4.Проверка гипотез. Критерий Неймана -Пирсона

При построении *критерия* для проверки *статистических гипотез*, как правило, исходят из необходимости максимизации его *мощности* $1-\beta$ (минимизации вероятности совершения *ошибки второго рода*) при фиксированном *уровне значимости* α *критерия* (вероятности совершения *ошибки первого рода*). Для упрощения дальнейших рассуждений будем считать, что $\vec{X}_n$ — *случайная выборка объема* n из *генеральной совокупности* непрерывной случайной величины X, плотность распределения вероятностей которой $p(t;\theta)$ зависит от неизвестного

параметра θ, и рассмотрим две *простые гипотезы* H_0: $\theta=\theta_0$ и H_1: $\theta=\theta_1$.

Введем функцию случайной выборки $\vec{X}_n$:

$$\varphi(\vec{X}_n)=\frac{L(\vec{X}_n;\theta_1)}{L(\vec{X}_n;\theta_0)}, \qquad L(\vec{X}_n;\theta)=\prod_{i=1}^{n}p(X_i;\theta).$$

Статистика $\varphi(\vec{X}_n)$ представляет собой отношение *функций правдоподобия* при истинности *альтернативной* и *основной гипотез* соответственно. Ее называют ***отношением правдоподобия.*** Для построения ***оптимального**** (наиболее мощного) при заданном уровне значимости α ***критерия Неймана — Пирсона*** в *критическое множество* W включают те элементы $\vec{x}_n$ *выборочного пространства* $\mathcal{X}_n$ случайной выборки $\vec{X}_n$, для которых выполняется неравенство

$$\varphi(\vec{x}_n)\geqslant C_\varphi,$$

где константу C_φ выбирают из условия

$$\mathbf{P}\{\varphi(\vec{X}_n)\geqslant C_\varphi \mid H_0\}=\alpha,$$

которое обеспечивает заданное значение уровня значимости α и может быть записано в виде

$$\int\limits_{\varphi(t_1,\ldots,t_n)\geqslant C_\varphi}\!\!\!\cdots\!\!\int L(t_1,\ldots,t_n;\theta_0)\,dt_1\ldots dt_n=\alpha.$$

При этом вероятность ошибки второго рода не может быть уменьшена при данном значении вероятности ошибки первого рода α.

Рассмотрим примеры построения оптимального критерия Неймана — Пирсона при проверке простых гипотез относительно параметров основных, наиболее часто используемых распределений.

Пример 4.4. Построение оптимального критерия Неймана — Пирсона для параметра μ нормального закона распределения с известной дисперсией σ^2 проведем для случая двух простых гипотез

$$H_0\colon\ \mu = \mu_0, \qquad H_1\colon\ \mu = \mu_1,$$

где μ_0 и μ_1 — некоторые заданные значения, связанные неравенством $\mu_0 < \mu_1$.

В рассматриваемом случае функция правдоподобия имеет вид

$$L(X_1, \ldots, X_n; \mu) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2\right),$$

а отношение правдоподобия —

$$\varphi(\vec{X}_n) = \frac{L(X_1, \ldots, X_n; \mu_1)}{L(X_1, \ldots, X_n; \mu_0)} = \\ = \exp\left(\frac{\mu_1 - \mu_0}{\sigma^2}\sum_{i=1}^{n} X_i\right) \exp\left(-\frac{n(\mu_1 - \mu_0)^2}{2\sigma^2}\right).$$

В данном случае неравенство

$$\varphi(\vec{x}_n) = \exp\left(\frac{\mu_1 - \mu_0}{\sigma^2}\sum_{i=1}^{n} x_i\right) \exp\left(-\frac{n(\mu_1 - \mu_0)^2}{2\sigma^2}\right) \geqslant C_\varphi$$

равносильно неравенству

$$\sum_{i=1}^{n} x_i \geqslant C, \tag{4.1}$$

где константу C выбирают из условия обеспечения заданного уровня значимости α:

$$\mathbf{P}\left\{\sum_{i=1}^{n} X_i \geqslant C \mid \mu = \mu_0\right\} = \alpha. \tag{4.2}$$

Действительно,

$$\ln\left(\exp\left(\frac{\mu_1-\mu_0}{\sigma^2}\sum_{i=1}^{n}x_i\right)\exp\left(-\frac{n(\mu_1-\mu_0)^2}{2\sigma^2}\right)\right)=$$
$$=\frac{\mu_1-\mu_0}{\sigma^2}\sum_{i=1}^{n}x_i-\frac{n(\mu_1-\mu_0)^2}{2\sigma^2}\geqslant\ln C_\varphi,$$

