



Электронный учебно-методический комплекс по учебной дисциплине

"Теория вероятностей и математическая статистика"

для специальности:

310304 «Информатика»

Оглавление | Программа | Теория | Практика | Контроль знаний | Об авторах

ТЕМА 12. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

12.1. Сущность задачи проверки статистических гипотез

Статистическая гипотеза представляет собой некоторое предположение о законе распределения случайной величины или о параметрах этого закона, формулируемое на основе выборки. Примерами статистических гипотез являются предположения: генеральная совокупность распределена по экспоненциальному закону; математические ожидания двух экспоненциально распределенных выборок равны друг другу. В первой из них высказано предположение о виде закона распределения, а во второй – о параметрах двух распределений. Гипотезы, в основе которых нет никаких допущений о конкретном виде закона распределения, называют непараметрическими, в противном случае – параметрическими.

Гипотезу, утверждающую, что различие между сравниваемыми характеристиками отсутствует, а наблюдаемые отклонения объясняются лишь случайными колебаниями в выборках, на основании которых производится сравнение, называют нулевой (основной) гипотезой и обозначают H_0 . Наряду с основной гипотезой рассматривают и альтернативную (конкурирующую, противоречащую) ей гипотезу H_1 . И если нулевая гипотеза будет отвергнута, то будет иметь место альтернативная гипотеза.

Различают простые и сложные гипотезы. Гипотезу называют простой, если она однозначно характеризует параметр распределения случайной величины. Например, если λ является параметром экспоненциального распределения, то гипотеза H_0 о равенстве $\lambda = 10$ – простая гипотеза. Сложной называют гипотезу, которая состоит из конечного или бесконечного множества простых гипотез. Сложная гипотеза H_0 о неравенстве $\lambda > 10$ состоит из бесконечного множества простых гипотез H_0 о равенстве $\lambda = b_i$, где b_i – любое число, большее 10.

Проверка гипотезы основывается на вычислении некоторой случайной величины – критерия, точное или приближенное распределение которого известно. Обозначим эту величину через z , ее значение является функцией от элементов выборки

$$z = z(x_1, x_2, \dots, x_n).$$

Процедура проверки гипотезы предписывает каждому значению критерия одно из двух решений – принять или отвергнуть гипотезу. Тем самым все выборочное пространство и соответственно множество значений критерия делятся на два непересекающихся подмножества S_0 и S_1 . Если значение критерия z попадает в область S_0 , то гипотеза принимается, а если в область S_1 , – гипотеза отклоняется. Множество S_0 называется областью принятия гипотезы или областью допустимых значений, а множество S_1 – областью отклонения гипотезы или критической областью. Выбор одной области однозначно определяет и другую область.

Принятие или отклонение гипотезы H_0 по случайной выборке соответствует истине с некоторой вероятностью и, соответственно, возможны два рода ошибок.

Ошибка первого рода возникает с вероятностью α тогда, когда отвергается верная гипотеза H_0 и принимается конкурирующая гипотеза H_1 .

Ошибка второго рода возникает с вероятностью β в том случае, когда принимается неверная гипотеза H_0 , в то время как справедлива конкурирующая гипотеза H_1 .

Доверительная вероятность – это вероятность не совершить ошибку первого рода и принять верную гипотезу H_0 .

Вероятность отвергнуть ложную гипотезу H_0 называется мощностью критерия. Следовательно, при проверке гипотезы возможны четыре варианта исходов, табл. 3.1.

Таблица 3.1.

Гипотеза H_0	Решение	Вероятность	Примечание
Верна	Принимается	$1-\alpha$	Доверительная вероятность
	Отвергается	α	Вероятность ошибки первого рода
Неверна	Принимается	β	Вероятность ошибки второго рода
	Отвергается	$1-\beta$	Мощность критерия

Целесообразно полагать одинаковыми значения вероятности выхода параметра θ^* за нижний и верхний пределы интервала. Суммарная вероятность того, что параметр θ^* выйдет за пределы интервала с границами $\theta^*_{1-\alpha/2}$ и $\theta^*_{\alpha/2}$, составляет величину α . Эту величину следует выбрать настолько малой, чтобы выход за пределы интервала был маловероятен. Если оценка параметра попала в заданный интервал, то в таком случае нет оснований подвергать сомнению проверяемую гипотезу, следовательно, гипотезу равенства $\theta^* = \theta$ можно принять. Но если после получения выборки окажется, что оценка выходит за установленные пределы, то в этом случае есть серьезные основания отвергнуть гипотезу H_0 . Отсюда следует, что вероятность допустить ошибку первого рода равна α (равна уровню значимости критерия).

