

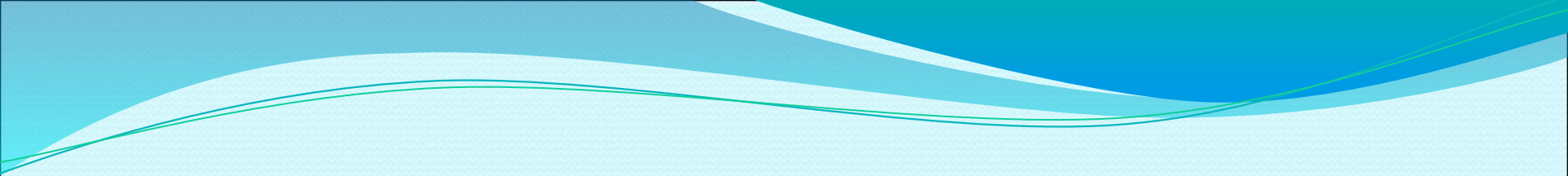
# **Моделирование объекта принятия решений на основе проверки многомерных статистических гипотез**

*Лекция 6*

## Учебные вопросы:

- 1 Математический аппарат моделирования на основе проверки многомерных статистических гипотез
- 2 Методика принятие решений на основе проверки статистических гипотез





# **Методические основы проверки многомерных статистических гипотез**

# **1 Методика принятие решений на основе проверки статистических гипотез**



# Понятия статистических гипотез

**Статистические гипотезы** – это выдвигаемые теоретические **предположения** относительно **параметров** статистического распределения или закона распределения случайных величин.

Гипотеза  $H_0$  (основная) – содержит утверждение об **отсутствии различий** между сравниваемыми величинами.

Гипотеза  $H_1$  – принимается если отвергнута основная гипотеза.

| Случайная величина | Нулевые гипотезы              |                              | Альтернативные гипотезы                        |
|--------------------|-------------------------------|------------------------------|------------------------------------------------|
| Одномерная         | $H_0 : \mu = \mu_0$           | $H_1 : \mu \neq \mu_0$       | $\mu < \mu_0, \mu > \mu_0$                     |
|                    | $H_0 : \sigma^2 = \sigma_0^2$ | $H_1 : \sigma \neq \sigma_0$ | $\sigma^2 < \sigma_0^2, \sigma^2 > \sigma_0^2$ |
| Многомерная        | $H_0 : \mu_j = \mu_{0j}$      |                              | $H_1 : \mu_j \neq \mu_{0j}$                    |
|                    | $H_0 : \Sigma = \Sigma_0$     |                              | $H_1 : \Sigma \neq \Sigma_0$                   |

Для проверки гипотез используют статистические критерии, позволяющие выяснить, следует принять или отвергнуть нулевую гипотезу

Проверка статистических гипотез всегда допускает определенную **вероятность ошибки** в выводах:

**$\alpha$**  - вероятность **отвергнуть** нулевую гипотезу, когда она **справедлива**;

**$\beta$**  - вероятность принять нулевую гипотезу, когда она ложна.

# Методика проверки гипотезы о равенстве вектора средних значений заданному вектору

Имеет место  $m$  - чисел выборочных средних значений анализируемых величин и вектор требуемых значений этих величин:

Для проверки многомерной гипотезы  $H_0 : \bar{Y} = M$  используется критерий Хотеллинга

$$T_p^2 = n(\bar{Y} - M)^T \Sigma^{-1} (\bar{Y} - M),$$

$$\Sigma = \frac{1}{n-1} (\hat{Y}^T \hat{Y}) \quad - \text{ковариационная матрица};$$

$$\hat{Y} \quad - \text{матрица с центрированными значениями переменной:} \quad \hat{y}_{ij} = y_{ij} - \bar{y}_{ij}$$

$$T_{кр(\alpha, m, n-m)}^2 = \frac{m(n-1)}{n-m} F_{\alpha, m, n-m}$$

$F_{\alpha, m, n-m}$  - табличное значение F-критерия Фишера для числа степеней свободы

