6. СТАТИСТИЧЕСКИЕ ГИПОТЕЗЫ В ЗАДАЧАХ ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

6.1. Понятие статистической гипотезы. Виды гипотез

Гипотезой принято называть предположение о некоторых свойствах изучаемых явлений. При обработке экспериментальных данных рассматриваются гипотезы о свойствах генеральной совокупности, например о виде закона распределения исследуемого признака, о параметрах закона распределения. Эти гипотезы проверяются путём обработки случайной выборки и в дальнейшем называются ***статистическими***.

Наряду с принятой гипотезой рассматривают и противоречащую ей гипотезу, которая может быть принята в том случае, если первая не подтвердится. Для отличия эти гипотез друг от друга выдвинутую гипотезу принято называть **нулевой** или **основной** и обозначать символом H_0, а гипотезу, противоречащую нулевой, – **конкурирующей** или **альтернативной** и обозначать символом H_1.

Для краткости записи гипотез используют специальное обозначение. Пусть нулевая гипотеза состоит в предположении, что математические ожидания двух нормально распределённых случайных величин $\hat{x}$ и $\hat{y}$ равны, а конкурирующая гипотеза состоит в том, что они не равны. Эти гипотезы записываются следующим образом:

$$H_0 : M_{\hat{x}} = M_{\hat{y}}; \quad H_1 : M_{\hat{x}} \neq M_{\hat{y}}.$$

Гипотезы принято подразделять на простые и сложные. Простой называют гипотезу, содержащую только одно предположение, например $H_0 : M_{\hat{x}} = M_{\hat{y}}$. Сложной называют гипотезу, содержащую конечное или бесконечное число предположений. Например, гипотеза $H_0 : M_{\hat{x}} > 10$ состоит из бесконечного множества простых гипотез вида $H_i : M_{\hat{x}} = b_i$, где b_i – любое число, превосходящее 10.

Задачи проверки гипотез можно разделить на несколько классов, отличающихся друг от друга как по форме, так и по методом решения. Прежде всего, эти задачи делятся на ***параметрические***, когда вид закона распределения известен, и ***непараметрические***, когда закон распределения неизвестен. В свою очередь, каждый из данных классов содержит следующие подклассы.

1. **Задачи согласия**. Данные задачи сводятся к проверке согласия (соответствия) вида закона распределения или значений параметров распределения, выдвинутых в качестве предполагаемых, с законом распределения или параметрами закона распределения исследуемой случайной величины.

Формулировка данных задач имеет следующий вид. Нулевая гипотеза:

$$H_0 : Q_{\hat{x}} = S_{\hat{x}}.$$

Конкурирующие гипотезы:

$$H_1^{(1)} : Q_{\hat{x}} < S_{\hat{x}}; \quad H_1^{(2)} : Q_{\hat{x}} > S_{\hat{x}}; \quad H_1^{(3)} : Q_{\hat{x}} \neq S_{\hat{x}}.$$

Здесь $Q_{\hat{x}}$ и $S_{\hat{x}}$ – символы сравниваемых случайных объектов. Так, если речь идёт о проверке согласия закона распределения, то $Q_{\hat{x}} = F_{\hat{x}}$, $S_{\hat{x}} = F_{\hat{x}\text{г}}$, где $F_{\hat{x}}$ – закон распределения исследуемой случайной величины; $F_{\hat{x}\text{г}}$ – гипотетический закон распределения. Для данного случая получим:

$$H_0 : \ F_{\hat{x}} = F_{\hat{x}\text{г}}; \quad H_1 : \ F_{\hat{x}} \neq F_{\hat{x}\text{г}}.$$

При проверке согласия параметров распределения альтернативные гипотезы могут выдвигаться в форме гипотез, содержащих бесчисленное множество предположений.

2. **Задачи независимости**. Эти задачи возникают в тех случаях, когда необходимо проверить, являются ли компоненты некоторого случайного вектора независимыми. Очевидно, что если компоненты вектора

$$\hat{X}_{<n>} = (\hat{x}_1, \hat{x}_2, ..., \hat{x}_n)^{\text{т}}$$

независимы, то

$$F_{\hat{X}_{<n>}}(X_{<n>}) = \prod_{i=1}^{n} F_{\hat{x}_i}(x_i),$$

где $F_{\hat{X}_{<n>}}(X_{<n>})$ – закон распределения случайного вектора $X_{<n>}$; $F_{\hat{x}_i}(x_i)$, $i = \overline{1, n}$ – закон распределения *i*–го компонента вектора $X_{<n>}$.

Поэтому задачу проверки гипотезы о независимости можно сформулировать следующим образом:

$$H_0 : \ F_{\hat{X}_{<n>}}(X_{<n>}) = \prod_{i=1}^{n} F_{\hat{x}_i}(x_i);$$

$$H_1 : \ F_{\hat{X}_{<n>}}(X_{<n>}) \neq \prod_{i=1}^{n} F_{\hat{x}_i}(x_i).$$

3. **Задачи проверки выборки**. Данные задачи появляются в случае необходимости проверки того факта, что полученная выборка является простой, т.е. варианты выборки подчинены одному и тому же закону распределения.

Задачи такого типа формулируются в виде соотношений:

$$H_0 : \ F_{\hat{X}_{<n>}}(x_1, x_2, .., x_n) = \prod_{i=1}^{n} F_{\hat{x}}(x_i);$$

$$H_1:\ F_{\hat{X}_{<n>}}(x_1,\ x_2,..,x_n) \neq \prod_{i=1}^{n} F_{\hat{x}}(x_i)\,.$$

6.2. Общий подход к проверке гипотез

Подход к решению задачи проверки гипотез рассмотрим на следующих двух примерах.

Пример 6.1. На склад готовой продукции микросхемы одного типа поступают с двух заводов, выпускающих продукцию разного качества, и такими же партиями микросхемы отпускаются со склада потребителю. Качество продукции заводов характеризуется вероятностью p того, что случайным образом выбранная микросхема является бракованной. Для одного завода $p = p_0$, для другого $p = p_1$ ($p_0 < p_1$). Потребитель произвольно выбирает одну партию микросхем. Необходимо на основании результатов контроля решить, на каком заводе изготовлена выбранная партия микросхем.