При заданном объеме выборки вероятность совершения ошибки первого рода можно уменьшить, снижая уровень значимости α . Однако при этом увеличивается вероятность ошибки второго рода β (снижается мощность критерия).

Единственный способ уменьшить обе вероятности состоит в увеличении объема выборки (плотность распределения оценки параметра при этом становится более "узкой"). При выборе критической области руководствуются правилом Неймана – Пирсона: следует так выбирать критическую область, чтобы вероятность α была мала, если гипотеза верна, и велика в противном случае. Однако выбор конкретного значения α относительно произволен. Употребительные значения лежат в пределах от 0,001 до 0,2. В целях упрощения ручных расчетов составлены таблицы интервалов с границами $\theta^*_{1-\alpha/2}$ и $\theta^*_{\alpha/2}$ для типовых значений α и различных способов построения критерия.

В зависимости от сущности проверяемой гипотезы и используемых мер расхождения оценки характеристики от ее теоретического значения применяют различные критерии. К числу наиболее

часто применяемых критериев для проверки гипотез о законах распределения относят критерии хи-квадрат Пирсона, Колмогорова, Мизеса, Вилкоксона, о значениях параметров – критерии Фишера, Стьюдента.

12.2. Проверка гипотез о законе распределения

Обычно сущность проверки гипотезы о законе распределения ЭД заключается в следующем. Имеется выборка ЭД фиксированного объема, выбран или известен вид закона распределения генеральной совокупности. Необходимо оценить по этой выборке параметры закона, определить степень согласованности ЭД и выбранного закона распределения, в котором параметры заменены их оценками. Пока не будем касаться способов нахождения оценок параметров распределения, а рассмотрим только вопрос проверки согласованности распределений с использованием наиболее употребительных критериев.

12.3.1. Критерий хи-квадрат К. Пирсона

Использование этого критерия основано на применении такой меры (статистики) расхождения между теоретическим $F(x)$ и эмпирическим распределением $F_n^*(x)$, которая приближенно подчиняется закону распределения χ^2 . Гипотеза H_0 о согласованности распределений проверяется путем анализа распределения этой статистики. Применение критерия требует построения статистического ряда.

Пусть выборка представлена статистическим рядом с количеством разрядов M . Наблюдаемая частота попаданий в i -й разряд n_i . В соответствии с теоретическим законом распределения ожидаемая частота попаданий в i -й разряд составляет F_i . Разность между наблюдаемой и ожидаемой частотой составит величину $(n_i - F_i)$. Для нахождения общей степени расхождения между $F(x)$ и $F_n^*(x)$ необходимо подсчитать взвешенную сумму квадратов разностей по всем разрядам статистического ряда

$$\chi^2 = \sum_{i=1}^M \frac{(n_i - F_i)^2}{F_i}. \quad (3.1)$$

Величина χ^2 при неограниченном увеличении n имеет χ^2 -распределение (асимптотически распределена как χ^2). Это распределение зависит от числа степеней свободы k , т.е. количества независимых значений слагаемых в выражении (3.1). Число степеней свободы равно числу u минус число линейных связей, наложенных на выборку. Одна связь существует в силу того, что любая частота может быть вычислена по совокупности частот в оставшихся $M-1$ разрядах. Кроме того, если параметры распределения неизвестны заранее, то имеется еще одно ограничение, обусловленное подгонкой распределения к выборке. Если по выборке определяются S параметров распределения, то число степеней свободы составит $k = M - S - 1$.

Область принятия гипотезы H_0 определяется условием $\chi^2 < \chi^2(k; a)$, где $\chi^2(k; a)$ – критическая точка χ^2 -распределения с уровнем значимости a . Вероятность ошибки первого рода равна a , вероятность ошибки второго рода четко определить нельзя, потому что существует бесконечно большое множество различных способов несовпадения распределений. Критерий рекомендуется применять при $n > 100$, допускается применение при $n > 40$, именно при таких условиях критерий состоятелен (как правило, отвергает неверную нулевую гипотезу).

Алгоритм проверки по критерию

1. Построить гистограмму равновероятностным способом.
2. По виду гистограммы выдвинуть гипотезу

$$H_0 : f(x) = f_0(x),$$

$$H_1 : f(x) \neq f_0(x),$$

где $f_0(x)$ – плотность вероятности гипотетического закона распределения.

