## Проверка гипотез

Есть желательно выборка $X_1 \ldots X_n$, $X_i \sim P_X$ неизвест. нам

**Ликбез:** • $H_0: P_X \in \mathcal{P}_0$, где $\mathcal{P}_0$ — некий класс распред., может состоять из одного распр.
основная гипотеза

**Пример** Пусть вероятность того, что Лусин подавится, когда я о нём вспомню, равна $\theta$. $H_0: \theta > 1/2$ (т.е. $\mathcal{P}_0 = \{Bern(\theta), \theta > \frac{1}{2}\}$).

• $H_1: P \in \mathcal{P}_1$   $\mathcal{P}_0 \cap \mathcal{P}_1 = \varnothing$
альтерн. гипотеза   тоже некий класс распред.

• Статистика критерия $S(X)$ — статистика, распред. которой мы знаем, если $P \in \mathcal{P}_0$.

• Критерий: как правило, выглядит так $\{S(X) > s_{1-\gamma}\}$. (А в общем случае — $B \subset \mathcal{X}$)
(это некое мн-во из выб. пространства)   (ур-ня $1-\gamma$, $\gamma$ мало)

Т.е. если статистика критерия больше ~~како~~ какого-то квантиля из того самого извест. нам распред. статистики $S(X)$, то гипотезу $H_0$ отвергаем. Если меньше — то не отвергаем и перех. к след. проверкам, если надо.

• уровень значимости критерия: такое число $\alpha$, что $\forall P_X \in \mathcal{P}_0$   $P_X(S(X) > s_{1-\gamma}) \le \alpha$. ($\alpha = \gamma$, если $\mathcal{P}_0$ из одного распред. состоит)
или $X \in B$, где $B$ — наше критич. мн-во

Миним. ур. значимости называется размером критерия

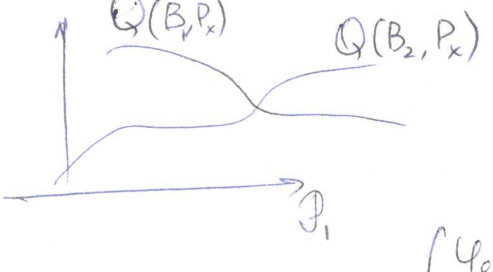
• ошибка I рода: отвергнуть ~~верную~~ основную гипотезу, если она верна.
Вер. ошибки первого рода меньше ур. значимости.

• ошибка II рода: ~~отверг~~ принять ~~неверную~~ основную гипотезу, если она неверна. Не так страшна, как ошибка I рода.
Уменьшая вер. ошибки I рода, мы неизменно увел. вер. ош. II рода, так что надо знать меру.

• Мощность критерия (ф-ция мощности):
$$Q(B, P_X) = P_X(X \in B), \quad \text{где } P_X \in \mathcal{P}_1.$$
наша выборка   критич. мн-во

• Сост. критерия: мощность $\to 1$, если $n \to \infty$

Т.е. мы зафиксир. уровень значимости, ограничив им вероятность ошибки I рода — а теперь хотим при этом условии ещё и критерий с максимальной мощностью выбрать, а то неясно, какой критерий лучше.




Т.е. хотим найти равномерно наиболее мощный критерий, ~~для~~ ф-ция мощности которого мажорирует ф-ции мощности всех остальных критериев ур-ня знач. $\alpha$.

(Чем больше ф-ция мощности, тем меньше вер. ошибки второго рода!)

## Введение | МСПС |



Увы, как мы знаем, такие крит. не всегда $\exists$.
Для проверки простых гипотез в параметрич. модели (т.е. $\mathcal{P}=\{P_\theta, \theta\in\Theta\}$, $H_0:\theta=\theta_0$ и $H_1:\theta=\theta_1$) использ. лемма

**Неймана-Пирсона:**

Пусть $f(X_1\dots X_n,\theta)=\prod_{i=1}^{n} p(X_i,\theta)$, где $p(x,\theta)$ – обобщ. плотность распр. $P_\theta$. Если $\exists\lambda_\alpha$ такое, что $P_{\theta_0}(B)=\alpha$,
Тогда р.н.м. критерий будет таким: $B=\left\{\frac{f(X_1\dots X_n,\theta_1)}{f(X_1\dots X_n,\theta_0)}>\lambda_\alpha\right\}$, где
(подставляем вместо сл. в. $X_i$ её реализацию $x_i$ и пользуемся).
$\lambda$ выбирается нами, исходя из зад. уровня значимости.

Ну и пример.