▼ Введём нулевую гипотезу H_0, состоящую в том, что выбранная партия микросхем изготовлена на одном заводе (вероятность брака равна p_0), и конкурирующую гипотезу H_1 о том, что партия микросхем изготовлена на другом заводе (вероятность брака равна p_1).

Отберём из партии случайным образом n изделий. Обозначим число бракованных микросхем среди отобранных символом $\hat{x}$. Очевидно, что $\hat{x}$ – дискретная случайная величина, множество значений которой

$$X_{\{n\}} = \{0,\ 1,\ 2,..,n\}.$$

Назовём ***решающим правилом*** или ***критерием проверки гипотезы*** совокупность условий, при которых нулевая гипотеза принимается или отвергается.

В рассматриваемом примере решающее правило будет состоять в некотором разбиении множества $X_{\{n\}}$ на два подмножества X_0 и X_1 ($X_0 \cup X_1 = X,\ \ X_0 \cap X_1 = \varnothing$) таких, что при попадании возможного значения случайной величины $\hat{x}$ в множество X_0 гипотеза H_0 принимается, а в множество X_1 – отвергается. ▲

Разбиение множества X на подмножества X_0 и X_1 можно осуществлять различным образом, поэтому прежде чем решать поставленную задачу, необходимо определить, какое из возможных разбиений множества X на подмножества X_0 и X_1 следует выбрать.

Пример 6.2. На вход приёмного устройства в некоторый момент времени поступает случайный сигнал $\hat{y}$, который представляет сумму известного сигнала x и случайной помехи $\hat{z}$, либо одну помеху $\hat{z}$. Измеряется величина $\hat{y}$, и на основании полученного числового значения y

необходимо установить, присутствовал ли на входе сигнал, т.е. выбрать одну из возможностей: $\hat{y} = x + \hat{z}$ или $\hat{y} = \hat{z}$.

▼ Введём нулевую гипотезу H_0, состоящую в том, что сигнал присутствует ($H_0 : \hat{y} = x + \hat{z}$) и конкурирующую гипотезу о том, что сигнал на входе отсутствует ($H_1 : \hat{y} = \hat{z}$).

Множество Y возможных значений случайной величины $\hat{y}$ представляет собой всю числовую ось. Решающее правило в данном случае будет состоять в разбиении множества Y на две части Y_0 и Y_1, такие, что при попадании возможного значения случайной величины $\hat{y}$ в множество Y_0 гипотеза H_0 принимается, а при попадании возможного значения $\hat{y}$ в множество Y_1 эта гипотеза отвергается. Как и в предыдущей задаче, необходимо решить вопрос о таком разбиении. ▲

Из приведённых примеров видно, что при наличии способов разбиения множества возможных значений наблюдаемой величины $\hat{x}$ или $\hat{y}$ на подмножества, соответствующие приёму и отклонению гипотезы H_0, общий подход к решению задачи проверки гипотез включает в себя следующие этапы.

1. Выдвигается нулевая и конкурирующая гипотезы.

2. Выбирается некоторая величина, которая представляет собой функцию элементов выборки, связана с нулевой и конкурирующей гипотезами и зависит от условий проведения эксперимента. В дальнейшем эту величину будем называть ***показателем согласованности гипотезы***.

3. Выбирается критерий проверки (критерий согласия, критерий соответствия), т.е. совокупность правил, указывающих, при каких значениях показателя согласованности гипотеза отвергается, а при каких не отвергается.

4. Множество возможных значений показателя согласованности в соответствии с принятым критерием разбивается на два подмножества таким образом, что попадание возможного значения данного показателя в одно из этих подмножеств означает принятие гипотезы H_0, а в другое – отклонение указанной гипотезы.

5. Проводится эксперимент, вычисляется величина показателя и определяется, к какому из подмножеств относится эта величина, на основании чего принимается решение о приёме или отклонении гипотезы H_0.

Из описанного выше общего подхода к решению задачи проверки гипотез следует, что это решение связано с предварительным выбором показателя согласованности и критерия проверки гипотез, которые должны обладать определёнными свойствами.

6.3. Показатель согласованности и его свойства

Показателем согласованности или ***статистической характеристикой*** гипотезы называется случайная величина $\hat{u}$, являющаяся функцией гипотетических данных и результатов наблюдений, предназначенная для проверки нулевой гипотезы.

Конкретный вид показателя согласованности для различных гипотез может быть различным. Так, при проверке гипотезы о законе распределения показатель согласованности может задаваться следующими способами:

– в виде зависимости от гипотетической функции распределения $F_{\hat{x}г}(x)$, т.е. функции распределения, выдвинутой в качестве нулевой гипотезы, и статистической функции распределения $F^*_{\hat{x}}(x)$, полученной экспериментально:

$$\hat{u} = f_1\left(F_{\hat{x}г}(x);\ F^*_{\hat{x}}(x)\right); \tag{6.3.1}$$

– в виде зависимости от гипотетической вероятности p и частоты p^*, полученной в результате проведения эксперимента:

$$\hat{u} = f_2\left(p;\ p^*\right).$$

При проверке гипотезы о равенстве математических ожиданий двух независимых случайных величин $\hat{x}$ и $\hat{y}$ показатель согласованности может выбираться в виде различного рода зависимостей от начальных и центральных моментов первого и второго порядков от случайных величин $\hat{x}$ и $\hat{y}$:

$$\hat{u} = f_3\left(\nu^*_1(\hat{x});\ \nu^*_2(\hat{y});\ \mu^*_2(\hat{x});\ \mu_2(\hat{y})\right).$$

Применяются также и другие виды зависимостей. Однако, несмотря на такое разнообразие, в любом случае показатель согласованности должен удовлетворять ряду требований. Поскольку это величина случайная, то и требования формулируются применительно к закону распределения показателя согласованности. Состоят они в следующем.

1. Показатель согласованности должен определяться нулевой и конкурирующей гипотезами, а также условиями проведения эксперимента. Так, в показателе согласованности, определяемом выражением (6.3.1), эта зависимость представлена наличием как гипотетической, так и статистической функций распределения в качестве аргументов функции f_1.