откуда следует, что

$$\sum_{i=1}^{n}x_i\geqslant\frac{\sigma^2}{\mu_1-\mu_0}\left(\ln C_\varphi-\frac{n(\mu_1-\mu_0)^2}{2\sigma^2}\right)=C.$$

Случайная величина $X_1+\ldots+X_n$ имеет нормальное распределение с математическим ожиданием $n\mu$ и дисперсией $n\sigma^2$ (см. **1.2**). Поэтому условие (4.2) можно записать в виде

$$1-\Phi\left(\frac{C-n\mu_0}{\sigma\sqrt{n}}\right)=\alpha, \tag{4.3}$$

или

$$\frac{C-n\mu_0}{\sigma\sqrt{n}}=u_{1-\alpha}.$$

Таким образом, константа C, задающая критическую область в (4.1), определяется равенством

$$C=n\mu_0+u_{1-\alpha}\sigma\sqrt{n}. \tag{4.4}$$

При этом вероятность совершения ошибки второго рода

$$\beta=\mathbf{P}\left\{\sum_{i=1}^{n}X_i<C\mid\mu=\mu_1\right\}=\Phi\left(\frac{C-n\mu_1}{\sigma\sqrt{n}}\right) \tag{4.5}$$

является минимально возможной при данном значении α.

Пример 4.6. Построение оптимального критерия Неймана — Пирсона в случае экспоненциального распределения с параметром λ проведем для двух простых гипотез

$$H_0\colon\ \lambda=\lambda_0,\qquad H_1\colon\ \lambda=\lambda_1,$$

где $\lambda_0<\lambda_1$. В этом случае функция правдоподобия

$$L(X_1,\ldots,X_n;\lambda)=\lambda^n\exp\Bigl(-\lambda\sum_{i=1}^{n}X_i\Bigr).$$

Таким образом,

$$\varphi(\vec{X}_n)=\Bigl(\frac{\lambda_1}{\lambda_0}\Bigr)^n\exp\Bigl(-(\lambda_1-\lambda_0)\sum_{i=1}^{n}X_i\Bigr).$$

Отсюда видно, что критическое множество можно задать неравенством

$$\sum_{i=1}^{n}x_i\leqslant C,$$

где константа C выбрана из условия обеспечения заданного уровня значимости α:

$$\mathbf{P}\Bigl\{\sum_{i=1}^{n}X_i\leqslant C\mid\lambda=\lambda_0\Bigr\}=\alpha.$$

Случайная величина $2\lambda(X_1+\ldots+X_n)$ при $\lambda=\lambda_0$ имеет χ^2-*распределение* с $2n$ степенями свободы (см. **Д.3.1**). Исходя из этого, получаем выражение для константы C:

$$C=\frac{\chi^2_\alpha(2n)}{2}\lambda_0,$$

где $\chi^2_\alpha(2n)$ — квантиль уровня α для χ^2-распределения с $2n$ степенями свободы. При этом вероятность совершения ошибки второго рода равна

$$\beta=\mathbf{P}\Bigl\{\sum_{i=1}^{n}X_i>C\mid\lambda=\lambda_1\Bigr\}=$$
$$=1-H_{2n}(2\lambda_1C)=1-H_{2n}\Bigl(\chi^2_\alpha(2n)\frac{\lambda_1}{\lambda_0}\Bigr),$$

где $H_{2n}(t)$ — функция распределения случайной величины, имеющей χ^2-распределение с $2n$ степенями свободы.

Пример 4.7. Построение оптимального критерия Неймана — Пирсона для параметра биномиального распределения проведем для случая двух простых гипотез

$$H_0:\ p=p_0,\qquad H_1:\ p=p_1,$$

где p — вероятность „успеха“ в одном испытании при реализации схемы независимых испытаний Бернулли, а p_0 и p_1 —

заданные значения параметра, удовлетворяющие неравенству $p_0 < p_1$.

