

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 :

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

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

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

где константу C_φ выбирают из условия

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

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

$$\int_{\varphi(t_1, \dots, t_n) \geq C_\varphi} \dots \int L(t_1, \dots, t_n; \theta_0) dt_1 \dots dt_n = \alpha.$$

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

$$H_0: \mu = \mu_0, \quad H_1: \mu = \mu_1,$$

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

В рассматриваемом случае функция правдоподобия имеет вид

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

а отношение правдоподобия —

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

В данном случае неравенство

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

равносильно неравенству

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

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

$$P\left\{\sum_{i=1}^n X_i \geq C \mid \mu = \mu_0\right\} = \alpha. \quad (4.2)$$

Действительно,

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

откуда следует, что

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

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

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

или

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

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

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

При этом вероятность совершения ошибки второго рода

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

является минимально возможной при данном значении α .

Пример 4.5. Если в условиях примера 4.4 неравенство $\mu_0 < \mu_1$ заменить неравенством $\mu_1 < \mu_0$, то в этом случае критическое множество W задается неравенством

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

где константу C выбирают из условия

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

Таким образом,

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

или, что то же самое,

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

Из последнего равенства находим $C = n\mu_0 - u_{1-\alpha}\sigma\sqrt{n}$.

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

$$H_0: \lambda = \lambda_0, \quad H_1: \lambda = \lambda_1,$$

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

$$L(X_1, \dots, X_n; \lambda) = \lambda^n \exp\left(-\lambda \sum_{i=1}^n X_i\right).$$

Таким образом,

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

Отсюда видно, что критическое множество можно задать неравенством

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

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

$$P\left\{\sum_{i=1}^n X_i \leq C \mid \lambda = \lambda_0\right\} = \alpha.$$

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

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

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

$$\begin{aligned} \beta &= P\left\{\sum_{i=1}^n X_i > C \mid \lambda = \lambda_1\right\} = \\ &= 1 - H_{2n}(2\lambda_1 C) = 1 - H_{2n}\left(\chi_\alpha^2(2n) \frac{\lambda_1}{\lambda_0}\right), \end{aligned}$$

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

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

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

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

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

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

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

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

Константу C выбирают исходя из условия

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

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

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

откуда, используя квантиль $u_{1-\alpha}$ стандартного нормального закона, получаем

$$C = np_0 + u_{1-\alpha} \sqrt{np_0(1-p_0)}.$$

При этом вероятность ошибки второго рода равна

$$\begin{aligned} \beta &= \mathbf{P}\{K(\vec{X}_n) < C \mid p = p_1\} \approx \Phi\left(\frac{C - np_1}{\sqrt{np_1(1-p_1)}}\right) = \\ &= \Phi\left(\frac{n(p_0 - p_1) + u_{1-\alpha} \sqrt{np_0(1-p_0)}}{\sqrt{np_1(1-p_1)}}\right). \quad (4.8) \end{aligned}$$

4.4. Определение объема выборки

Выше (см. 4.3) при построении *оптимального критерия Неймана — Пирсона* с заданным уровнем значимости α предполагалось, что объем n случайной выборки \vec{X}_n известен и фиксирован. Но возможной является ситуация, когда возникает необходимость в определении (заранее, до проведения наблюдений) такого объема n^* случайной выборки, при котором может быть построен *критерий* для проверки двух *простых гипотез* $H_0: \theta = \theta_0$ и $H_1: \theta = \theta_1$ с заданными или меньшими значениями вероятностей α и β совершения *ошибок первого и второго рода* соответственно.

В рассматриваемой ситуации величину n^* определяют как минимальное целое значение n , для которого система неравенств

$$\begin{cases} \mathbf{P}\{\varphi(X_1, \dots, X_n) \geq C_\varphi \mid \theta = \theta_0\} \leq \alpha, \\ \mathbf{P}\{\varphi(X_1, \dots, X_n) < C_\varphi \mid \theta = \theta_1\} \leq \beta \end{cases} \quad (4.9)$$

может быть выполнена при некотором значении константы $C = C^*$. При этом соответствующий оптимальный критерий Неймана — Пирсона, обеспечивающий заданные значения α, β

будет иметь критическое множество, определяемое неравенством

$$\varphi(x_1, \dots, x_n) \geq C_\varphi^*.$$

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

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

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

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

Используя квантили стандартного нормального распределения, запишем эти уравнения в виде

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

Исключая из уравнений константу C , находим необходимый объем выборки

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

Пусть, например, требуется проверить гипотезы

$$H_0: \mu = \mu_0 = 3,5, \quad H_1: \mu = \mu_1 = 3,8$$

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 $n^* = 61$.