$X_1\dots X_n\sim Bern(\theta)$; $H_0:\theta=\theta_0$; $H_1:\theta=\theta_1$, $\theta_1>\theta_0$

Статистика $\frac{1}{n}\sum_{i=1}^{n}X_i$ – хотим такую (вытекает из вида правдоподобия). Тогда выберем крит.: $\sum_{i=1}^{n}X_i\geq k_\alpha$, где
$P_{\theta_0}\left(\sum_{i=1}^{n}X_i\geq k_\alpha\right)=\sum_{i=k_\alpha}^{n}C_n^i\theta_0^i(1-\theta_0)^{n-i}\leq\alpha$. Ну и мощность $P_{\theta_1}\left(\sum_{i=1}^{n}X_i\geq k_\alpha\right)=\sum_{i=k_\alpha}^{n}C_n^i\theta_1^i(1-\theta_1)^{n-i}$

**Пример на Неймана-Пирсона**

Пусть $\mathcal{P}=\{N(\theta,1),\ \theta\in\mathbb{R}\}$, $H_0:\theta=\theta_0$; $H_1:\theta=\theta_1$, $\theta_1>\theta_0$

Правдоподобие: $f(X,\theta)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{(X_i-\theta)^2}{2}}=\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{\sum X_i^2+2\sum\theta X_i-n\theta^2}{2}}$.

Критерий: найдём
$$\left\{\frac{f(X,\theta_1)}{f(X,\theta_0)}>\lambda_\alpha\right\}=\left\{\frac{e^{-\sum X_i^2+2\sum\theta_1X_i-n\theta_1^2}}{e^{-\sum X_i^2+2\sum\theta_0X_i-n\theta_0^2}}>\lambda\right\}=\left\{e^{2\sum(\theta_1-\theta_0)X_i}>\tilde{\lambda}_\alpha\right\}=$$
$$=\left\{(\theta_1-\theta_0)\cdot\sum X_i>\frac{\ln\tilde{\lambda}_\alpha}{2}\right\}=\left\{\frac{1}{n}\sum X_i>\hat{\lambda}_\alpha\right\}.$$

Но $\frac{1}{n}\sum X_i\sim N(\cancel{n}\theta_0,\frac{1}{n})$ при верной $H_0$ $\Rightarrow$ $\hat{\lambda}_\alpha$ – $N(\theta_0,\frac{1}{n})$ квантиль ур-ня $1-\alpha$.

Монот. отн. пр-е.

**Пример на проверку гипотез с помощью пред. теорем**

Пусть $\{X_i\}_{i=1}^{n}$ – выборка из $Exp(\alpha)$ с неизвест. $\alpha$
Проверим $H_0:\alpha=\alpha_0$ против $H_1:\alpha>\alpha_0$.
Для этого берём ЦПТ: $\frac{\sum X_i-nEX_1}{\sqrt{nDX_1}}\xrightarrow{d}N(0,1)$, где $EX_1=\frac{1}{\alpha}$; $DX_1=\frac{1}{\alpha^2}$.
~~Заменим, скажем, $\frac{1}{\alpha}$ в знам.~~ Имеем: $\frac{\sum X_i-n\frac{1}{\alpha_0}}{\sqrt{n}\frac{1}{\alpha_0}}\to N(0,1)$

Критерий: если $\left|\frac{\sum X_i-n\frac{1}{\alpha_0}}{\sqrt{n}\frac{1}{\alpha_0}}\right|>z_{1-\gamma}$, где $z_{1-\gamma}$ – квантиль ур. $1-\gamma$ из $N(0,1)$, то отвергаем $H_0$.

p-знач.!   можно дисп. заменить на её оценку.

<u>Опр.</u> p-значение критерия: ~~пусть значение нашей статистики~~
Пусть статистика критерия $- K(X_1, \dots X_n)$, и пусть $K(X_1, \dots X_n) \geq 0$, имеющая распределение $F_K(x)$ и пусть значение статистики на наших данных равно $y$, т.е.
$y = K(\underbrace{x_1, \dots x_n}_{\text{данные, т.е. реализация выборки } X_1, \dots X_n})$ / Одностор. и двустор. p-знач.!

Тогда число $p = 1 - F_K(y)$, называется p-значением критерия.

<u>Пример</u> Пусть мы отвергаем гипотезу $H_0$ с пом. критерия $K$ на уровне значимости $\overset{0,05}{\alpha}$, причём критерий выглядит так: $K(x_1 \dots x_n) > z_{1-\alpha}$, то отвергаем, где $z_{1-\alpha}$ — $(1-\alpha)$-квантиль $F_K$. И пусть p-значение критерия $K$ равно 0,08. Что делать? Не отвергать.

Часто компьютер нам выдаёт вместо знач. критериев их p-значения потому что подставлять в таблицы долго и проблематично.

<u>Вопрос</u> Пусть проверяем гипотезу $H_0$ двумя критериями. У одного p-знач. выше уровня значимости, у второго — ниже. Что делать?

<u>Ответ</u> Отвергнуть гипотезу, если уровень знач. низок. Иначе продолжать исследования. Хотя лучше отвергнуть и уточнить гипотезу. См. множеств. проверку гипотез.

<u>Проверка гипотез в байесовском подходе.</u>

Итак, пусть $H_0 : \cancel{\exists} P \in \mathcal{P}_0$, где $\mathcal{P}_0 = \{P_\theta, \theta \in \Theta_0\}$
$H_1 : P \in \mathcal{P}_1$, где $\mathcal{P}_1 = \{\hat{P}_\gamma, \gamma \in \Gamma\}$

(Например, гаусс. семейство и экспоненциальное; или $\{N(a, 1), a \in \mathbb{R}\}$ и $\{N(a, 2), a \in \mathbb{R}\}$)

Рассмотрим $K = \dfrac{P(\mathcal{P}_0 | y)}{P(\mathcal{P}_1 | y)} = \cancel{P(y|\mathcal{P}_0)\, P(\mathcal{P}_0)}$

фактически, что вероятнее при наших данных!

$$= \frac{\int P(y|\theta)\, P(\theta)\, d\theta}{\int P(y|\gamma)\, P(\gamma)\, d\gamma}$$

Ну и как же определить?