Пусть объем испытаний достаточно велик и X_j — результат j-го испытания. Случайная величина X_j принимает значения 0 и 1 с вероятностями $1-p$ и p соответственно. Функция правдоподобия в этом случае имеет вид

$$L(X_1,\ldots,X_n;p) = C_n^{K(\vec{X}_n)} p^{K(\vec{X}_n)} (1-p)^{n-K(\vec{X}_n)},$$

где $K(\vec{X}_n) = X_1 + \ldots + X_n$ — общее число „успехов“ в серии из n испытаний. Отношение правдоподобия определяется равенством

$$\varphi(\vec{X}_n) = \frac{L(X_1,\ldots,X_n;p_1)}{L(X_1,\ldots,X_n;p_0)} = \left(\frac{p_1}{p_0}\right)^{K(\vec{X}_n)} \left(\frac{1-p_0}{1-p_1}\right)^{n-K(\vec{X}_n)}.$$

Значит, критическое множество для оптимального критерия Неймана — Пирсона в данном случае имеет вид

$$K(\vec{x}_n) = \sum_{i=1}^{n} x_i \geqslant C. \qquad (4.6)$$

Константу C выбирают исходя из условия

$$\mathbf{P}\{X_1 + \ldots + X_n \geqslant C \mid p = p_0\} = \alpha.$$

Распределение случайной величины $K(\vec{X}_n)$ при достаточно больших n в соответствии с известной интегральной теоремой Муавра — Лапласа имеет асимптотически нормальное распределение с математическим ожиданием $\mu = np$ и дисперсией $\sigma^2 = np(1-p)$. Используя указанное распределение, выберем константу C в (4.6) из условия обеспечения заданного уровня значимости α, т.е. из условия

$$\mathbf{P}\{K(\vec{X}_n) \geqslant C \mid p = p_0\} \approx 1 - \Phi\left(\frac{C - np_0}{\sqrt{np_0(1-p_0)}}\right) = \alpha, \qquad (4.7)$$

Пример 4.8. Определим объем выборки для случая нормальной модели.

Для ситуации, рассмотренной в примере 4.4, из выражений (4.3), (4.5) получаем, что система неравенств (4.9) в этом случае имеет вид

$$1-\Phi\left(\frac{C-n\mu_0}{\sigma\sqrt{n}}\right)\leqslant\alpha,\qquad \Phi\left(\frac{C-n\mu_1}{\sigma\sqrt{n}}\right)\leqslant\beta.$$

Следовательно, для обеспечения заданных значений α, β вероятностей совершения ошибок первого и второго рода минимально необходимый объем n^* выборки и соответствующую константу C^* можно определить из системы уравнений

$$1-\Phi\left(\frac{C-n\mu_0}{\sigma\sqrt{n}}\right)=\alpha,\qquad \Phi\left(\frac{C-n\mu_1}{\sigma\sqrt{n}}\right)=\beta.$$

Используя квантили стандартного нормального распределения, запишем эти уравнения в виде

$$\frac{C-n\mu_0}{\sigma\sqrt{n}}=u_{1-\alpha},\qquad \frac{C-n\mu_1}{\sigma\sqrt{n}}=u_\beta=-u_{1-\beta}. \tag{4.10}$$

Исключая из уравнений константу C, находим необходимый объем выборки

$$n^*=\frac{\sigma^2\bigl(u_{1-\alpha}+u_{1-\beta}\bigr)^2}{\bigl(\mu_1-\mu_0\bigr)^2}. \tag{4.11}$$

Пусть, например, требуется проверить гипотезы

$$H_0\colon\ \mu=\mu_0=3{,}5,\qquad H_1\colon\ \mu=\mu_1=3{,}8$$

при $\sigma=0{,}8$ и заданных значениях вероятностей $\alpha=0{,}05$, $\beta=0{,}1$. Применяя формулу (4.11) и учитывая, что $u_{1-\alpha}=u_{0,95}=1{,}64$,

$u_{1-\beta}=u_{0,9}=1{,}28$, получаем необходимый в этом случае объем выборки $n^*=61$.

Пример 4.9. Определим объем выборки для схемы испытаний Бернулли.

Для задачи проверки гипотез, рассмотренной в примере 4.7, вновь используем возможность аппроксимации биномиального распределения нормальным распределением с параметрами $\mu=np$ и $\sigma^2=np(1-p)$. После этого, согласно (4.7), (4.8), приходим к системе уравнений для определения n^* и C^*:

$$1-\Phi\left(\frac{C-np_0}{\sqrt{np_0(1-p_0)}}\right)=\alpha,\qquad \Phi\left(\frac{C-np_1}{\sqrt{np_1(1-p_1)}}\right)=\beta,$$

которая может быть представлена в следующем виде:

$$\frac{C - np_0}{\sqrt{np_0(1-p_0)}} = u_{1-\alpha}, \qquad \frac{C - np_1}{\sqrt{np_1(1-p_1)}} = u_\beta = -u_{1-\beta}.$$