уменьшен при тех же значениях вероятностей совершения ошибок первого и второго рода в последовательной схеме наблюдений, когда решение об остановке наблюдений принимается по ходу процесса наблюдений, в зависимости от получаемых данных (см. 4.6).

4.5. Сложные параметрические гипотезы

Предположим, что требуется проверить две *сложные гипотезы*

$$H_0: \theta \in \Theta_0, \quad H_1: \theta \in \Theta_1, \quad (4.13)$$

где Θ_0, Θ_1 — некоторые непересекающиеся области значений параметра θ . Например, области Θ_0, Θ_1 могут быть заданы неравенствами $\theta \leq \theta_0$ и $\theta \geq \theta_1$, где θ_0 и θ_1 — некоторые фиксированные значения параметра, удовлетворяющие неравенству $\theta_0 < \theta_1$.

Критерий проверки сложных гипотез (4.13) по-прежнему задается с помощью *критического множества W реализаций случайной выборки \vec{X}_n* , на основе которого решение принимают следующим образом:

– если реализация \vec{x}_n случайной выборки \vec{X}_n принадлежит критическому множеству W , тогда *основную гипотезу H_0* отвергают и принимают *альтернативную гипотезу H_1* ;

– если реализация \vec{x}_n случайной выборки \vec{X}_n не принадлежит критическому множеству W , тогда отвергают альтернативную гипотезу H_1 и принимают основную гипотезу H_0 .

Вероятности совершения *ошибок первого и второго рода* в случае сложных гипотез имеют прежний смысл и определяются выражениями

$$\alpha(\theta) = \mathbf{P}\{(X_1, \dots, X_n) \in W \mid \theta\}, \quad \theta \in \Theta_0;$$

$$\beta(\theta) = \mathbf{P}\{(X_1, \dots, X_n) \in \bar{W} \mid \theta\}, \quad \theta \in \Theta_1.$$

В отличие от случая *простых гипотез*, величины $\alpha(\theta)$, $\beta(\theta)$ являются некоторыми функциями от параметра θ .

Максимально возможное значение вероятности совершения ошибки первого рода

$$\alpha = \max_{\theta \in \Theta_0} \alpha(\theta)$$

называют *размером критерия*.

Функцию

$$M(\theta) = \mathbf{P}\{(X_1, \dots, X_n) \in W \mid \theta\},$$

определяющую значение вероятности отклонения основной гипотезы H_0 в зависимости от истинного значения параметра θ , называют *функцией мощности критерия*. Если существует критерий, который при данном фиксированном размере α максимизирует функцию мощности $M(\theta)$ по всем возможным критериям одновременно при всех θ из множества Θ_1 , то такой *критерий* называют *равномерно наиболее мощным*. Равномерно наиболее мощные критерии существуют лишь в некоторых частных случаях при проверке гипотез относительно одномерных параметров (см. примеры 4.10–4.12).

Вероятности совершения ошибок первого и второго рода связаны с функцией мощности следующими соотношениями:

$$\alpha(\theta) = M(\theta), \quad \theta \in \Theta_0; \quad (4.14)$$

$$\beta(\theta) = 1 - M(\theta), \quad \theta \in \Theta_1. \quad (4.15)$$

Тем самым равномерно наиболее мощный критерий, если он существует, минимизирует вероятность совершения ошибки второго рода $\beta(\theta)$ (при фиксированном размере α) одновременно при всех $\theta \in \Theta_1$.

Замечание 4.1. Формально равенства (4.14), (4.15) справедливы при всех возможных значениях θ , но при значениях θ , отличных от указанных в (4.14), (4.15), величины $\alpha(\theta)$, $\beta(\theta)$

теряют свой смысл — вероятностей совершения соответствующих ошибок. #

Иногда наряду с функцией мощности используется также *оперативная характеристика критерия*

$$s(\theta) = P\{(X_1, \dots, X_n) \in \bar{W} | \theta\},$$

представляющая собой вероятность принятия основной гипотезы H_0 при условии, что истинное значение параметра равно θ . Нетрудно увидеть, что оперативная характеристика и функция мощности связаны соотношением $s(\theta) = 1 - M(\theta)$.

Построение критериев для проверки сложных параметрических гипотез проиллюстрируем далее для случая нормальной модели.

