Спецкурс "Большие уклонения"

Шкляев А.В.

21 ноября 2017 г.

Теорема Гартнера-Эллиса

Сперва рассмотрим еще один пример:

Пример 5.1. Применим теорему Санова к задаче проверки гипотезы согласия $H_0 : p_1 = p_1^0, ..., p_k = p_k^0$, где p_i^0 заданы, с помощью критерия хи-квадрат. Критерий хи-квадрат основан на том факте, что статистика

$$T(X_1, ..., X_n) = \sum_{i=1}^{k} \frac{(N_i - np_i^0)^2}{np_i^0}$$

имеет асимптотическое распределение χ^2_{k-1} при выполнении гипотезы. Тем самым вероятность $\mathbf{P}_0(T(X_1, ..., X_n) > c)$ аппроксимируется вероятностью $1 - F_{\chi^2_{k-1}}()$. Эта аппроксимация довольно точна при c порядка константы, но, скажем, при c порядка n она не является удовлетворительной. В этом случае фактический уровень значимости критерия можно оценить с помощью теоремы Санова

$$\frac{1}{n} \ln \mathbf{P}(T(X_1, ..., X_n) > nc) = \frac{1}{n} \ln \mathbf{P}\left(\sum_{i=1}^{k} \frac{(\nu_i - p_i^0)^2}{p_i^0} > c\right) \to -\inf_{\vec{\theta} \in A} \sum_{i=1}^{k} \theta_i \ln \frac{\theta_i}{p_i^0},$$

где $A = \{\theta_i : \sum_{i=1}^k \theta_i = 1,\ \sum_{i=1}^k \frac{(\theta_i - p_i^0)^2}{p_i^0} > c\}$.
При этом A — внешность эллипсоида

$$\sum_{i=1}^{k} \frac{(\theta_i - p_i^0)^2}{p_i^0} > c,$$

пересеченная с плоскостью $\sum_{i=1}^k \theta_i$. Нетрудно убедиться, что минимум достигается на границе, то есть

$$\sum_{i=1}^{k} \frac{(\theta_i - p_i^0)^2}{p_i^0} = c.$$

Полученная задача достаточно сложна для аналитической работы, но может быть решена численно.

Принцип больших уклонений можно сформулировать и для значительно более общей модели. Пусть Z_1,...,Z_n — некоторые случайные векторы, $R_{Z_n}(\vec{h}) = \mathbf{E}e^{(\vec{h}, Z_n)}$. Предположим, что существует предел

$$\ln R(\vec{h}) = \lim_{n\to\infty} \frac{1}{n} \ln R_{Z_n}(n\vec{h}),$$

причем $R(\vec{h})$ конечна в окрестности нуля. Пусть $\Lambda(\vec{\theta}) = (\vec{\theta}, \vec{h}) - \ln R(\vec{h})$.
Назовем функцию Λ существенно гладкой, если $\Lambda(\vec{\theta})$ конечна на множестве D_Λ, внутренность которого D_Λ^{int} непуста, Λ дифференцируема в этом множестве, $|grad \Lambda(\vec{\theta})|$ сходится к бесконечности при $\vec{\theta} \to \vec{\theta}_0$, $\vec{\theta}_0$ на границе D_Λ^{int}. Тогда справедлива следующая теорема.

Теорема 5.1. Гартнер-Эллис. Если Λ существенно гладкая, непрерывная снизу функция, то $P(\vec{Z_n} \in A)$ удовляет ПБУ с $\Lambda(\vec{\theta})$.

Прежде чем переходить к доказательству теоремы, рассмотрим несколько примеров

Пример 5.2. Если $\vec{Z}_n = \vec{S}_n/n$, то $R_{\vec{Z}_n}(\vec{h}n) = R^n_{\vec{X}_1}(\vec{h})$, откуда $R(\vec{h}) = R_{\vec{X}_1}(\vec{h})$ и мы получаем теорему Крамера с дополнительным условием $R(\vec{h})$ конечна в окрестности нуля.