Важное св-во: $K$ не зависит от $\theta$ и $\gamma$ — пар-ров распред., т.е. мы можем найти распред. $K$ при верной $H_0$!

Шкала Джеффри:

| $K$ | верна ли $H_0$? |
|---|---|
| 1-3 | нельзя определённо сказать |
| 3-10 | большие основания принять $H_0$ |
| 10-30 | почти наверняка |
| >30 | ~~наверняка~~ точно. |

* <u>Опр.</u> Пусть $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ и $F_\theta$ — ф.р. $P_\theta$. Пусть $F_\theta(x) = F_0(x - \theta)$.
или плотность
Тогда $\theta$ наз. пар-ром сдвига. Пример $N(a, 1)$.

<u>Опр.</u> Пусть —//— и $F_\theta(x) = F_1\left(\frac{x}{\theta}\right)$. Тогда $\theta$ наз. пар-ром масштаба
Пример $N(0, \sigma^2)$.

Все остальные пар-ры наз. пар-рами формы. Напр., Pareto($\alpha$)

Задача на Байеса (гипотезы) | МСПС | Введение 

Пусть $\mathcal{P}_1 = \{Exp(\alpha)\}$; $\mathcal{P}_2 = \{R[0,\theta]\}$

Есть выборка $X_1 \ldots X_n$. Тогда

$$K = \frac{\int \alpha^n \exp(-\alpha \sum_{i=1}^n X_i)\, \beta(\alpha)\, d\alpha}{\int \frac{1}{\theta^n} I(X_{(n)} \le \theta)\, \beta(\theta)\, d\theta} =$$

Пусть $\beta(\theta)$ – степенное, типа $\frac{1}{\theta^2} \cdot I(\theta \geq 1)$, а $p(\alpha)$ – ~~тоже степенное~~ – экспон.: $\gamma e^{-\gamma\alpha}$

$$= \frac{\int_0^{+\infty} \alpha^{n} \cdot e^{-\alpha \sum X_i - \alpha\gamma} \cdot \gamma\, d\alpha}{\int_{X_{(n)}}^{+\infty} \frac{1}{\theta^{n+2}}\, d\theta} = \frac{\frac{n!\,\alpha^n \gamma}{(\sum X_i + \gamma)^n}}{\frac{1}{(n+1) X_{(n)}^{n+1}}} = \frac{(n+1)!\,\alpha^n \cdot \gamma (X_{(n)})^{n+1}}{(\sum X_i + \gamma)^n}$$

---

## Теорема о монот. отн. правдоподобия

**Опр.** $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ – (доминир.) сем-во распред. с монот. отн. правдоподобия по статистике $T(X)$, если $\forall \theta_0 < \theta_1$ $\frac{f_{\theta_1}(X)}{f_{\theta_0}(X)}$ ← пр-я явл. монотонной ф-цией от $T(X)$, причём монотонность одна и та же $\forall \theta_1 > \theta_0$

**Теорема** Пусть стоит задача проверки гипотезы $H_0: \theta \le \theta_0$ (или $\theta = \theta_0$) и $H_1: \theta > \theta_0$, а $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ – сем-во с монотон. отнош. правдоподобия, причём $\frac{f_{\theta_1}(X)}{f_{\theta_2}(X)}$ возрастает по $T(X)$ при $\forall \theta_1 > \theta_2$. Если $\exists c_\alpha$ т.ч. $P_{\theta_0}(T(X) \geq c_\alpha) = \alpha$, то $S = \{T(X) \geq c_\alpha\}$ – р.н.м.к. ур-ня знач. $\alpha$ для проверки $H_0$ против $H_1$.

---

На практике часто возникает задача различения одного сем-ва распред. от другого. Для этих целей используется критерий отношения правдоподобия (RML-test)

Пусть $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_1$, $\Theta_0 \cup \Theta_1 \subset \Theta$.

Обозначим $\lambda_n(X, \Theta_0) = \frac{\sup_{\theta \in \Theta_0} f_\theta(X)}{\sup_{\theta \in \Theta} f_\theta(X)}$

**Теорема** (Уилкс) Если $\Theta \subset \mathbb{R}^k$, $\dim(\Theta) = k$, $\dim(\Theta_0) = s < k$, то при вер. $H_0$:

$$-2 \ln \lambda_n(X, \Theta_0) \to \chi^2_{k-s} \text{ при } n \to \infty.$$

Кроме того, критерий $S = \{x: -2 \ln \lambda_n(x, \Theta_0) \geq \chi^2_{1-\alpha, k-s}\}$ будет состоятельным.

---

В остальных случаях таблицы квантилей составляются вручную, с помощью моделирования.