Решая эту систему, находим

$$n^* = \frac{\left(u_{1-\alpha}\sqrt{p_0(1-p_0)} + u_{1-\beta}\sqrt{p_1(1-p_1)}\right)^2}{(p_1 - p_0)^2}. \tag{4.12}$$

Равенство (4.12) (приближенно) определяет минимально необходимый объем выборки, позволяющий обеспечить заданные значения вероятностей совершения ошибок первого и второго рода при проверке простых гипотез вида $H_0: p = p_0$, $H_1: p = p_1$ в схеме Бернулли. Поскольку в (4.12) величина n^* не обязательно целая, то на практике в качестве объема выборки берут наименьшее целое число, большее или равное n^*. #

Пример 4.19. Для *выборки объема* $n = 9$ построить *оптимальный критерий Неймана — Пирсона* для проверки двух *простых гипотез* относительно параметра μ нормального распределения

$$H_0\colon\ \mu = \mu_0 = 53, \qquad H_1\colon\ \mu = \mu_1 = 54$$

с заданным *уровнем значимости* (вероятностью *ошибки первого рода*) $\alpha = 0{,}1$ при известной дисперсии $\sigma^2 = 16$. Для построенного критерия найти вероятность *ошибки второго рода* β и *мощность критерия.*

Решение. В соответствии с результатами примера 4.4 *критическое множество* задается неравенством

$$\sum_{i=1}^{n} X_i \geqslant C, \tag{4.56}$$

где константа C выбирается из условия обеспечения заданного уровня $\alpha = 0{,}1$:

$$C = n\mu_0 + u_{1-\alpha}\sigma\sqrt{n} = 9 \cdot 53 + 1{,}28 \cdot 4 \cdot 3 = 492{,}4.$$

Для построенного критерия с критическим множеством (4.56) вероятность ошибки второго рода равна

$$\beta = \Phi\left(\frac{C - n\mu_1}{\sigma\sqrt{n}}\right) = \Phi\left(\frac{492{,}4 - 9 \cdot 54}{4 \cdot 3}\right) = 0{,}76.$$

Мощность критерия равна $1 - \beta = 0{,}24$. Значение мощности невелико, что объясняется в данном случае относительно малым объемом выборки $n=9$.

Пример 4.20. В предыдущей задаче найти минимально необходимый объем выборки n^*, позволяющий обеспечить заданные значения вероятностей ошибок $\alpha=0{,}1$, $\beta=0{,}1$. Построить соответствующий оптимальный критерий Неймана — Пирсона в этой ситуации.

Решение. В соответствии с результатами примера 4.8

$$n^* = \frac{\sigma^2(u_{1-\alpha}+u_{1-\beta})^2}{(\mu_1-\mu_0)^2} = \frac{16\cdot(1{,}28+1{,}28)^2}{1^2} = 105.$$

Оптимальный критерий Неймана — Пирсона в этом случае задается с помощью критического множества

$$\sum_{i=1}^{n} x_i \geqslant C,$$

где константа C определяется из равенства (4.4):

$$C = n^*\mu_0 + u_{1-\alpha}\sigma\sqrt{n^*} = 570{,}4.$$

Пример 4.21. Партии волокна испытываются на прочность, при этом предел прочности X распределен по нормальному закону с дисперсией $\sigma^2=9$. Партия считается удовлетворительной, если среднее значение предела прочности входящих в партию образцов $\mu=\mathbf{M}X\geqslant 14$, и неудовлетворительной, если $\mu\leqslant 10$. Из каждой партии на испытание ставится n образцов, для которых измеряют значения их прочности $x_1, \ldots, x_n$. Требуется проверить по результатам испытаний две *сложные гипотезы* H_0: $\mu\geqslant 14$, H_1: $\mu\leqslant 10$ с заданными максимальными вероятностями ошибок $\alpha=0{,}1$, $\beta=0{,}05$. Для этой ситуации необходимо решить следующие задачи:

– найти необходимый объем выборки n^*, при котором могут быть обеспечены данные значения α, β;

– построить *равномерно наиболее мощный критерий* при найденном объеме выборки;

– для построенного критерия найти *функцию мощности* и *оперативную характеристику*.

