# Проверка гипотез. Критерий Пирсона

Занятие 10

# Статистическая гипотеза

\- предположение о виде законов распределения случайных величин или о соотношении между их числовыми характеристиками.

Приняв ту или иную гипотезу, из нее выводят определенное следствие и рассматривают, насколько оно оправдывается на опыте (проверяют согласие принятой гипотезы с опытом).

# Критерий Пирсона – критерий $\chi^2$

Применяется при проверке гипотезы о виде закона распределения случайных величин. Он позволяет производить проверку гипотезы соответствия **опытного** *(практического)* закона распределения **теоретическому** *(предполагаемому)* не только в случаях, когда последний известен полностью, но и тогда, когда параметры предполагаемого закона распределения определяются на основании опытных данных.

# Предположение

Допустим произведено $n$ независимых опытов, в каждом из которых случайная величина $X$ приняла определенное значение. Результаты опытов сведены в $k$ разрядов и оформлены в виде таблицы

| | | | | |
|---|---|---|---|---|
| $I_i$ | | | … | |
| | | | … | |
| | | | … | |

Мера расхождения между теоретическим и практическим законом распределения определяется по формуле

$$\chi^2_{эксп} = \sum_{i=1}^{k} \frac{\left(n_i - np_i^*\right)^2}{np_i^*}$$

$n_i$ - частота случайной величины, полученная в ходе эксперимента

$n$ - объем выборки

$$p_i^* = \frac{n_i^*}{n}$$

- частость, вычисленная в предположении известного распределения (теоретическая)

# Критерий Пирсона

Распределение $\chi^2$ зависит от параметра (*число степеней свободы)* распределения

$$r = k - s$$

$k$ – число разрядов , минус число $s$ независимых условий (связей), наложенных на частоты

# Примеры связей

1. $\sum_{i=1}^{k} p_i^* = 1$, сумма всех частот (накладывается во всех случаях)

2. $\sum_{i=1}^{k} \widetilde{x}_i \, p_i^* = \overline{x}$, теоретическое распределение подбирается т.о., чтобы совпадали теоретическое и статистическое средние значения

3. $\sum_{i=1}^{k} \left(\widetilde{x}_i - \overline{x}^*\right)^2 p_i^* = \sigma^2$, совпадения теоретической и статистической дисперсий

# Схема применения критерия Пирсона

1.Исходя из теоретического закона распределения, находят вероятности попадания случайной величины в каждый из заданных $k$ интервалов таблицы

2. Вычисляют по формуле меру расхождения $\chi^2$

3. Определяют число степеней свободы

# Схема применения критерия Пирсона

4. По значениям $r$, $\chi^2$ с помощью таблицы (прил.4) определяют вероятность того, что величина, имеющая распределение со степенями свободы $r$, превзойдет данное значение.

Если эта вероятность весьма мала, гипотеза отбрасывается как неправдоподобная. Если эта вероятность относительно велика, гипотезу можно признать не противоречащей опытным данным.

Вероятность $\alpha$, при которой гипотезу о виде распределения принимают или отбрасываю называется ***уровнем значимости критерия***, а соответствующая ей область больших отклонений - ***критической областью***.

# Пример

На экзамене по некоторому предмету экзаменатор задает студенту только один вопрос по одной из четырех частей курса. Из 100 студентов 26 получили вопрос по первой части, 32 - по второй, 17 - по третьей, остальные - по четвертой. Можно ли по этим результатам принять гипотезу, что для пришедшего на экзамен имеется одинаковая вероятность получить вопрос по любой из четырех частей? Уровень значимости равен 0,05

# Решение

По условию задачи имеем

| $n_i$ | 26 | 32 | 17 | 25 |
|---|---|---|---|---|
| $n_i^*$ | 25 | 25 | 25 | 25 |

Вычисляем меру расхождения

$$\chi^2_{набл} = \frac{(26-25)^2}{25} + \frac{(32-25)^2}{25} + \frac{(17-25)^2}{25} + \frac{(25-25)^2}{25} = 4,56$$

По таблице определяем границу критической области

$$\chi^2_{1-\alpha}(k-1) = \chi^2_{0.95}(3) = 7.815$$

# Ответ

Так как вычисленное значение меры расхождения меньше границы критической области **(4,56 < 7,815)** , то гипотеза о равномерном распределении подтверждается.

<u>**ЗАМЕЧАНИЕ**</u> При изменении уровня значимости, меняется граница области больших отклонений (критической области).

# СПАСИБО ЗА ВНИМАНИЕ