Пример 4.10. Рассмотрим проверку простой гипотезы $H_0: \mu = \mu_0$ против сложной гипотезы $H_1: \mu > \mu_0$ относительно параметра — среднего μ нормального распределения при известной дисперсии σ^2 .

При любом $\mu_1 > \mu_0$ критическая область оптимального наиболее мощного критерия Неймана — Пирсона размера α для простых гипотез $\mu = \mu_0$ против $\mu = \mu_1$ имеет вид (4.1), где константу C выбирают из условия (4.2) или (4.3). Поэтому она не зависит от μ_1 . Это означает, что построенный уже выше для указанных простых гипотез критерий с критическим множеством, задаваемым неравенством (4.1)

$$\sum_{i=1}^n x_i \geq C = n\mu_0 + u_{1-\alpha}\sigma\sqrt{n}, \quad (4.16)$$

является равномерно наиболее мощным критерием размера α для данной задачи со сложной альтернативной гипотезой $H_1: \mu > \mu_0$.

4.6. Последовательный критерий отношения правдоподобия

Во многих случаях на практике наблюдения проводят последовательно. При этом статистическая информация поступает не один раз, а последовательными порциями данных. Предпо-

ложим, что наблюдается последовательность независимых одинаково распределенных непрерывных случайных величин X_1, \dots, X_n, \dots , каждая из которых имеет плотность распределения $p(x; \theta)$, где θ — некоторый параметр, значение которого неизвестно. На основе результатов наблюдений x_1, \dots, x_n, \dots нужно проверить две простые гипотезы $H_0: \theta = \theta_0$ и $H_1: \theta = \theta_1$, где θ_0, θ_1 — некоторые заданные значения параметра.

В рассматриваемой ситуации количество наблюдаемых случайных величин (и тем самым объем выборки) не фиксируется заранее, а определяется по ходу наблюдений, в зависимости от получаемых данных. **Последовательный критерий отношения правдоподобия (критерий Вальда*)** строят следующим образом. На очередном n -м шаге наблюдений, исходя из полученных результатов наблюдений x_1, \dots, x_n , вычисляют

$$\varphi_n(x_1, \dots, x_n) = \frac{p(x_1; \theta_1) p(x_2; \theta_1) \dots p(x_n; \theta_1)}{p(x_1; \theta_0) p(x_2; \theta_0) \dots p(x_n; \theta_0)}.$$

Эта величина имеет смысл значения отношения правдоподобия для гипотез H_0 и H_1 на n -м шаге наблюдений. На каждом n -м шаге проверяют следующие два неравенства:

$$B < \varphi_n(x_1, \dots, x_n) < A, \quad (4.26)$$

где B и A — некоторые заданные константы, удовлетворяющие условию $0 < B < 1 < A$. Если оба неравенства (4.26) выполняются, то наблюдения продолжают, т.е. осуществляют наблюдение следующей случайной величины X_{n+1} . Другими словами, неравенства (4.26) задают „область продолжения наблюдений“ для критерия Вальда.

Наблюдения прекращают при первом нарушении хотя бы одного неравенства (4.26). При нарушении левого неравенства принимают гипотезу H_0 . При нарушении правого неравенства принимают гипотезу H_1 . Таким образом, номер ν шага, на котором прекращают наблюдения для критерия Вальда, определяют из равенства

$$\nu = \min \{n: \varphi_n(x_1, \dots, x_n) \notin (B, A)\}. \quad (4.27)$$

Вектор результатов наблюдений для любого последовательного критерия, и в том числе для критерия Вальда, имеет вид (ν, x_1, \dots, x_ν) , где ν — номер шага, на котором прекращены наблюдения, x_1, \dots, x_ν — совокупность всех результатов наблюдений.

Для критерия Вальда правило принятия решения по результатам испытаний x_1, \dots, x_n имеет следующий вид:

- если $\varphi_\nu(x_1, \dots, x_\nu) \leq B$, то принять гипотезу H_0 ;
- если $\varphi_\nu(x_1, \dots, x_\nu) \geq A$, то принять гипотезу H_1 .

Вероятности совершения *ошибок первого и второго рода* (*риски первого и второго рода*) для этого критерия равны соответственно

$$\alpha = P\{\varphi_\nu(X_1, \dots, X_\nu) \geq A | H_0\}, \quad \beta = P\{\varphi_\nu(X_1, \dots, X_\nu) \leq B | H_1\}.$$