2. Показатель согласованности должен представлять собой случайную величину, точное или приближённое распределение которой известно. В настоящее время наиболее распространён выбор показателей согласованности, распределённых по нормальному закону, законам хи-квадрат, Стьюдента, Фишера. Причём показатели согласованности с различными законами распределения обозначаются разными символами.

Так, показатели, распределённые по нормальному закону, обозначают через u или z, по закону хи-квадрат – через χ^2, по закону Стьюдента – через t, по закону Фишера – через F.

3. Закон распределения показателя согласованности должен быть инвариантен к виду закона распределения исследуемой случайной величины. Именно данное обстоятельство и определило широкое распространение показателей согласованности, имеющих указанные выше законы распределения.

4. Для построения закона распределения показателя согласованности должен быть востребован минимум априорных сведений, так как возможность получения достоверных сведений до опыта существенно ограничена.

5. Закон распределения показателя согласованности должен быть критичен по отношению к проверяемой гипотезе. Указанное требование означает, что условные плотности распределения $\varphi_{\hat{u}/H_0}(u)$ и $\varphi_{\hat{u}/H_1}(u)$ должны существенно отличаться друг от друга.

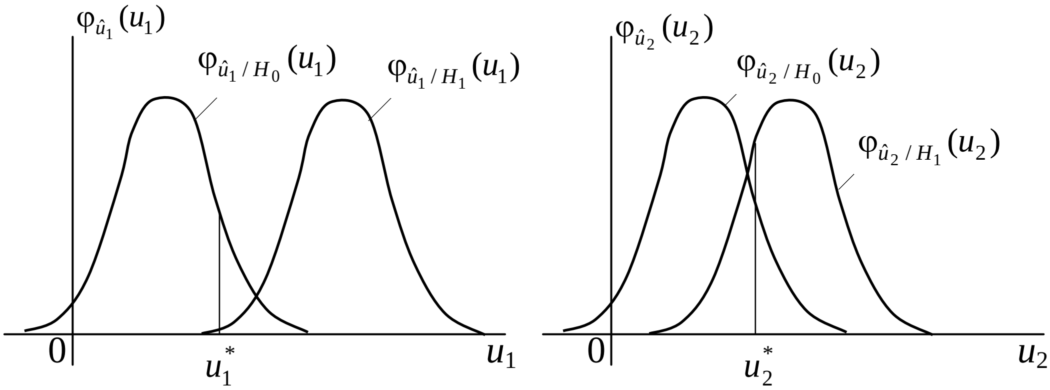


Рис.6.1. Условные плотности распределения показателей согласованности

На рис.6.1 изображены кривые условных плотностей распределения двух различных показателей согласованности $\hat{u}_1$ и $\hat{u}_2$ при нулевой и конкурирующей гипотезах. Из сравнения кривых видно, что применение показателя $\hat{u}_1$ предпочтительнее, так как он обеспечивает более высокую степень уверенности различения гипотез H_0 и H_1, чем показатель $\hat{u}_2$. Действительно, при одном и том же значении $u = u_1^* = u_2^*$, т.е. при наступлении одного и того же события $\hat{u} = u$, вероятность отнесения его к нулевой гипотезе значительно выше, когда используется показатель $\hat{u}_1$.

В заключение следует отметить, что для проверки гипотезы по данным выборки вычисляют частные значения входящих в показатель согласованности величин и, таким образом, получают частное значение показателя согласованности гипотезы. Это значение, вычисленное по данным выборки, в дальнейшем будем называть ***наблюдаемым значением*** показателя согласованности $\hat{u}$ и обозначать через u.

6.4. Методы задания критической области

Как отмечалось выше, проверка гипотез требует задания решающего правила, т.е. метода разбиения множества U возможных значений показателя согласованности $\hat{u}$ на два подмножества: подмножество U_0, при попадании в которое наблюдаемого значения u показателя $\hat{u}$ нулевая гипотеза принимается, и подмножество U_1, при попадании в которое наблюдаемого значения u нулевая гипотеза отвергается. В дальнейшем область, соответствующую U_0, будем называть ***областью допустимых значений*** (областью принятия гипотезы H_0) и обозначать символом D, а область, соответствующую подмножеству U_1 – ***критической областью*** показателя $\hat{u}$ и обозначать символом Q.

Поскольку $\hat{u}$ – одномерная случайная величина, то все её возможные значения принадлежат некоторому интервалу. Поэтому области Q и D являются интервалами, следовательно, существуют разделяющие их точки. Они называются ***критическими точками*** (***границами***). Задание критической области и сводится к заданию критических точек.

Сущность задания критических точек состоит в следующем. Рассмотрим события:

A – верна гипотеза H_0;

$\overline{A}$ – верна гипотеза H_1;

B – наблюдаемое значение u показателя согласованности $\hat{u}$ попало в область D;

$\overline{B}$ – наблюдаемое значение u попало в область Q.

Тогда в процессе принятия решения возможен один из следующих исходов:

$A \cap B$ – верна гипотеза H_0 и принято решение о её справедливости;

$A \cap \overline{B}$ – верна гипотеза H_0, а принято решение о справедливости гипотезы H_1;

$\overline{A} \cap B$ – верна гипотеза H_1, а принято решение о справедливости гипотезы H_0;

$\overline{A} \cap \overline{B}$ – верна гипотеза H_1 и принято решение о её справедливости.

Очевидно, что исходы $A \cap \overline{B}$ и $\overline{A} \cap B$ являются ошибочными, первому из них соответствует ошибка первого рода, а второму – ошибка второго рода.

Таким образом, под ошибкой первого рода понимается принятие решения об отклонении нулевой гипотезы в случае, если в действительности она является правильной, а под ошибкой второго рода – решение о принятии нулевой гипотезы, если в действительности она не верна.

Поскольку рассмотренные события являются случайными, то им могут быть поставлены в соответствие вероятности наступления данных событий, а именно:

p_{11} – вероятность наступления события $A \cap B$;

p_{12} – вероятность наступления события $A \cap \overline{B}$;

p_{21} – вероятность наступления события $\overline{A} \cap B$;

p_{22} – вероятность наступления события $\overline{A} \cap \overline{B}$.

Значения p_{11}, p_{12}, p_{21}, p_{22} можно вычислить как вероятности попадания случайной величины $\hat{u}$ в области D и Q следующим образом.

