Проверка статистических гипотез

1. Формулировка задачи. Термины и определения.
2. Схема проверки статистической гипотезы.
3. Мощность критерия.
4. Проверка гипотез о распределении случайных величин. Критерий хи-квадрат.

Проверка статистических гипотез

- ***Статистической гипотезой*** называется всякое непротиворечивое множество утверждений $\{H_0, H_1, \ldots, H_{k-1}\}$ относительно свойств распределения случайной величины.

- Любое из утверждений H_i называется альтернативой гипотезы.

- Простейшей гипотезой является двухальтернативная: $\{H_0, H_1\}$. В этом случае альтернативу H_0 называют нулевой гипотезой, а H_1- конкурирующей гипотезой.

Проверка статистических гипотез

- ***Критерием*** (статистикой) называется случайная величина

$$U=f(x_1, x_2, \dots, x_n),$$

где x_i – значения выборки

- При некторых значениях U гипотеза H_0 отвергается (эти значения образуют ***критическую область проверяемой гипотезы***; значения критерия, при которых гипотезу принимают, образуют область принятия гипотезы - ***область допустимых значений***).

- ***Критические точки*** отделяют критическую область от области принятия гипотезы.

Проверка статистических гипотез

- ***Ошибка I рода***: отклонить верную гипотезу H_0 («пропуск цели»)

- ***Ошибка II рода***: принять ложную гипотезу H_0 («ложное срабатывание»)

- ***Вероятность*** совершить ошибку первого рода – ***уровень значимости*** α (обычно принимают 5% или 1%). ***Вероятность*** совершить ошибку второго рода β.

- ***Вероятность*** не допустить ошибку второго рода $(1-\beta)$ – ***мощность критерия.***

Критерии согласия:

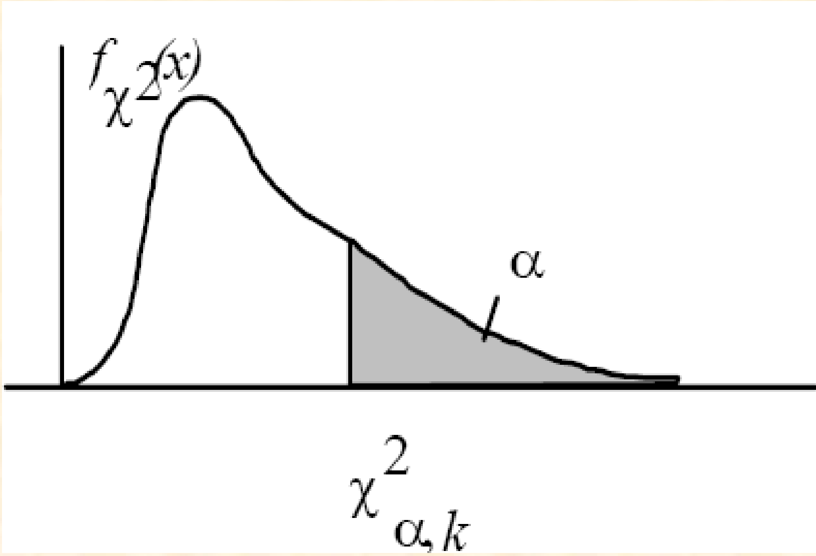
используются для проверки гипотез о предполагаемом законе распределения случайной величины

Распределение хи-квадрат

Пусть X_1, X_2, X_n - независимые случайные нормально распределенные величины с нулевым средним и единичной дисперсией.

Случайная величина $\chi^2 = \sum_{i=1}^{n} X_i^2$

имеет распределение хи-квадрат.



Критерий согласия Пирсона (хи-квадрат).

1. Задать уровень значимости.
2. Построить интервальный статистический ряд. Рассчитать частоты попадания случайной величины в каждый из ***M*** бинов. Число наблюдений в одном бине не должно быть меньше 4-5. Построить гистограмму.
3. По виду гистограммы выдвинуть гипотезу:

H_0 – величина X распределена по такому-то закону $f(x) = f_0(x)$,

H_1 – величина X не распределена по такому-то закону $f(x) \neq f_0(x)$,

$f_0(x)$, $F_0(x)$ – плотность и функция гипотетического закона распределения.

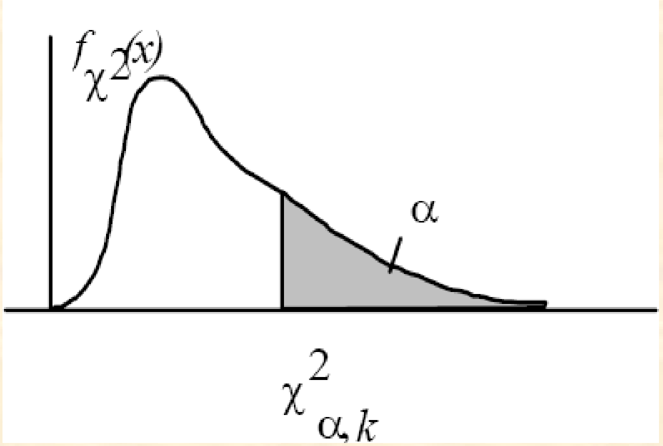
4. Найти оценки ***s*** неизвестных параметров гипотетического закона распределения.

5. Рассчитать теоретические вероятностей попадания случайной величины в каждый из бинов (p_j).

6. Найти значение статистики хи-квадрат

$$\chi^2 = n\sum_{j=1}^{M}\frac{\left(p_j - p_j^*\right)^2}{p_j} = \sum_{j=1}^{M}\frac{\left(\nu_j - np_j\right)^2}{np_j}$$

7. Сравнить значение статистики с критическим значением $\chi^2_{\alpha,k}$



Число степеней свободы $k = M - 1 - s$.

8. Принять или отвергнуть гипотезу H_0.

Критерий согласия Колмогорова. Алгоритм проверки следующий:

1. Построить вариационный ряд и график эмпирической функции распределения $F^*(x)$.

2. По виду графика $F^*(x)$ выдвинуть гипотезу:

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

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

где $F_0(x)$ – функция гипотетического закона распределения.

3. Используя метод моментов или максимального правдоподобия определить оценки неизвестных параметров $\hat{Q}_1, \; ..., \; \hat{Q}_m$ гипотетического закона распределения.

4. Рассчитать 10...20 значений функции $F_0(x)$ и построить ее график в одной системе координат с функцией $F^*(x)$.

5. По графику определить максимальное по модулю отклонение между функциями $F^*(x)$ и $F_0(x)$.

$$Z = \max_{i=1}^{n} \left| F^*(x_i) - F_0(x_i) \right|.$$

6. Вычислить значение критерия Колмогорова

$$\lambda = \sqrt{n} \cdot Z \, .$$

7. Из таблицы распределения Колмогорова выбрать критическое значение λ_{γ}, $\gamma = 1 - \alpha$. α – заданный уровень значимости ($\alpha = 0{,}05$ или $\alpha = 0{,}01$).

8. Если $\lambda > \lambda_{\gamma}$, то нулевая гипотеза H_0 отклоняется, в противном случае нет оснований ее отклонить.