26. Элементы математической статистики-4

26.1. Предварительные сведения

Пусть $H_0 : \theta = \theta_0$ — *основная гипотеза*, $H_1 : \theta = \theta_1$ — *альтернативная гипотеза.*

S*-критерий* проверки H_0 против H_1 состоит в нахождении такого критического множества S, что при попадании выборочных значений в это множество H_0 отвергается, а H_1 — принимается.

φ*-критерий* проверки H_0 против H_1 состоит в нахождении по выборочным данным x значения функции $\varphi(x)$ и принятии гипотезы H_1 с вероятностью $\varphi(x)$ и гипотезы H_0 с вероятностью $1 - \varphi(x)$.

Если $\varphi(x) = \mathbf{1}(x \in S)$, то φ-критерий превращается в соответствующий S-критерий.

Обозначим $p_0(x) = p(x; \theta_0)$, $p_1(x) = p(x; \theta_1)$, $\mathbb{E}_0\varphi = \int\limits_\Omega \varphi(x)p_0(x)\,dx$, $\mathbb{E}_1\varphi = \int\limits_\Omega \varphi(x)p_1(x)\,dx$.

Значение $\mathbb{E}_0\varphi$ есть *вероятность ошибки I рода* (принятие H_1, когда на самом деле верна H_0). Значение $1 - \mathbb{E}_1\varphi$ есть *вероятность ошибки II рода* (принятие H_0, когда на самом деле верна H_1).

Критерий имеет *уровень значимости* α, если вероятность ошибки I рода равна α. *Оптимальный критерий* уровня значимости α обладает минимальной вероятностью ошибки II рода.

Теорема 1 (Нейман-Пирсон). *Для любого* $0 < \alpha < 1$ *существуют такие числа* $c \geq 0$ *и* $0 \leq \varepsilon \leq 1$, *что* φ^**-критерий с функцией*

$$\varphi^*(x) = \begin{cases} 1, & p_1(x) > cp_0(x), \\ \varepsilon, & p_1(x) = cp_0(x), \\ 0, & p_1(x) < cp_0(x), \end{cases}$$

определяет оптимальный критерий с уровнем значимости α.

26.2. Практическое занятие

1. Генеральная совокупность имеет нормальное распределение с известной дисперсией. Предложите способ проверки гипотезы $H_0 : \ a = a_0$ против гипотезы $H_1 : \ a \neq a_0$, основанный на построении критического множества, при заданном уровне значимости α по выборочным данным $x_1, \ldots, x_n$. Что изменится, если альтернативная гипотеза имеет вид $H_1 : \ a > a_0$? Что изменится, если дисперсия неизвестна?

Ответ. В первом случае принимается H_0, если $\bar{x} \in (a_0 - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, a_0 + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}})$, иначе принимается H_1. Во втром случае принимается H_0, если $\bar{x} < a_0 + u_{1-\alpha} \frac{\sigma}{\sqrt{n}}$, иначе принимается H_1. Здесь u_γ — квантиль нормального распределения уровня γ. Если дисперсия неизвестна, то нужно вместо σ использовать выборочное отклонение s, а вместо квантилей нормального распределения — квантили распределения Стьюдента.

2. Пусть X — число изюминок, содержащихся в булочке. Как следует из ранее решенной задачи (см. тему «Схема Бернулли-2»), распределение X подчиняется закону Пуассона, т.е. $\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ с некоторым $\lambda > 0$. Требуется проверить основную гипотезу $H_0 : \lambda = 4$ против альтернативной гипотезы $H_1 : \lambda = 2$ на уровне значимости $\alpha = 0.3$. Постройте соответствующий оптимальный критерий Неймана-Пирсона, найдите вероятность ошибки II рода и четко сформулируйте алгоритм проверки гипотез, основанный на этом критерии.

Ответ. $\varphi(x) = \begin{cases} 1, & x < 3, \\ 0.317, & x = 3, \\ 0, & x > 3. \end{cases}$ Вероятность ошибки II рода равна 0.266.

26.3. Домашнее задание

3. Для проверки того, повышается ли скорость чтения учеников младших классов в зависимости от содержания текста, учитель проверил скорость чтения 7 учеников, когда они читали текст по своему выбору и текст, предложенный учителем. Число слов, прочитанных за 6 минут каждым из учеников, дано в таблице.

Ученик	1	2	3	4	5	6	7
Свой текст	549	555	560	521	544	552	542
Текст учителя	534	549	555	525	542	540	534

При уровне значимости 5% проверьте гипотезу о том, что ученики читают быстрее, если они сами выбирают текст.
Ответ. Гипотеза принимается.
4. По мишени, вероятность попадания в которую равна p, производятся выстрелы до первого попадания. Пусть X — число сделанных выстрелов. Постройте оптимальный критерий Неймана-Пирсона для проверки нулевой гипотезы $H_0 : p = \frac{1}{3}$ против альтернативной гипотезы $H_1 : p = \frac{1}{2}$ для уровня значимости $\alpha = 0.3$, найдите вероятность ошибки II рода и четко сформулируйте алгоритм проверки гипотез, основанный на этом критерии.
Ответ. $\varphi(x) = \begin{cases} 0.9, & x = 1, \\ 0, & x = 2, 3, \ldots \end{cases}$ Вероятность ошибки II рода равна 0.55.