3. Вычислить значение критерия по формуле

$$\chi^2 = n \sum_{i=1}^M \frac{(p_i - p_i^*)^2}{p_i} = \sum_{i=1}^M \frac{(v_i - np_i)^2}{np_i},$$

где $p_i^* = v_i / n$ – частота попадания в i -тый интервал;

p_i – теоретическая вероятность попадания случайной величины в i -тый интервал при условии, что гипотеза H_0 верна.

Замечания. После вычисления всех вероятностей p_i проверить, выполняется ли контрольное соотношение

$$\left| 1 - \sum_{i=1}^M p_i \right| \leq 0,01.$$

4. Из таблицы "Хи-квадрат" Приложения выбирается значение $\chi^2_{\alpha, k}$, где α – заданный уровень значимости, а k – число степеней свободы, определяемое по формуле

$$k = M - 1 - S.$$

Здесь S – число параметров теоретического распределения, значения которых были получены из экспериментальных данных.

5. Если $\chi^2 > \chi^2_{\alpha, k}$, то гипотеза H_0 отклоняется. В противном случае нет оснований ее отклонить: с вероятностью 1 – она верна, а с вероятностью α – неверна, но величина неизвестна.

Пример 3.1 По выборке СВ объемом $n=200$ и построена гистограмма равновероятностным способом. Проверить гипотезу о соответствии выборке закону, имеющему следующую функцию распределения:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{2}{\pi} * \arcsin(x), & x \in [0,1] \\ 1, & x > 1 \end{cases}$$

Доверительная вероятность равна 0.01.

i	A_i	B_i	h_i	v_i	f_i^*

0	0	0.15	0.15	20	0.64
1	0.15	0.33	0.18	20	0.54
2	0.33	0.48	0.15	20	0.66
3	0.48	0.60	0.11	20	0.90
4	0.60	0.69	0.09	20	1.08
5	0.69	0.76	0.07	20	1.36
6	0.76	0.86	0.09	20	1.05
7	0.86	0.92	0.06	20	1.54
8	0.92	0.98	0.05	20	1.80
9	0.98	1	0.02	20	5.11

Гистограмма и кривая теоретического распределения приведены на рис. 3.1.

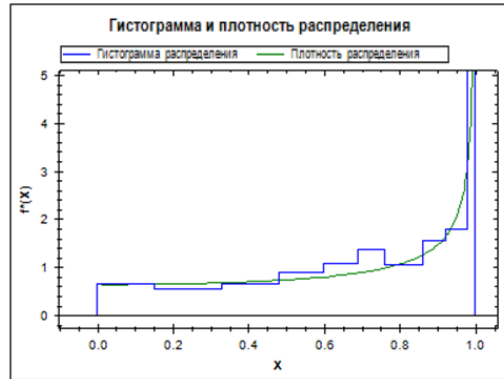


Рис. 3.1. Гистограмма и кривая теоретического распределения

Рассчитаем теоретическую вероятность попадания случайной величины в i -тый интервал при условии, что гипотеза верна.

$$p_i = F(B_i) - F(A_i) = \frac{2}{\pi} * (\arcsin(B_i) - \arcsin(A_i))$$

Вычислим значение критерия по формуле

$$\chi^2 = n \sum_{i=1}^M \frac{(p_i - p_i^*)^2}{p_i} = \sum_{i=1}^M \frac{(v_i - np_i)^2}{np_i}$$

Результаты промежуточных вычислений представлены в таблице:

	$F(A_i)$	$F(B_i)$	p_i	p_i^*	$n * \frac{(p_i - p_i^*)^2}{p_i}$
1	0	0.095903	0.095903	0.1	0.034999
2	0.095903	0.214206	0.118303	0.1	0.566327
3	0.214206	0.318888	0.104682	0.1	0.041886
4	0.318888	0.409873	0.090985	0.1	0.178647
5	0.409873	0.485025	0.075152	0.1	1.6432
6	0.485025	0.549881	0.064856	0.1	3.808723
7	0.549881	0.659407	0.109526	0.1	0.16572
8	0.659407	0.744	0.084593	0.1	0.561228
9	0.744	0.872905	0.128905	0.1	1.29631
10	0.872905	1.000507	0.127602	0.1	1.194121

Полученное значение $\chi^2 = 5,36$

Количество степеней свободы в нашем примере равно:

$$k = M - 1 - S = 10 - 1 - 0 = 9$$

Значение $S = 0$, так как закон распределения не зависит ни от каких параметров. Из таблицы "

Хи-квадрат" Приложения выберем значение $\chi_{\alpha, k}^2$, где
 $= 0.01$

$$\chi_{\alpha, k}^2 = 21,07$$

$\chi^2 < \chi_{\alpha, k}^2$ следовательно, нет оснований отклонять выдвинутую гипотезу.