Пусть законы распределения показателя согласованности $\hat{u}$ при условии, что справедлива нулевая или конкурирующая гипотеза заданы в форме плотности распределения $\varphi_{\hat{u}/H_0}(u)$ и $\varphi_{\hat{u}/H_1}(u)$, а границей данных областей является точка u^*. Предположим, что взаимное расположение кривых распределения $\varphi_{\hat{u}/H_0}(u)$ и $\varphi_{\hat{u}/H_1}(u)$, а также областей D и Q имеет вид, изображённый на рис.6.2.

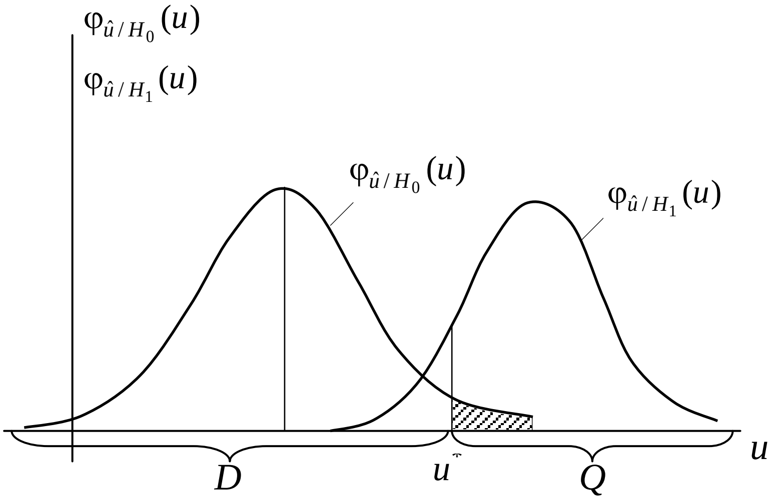


Рис.6.2. Кривые распределения показателя согласованности при различных гипотезах

В этом случае вероятности p_{11}, p_{12}, p_{21}, p_{22} определяются следующими соотношениями:

$$p_{11} = P(A/B) = \int_D \varphi_{\hat{u}/H_0}(u)du; \quad (6.4.1)$$

$$p_{12} = P(A/\overline{B}) = \int_Q \varphi_{\hat{u}/H_0}(u)du; \quad (6.4.2)$$

$$p_{21} = P(\overline{A}/B) = \int_D \varphi_{\hat{u}/H_1}(u)du; \quad (6.4.3)$$

$$p_{22} = P(\overline{A}/\overline{B}) = \int_Q \varphi_{\hat{u}/H_1}(u)du. \quad (6.4.4)$$

Из выражений (6.4.1) – (6.4.4) следует, что значения p_{11}, p_{12}, p_{21}, p_{22} зависят от размеров и расположения области D допустимых значений и критической области Q. Поэтому, предъявляя соответствующие требования к вероятностям p_{11}, p_{12}, p_{21}, p_{22}, можно определить расположение и размеры данных областей, т.е. критические границы.

Так как проверка гипотезы связана с принятием решения о справедливости или несправедливости выдвинутой гипотезы, то при нахождении областей D и Q целесообразно опираться на результаты теории статистических решений (см. § 2.3). Практическое применение данных результатов зависит от объёма априорной информации, которая может быть использована при проверке гипотез. В связи с этим методы задания критической области принято делить на две группы [5]:

– методы, опирающиеся на оценки потерь от неправильного решения;

– методы, опирающиеся на оценки вероятностей ошибок при принятии решений.

Наибольшее практическое распространение получили методы второй группы, так как их применение требует минимальной априорной информации при проверке гипотез, однако, достоверность принятия решений с помощью данных методов ниже, чем для методов первой группы.

6.5. Проверка гипотез как задача принятия решений

Чтобы формализовать задачу проверки статистических гипотез в виде задачи принятия решения, опишем эту задачу в терминах теории статистических решений.

В качестве объекта наблюдения здесь выступает гипотеза H. Будем полагать, что два возможных варианта данной гипотезы – нулевая H_0 и конкурирующая H_1 представляют собой простые гипотезы.

В качестве статистической характеристики гипотезы используется показатель согласованности $\hat{u}$, являющийся некоторой функцией результатов наблюдения.

Множество решений включает в себя: решение $\tilde{E}_1$, состоящее в принятии гипотезы H_0, и решение $\tilde{E}_2$, состоящее в отклонении гипотезы H_0 (т.е. принятии гипотезы H_1).

Объём априорной информации в процессе принятия гипотез может быть различен. Так, при минимальной неопределённости она включает в себя (см. § 2.3):

– вероятности наступления гипотез H_0 и H_1, данные вероятности запишем как $P(H = H_0)$ и $P(H = H_1)$;

– законы распределения показателя согласованности $\hat{u}$ при условии справедливости нулевой и конкурирующей гипотез, т.е. условные плотности распределения $\varphi_{\hat{u}/H_0}(u)$ и $\varphi_{\hat{u}/H_1}(u)$;

– функцию потерь π, задаваемую в виде матрицы потерь:

$$\pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix}.$$

Решающее правило для данного случая состоит в разделении множества возможных значений показателя согласованности на два подмножества D и Q. Попадание наблюдаемого значения в первое из них означает принятие решения о справедливости гипотезы H_0, а второе - принятие решения об отклонении H_0. Таким образом, решающее правило определяет выбор критической границы или границ, если их несколько, и, таким образом задаёт критическую область Q.

Решающее правило может быть сформулировано на основе принципов принятия статистических решений. При проверке гипотез используется четыре вида правил, причём применение того или иного вида зависит от полноты априорных данных.

Если задача проверки гипотез сформулирована как задача выбора решений и матрица потерь π определена, то оптимальное решение может быть получено на основе байесовского или минимаксного правила.

Если функция потерь π не определена, то для однозначного выбора решения при проверке гипотез можно использовать два подхода. Применительно к задачам с известным априорным распределением гипотез наиболее полной характеристикой степени соответствия каждой из них результатам произведённого испытания является апостериорная вероятность этой гипотезы. При этом истинной считается апостериорная вероятность. Указанное правило называется ***правилом апостериорной вероятности.***

При отсутствии данных об априорном распределении гипотез единственной характеристикой степени соответствия той или иной гипотезы результатам наблюдения является функция правдоподобия.

