Согласны? Проверяли? Отвергали?

Критерии согласия. pvalue

Чертоги разума

Определение. Пусть $R_\alpha \subset \mathcal{X}$ — критерий с ошибкой I рода α, причём $R_\alpha \subset R_\beta$ для $\alpha < \beta$. *pvalue* называется статистика, равна минимальному уровню значимости, на котором гипотеза отвергается:

$$\mathrm{pvalue}(\mathbf{X}) = \inf\{\alpha \colon \mathbf{X} \in R_\alpha\}.$$

По определению, H_0 отвергается $\iff \mathrm{pvalue}(\mathbf{X}) \leqslant \alpha$.

Пример. Для критериев вида $R_\alpha = \{\mathbf{x} \colon T(\mathbf{x}) > c_\alpha\}$, где распределение F_T статистики $T(\mathbf{X})$ одно и то же при верности H_0, pvalue считается как

$$\mathrm{pvalue}(\mathbf{x}) = \mathsf{P}_0(T(\mathbf{X}) > T(\mathbf{x})) = 1 - F_T(T(\mathbf{x})), \quad \mathsf{P}_0 \in \mathcal{P}_0,$$

откуда следует, что $\mathrm{pvalue}(\mathbf{X}) \overset{H_0}{\sim} \mathrm{U}[0,1]$ для непрерывной F_T. Часто F_T выше заменяют на асимптотическое распределение $T(\mathbf{X})$, такую величину также принято называть pvalue.

Далее мы будем изучать *критерии согласия*, то есть критерии для проверки $H_0 \colon \mathsf{P} = \mathsf{P}_0$ против общей альтернативы $H_1 \colon \mathsf{P} \neq \mathsf{P}_0$.

Критерий Колмогорова проверяет гипотезу для непрерывного P_0 с ф.р. F_0 посредством статистики

$$D_n = \sup_{x \in \mathbb{R}} |\widehat{F}_n(x) - F_0(x)|$$

Теорема. Распределение $\sqrt{n} D_n$ при верности H_0 не зависит от F_0 и сходится при $n \to \infty$ к распределению Колмогорова:

$$K(x) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2x^2}.$$

Критерий же имеет вид $\{\mathbf{x} \colon \sqrt{n} D_n(\mathbf{x}) \geqslant k_{1-\alpha}\}$ и обладает асимпт. ур. зн. α.

Аналогично строится *критерий омега-квадрат* со статистикой

$$\omega^2(\mathbf{X}) = \int_{\mathbb{R}} \left(\widehat{F}_n(x) - F_0(x)\right)^2 \psi(F_0(x))\, dF_0(x),$$

где $\psi(t)$ — некоторая весовая функция (например, $\psi(t) \equiv 1$ или $\psi(t) = \frac{1}{t(1-t)}$).

Известно, что $n\omega^2(\mathbf{X}) \xrightarrow{d} F_\psi$.

Критерий хи-квадрат Пирсона проверяет гипотезу для категориального распределения $\mathsf{P}_0 = \mathrm{Cat}(p_1^0, \ldots, p_k^0)$ посредством статистики

$$\chi^2(\mathbf{X}) = \sum_{i=1}^{k} \frac{(\nu_i - np_i^0)^2}{np_i^0}, \quad \nu_i = \sum_{j=1}^{n} I(X_j = i).$$

При верности H_0, $\chi^2(\mathbf{X}) \xrightarrow{d} \chi^2_{k-1}$. Сам критерий имеет вид $\{\mathbf{x} \colon \chi^2(\mathbf{x}) \geqslant \chi^2_{k-1,1-\alpha}\}$.

数学题

1. Пусть $X_1, \ldots, X_n \sim \mathcal{N}(a, \sigma^2)$, где оба параметра неизвестны. Найдите pvalue для
(а) точного правостороннего критерия $\left\{\mathbf{x}\colon \sqrt{n} \cdot \frac{\overline{\mathbf{x}} - a_0}{S(\mathbf{x})} > t_{n-1, 1-\alpha}\right\}$;
(б) асимптотического двустороннего критерия $\left\{\mathbf{x}\colon \left|\sqrt{n} \cdot \frac{\overline{\mathbf{x}} - a_0}{S(\mathbf{x})}\right| > z_{1-\alpha/2}\right\}$
для проверки гипотезы $H_0\colon a = a_0$.

2. Придумайте способ, как посчитать статистику **(а)** Колмогорова; **(б)** ω^2 для $\psi \equiv 1$, не считая супремум/интеграл напрямую.

3. Согласно *закону Бенфорда*, первая цифра ξ_1 случайного числа с десятичной записью $\overline{\xi_1 \ldots \xi_n}$ из достаточно широко диапазона имеет распределение

$$\mathsf{P}(\xi_1 \leqslant d) = \log_{10}(d + 1), \quad d \in \{1, \ldots, 9\}.$$

Для выборки из стран мира (данные можно взять, например, отсюда) и уровня значимости 0.05 проверить гипотезу о том, что численность населения подчиняется закону Бенфорда.

4. Критерий χ^2 подходит для проверки не только большого уклонения от нулевой гипотезы, но и подозрительно точного соответствия ей. В этом случае можно взять критерий вида

$$R'_\alpha = \{\mathbf{x}\colon \chi^2(\mathbf{x}) \geqslant \chi^2_{k-1, 1-\alpha/2} \vee \chi^2(\mathbf{x}) \leqslant \chi^2_{k-1, \alpha/2}\}.$$

Критерий выше будет отклонять нулевую гипотезу и когда данные далеки от теории, и когда они слишком хорошо ею описываются. Найдите pvalue$(\mathbf{X})$ такого критерия.

5. *Дивергенцией Кульбака-Лейблера* и *дивергенцией-*χ^2 двух дискретных распределений $\mathsf{P} = (p_1, \ldots, p_k)$ и $\mathsf{Q} = (q_1, \ldots, q_k)$ называются соответственно величины

$$KL(\mathsf{P} \parallel \mathsf{Q}) = \sum_{i=1}^{k} p_i \cdot \ln \frac{p_i}{q_i} \text{ и } \chi^2(\mathsf{P} \parallel \mathsf{Q}) = \sum_{i=1}^{k} \frac{(p_i - q_i)^2}{q_i}.$$

Пусть k-гранный кубик с вероятностями выпадения i-ой грани p_i^0 подкидывают n раз, $(\nu_1, \ldots, \nu_k)$ — наблюдаемые частоты, а $p_i = \nu_i / n$. Докажите, что

$$\frac{\chi^2(\mathsf{P} \parallel \mathsf{P}^0)}{KL(\mathsf{P} \parallel \mathsf{P}^0)} \xrightarrow{d} 2,$$

откуда выведите, что

$$2n \cdot KL(\mathsf{P} \parallel \mathsf{P}^0) = \sum_{i=1}^{k} 2\nu_i \cdot \ln \frac{\nu_i}{np_i^0} \xrightarrow{d} \chi^2_{k-1}, \quad n \to \infty.$$

Что можно сказать про сходимость статистики $2n \cdot D(\mathsf{P}^0 \parallel \mathsf{P})$? Какую из них лучше использовать в качестве основы критерия?

6. Модифицируйте критерий Колмогорова таким образом, чтобы, во-первых, в качестве основной гипотезы $H_0\colon \mathsf{P} = \mathsf{P}_0$ можно было выбрать произвольное распределение P_0 (необязательно непрерывное), и, во-вторых, критерий остался состоятельным против альтернативы $H_1\colon \mathsf{P} \neq \mathsf{P}_0$.