H_0: Этот листочек — простой

Проверка гипотез

Чертоги разума

Определение. Пусть $\mathcal{P}$ — пространство распределений из нашей модели. *Гипотезой* называют какое-то её подмножество $\mathcal{P}_0 \subset \mathcal{P}$. Обозначают как $H_0\colon \mathsf{P} \in \mathcal{P}_0$.

Основная задача. По наблюдениям понять, можно ли считать допустимой *нулевую гипотезу* $H_0\colon \mathsf{P} \in \mathcal{P}_0$ или есть основания считать, что на самом деле верна *альтернатива* $H_1\colon \mathsf{P} \in \mathcal{P}_1$, где $\mathcal{P}_1 \subset \mathcal{P} \setminus \mathcal{P}_0$.

Процедура. Отклоняем или не отклоняем H_0 в зависимости от попадания выборки в подмножество $R \subset \mathcal{X}$ (его называют *критерием* или *критическим множеством*):

$$\mathbf{X} \in R \Longrightarrow \text{ отвергаем } H_0;$$
$$\mathbf{X} \notin R \Longrightarrow \text{ не отвергаем } H_0.$$

Пример. $R = \{\mathbf{x} \in \mathcal{X}\colon T(\mathbf{x}) \geqslant c\}$, где $T(\mathbf{X})$ — некоторая статистика.

Обычно критерий подбирают так, чтобы:

- вероятность отклонить H_0, когда она верна (*ошибка I рода*), была не больше определённого порога α — *уровня значимости* критерия:

  $$\forall \mathsf{P} \in \mathcal{P}_0\colon \mathsf{P}(\mathbf{X} \in R) \leqslant \alpha;$$

  если неравенство выше выполнено в пределе по размеру выборки, то говорят, что R $(= R_n)$ обладает а*симптотическим уровнем значмости* α;
- вероятность не отвергнуть H_0, когда она не верна (*ошибка II рода*), была просто как можно меньше или, что то же самое, *мощность критерия* $\beta(\mathsf{P}, R) := \mathsf{P}(\mathbf{X} \in R)$ была как можно больше.

Критерий уровня значимости α называют *равномерно наиболее мощным* (р.н.м.к.), если для любого $\mathsf{P} \in \mathcal{P}_1$ его мощность не меньше мощности любого другого критерия уровня значимости α.

Лемма (Нейман, Пирсон). Для проверки

$$H_0\colon \rho(x) = \rho_0(x) \quad \text{versus} \quad H_1\colon \rho(x) = \rho_1(x)$$

критерий вида

$$R_\lambda = \{\mathbf{x}\colon \rho_1(\mathbf{x}) > \lambda \rho_0(\mathbf{x})\}, \quad \lambda > 0$$

является р.н.м.к. уровня значимости $\alpha = \mathsf{P}_0(\mathbf{X} \in R_\lambda)$.

Если р.н.м.к. построить не удалось, можно наложить чуть менее крутые, но полезные ограничения на мощность:

- *состоятельность*: $\forall \mathsf{P} \in \mathcal{P}_1\colon \mathsf{P}(\mathbf{X} \in R_n) \to 1$, $n \to \infty$;
- *несмещённость*: $\sup\limits_{\mathsf{P} \in \mathcal{P}_0} \mathsf{P}(\mathbf{X} \in R) \leqslant \inf\limits_{\mathsf{P} \in \mathcal{P}_1} \mathsf{P}(\mathbf{X} \in R)$;

Ekzercoj

1. Пусть $\widehat{\theta}$ — асимп. нормальная оценка θ с асимпт. дисперсией с состоятельной оценкой $\widehat{\sigma}^2$. Постройте на её основе состоятельный критерий асимпт. уровня значимости α для проверки

$$H_0\colon \theta = \theta_0 \quad \text{versus} \quad H_1\colon \theta \neq \theta_0.$$

2. Партия в преферанс предназначена для трёх игроков, каждому из которых раздаётся случайным образом по 10 карт, а остальные 2 карты скидываются в прикуп (итого, 32 карты — от семёрок до тузов). Двое игроков заметили, что третьему за 100 партий на руки выпал хотя бы один туз 87 раз. На уровне значимости 0.01 проверьте гипотезу о том, что он играет честно, против альтернативы, что он подмешивает себе тузов, с помощью какого-то критерия (быть может асимптотического).

3. Случайно равновероятно возьмём M — произвольного человека из некоторой страны (какой — мы не знаем) и поставим на проверку гипотезу $H_0\colon M$ — американец. Критерий предлагается построить на основе его профессии, возьмём множество

$$R = \{m \in \{\text{Люди}\}\colon m \text{ — конгрессмен}\}.$$

Очевидно, при верности H_0 вероятность $\mathsf{P}_0(M \in R)$ крайне мала, поэтому данный критерий обладает разумным уровнем значимости, например, $\alpha = 0.01$. Таким образом, если случайный человек оказался конгрессменом, то в соответствии с критерием мы должны отклонить гипотезу H_0. Всё ли корректно в данной процедуре?

4. Пусть $X_1, \ldots, X_n$ — выборка из распределения $\mathcal{N}(a, \sigma^2)$. Предложите (не асимптотический) критерий для проверки гипотезы $H_0\colon a = 0$.

5. Пусть $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_m$ — две независимые выборки из $\mathbb{R}_+$ со средними μ_1, μ_2 и дисперсиями σ_1^2, σ_2^2 соответственно (все параметры неизвестны). Покажите, что $\overline{X}/\overline{Y}$ является асимптотически нормальной оценкой μ_1/μ_2 и на основании этой статистики постройте асимптотический критерий для проверки гипотезы $H_0\colon \mu_1 = \mu_2$. Чем может быть полезен такой способ проверки?

6. В модели $X_1, \ldots, X_n \sim \mathcal{N}(\theta, 1)$ постройте р.н.м.к. для проверки $H_0\colon \theta = \theta_0$ против альтернативы $H_1\colon \theta = \theta_1$, где $\theta_1 > \theta_0$. Убедитесь, что критерий состоятелен.

7. Прочитайте доказательство леммы Неймана-Пирсона и предложите достаточное условие, при котором р.н.м.к. из этой леммы будет единственным с точностью до μ-п.н., где μ — мера, по которой берутся плотности $\rho_0(x)$ и $\rho_1(x)$ из условия.

8. Докажите, что в модели сдвига $X_1, \ldots, X_n \sim \mathcal{N}(\theta, 1)$, $\theta \in \mathbb{R}$, не существует р.н.м.к. для проверки гипотезы $H_0\colon \theta = 0$ vs. $H_1\colon \theta \neq 0$.