

Семинар №11.

Проверка статистических гипотез.

Пусть X – выборка из неизвестного распределения P , которое лежит в некоем семействе распределений \mathcal{P} . (Пример: $P = N(0, 1)$, \mathcal{P} – это все нормальные распределения.)

Определение 1. Статистическая гипотеза – это предположение общего вида о распределении P :

$$H_0 : P \in \mathcal{P}_0 \subset \mathcal{P}.$$

Рассматриваемая (в данный момент) гипотеза называется основной. (Пример: по-прежнему \mathcal{P} – это все нормальные распределения с дисперсией 1, выдвигаем гипотезу, что P имеет нулевое матожидание, т.е. $\mathcal{P}_0 = \{N(0, 1)\}$.)

Определение 2. Если H_0 отвергается, то рассматриваем альтернативную гипотезу (или альтернативу) $H_1 : P \in \mathcal{P}_1$, где $\mathcal{P}_1 \subset \mathcal{P}$ и $\mathcal{P}_1 \cap \mathcal{P}_0 = \emptyset$. (Пример: \mathcal{P} – это все нормальные распределения с единичной дисперсией, основная гипотеза: P имеет нулевое матожидание, альтернатива: матожидание P не равняется 0.)

Определение 3. Если семейство \mathcal{P} параметризовано, $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$, то мы рассматриваем параметрические гипотезы: $H_0 : \theta \in \Theta_0$, $H_1 : \theta \in \Theta_1$, $\Theta_0 \cap \Theta_1 = \emptyset$.

Пусть \mathcal{X} – выборочное пространство и пусть $\mathcal{X} = S \sqcup (\mathcal{X} \setminus S)$.

Определение 4. Пусть правило принятия H_0 по выборке X выглядит так:

- 1) если $X \notin S$, то H_0 принимается,
- 2) если $X \in S$, то H_0 отвергается,

тогда S называется критерием (или критическим множеством) для проверки H_0 (против альтернативы H_1), а $\mathcal{X} \setminus S$ называется областью принятия гипотезы. (В качестве критического множества практически всегда берётся отрицание доверительного интервала по какой-нибудь удобной статистике $T(X)$.)

Определение 5.

- 1) Ошибкой первого рода называется ситуация, когда мы отвергли H_0 при том, что она верна. (Это самая плохая ошибка, её вероятность следует делать как можно меньше.)
- 2) Ошибкой второго рода называется ситуация, когда H_0 не отвергается при том, что она неверна (эта ошибка менее опасна, потому что мы можем отвергнуть неверную гипотезу H_0 в результате последующих проверок).

Пусть $\{P_\theta, \theta \in \Theta\}$ – параметризованное семейство распределений, а S – критерий для проверки $H_0 : \theta \in \Theta_0$ против $H_1 : \theta \in \Theta_1$.

Определение 6. Функцией мощности критерия S называется функция $\beta(\theta, S) = P_\theta(X \in S)$. Величина α называется уровнем значимости для критерия S , если $\alpha \geq \beta(\theta, S)$ для $\forall \theta \in \Theta_0$. Минимальный уровень значимости $\alpha_0 = \sup_{\theta \in \Theta_0} \beta(\theta, S)$ называется размером критерия S .

Определение 7. Если S и R – два критерия уровня значимости α , то S мощнее R , если

$$\beta(\theta, S) \geq \beta(\theta, R) \quad \forall \theta \in \Theta_1.$$

(Заметьте, здесь уже сравнение идёт не по Θ_0 , а по Θ_1 .) Критерий S называется равномерно наиболее мощным критерием (р.н.м.к.) уровня значимости α , если он мощнее любого другого критерия уровня значимости α . Проблема в том, что р.н.м.к. можно построить довольно редко (часто бывает так, что сложно построить хоть какой-нибудь критерий).

Как строить р.н.м.к.

Пусть $\{P_\theta, \theta \in \Theta\}$ - доминируемое (т.е. либо все распределения абсолютно непрерывные, либо все дискретные) семейство с плотностью $p_\theta(x)$. Пусть $H_0 : \theta = \theta_0$ и $H_1 : \theta = \theta_1$ (такие гипотезы называются простыми), $R(X) = \frac{p_{\theta_1}(X)}{p_{\theta_0}(X)}$ — отношение правдоподобий.