Поэтому в таких случаях выбор решений производится на основе ***правила максимума правдоподобия***, в соответствии с которым истинной считается гипотеза с наибольшим значением функции правдоподобия.

6.6. Проверка гипотез классическим методом

Применение рассмотренных выше методов, как уже отмечалось, связано с использованием априорной информации, которая далеко не всегда имеется в распоряжении исследователя при обработке экспериментальных данных. В связи с этим на практике наибольшее распространение получили методы проверки гипотез, опирающиеся при назначении критических границ только на информацию о вероятностях ошибок первого и второго рода.

Сущность данных методов состоит в задании вероятности p_{12} ошибки первого рода. Зная эту величину и зависимость, связывающую вероятность p_{12} с показателем согласованности $\hat{u}$, можно определить границы и расположение критической области Q.

В связи с тем, что вероятность p_{12} ошибки первого рода играет в данных методах ведущую роль, она имеет специальное название – ***уровень значимости критерия проверки гипотезы*** и специальное обозначение α. Таким образом, в соответствии с выражением (6.4.2)

$$\alpha = p_{12} = P(A/\overline{B}) = \int\limits_{Q} \varphi_{\hat{u}/H_0}(u)du. \qquad (6.6.1)$$

Если известны $\varphi_{\hat{u}/H_0}(u)$ и α, то выражение (6.6.1) позволяет определить критические границы, т.е. границы области Q. Данные границы в дальнейшем будем обозначать символами $u_{\alpha 1}$, $u_{\alpha 2}$,…, если область Q состоит из нескольких подобластей, или символом u_α, если она представляет собой одну область, т.е. возможные значения показателя согласованности $\hat{u}$ делятся границей u_α на две полупрямые.

При конкретном выборе критических границ u_α необходимо учитывать два дополнительных обстоятельства, а именно:

– соотношение между условными законами распределения показателя согласованности $\hat{u}$, соответствующими нулевой и альтернативной гипотезам;

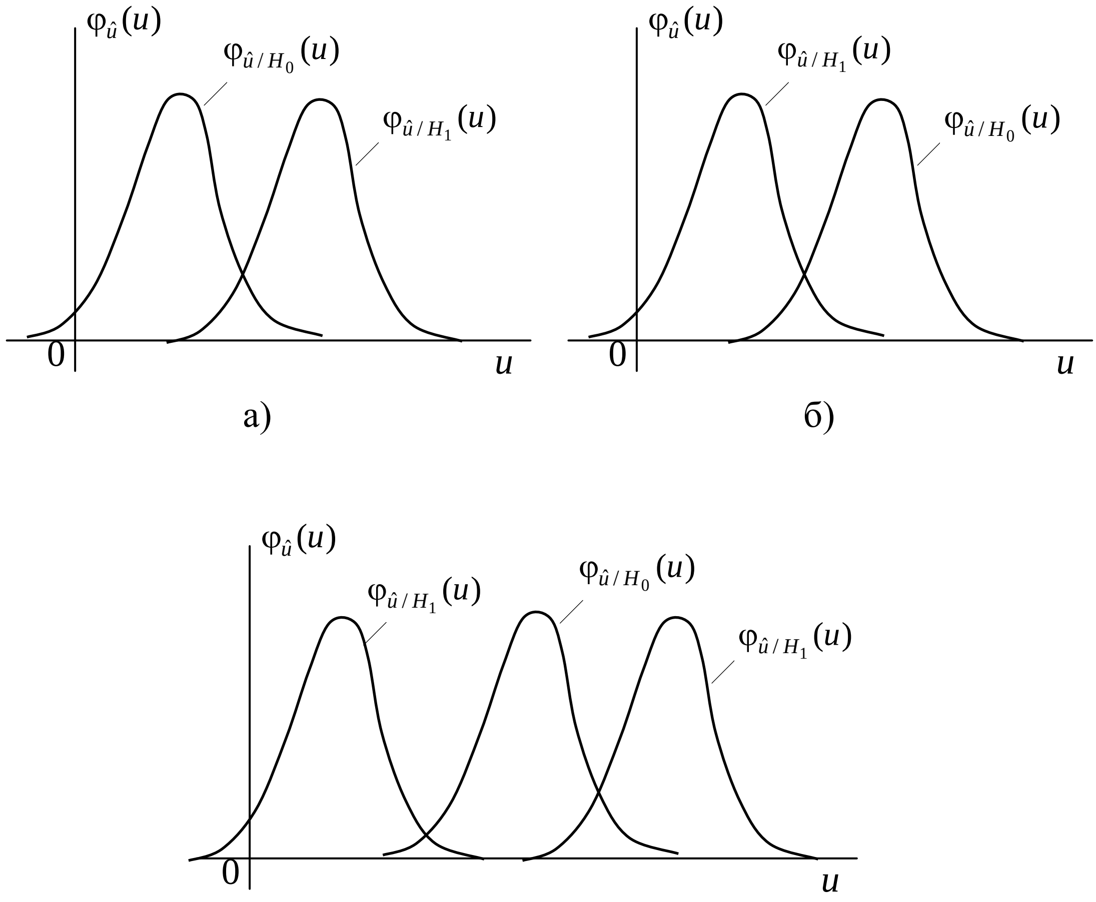
– взаимозависимость ошибок первого и второго рода.

Поясним эти обстоятельства более подробно. Соотношение между условными законами распределения показателя согласованности $\hat{u}$, соответствующими нулевой и конкурирующей гипотезам, выражается в виде взаимного расположения их кривых распределения на оси абсцисс. Особенности характеристики $\hat{u}$, а также нулевой и конкурирующей гипотез приводят к тому, что кривые распределения $\varphi_{\hat{u}/H_0}(u)$, $\varphi_{\hat{u}/H_1}(u)$ могут располагаться друг относительно друга тремя различными способами:

– известно, что кривая распределения $\varphi_{\hat{u}/H_1}(u)$ сдвинута относительно $\varphi_{\hat{u}/H_0}(u)$ вправо (рис.6.3,а);

– известно, что кривая распределения $\varphi_{\hat{u}/H_1}(u)$ сдвинута относительно $\varphi_{\hat{u}/H_0}(u)$ влево (рис.6.3,б);

– сдвиг кривой распределения $\varphi_{\hat{u}/H_1}(u)$ неизвестен, т.е. она может быть сдвинута относительно $\varphi_{\hat{u}/H_0}(u)$ как вправо, так и влево (рис.6.3,в).



