |     |     |     |
|--------------------------------------------------|----------------------|-----|-----|-----|---------------------|-----|-----|-----|-----|
| приведены в таблице. Предполагается, что генерал |                      |     |     |     |                     |     |     |     |     |
| нормальное                                       | По старой технологии |     |     |     | По новой технологии |     |     |     |     |
| Расход сырья                                     | $x_i$                | 304 | 307 | 308 | $y_j$               | 303 | 304 | 306 | 308 |
| Число изделий                                    | $n_i$                | 1   | 4   | 4   | $n_j$               | 2   | 6   | 4   | 1   |

распределение с одинаковыми дисперсиями и средними  $a_1$  и

Проверить гипотезы о

равенстве

средних и о равенстве



Московский  
Технологический  
Институт

## Пример 6 (решение)

По условию задачи  $n_1=1+4+4=9$ ,  $n_2=2+6+4+1=13$ ,  $\alpha=0,1$ . Вычислим выборочные

средние:

$$\bar{x} = \frac{304 \cdot 1 + 307 \cdot 4 + 308 \cdot 4}{9} = \frac{2764}{9} = 307,11$$

$$\bar{y} = \frac{303 \cdot 2 + 304 \cdot 6 + 306 \cdot 4 + 308 \cdot 1}{13} = \frac{3962}{13} = 304,77$$

Вычислим выборочные дисперсии, для этого сначала вычислим

$$\overline{x^2} = \frac{304^2 \cdot 1 + 307^2 \cdot 4 + 308^2 \cdot 4}{9} = \frac{848868}{9} = 94318,667$$

$$\overline{y^2} = \frac{303^2 \cdot 2 + 304^2 \cdot 6 + 306^2 \cdot 4 + 308^2 \cdot 1}{13} = \frac{1207522}{13} = 92886,307$$

тогда получим

$$S_1^2 = \overline{x^2} - (\bar{x})^2 = 94318,667 - (307,11)^2 = 2,115$$

$$S_2^2 = \overline{y^2} - (\bar{y})^2 = 92886,307 - (304,77)^2 = 1,555$$

1) Проверим гипотезу  $H_0: a_1 = a_2$  против гипотезы  $H_1: a_1 \neq a_2$  для  $\gamma = 1 - \alpha = 1 - 0,1 = 0,9$ .

Вычислим статистику  $t$  по формуле:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$



## Пример 6 (продолжение)

$$t = \frac{307,11 - 304,77}{\sqrt{\frac{9 \cdot 2,115 + 13 \cdot 1,555}{9 + 13 - 2} \left( \frac{1}{9} + \frac{1}{13} \right)}} = \frac{2,34}{\sqrt{\frac{39,25}{20} \left( \frac{22}{117} \right)}} = \frac{2,34}{0,6075} = 3,852.$$

По таблице распределения Стьюдента находим критическое значение для  $\gamma=0,9$  и степени свободы  $n_1+n_2-2=9+13-2=20$ :  $t_{кр} = 1,72$ . Так как  $|t| > t_{кр}$ , то гипотеза  **$H_0: a_1=a_2$**  **не принимается**, т.е. при переходе на новую технологию **происходит изменение** величины среднего расхода сырья.

| Число<br>степеней<br>свободы | Вероятность $\gamma$ |      |      |      |      |      |      |      |      |      |      |      |
|------------------------------|----------------------|------|------|------|------|------|------|------|------|------|------|------|
|                              | 0,1                  | 0,2  | 0,3  | 0,4  | 0,5  | 0,6  | 0,7  | 0,8  | 0,90 | 0,95 | 0,98 | 0,99 |
| 20                           | 0,13                 | 0,26 | 0,39 | 0,53 | 0,69 | 0,86 | 1,06 | 1,32 | 1,72 | 2,09 | 2,53 | 2,84 |

2) Проверим гипотезу  **$H_0: \sigma_1^2=\sigma_2^2$**  против гипотезы  **$H_1: \sigma_1^2 \neq \sigma_2^2$** . Вычислим исправленные выборочные дисперсии

$$\overline{S_1^2} = \frac{n_1 S_1^2}{n_1 - 1} = \frac{9 \cdot 2,115}{8} = 2,379, \quad \overline{S_2^2} = \frac{n_2 S_2^2}{n_2 - 1} = \frac{13 \cdot 1,555}{12} = 1,685$$