Многомерная гипотеза о равенстве вектора средних величин заданному вектору **подтверждается** при

$$T_p^2 < T_{кр(\alpha, m, n-m)}^2$$



# Методика проверки гипотезы о равенстве двух векторов средних значений

$$H_0 : (\bar{X}_{11} \bar{X}_{12} \bar{X}_{13} \dots \bar{X}_{1m}) = (\bar{X}_{21} \bar{X}_{22} \bar{X}_{23} \dots \bar{X}_{2m}),$$

$$H_1 : (\bar{X}_{11} \bar{X}_{12} \bar{X}_{13} \dots \bar{X}_{1m}) \neq (\bar{X}_{21} \bar{X}_{22} \bar{X}_{23} \dots \bar{X}_{2m}),$$

Для проверки гипотезы применяется многомерный критерий Хотеллинга

$$T_p^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)^T \Sigma_*^{-1} (\bar{X}_1 - \bar{X}_2)$$

где  $\bar{X}_1, \bar{X}_2$  – векторы средних значений;  
 $n_1, n_2$  – количество отсчетов в первой и второй выборке соответственно;  
 $\Sigma_*^{-1}$  – обратная матрица, рассчитанная для объединенной ковариационной матрицы вида

$$\Sigma_* = \frac{1}{n_1 + n_2 - 2} (\hat{X}_1^T \hat{X}_1 + \hat{X}_2^T \hat{X}_2)$$

Критические значения для критерия

$$T_{kp(\alpha, m, n_1 + n_2 - m - 2)}^2 = \frac{(n_1 + n_2 - 2)m}{n_1 + n_2 - m - 2} \times F_{\alpha, m, n_1 + n_2 - m - 2}$$

При  $T_p^2 < T_{kp(\alpha, m, n_1 + n_2 - m - 2)}^2$  нулевая гипотеза принимается с вероятностью  $(1-\alpha)$ .

## **2 Пример принятия решений на основе проверки статистических гипотез**



## Условия задачи

Для предприятия розничной торговли в административном районе установлены следующие нормативные показатели: эффективности деятельности (рентабельность) – 20 % и средняя продолжительность оборота материальных средств – 12 дней.

Предполагается (гипотеза  $H_0$ ), что более низкие значения показателей означают **нарушение ритмичности товарно-денежных операций и снижение конкурентоспособности предприятий торговли.**

# Рентабельность торговой деятельности и оборачиваемость материальных средств предприятий

| Номер объекта                    | Рентабельность, % | Продолжительность оборота, дней |
|----------------------------------|-------------------|---------------------------------|
| 01                               | 14                | 19                              |
| 02                               | 12                | 15                              |
| 03                               | 16                | 19                              |
| 04                               | 14                | 17                              |
| 05                               | 15                | 24                              |
| 06                               | 18                | 12                              |
| 07                               | 22                | 10                              |
| 08                               | 20                | 15                              |
| 09                               | 13                | 18                              |
| 10                               | 19                | 20                              |
| 11                               | 12                | 22                              |
| 12                               | 14                | 23                              |
| Среднее значение ( $\bar{X}_j$ ) | 15,8              | 17,8                            |



# Последовательность решения

1. Ковариационная матрица

$$\Sigma = \frac{1}{n-1} (\hat{X}^T \hat{X})$$

$$\Sigma = \frac{1}{11} \times \begin{pmatrix} 14-15,8 & 12-15,8 & 16-15,8 & 14-15,8 & 15-15,8 & \boxtimes & 14-15,3 \\ 19-17,8 & 15-17,8 & 19-17,8 & 17-17,8 & 24-17,8 & \boxtimes & 23-17,8 \end{pmatrix} \times$$

$$\times \begin{pmatrix} 14-15,8 & 19-17,8 \\ 12-15,8 & 15-17,8 \\ 16-15,8 & 19-17,8 \\ 14-15,8 & 17-17,8 \\ 15-15,8 & 24-17,8 \\ \boxtimes & \boxtimes \\ 14-15,8 & 23-17,8 \end{pmatrix} = \frac{1}{11} \times \begin{pmatrix} 118,3 & -86,5 \\ -86,5 & 210,7 \end{pmatrix} = \begin{pmatrix} 10,75 & -7,86 \\ -7,86 & 18,34 \end{pmatrix}.$$

Обратная ковариационная матрица

$$\Sigma^{-1} = \frac{1}{|\Sigma|} a_{ij} \Sigma = \begin{pmatrix} 0,1355 & 0,0581 \\ 0,0581 & 0,1093 \end{pmatrix}$$