Лемма 1 (Неймана-Пирсона)

Если $\exists c > 0$ такое, что

$$P_{\theta_0}(R(X) \geq c) = \alpha,$$

то $S = \{X \in \mathcal{X} : R(X) \geq c\}$ является р.н.м.к. уровня значимости α для проверки H_0 против H_1 .

Пример. Пусть X_1, \dots, X_n — выборка из бернуллиевского закона с вероятностью успеха θ . Построить р.н.м.к. для проверки гипотезы $H_0 : \theta = \theta_0$ против альтернативы $H_1 : \theta = \theta_1$, если а) $\theta_0 < \theta_1$, б) $\theta_1 < \theta_0$.

Δ Распишем отношение правдоподобий: $R(X_1, \dots, X_n) = \frac{\theta_1^{\sum X_i} (1-\theta_1)^{n-\sum X_i}}{\theta_0^{\sum X_i} (1-\theta_0)^{n-\sum X_i}}$. Мы знаем, что равномерно наиболее мощный критерий должен выглядеть так: $\{R(X) \geq c\}$. Так как мы знаем θ_0 , θ_1 , и n , то члены, зависящие только от них, мы можем перенести в константу c . Тогда

$$R(X) \geq c \iff \left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)} \right)^{\sum X_i} \geq c_1 \quad (1)$$

для какой-то константы c_1 . В случае а) $\theta_0 < \theta_1$ выражение (1) эквивалентно $\sum X_i \geq c_2$ (взяли логарифм от (1)). Но $\sum_{i=1}^n X_i \sim \text{Bin}(n, \theta_0)$, поэтому если мы хотим, чтобы уровень значимости нашего критерия равнялся α , то должно выполняться $P(\sum X_i > c_2) = \alpha$, т.е. c_2 — $(1-\alpha)$ -квантиль распределения $\text{Bin}(n, \theta_0)$. Если же мы рассматриваем случай б) $\theta_1 < \theta_0$, то выражение (1) эквивалентно $\sum X_i \leq c_2$, и в качестве c_2 берём α -квантиль $\text{Bin}(n, \theta_0)$. \square

Как проверять сложные гипотезы.

Определение 8. Семейство $\{P_\theta, \theta \in \Theta\}$ имеет монотонное отношение правдоподобия по статистике $T(X)$, если функция $L(X) = L(T(X)) = \frac{p_{\theta'}(X)}{p_{\theta''}(X)}$ является неубывающей (невозрастающей) функцией от статистики $T(X)$ для всех $\theta'' < \theta'$.

Замечание. Как правило, эта статистика $T(X)$ является достаточной статистикой для данного семейства распределений.

Теорема 1 (О монотонном отношении правдоподобий)

Пусть $\{P_\theta, \theta \in \Theta\}$ — семейство с монотонным отношением правдоподобия по статистике $T(X)$. Пусть $\frac{p_{\theta'}(X)}{p_{\theta''}(X)}$ возрастает по $T(X)$ при любых $\theta' > \theta''$. Пусть $H_0 : \theta \leq \theta_0$ (или $\theta = \theta_0$), $H_1 : \theta > \theta_0$. Если $\exists c_\alpha$ такое, что $P_{\theta_0}(T(X) \geq c_\alpha) = \alpha$, то $S = \{T(X) \geq c_\alpha\}$ есть р.н.м.к. уровня значимости α для проверки H_0 против H_1 .

Замечание. В условиях теоремы р.н.м.к. для проверки гипотезы $H_0 : \theta \geq \theta_0$ (или $\theta = \theta_0$) против $H_1 : \theta < \theta_0$ выглядит так: $S = \{T(X) \leq c_\alpha\}$ (достаточно в параметрическом множестве сделать замену переменной $\theta_1 = -\theta$, и мы вернёмся к первоначальной гипотезе). Если же в условиях теоремы отношение правдоподобий убывает по $T(X)$ (проверяем гипотезу $H_0 : \theta \leq \theta_0$ против $H_1 : \theta > \theta_0$), то р.н.м.к. будет таким: $S = \{T(X) \leq c_\alpha\}$. Действительно, достаточно рассмотреть $\hat{T}(X) = -T(X)$, тогда по статистике $\hat{T}(X)$ правдоподобие будет возрастать, и критерий по теореме о монотонном отношении правдоподобий будет таким: $\{\hat{T}(X) \geq \hat{c}_\alpha\} = \{-T(X) \geq \hat{c}_\alpha\} = \{T(X) \leq -\hat{c}_\alpha\}$, что и требовалось.

Но, к сожалению, р.н.м.к. в других случаях существует довольно редко.