Найдем статистику

$$t = \frac{\overline{S_1^2}}{\overline{S_2^2}} = \frac{2,379}{1,685} = 1,41$$



## Пример 6 (окончание)

По таблице распределения Фишера найдем  $U$  и  $V$  для степеней свободы  $n_1-1=8$  и  $n_2=13-1=12$ . Величина  $U$  определяется при  $1-\alpha/2=1-0,1/2=1-0,05=0,95$ , а величина  $V$  при  $\alpha/2=0,05$ . Тогда  $U=1/3,28=0,305$  и  $V=2,85$ . Так как  $t=1,41$  удовлетворяет интервалу  $0,305 < t < 2,85$ , то гипотеза  $H_0: \sigma_1^2 = \sigma_2^2$  принимается.

| $k_2 \backslash k_1$ | $\alpha = 0,05$ |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
|----------------------|-----------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
|                      | 1               | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 12   | 15   | 20   | 24   | 30   |
| 8                    | 5,32            | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,50 | 3,44 | 3,39 | 3,35 | 3,28 | 3,22 | 3,15 | 3,12 | 3,08 |
| 9                    | 5,12            | 4,26 | 3,86 | 3,63 | 3,48 | 3,37 | 3,29 | 3,23 | 3,18 | 3,14 | 3,07 | 3,01 | 2,94 | 2,90 | 2,86 |
| 10                   | 4,96            | 4,10 | 3,71 | 3,48 | 3,33 | 3,22 | 3,14 | 3,07 | 3,02 | 2,98 | 2,91 | 2,85 | 2,77 | 2,74 | 2,70 |
| 11                   | 4,84            | 3,98 | 3,59 | 3,36 | 3,20 | 3,09 | 3,01 | 2,95 | 2,90 | 2,85 | 2,79 | 2,72 | 2,65 | 2,61 | 2,57 |
| 12                   | 4,75            | 3,89 | 3,49 | 3,26 | 3,11 | 3,00 | 2,91 | 2,85 | 2,80 | 2,75 | 2,69 | 2,62 | 2,54 | 2,51 | 2,47 |
| 13                   | 4,67            | 3,81 | 3,41 | 3,18 | 3,03 | 2,92 | 2,83 | 2,77 | 2,71 | 2,67 | 2,60 | 2,53 | 2,46 | 2,42 | 2,38 |

Замечание. Так как в таблице Фишера есть значения только для  $\alpha=0,05$  со степенями свободы  $(k_1, k_2)$ , то чтобы найти значение для  $\alpha=0,95$  нужно по данной таблице найти значение  $F$  со степенями  $(k_2, k_1)$  и затем вычислить  $1/F$ . Этот прием мы использовали при определении величины  $U$ .



# Тестовые вопросы:

1. Любое предположение о свойствах распределения вероятностей, лежащего в основе наблюдаемых явлений, называется

- а) статистической гипотезой
- б) статистическим критерием
- в) статистическим показателем
- г) статистической величиной

2. Гипотеза называется **простой**, если ...

- а) ее желательно проверить
- б) если ей нельзя противопоставить другую гипотезу
- в) если она полностью определяет теоретическое распределение случайной величины по имеющейся выборке ее значений
- г) она частично определяет теоретическое распределение случайной величины по имеющейся выборке ее значений

3. Гипотеза называется **сложной**, если ...

- а) ее желательно проверить
- б) если ей нельзя противопоставить другую гипотезу
- в) если она полностью определяет теоретическое распределение случайной величины по имеющейся выборке ее значений
- г) она не полностью определяет теоретическое распределение случайной величины по имеющейся выборке ее значений



## Тестовые вопросы:

4. Правило, по которому принимается или отвергается гипотеза, называется

- а) статистикой
- б) критерием
- в) ограничением
- г) законом

5. Критерии, устанавливающие закон распределения случайной величины, называются

- а) критерии определенности
- б) критерии согласия
- в) критерии непротиворечия
- г) критерии оптимальности

6. Если основная гипотеза имеет вид  $H_0: p=0,5$ , то конкурирующей может быть гипотеза

- а)  $H_1: p \geq 0,5$
- б)  $H_1: p \neq 0,5$
- в)  $H_1: p \leq 0,5$
- г)  $H_1: p \neq 0,6$



## Тестовые вопросы:

7. Если основная гипотеза имеет вид  $H_0: \sigma=8$ , то конкурирующей может быть гипотеза