## Последовательность решения (продолжение)

### 3. Расчетное значение критерия Хотеллинга

$$T_p^2 = n(\bar{X} - \mu)^T \Sigma^{-1}(\bar{X} - \mu) = 12 * (15,8 - 20 \quad 17,8 - 12) \times \begin{pmatrix} 0,1355 & 0,0581 \\ 0,0581 & 0,1093 \end{pmatrix} \times \\ \times \begin{pmatrix} 15,8 & -20 \\ 17,8 & -12 \end{pmatrix} = 32,16 .$$

### 4. Критическое значение критерия Хотеллинга

$$T_{кр}^2 = \frac{m(n-1)}{n-m} F_{0,05;2;10} = \frac{2(12-1)}{12-2} \times 4,459 = 9,8$$



**Пример 2.** Для того, чтобы оценить уровень различия двух групп инженерно-технических работников на предприятиях легкой промышленности (мужчины и женщины) по двум признакам, было проведено выборочное обследование. В выборку попали 20 человек (10 мужчин и 10 женщин). Результаты наблюдения представлены в таблице

| Мужчины |       | Женщины |       |
|---------|-------|---------|-------|
| $X_1$   | $X_2$ | $X_1$   | $X_2$ |
| 4       | 9,0   | 20      | 7,2   |
| 15      | 8,2   | 8       | 6,0   |
| 17      | 10,0  | 12      | 9,2   |
| 20      | 8,0   | 6       | 6,5   |
| 22      | 6,5   | 5       | 7,5   |
| 30      | 8,5   | 18      | 8,0   |
| 20      | 7,5   | 15      | 8,4   |
| 7       | 7,0   | 25      | 9,0   |
| 18      | 9,7   | 10      | 8,1   |
| 4       | 8,4   | 17      | 7,8   |

$X_1$  - стаж работы, лет;  $X_2$  - средняя дневная заработная плата, ден. ед.

Нулевая гипотеза – **расхождения** между работающими мужчинами и женщинами по анализируемым признакам **несущественны**.

## Решение

1. Для каждой группы отдельно и для совокупности в целом выполнен расчет средних значений и дисперсии анализируемых признаков.

$$\begin{array}{l} \text{I группа (мужчины)} \\ \bar{X}_1 = 15,7 \text{ лет; } \sigma_1^2 = 63,810; \\ \bar{X}_2 = 8,28 \text{ ден. ед.; } \sigma_2^2 = 1,106. \end{array}$$

$$S_1 = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 63,810 & -0,576 \\ -0,576 & 1,106 \end{pmatrix}$$

$$\begin{array}{l} \text{II группа (женщины)} \\ \bar{X}_1 = 13,6 \text{ лет; } \sigma_1^2 = 38,24; \\ \bar{X}_2 = 7,77 \text{ ден. ед.; } \sigma_2^2 = 0,926. \end{array}$$

$$S_2 = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 38,240 & 3,078 \\ 3,078 & 0,926 \end{pmatrix}$$

2 Расчет совместной ковариационной матрицы для двух групп

$$S_* = \frac{1}{n_1 + n_2 - 2} (S_1 n_1 + S_2 n_2) = \frac{1}{10 + 10 - 2} \begin{pmatrix} 1020,5 & 25,0 \\ 25,0 & 20,3 \end{pmatrix} = \begin{pmatrix} 56,7 & 1,4 \\ 1,4 & 1,1 \end{pmatrix}$$

$$S^{-1} = \begin{pmatrix} 0,018 & -0,022 \\ -0,022 & 0,913 \end{pmatrix}$$



### 3 Определение расчетного значения критерия Хотеллинга

$$T_p^2 = \frac{n_1 n_2}{n_1 + n_2 - 2} (\bar{X}_1 - \bar{X}_2)' S_*^{-1} (\bar{X}_1 - \bar{X}_2) =$$
$$= \frac{10 \cdot 10}{10 + 10 - 2} (2,1 \quad 0,51) \times \begin{pmatrix} 0,018 & -0,022 \\ -0,022 & 0,913 \end{pmatrix} \times \begin{pmatrix} 2,10 \\ 0,51 \end{pmatrix} = 1,49.$$

### 4 Определение табличного значения критерия

$$T_{кр}^2 = \frac{(10 + 10 - 2) \cdot 2}{(10 + 10 - 2 - 2)} \cdot 3,592 = 8,082$$

Следовательно, нулевая гипотеза о равенстве векторов средних значений двух множеств принимается, т.е. расхождения между работающими мужчинами и женщинами по изучаемым признакам несущественны.