в)

Рис.6.3.Варианты расположения условных законов распределения показателя согласованности

Очевидно, что вид критической области для каждого способа должен быть различен, а именно в первом случае должна быть выбрана правосторонняя критическая область, во втором – левосторонняя и в третьем – двусторонняя. Правосторонней называют критическую область, определяемую неравенством $u > u_\alpha$, левосторонней – неравенством $u < u_\alpha$, двусторонней – неравенствами $u < u_{\alpha 1}$, $u > u_{\alpha 2}$, где $u_{\alpha 2} > u_{\alpha 1}$.

При отыскании критической области достаточно найти критические точки. Методика их отыскания состоит в следующем. Задаются уровнем значимости α и ищут критическую точку u_α, исходя из требования, чтобы при условии справедливости нулевой гипотезы вероятность попадания показателя согласованности $\hat{u}$ в критическую область была равна принятому уровню значимости.

На основании (6.6.1) для правосторонней критической области это условие имеет вид

$$P(\hat{u} \geq u_\alpha) = \int_{u_\alpha}^{\infty} \varphi_{\hat{u}/H_0}(u)du = \alpha,$$

для левосторонней

$$P(\hat{u} < u_{\alpha}) = \int_{-\infty}^{u_{\alpha}} \varphi_{\hat{u}/H_0}(u)du = \alpha,$$

для двусторонней

$$P(\hat{u} < u_{\alpha 1}) + P(\hat{u} \ge u_{\alpha 2}) = \int_{-\infty}^{u_{\alpha}} \varphi_{\hat{u}/H_0}(u)du + \int_{u_{\alpha}}^{\infty} \varphi_{\hat{u}/H_0}(u)du = \alpha.$$

В последнем случае чаще всего выбирают симметрично расположенные критические точки:

$$P(\hat{u} < u_{\alpha 1}) = P(\hat{u} \ge u_{\alpha 2}) = \frac{\alpha}{2}.$$

Критические точки, удовлетворяющие приведённым выше условиям, находят по соответствующим таблицам (см., например, приложения 5 и 6).

Из вышеизложенного следует, что при выборе критических областей необходимо учитывать не только свойства нулевой, но и свойства конкурирующей гипотезы.

Для пояснения взаимозависимости ошибок первого и второго рода рассмотрим условные плотности $\varphi_{\hat{u}/H_0}(u)$, $\varphi_{\hat{u}/H_1}(u)$ и правостороннюю критическую область с критической границей u_{α}, см. рис.6.4.

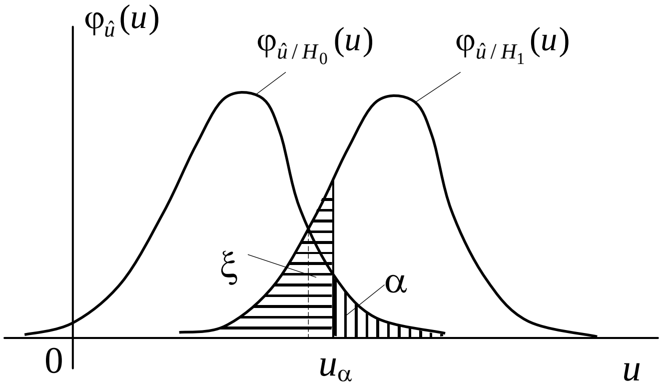


Рис.6.4. Взаимозависимость ошибок первого и второго рода

Следует напомнить, что вероятность ошибки первого рода равняется вероятности α попадания показателя $\hat{u}$ в критическую область. Чем меньше уровень значимости α, тем реже будет допускаться ошибка первого рода, а, следовательно, отвергаться правильная нулевая гипотеза H_0. Однако было бы неверно на основании этого делать вывод о том, что значение вероятности α должен быть выбрано как можно меньшим.

Из рис. 6.4 видно, чем меньше α, тем больше вероятность

$$\xi = \int_{-\infty}^{u_{\alpha}} \varphi_{\hat{u}/H_1}(u)du \qquad (6.6.2)$$

ошибки второго рода.

Указанная взаимосвязь ошибок первого и второго рода позволяет сделать следующий важный вывод: для уменьшения вероятности ошибки при принятии гипотезы критическую границу необходимо выбирать таким образом, чтобы сумма вероятностей ошибок первого и второго рода была минимальной (см. п.п.2.3.3). Если показатель согласованности подчинён нормальному закону, то минимум суммы вероятностей ошибок первого и второго рода достигается при выборе критической границы в абсциссе точки пересечения кривых распределения $\varphi_{\hat{u}/H_0}(u)$ и $\varphi_{\hat{u}/H_1}(u)$. На рис.6.4. указанное деление показано пунктиром.

Вместе с тем следует заметить, что не во всех случаях подход к выбору u_α с учётом минимума суммы вероятностей ошибок первого и второго рода целесообразен. На практике выбор целесообразной величины α зависит от «тяжести» последствий ошибок первого и второго рода для каждой конкретной задачи. Например, если ошибка первого рода повлечёт большие потери, а второго рода – малые, то целесообразно принять возможно меньшее α. Вернёмся к примеру 6.1 и рассмотрим, какую величину уровня значимости критерия проверки целесообразно выбрать с точки зрения потребителя микросхем. Заметим, что в задачах подобного типа вероятность приёма бракованной партии изделий (в нашем примере – микросхем), т.е. вероятность ошибки первого рода, принято называть *риском потребителя*, а вероятность признать бракованной партию качественных изделий (вероятность ошибки второго рода) – *риском производителя*. С точки зрения потребителя желательно уменьшать вероятность приёма бракованной партии микросхем, т.е. уменьшать вероятность ошибки первого рода, в связи с чем величину α целесообразно выбирать возможно меньшей.

Поскольку вероятность ошибки второго рода также играет важную роль при выборе критической области, то критерий проверки принято характеризовать так называемой мощностью показателя согласованности.

Мощностью γ ***показателя согласованности*** называют вероятность попадания показателя согласованности $\hat{u}$ в критическую область при условии, что справедлива конкурирующая гипотеза.