Пример 5.3. Пусть $\vec{S}_n = \sum_{i=1}^n f(i/n)\vec{X}_i$, где f — интегрируемая на $[0,1]$ функция, $\vec{X}_i$ — н.о.р. с $R(\vec{h}) = \mathbf{E}e^{(\vec{h}, X)}$. Тогда

$$\frac{1}{n} \ln R_{\vec{S}_n/n}(\vec{h}n) = \frac{1}{n} \sum_{i=1}^{n} \ln R(a_i \vec{h}n) \to \int_0^1 \ln R(f(x)\vec{h})dx.$$

Следовательно, $\vec{S}_n/n$ удовлетворяет ПБУ с $\Lambda(\theta) = \sup(h\theta - \int_0^1 \ln R(f(x)\vec{h}))$

Пример 5.4. Пусть $Z_n = S_{N_n}$, где $S_n = X_1 + ... + X_n$, X_i н.о.р., а N_n не зависит от них и удовлетворяет $\frac{1}{n} \ln R_{N_n}(\lambda) \to f(\lambda)$. Действительно,

$$\frac{1}{n} \ln R_{S_{N_n}/n}(nh) = \frac{1}{n} \ln \left(\sum_{k=0}^{\infty} P(N_n = k) E \exp(hS_k)\right) = \frac{1}{n} \ln R_{N_n}(\ln R_{X_1}(h)) \to f(\ln R_{X_1}(h)).$$

Из теоремы Гартнера-Эллиса следует, что $\frac{1}{n} Z_n$ удовлетворяет ПБУ с $\Lambda(\theta) = \sup(h\theta - f(\ln R_{X_1}(h)))$.

Задача 5.1. Пусть $X_1, ..., X_n, ...$ — гауссовская стационарная последовательность с нулевым средним и ковариационной функцией $R(i)$, т.е. $(X_1, ..., X_k)$ при любом k гауссовский вектор с нулевым средним и ковариацией $cov(X_i, X_{i+j}) = R(j)$. Предположим, что R удовлетворяет условию

$$\sum_{i=-n}^{n} (1 - |i|/n) R(i) \to R,\ n \to \infty.$$

Тогда S_n/n удовлетворяют ПБУ с $\Lambda(\theta) = R\theta^2/2$.
Пример 5.5. Применим теорему Гартнера-Эллиса к конечной цепи Маркова с состояниями $S = \{1, ..., k\}$. Конечной цепью Маркова называют случайную последовательность для которой

$$\mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1})$$

при всех $i_0, ..., i_n$: $\mathbf{P}(X_{n-1} = i_{n-1}, ..., X_0 = i_0) > 0$. Будем рассматривать однородные цепи, то есть те, для которых $p_{i,j} = \mathbf{P}(X_n = j | X_{n-1} = i)$ одно и то же для всех n. Такая цепь задается вектором $\vec{p}_0 = (\mathbf{P}(X_0 = 1), ..., \mathbf{P}(X_0 = k))$ начального распределения и матрицей перехода $P = (p_{i,j})$. Тогда

$$(\mathbf{P}(X_n = 1), ..., \mathbf{P}(X_n = k)) = \vec{p}_0 P^n.$$

Нам понадобится следующая теорема, называемая теоремой Перрона (Перрона-Фробениуса). **Теорема** Если P — матрица с положительными элементами, то существует собственное значение $\lambda > 0$ (оно называется главным), такое что
1) Оно имеет кратность 1.
2) Остальные собственные значения не превосходят его по модулю.
3) У соответствующего ему собственного вектора $\vec{e}_\lambda$ (такой вектор называется главным) все координаты положительны. Это единственный собственный вектор с положительными координатами.

Это достаточно удобная теорема из которой, в частности, следует эргодическая теорема для цепей Маркова. Предположим, что матрица перехода P имеет положительные элементы. У нее заведомо есть правый собственный вектор $(1, 1, ..., 1)$ с собственным значением 1. Значит в силу теоремы Перрона это собственное значение главное. Следовательно, если $\vec{e}_i$ — базис корневого пространства, то

$$\vec{p}_0 P^n = \sum_{i=1}^{n} a_i \vec{e}_i P^n \to a_1 \vec{e}_1,$$

где a_i — координаты разложения $\vec{p}_0$ по $\vec{e}_i$, поскольку корневое подпространство с с.з. 1 одномерно, а все остальные корневые подпространства соответствуют собственные значения с модулем меньше 1, а значит $\vec{e}_i P^n \to 0$ при $i > 1$. При этом $a_1 \vec{e}_i$ имеет сумму координат равную 1, поэтому это левый собственный вектор с суммой координат 1. Тем самым, предельное распределение нашей цепи будет вектором $\pi = (\pi_1, ..., \pi_k)$, т.ч. $\pi P = \pi$, $\sum_{i=1}^{k} \pi = 1$.