12.3.2. Критерий А.Н. Колмогорова

Для применения критерия А.Н. Колмогорова ЭД требуется представить в виде вариационного ряда (ЭД недопустимо объединять в разряды). В качестве меры расхождения между теоретической $F(x)$ и эмпирической $F_n^*(x)$ функциями распределения непрерывной случайной величины X используется модуль максимальной разности

$$d_n = \max |F(x) - F_n(x)|$$

А.Н. Колмогоров доказал, что какова бы ни была функция распределения $F(x)$ величины X при неограниченном увеличении количества наблюдений n функция распределения случайной величины d_n асимптотически приближается к функции распределения

$$K(\lambda) = P(d\sqrt{n} > \lambda) = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2\lambda^2 k^2).$$

Иначе говоря, критерий А.Н. Колмогорова характеризует вероятность того, что величина d_n не будет превосходить параметр l для любой теоретической функции распределения. Уровень значимости α выбирается из условия

$$P(d\sqrt{n} > \lambda) = \alpha = 1 - K(\lambda)$$

в силу предположения, что почти невозможно получить это равенство, когда существует соответствие между функциями $F(x)$ и $F_n^*(x)$. Критерий А.Н. Колмогорова позволяет проверить согласованность распределений по малым выборкам, он проще критерия хи-квадрат, поэтому его часто применяют на практике. Но требуется учитывать два обстоятельства.

1. В соответствии с условиями его применения необходимо пользоваться следующим соотношением

$$d = \max(d_n^+, d_n^-)$$

где

$$d_n^+ = \max \left| \frac{i}{n} - F(x_i) \right|; \quad d_n^- = \max \left| F(x_i) - \frac{i-1}{n} \right|$$

2. Условия применения критерия предусматривают, что теоретическая функция распределения известна полностью – известны вид функции и значения ее параметров.

Последовательность действий при проверке гипотезы следующая.

1. Построить вариационный ряд.
2. Построить график эмпирической функции распределения $F^*(x)$.
3. Выдвинуть гипотезу:

$$H_0 : F(x) = F_0(x),$$

$$H_1 : F(x) \neq F_0(x),$$

где $F_0(x)$ – теоретическая функция распределения.

4. Построить график функции $F_0(x)$ в одной системе координат с функцией $F_n^*(x)$.
5. Определить максимальное по модулю отклонение между функциями $F_n^*(x)$ и $F_0(x)$.
6. Вычислить значение критерия

$$\lambda = \sqrt{n} \cdot \max |F^*(x) - F_0(x)|.$$

7. Принимают тот или иной уровень значимости (чаще всего 0,05 или 0,01). Тогда доверительная вероятность $= 1 - \alpha$.

8. Из таблицы вероятностей Колмогорова выбрать критическое значение.

9. Если $\lambda > \lambda_{\alpha}$, то нулевая гипотеза H_0 отклоняется, в противном случае – принимается, хотя она может быть неверна.

Достоинства критерия Колмогорова по сравнению с критерием 2: возможность применения при очень маленьких объемах выборки ($n < 20$).

Недостаток: критерий можно использовать в том случае, если параметры Q_1, \dots, Q_k распределения заранее известны, а эмпирическая функция распределения $F^*(x)$ должна быть построена по *несгруппированным* выборочным данным.

Пример 3.2. Дана выборка СВ объемом $n=30$. Проверить гипотезу о соответствии выборке закону, имеющему следующую функцию распределения:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{2}{\pi} * \arcsin(x), & x \in [0,1] \\ 1, & x > 1 \end{cases}$$

Построим вариационный ряд

$X_i = [0.04; 0.05; 0.06; 0.08; 0.09; 0.11; 0.18; 0.22; 0.26; 0.28; 0.34; 0.42; 0.49; 0.51; 0.56; 0.58; 0.62; 0.64; 0.65; 0.66; 0.68; 0.73; 0.74; 0.75; 0.80; 0.85; 0.87; 0.87; 1.00; 1.00]$

и эмпирическую функцию распределения

$F_0(X_i) = [0.03; 0.07; 0.1; 0.13; 0.17; 0.2; 0.23; 0.27; 0.3; 0.33; 0.37; 0.4; 0.43; 0.47; 0.5; 0.53; 0.57; 0.6; 0.63; 0.67; 0.7; 0.73; 0.77; 0.8; 0.83; 0.87; 0.9; 0.93; 0.97; 1.1;]$

Графики эмпирической и теоретической функций распределения приведены на рис. 3.2.