На основе определения можно записать

$$\gamma = \int_{u_\alpha}^{\infty} \varphi_{\hat{u}/H_1}(u)du. \qquad (6.6.3)$$

С учётом выражения (6.6.2) равенство (6.6.3) примет вид

$$\gamma = 1 - \int_{-\infty}^{u_\alpha} \varphi_{\hat{u}/H_1}(u)du = 1 - \xi.$$

Таким образом, мощность показателя согласованности – вероятность того, что не будет допущена ошибка второго рода. Поэтому для уменьшения ошибки второго рода критическую область необходимо строить так, чтобы мощность показателя согласованности при заданном уровне значимости была максимальной.

Мощность показателя согласованности позволяет обоснованно подойти к выбору односторонних критических областей. Предположим, что в качестве критической выбрана правосторонняя область (рис.6.4), но кривая условной плотности распределения для конкурирующей гипотезы $\varphi_{\hat{u}/H_1}(u)$ смещена относительно кривой $\varphi_{\hat{u}/H_0}(u)$ влево (рис.6.5).

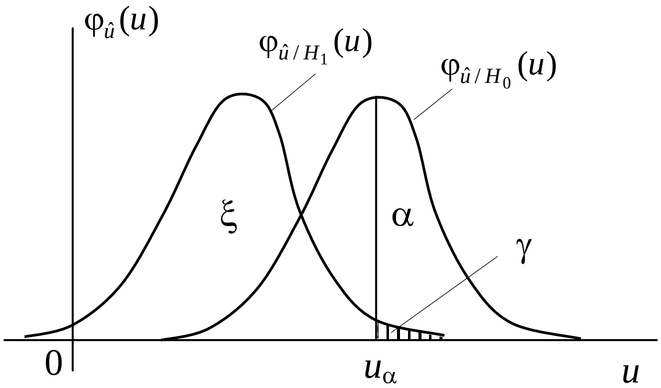


Рис.6.5. Условные распределения показателя согласованности гипотез

Найденные в этих условиях вероятности

$$\xi = \int_{-\infty}^{u_\alpha} \varphi_{\hat{u}/H_1}(u)du \approx 1; \qquad \gamma = 1 - \xi \approx 0$$

показывают, что при таком выборе критической области показатель согласованности $\hat{u}$ становится непригодным для статистической проверки гипотезы H_0, так как его мощность близка к нулю, а ошибки второго рода становятся практически достоверными. Очевидно, что для увеличения мощности показателя согласованности необходимо выбрать левостороннюю критическую область.

Если характер конкурирующей гипотезы неясен, то в качестве критической целесообразно выбирать двустороннюю симметричную область. Но следует отметить, что единственный способ одновременного уменьшения вероятностей ошибок первого и второго рода состоит в увеличении объёма выборки.

6.7. Проверка гипотез об аномальности результатов наблюдений

При обработке экспериментальных данных существенное значение имеет процесс предварительной обработки, одним из этапов которого является исключение результатов, содержащих грубые ошибки, т.е. аномальных результатов.

В любой выборке сомнительными являются, как правило, наибольший и наименьший элементы, которые и подлежат проверке. Обозначим через $\hat{x}_1$ и $\hat{x}_n$ наименьший и наибольший элементы случайной выборки $\hat{x}_1$, $\hat{x}_2, \ldots,$ $\hat{x}_n$, а через x_1 и x_n – реализации этих элементов в данном эксперименте.

Предположим, что сомнительным является наибольший элемент x_n случайной выборки. При этом будем полагать, что наблюдаемая величина $\hat{x}$ подчинена нормальному закону распределения с известными числовыми характеристиками $M_{\hat{x}}$ и $\sigma_{\hat{x}}$. Уровень значимости примем равным некоторой достаточно малой вероятности α. Для выборки, состоящей из одного элемента x, можно утверждать, что он является следствием грубой ошибки, если

$$x > M_{\hat{x}} + \sigma_{\hat{x}} t_{1-2\alpha} = x_\alpha \qquad (6.7.1)$$

или

$$\frac{x - M_{\hat{x}}}{\sigma_{\hat{x}}} > t_{1-2\alpha}.$$

Следовательно, при $n = 1$ в качестве показателя согласованности гипотезы целесообразно использовать случайную величину

$$\hat{u} = \frac{\hat{x} - M_{\hat{x}}}{\sigma_{\hat{x}}}, \qquad (6.7.2)$$

а критическую границу определять на основе выражения

$$u_\alpha = t_{1-2\alpha}. \qquad (6.7.3)$$

Однако при проверке аномальности крайнего элемента случайной выборки объёма $n > 1$ использование показателя согласованности вида (6.7.2) может привести к грубым ошибкам. Покажем это на примере.

<u>П р и м е р 6.3</u>. Пусть $n = 1$ и $\alpha = 0{,}05$.

▼ В приложении 4 находим

$$u_\alpha = t_{1-2\alpha} = t_{0,9} = 1{,}64. \qquad (6.7.4)$$

Подставив найденное значение u_α в формулу (6.7.1), получим

$$x > x_\alpha = M_{\hat{x}} + 1{,}64\,\sigma_{\hat{x}}. \qquad (6.7.5)$$

Таким образом, при одном испытании будем констатировать факт грубой ошибки, если наблюдаемое значение случайной величины удовлетворяет неравенству (6.7.5).

Увеличим теперь число испытаний до 20 и найдём вероятность того, что наибольший член $\hat{x}_{20}$ выборки превзойдёт величину x_α:

$$P(\hat{x}_{20} > x_\alpha) = 1 - P(\hat{x}_{20} < x_\alpha) = 1 - P\left(\bigcap_{i=1}^{20}(\hat{x}_i < x_\alpha)\right). \qquad (6.7.6)$$

Предполагая, что испытания независимы, и учитывая, что

$$P(\hat{x}_i < x_\alpha) = 1 - P(\hat{x}_i > x_\alpha) = 1 - \alpha, \quad i = \overline{1, n},$$

из выражения (6.7.6) получим

$$P(\hat{x}_{20} > x_\alpha) = 1 - (1-\alpha)^n = 1 - (0{,}95)^{20} \approx 0{,}65.$$

Таким образом, при использовании критической границы, определяемой выражением (6.7.4), 65% нормальных наибольших элементов выборки следует признать аномальными. Иначе, вероятность ошибки первого рода при 20 испытаниях увеличится до 0,65.