а)  $H_1: \sigma \geq 8$

б)  $H_1: \sigma \neq 9$

в)  $H_1: \sigma \leq 8$

г)  $H_1: \sigma > 8$

8. Если основная гипотеза имеет вид  $H_0: \sigma^2=5$ , то конкурирующей может быть гипотеза

а)  $H_1: \sigma^2 < 9$

б)  $H_1: \sigma^2 \neq 9$

в)  $H_1: \sigma^2 \leq 5$

г)  $H_1: \sigma^2 \geq 5$

9. Ошибкой первого рода называется вероятность

а)  $P(H_0 | H_0)$

б)  $P(H_1 | H_0)$

в)  $P(H_0 | H_1)$

г)  $P(H_1 | H_1)$



## Тестовые вопросы:

10. Ошибкой второго рода называется вероятность

а)  $P(H_0 | H_0)$

б)  $P(H_1 | H_0)$

в)  $P(H_0 | H_1)$

г)  $P(H_1 | H_1)$

11. Мощностью критерия называется вероятность

а)  $P(H_0 | H_0)$

б)  $P(H_1 | H_0)$

в)  $P(H_0 | H_1)$

г)  $P(H_1 | H_1)$

12. Уровнем доверия называется вероятность

а)  $P(H_0 | H_0)$

б)  $P(H_1 | H_0)$

в)  $P(H_0 | H_1)$

г)  $P(H_1 | H_1)$



## Тестовые вопросы:

13. Статистика  $t = \frac{\bar{S}_1^2}{\bar{S}_2^2}$  имеет распределение Фишера  $F_{n_1-1, n_2-1}$ , если верна гипотеза

а)  $H_1: p_1 = p_2$

б)  $H_1: a_1 = a_2$

в)  $H_1: a_1 < a_2$

г)  $H_1: \sigma_1^2 = \sigma_2^2$

14. Статистика  $t = \frac{w_1 - w_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$  имеет распределение  $N(0,1)$ , если верна гипотеза

а)  $H_1: p_1 = p_2$

б)  $H_1: a_1 = a_2$

в)  $H_1: a_1 < a_2$

г)  $H_1: \sigma_1^2 = \sigma_2^2$

15. Статистика  $t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$  имеет распределение  $N(0,1)$ , если верна гипотеза

а)  $H_1: p_1 = p_2$

б)  $H_1: a_1 = a_2$

в)  $H_1: a_1 < a_2$

г)  $H_1: \sigma_1^2 = \sigma_2^2$



## Задачи для самостоятельного решения

1. Учет времени сборки узла машины бригадой из 10 слесарей показал, что среднее время сборки узла равно 76 с выборочной дисперсией 15.

Предполагая распределение времени сборки нормальным, проверить на уровне значимости 0,01 гипотезу о том, что 75 мин является нормативом трудоемкости.

2. В прошлом году доля бракованных изделий, выпускаемым предприятием, равнялась 0,04. В этом году было проверено 300 изделий, из которых 95 оказались бракованными. Можно ли на уровне значимости 0,01 считать, что качество продукции осталось на прежнем уровне?

3. Номинальная точность прибора равна  $\sigma_0 = \text{мкм}$ . Из 10 замеров детали была получена выборочная дисперсия показаний прибора, равная 0,9. На уровне доверия 0,9 проверить гипотезу  $H_0: \sigma^2 = \sigma_0^2$ .

4. Для проверки эффективности нового лекарства были отобраны две случайные группы по 15 человек, страдающих гриппом. При применении старого лекарства средний срок выздоровления составлял 11 дней с выборочной дисперсией 3, при применении нового – срок выздоровления составил 8 дней с выборочной дисперсией 4. Проверить на уровне доверия 0,99 гипотезу о преимуществе нового лекарства.

## Задачи для самостоятельного решения

5. При 4000 бросаниях монеты получили 2048 выпадений герба. На каком уровне значимости можно принять гипотезу о том, что вероятность выпадения герба равна 0,5?
6. Точность работы двух станков оценивалась отклонениями от номинала производимой продукции. Из 10 единиц продукции первого станка отклонение от номинала (выборочная дисперсия) составила 3,5, а из 15 единиц продукции второго станка – 4,5. Можно ли считать на уровне значимости 0,1, что станки имеют одинаковую точность?
7. Стрелок по летающим тарелкам попадал в цель в 80% случаев. После тренировок он стал поражать мишень в 89 случаях из 100. Можно ли считать на уровне доверия 0,95, что его квалификация улучшилась?