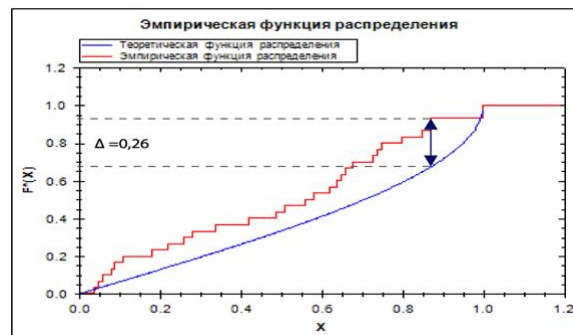


Рис. 3.2. Графики эмпирической и теоретической функций распределения

Максимальная разность по модулю между функцией $F_0(X_i)$ и $F(X)$ равна 0,14 при $x = 0.87$. Вычислим значение статистики

$$\lambda = \sqrt{n} \max |F(x) - F_0(x)| = 0.8$$

Из таблицы функции Колмогорова выберем критическое значение

$$\lambda_{\alpha} = 1.63.$$

Поскольку $\lambda < \lambda_{\alpha}$ следовательно у нас нет оснований отклонять выдвинутую гипотезу.

12.3.3. Критерий Р. Мизеса

В качестве меры различия теоретической функции распределения $F(x)$ и эмпирической $F_n^*(x)$ по критерию Мизеса (критерию w^2) выступает средний квадрат отклонений по всем значениям аргумента x

$$\sigma_n^2 = \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 dF(x) \quad (3.4)$$

Статистика критерия

$$n\omega_n^2 = \frac{1}{12n} + \sum_{j=1}^n \left(F(x_j) - \frac{j-0.5}{n} \right)^2 \quad (3.5)$$

При неограниченном увеличении n существует предельное распределение статистики $n\omega_n^2$. Задав значения вероятности a можно определить критические значения $n\omega_n^2(a)$. Проверка гипотезы о законе распределения осуществляется обычным образом: если фактическое значение $n\omega_n^2$ окажется больше критического или равно ему, то согласно критерию Мизеса с уровнем значимости a гипотеза H_0 о том, что закон распределения генеральной совокупности соответствует $F(x)$, должна быть отвергнута.

Пример 3.3. Дана выборка СВ объемом $n=30$. Проверить гипотезу о соответствии выборке закону, имеющему следующую функцию распределения:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{2}{\pi} * \arcsin(x), & x \in [0,1] \\ 1, & x > 1 \end{cases}$$

Исходные данные и результаты вычислений представлены в таблице, где приняты следующие обозначения:

$F_n(x_i) = (i-0.5)/30$ – значение эмпирической функции распределения;

$F(x_i)$ – значение теоретической функции распределения, соответствует значению функции нормального распределения в точке x_i ;

$$D_i = [F_n(x_i) - F(x_i)]^2$$

Таблица

i	x_i	$F_n(x_i)$	$F(x_i)$	D_i
1	0.023	0.017	0.015	0.001
2	0.092	0.05	0.059	0.002
3	0.218	0.083	0.140	0.008
4	0.358	0.117	0.233	0.022
5	0.413	0.15	0.271	0.024
6	0.458	0.183	0.303	0.023
7	0.481	0.217	0.320	0.019
8	0.507	0.25	0.339	0.015
9	0.518	0.283	0.347	0.009
10	0.526	0.317	0.353	0.005
11	0.569	0.35	0.385	0.033
12	0.576	0.383	0.391	0.002
13	0.611	0.417	0.419	0.001
14	0.636	0.45	0.439	0.0
15	0.673	0.483	0.470	0.0
16	0.751	0.517	0.541	0.003
17	0.754	0.55	0.744	0.001
18	0.804	0.583	0.595	0.002
19	0.853	0.617	0.650	0.004
20	0.863	0.65	0.662	0.002
21	0.864	0.683	0.664	0.0
22	0.898	0.717	0.710	0.001
23	0.901	0.75	0.714	0.0
24	0.937	0.783	0.773	0.001
25	0.948	0.817	0.794	0.0
26	0.973	0.85	0.852	0.001
27	0.994	0.883	0.930	0.006
28	0.997	0.917	0.954	0.005
29	1.0	0.95	0.987	0.005
30	1.0	0.983	0.987	0.002

Критическое значение статистики критерия Мизеса при заданном уровне значимости равно 0,744. Фактическое значение статистики

$$n\omega_n^2 = \frac{1}{12 \cdot 30} + \sum_{i=1}^{30} \Delta_i = 0,083$$

что меньше критического значения. Следовательно, гипотеза H_0 не противоречит имеющимся данным.