▲

С целью устранения указанного недостатка необходимо по мере увеличения объёма выборки сдвигать критическую границу вправо относительно значения u_α, определяемого равенством (6.7.4).

Для построения критической области, удовлетворяющей указанному требованию, рассмотрим функцию распределения $F_{\hat{x}_n}(x)$. Примем во внимание то обстоятельство, что для наступления события ($\hat{x}_n < x$) необходимо, чтобы все элементы выборки были меньше x:

$$\hat{x}_n < x = \bigcap_{i=1}^{n}(\hat{x}_i < x). \qquad (6.7.7)$$

Далее учитываем, что

$$P(\hat{x}_i < x) = F_{\hat{x}_i}(x) = F_{\hat{x}}(x) = \Phi_1\left(\frac{x - M_{\hat{x}}}{\sigma_{\hat{x}}}\right). \qquad (6.7.8)$$

К правой части выражения (6.7.7) применяем теорему умножения вероятностей независимых событий ($\hat{x}_i < x$), $i = \overline{1, n}$ и принимаем во внимание (6.7.8). В результате получаем

$$F_{\hat{x}_n}(x) = P(\hat{x}_n < x) = \prod_{i=1}^{n} P(\hat{x}_i < x) = \left(F_{\hat{x}}(x)\right)^n = \left(\Phi_1\left(\frac{x - M_{\hat{x}}}{\sigma_{\hat{x}}}\right)\right)^n. \qquad (6.7.9)$$

Границу u_α критической области, отвечающей уровню значимости α, можно найти как квантиль случайной величины $\hat{x}_n$ при аргументе $1-\alpha$. Зависимость между u_α и α определяется равенствами:

$$u_\alpha = F_{\hat{x}_n}^{-1}(1-\alpha);$$

$$F_{\hat{x}_n}(u_\alpha) = 1 - \alpha. \qquad (6.7.10)$$

Подставляем $x = u_\alpha$ в формулу (6.7.9) и на основании равенства (6.7.10) получим уравнение

$$\left(\Phi_1\left(\frac{u_\alpha - M_{\hat{x}}}{\sigma_{\hat{x}}}\right)\right)^n = 1-\alpha. \qquad (6.7.11)$$

Решаем (6.7.11) относительно u_α и находим

$$u_\alpha = M_{\hat{x}} + \sigma_{\hat{x}} t_{2(1-\alpha)^{1/n}-1}. \qquad (6.7.12)$$

Так как вероятность α обычно мала, то

$$(1-\alpha)^{\frac{1}{n}} \approx 1 - \frac{\alpha}{n}$$

и, следовательно,

$$t_{2(1-\alpha)^{1/n}-1} \approx t_{1-\frac{2\alpha}{n}}. \qquad (6.7.13)$$

Подставляя соотношение (6.7.13) в (6.7.12), получим

$$u_\alpha = M_{\hat{x}} + \sigma_{\hat{x}} t_{1-\frac{2\alpha}{n}}. \qquad (6.7.14)$$

Выражение (6.7.12) и применяется для определения критической границы в случае, если $n > 1$.

Порядок проверки гипотезы об аномальном значении наибольшего элемента выборки состоит в следующем.

1. Элемент, относительно которого выдвигается гипотеза, исключается из выборки, т.е. её объём уменьшается на единицу.

2. Назначается уровень значимости α и по приложению 4 определяется значение $t_{1-\frac{2\alpha}{n}}$.

3. По формуле (6.7.14) определяется критическая граница. При отсутствии априорных значений $M_{\hat{x}}$ и $\sigma_{\hat{x}}$ в данной формуле используются оценки $\tilde{M}_{\hat{x}}$ и $\tilde{\sigma}_{\hat{x}}$. При их вычислении предполагаемый аномальный результат из выборки исключается.

4. Наблюдаемое значение показателя согласованности гипотезы определяется по формуле

$$u = \frac{x_n - M_{\hat{x}}}{\sigma_{\hat{x}}}. \qquad (6.7.15)$$

5. Проверяется условие $u > u_\alpha$. Если оно выполняется, то наибольший элемент выборки, по отношению к которому выдвигалось предположение о наличии грубой ошибки, отбрасывается. При $u \le u_\alpha$ этот элемент сохраняется в выборке, поскольку данные эксперимента не подтверждают гипотезы о наличии грубой ошибки. Для подтверждения полученного вывода необходимо повторить проверку по пунктам 1–5, но с включением сомнительного элемента в выборку.

При рассмотрении в качестве аномального наименьшего элемента $\hat{x}_1$ случайной выборки, порядок проверки гипотезы сохраняется, но в качестве показателя согласованности последней используется случайная величина

$$\hat{u} = \frac{M_{\hat{x}} - \hat{x}_1}{\sigma_{\hat{x}}}.$$

При этом проверяемый элемент отбрасывается, если $u > u_\alpha$, и сохраняется в противном случае.

Пример 6.4. С помощью радиодальномера производятся 20 измерений дальности $\hat{z}$ до объекта. Точность радиодальномера характеризуется среднеквадратическим отклонением $\sigma_{\hat{x}}$ = 50 м. Имеются ли основания полагать, что наибольшее отклонение

$$\hat{z}_{20} - M_{\hat{z}} = 180 \text{ м},$$

зафиксированное в данной серии наблюдений, содержит грубую ошибку? Уровень значимости критерия проверки гипотезы принять равным 0,05.

▼ По условию задачи n = 20, α = 0,05. В соответствии с выражением (6.7.15) значение показателя согласованности

$$u = \frac{z_{20} - M_{\hat{z}}}{\sigma_{\hat{z}}} = \frac{180}{50} = 3{,}6.$$

Границу критической области находим в приложении 4:

$$u_\alpha = t_{1-\frac{2\alpha}{n}} = t_{1-\frac{0,1}{20}} = t_{0,995} = 2{,}82.$$

Так как $u > u_\alpha$, то наибольшее отклонение содержит грубую ошибку и его следует из дальнейшего рассмотрения исключить.

▲