Решение. Для решения поставленных задач используем результаты **4.3**–**4.5**. Необходимый объем выборки находим по формуле (4.11):

$$n^* = \frac{\sigma^2(u_{1-\alpha} + u_{1-\beta})^2}{(\mu_1 - \mu_0)^2} = 5.$$

Равномерно наиболее мощный критерий совпадает с оптимальным критерием Неймана — Пирсона для двух простых гипотез H_0: $\mu = 14$, H_1: $\mu = 10$. Соответствующее критическое множество задается неравенством (4.16):

$$\sum_{i=1}^{n^*} x_i \leqslant C,$$

где $C = n^*\mu_0 - u_{1-\alpha}\sigma\sqrt{n^*}$, откуда $C = 61{,}4$.

Функция мощности (вероятность отвергнуть гипотезу H_0) в данном случае имеет вид

$$M(\mu) = \mathbf{P}\Big\{\sum_{i=1}^{n^*} X_i \leqslant C \mid \mu\Big\} = \Phi\Big(\frac{C - n^*\mu}{\sigma\sqrt{n^*}}\Big).$$

Оперативная характеристика критерия $S(\mu) = 1 - M(\mu)$.

Пример 4.22. В условиях примера 4.7 найдем минимальный объем выборки, если H_0: $p = p_0 = 0{,}1$, H_1: $p = p_1 = 0{,}2$, $\alpha = 0{,}01$ и $\beta = 0{,}05$.

Решение. По таблице квантилей нормального распределения (см. табл. П.2) находим $u_{1-\alpha} = u_{0,99} = 2{,}33$, $u_\beta = u_{0,05} = -u_{0,95} = -1{,}65$. Далее, используя (4.12), получаем

$$n^* \geqslant \frac{\big(2{,}33\sqrt{0{,}1\cdot 0{,}9} + 1{,}65\sqrt{0{,}2\cdot 0{,}8}\big)^2}{(0{,}2 - 0{,}1)^2} \approx 185.$$

Условия задач.

4.18. В соответствии с техническими условиями среднее время безотказной работы приборов из большой партии должно составлять не менее 1000 ч со средним квадратичным отклонением 100 ч. Значение выборочного среднего времени безотказной работы для случайно отобранных 25 приборов оказалось равным 970 ч. Предположим, что среднее квадратичное время безотказной работы для приборов в выборке совпадает со средним квадратичным во всей партии, а контролируемая характеристика имеет нормальное распределение. Выясните, можно ли считать, что вся партия приборов не удовлетворяет техническим условиям, если: а) $\alpha = 0{,}1$; б) $\alpha = 0{,}01$.

Ответ: а) да; б) нет.

4.19. Решите предыдущую задачу при условии, что среднее квадратичное отклонение времени безотказной работы, вычисленное по выборке, равно 115 ч.

Ответ: а) нет; б) да.

4.20. Утверждается, что шарики, изготовленные станком-автоматом, имеют средний диаметр $d_0 = 10$ мм. Используя односторонний критерий при $\alpha = 0{,}05$, проверьте эту гипотезу, если в выборке из $n = 16$ шариков средний диаметр оказался равным 10,3 мм, считая, что: а) дисперсия σ^2 известна и равна $\sigma^2 = 1$ мм^2; б) значение оценки дисперсии, определенное по выборке, составляет $S^2 = 1{,}21$ мм^2. Контролируемый размер имеет нормальное распределение.

Ответ: а) гипотеза принимается; б) гипотеза принимается.

4.22. Из большой партии резисторов одного типа и номинала случайным образом отобраны 37 шт. Значение выборочного среднего величины сопротивления при этом оказалось равным 9,3 кОм. Используя двусторонний критерий при $\alpha = 0{,}05$, проверьте гипотезу о том, что выборка взята из партии с номинальным значением 10 кОм при альтернативной гипотезе, согласно которой номинальное значение не равно 10 кОм, если: а) дисперсия рассматриваемой случайной величины известна и равна 4 кОм^2; б) дисперсия значения сопротивления неизвестна, а значение выборочной дисперсии равно 6,25 кОм. Распределение контролируемого признака нормальное.

Ответ: а) гипотеза отклоняется; б) гипотеза принимается.

4.23. Установка имеет среднюю производительность 1000 кг вещества в сутки со средним квадратичным отклонением, равным 80 кг^2. При изменении технологии производительность возрастает до 1100 кг вещества в сутки с тем же средним квадратичным отклонением. Можно ли считать, что новая технология обеспечивает повышение производительности, если: а) $\alpha = 0{,}05$; б) $\alpha = 0{,}1$? Контролируемый признак имеет нормальное распределение.

Ответ: а) да; б) да.