

Seminar 5

Carlos Buitrago

Table of contents

1	Оптимальные оценки	1
2	Достаточные статистики	2
2.1	Критерий факторизации	2
3	Улучшение оценок с помощью достаточной статистики	2
3.1	Теорема Рао–Блэквелла–Колмогорова	2
3.1.1	Почему дисперсия уменьшается?	2
3.2	Следствия	3
4	Полные статистики	3
5	Теорема Лемана–Шеффе	4
6	Экспоненциальные семейства	4
6.1	Полнота в экспоненциальных семействах	5
7	Анциллярные статистики и теорема Басу	5
8	Основная схема	5

1 Оптимальные оценки

i Определение

Оценка $\hat{\theta}$ называется **оптимальной**, если она является наилучшей в классе несмещённых оценок в среднеквадратичном смысле.

Иными словами, если $\hat{\theta}$ — несмещённая оценка $\tau(\theta)$, то она оптимальна, если для любой другой несмещённой оценки θ^* величины $\tau(\theta)$ выполнено

$$\mathbb{E}_{\theta} (\hat{\theta} - \tau(\theta))^2 \leq \mathbb{E}_{\theta} (\theta^* - \tau(\theta))^2, \quad \theta \in \Theta.$$

Поскольку обе оценки несмещённые, это эквивалентно сравнению дисперсий:

$$\text{Var}_{\theta}(\hat{\theta}) \leq \text{Var}_{\theta}(\theta^*).$$

2 Достаточные статистики

Определение

Статистика $S(X)$ называется **достаточной** для семейства распределений $\mathcal{P}_\Theta$, если существует вариант условного распределения

$$\mathbb{P}_\theta(X \in B \mid S(X) = t),$$

который не зависит от параметра θ .

Таким образом, после того как известно значение $S(X)$, оставшаяся информация в выборке X не несёт дополнительной информации о θ .

2.1 Критерий факторизации

Теорема (критерий факторизации Неймана–Фишера)

Пусть $\mathcal{P}_\theta$ — доминируемое семейство распределений с обобщённой плотностью ρ_θ .

Тогда $S(X)$ является достаточной статистикой **тогда и только тогда**, когда обобщённая плотность допускает представление

$$\rho_\theta(x) = h(x) g_\theta(S(x)),$$

где для всех θ функции g_θ и h — борелевские и неотрицательные.

Критерий факторизации позволяет проверять достаточность, не вычисляя условное распределение выборки при фиксированном значении статистики.

3 Улучшение оценок с помощью достаточной статистики

3.1 Теорема Рао–Блэквелла–Колмогорова

Теорема (Колмогоров–Блэквелл–Рао)

Пусть $S(X)$ является достаточной статистикой для семейства

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\},$$

а $\theta^*(X)$ — несмещённая оценка $\tau(\theta)$.

Тогда

$$T(X) = \mathbb{E}_\theta(\theta^*(X) \mid S(X))$$

является статистикой и несмещённой оценкой $\tau(\theta)$, причём она не хуже исходной оценки в среднеквадратичном смысле:

$$\mathbb{E}_\theta (T - \tau(\theta))^2 \leq \mathbb{E}_\theta (\theta^* - \tau(\theta))^2, \quad \theta \in \Theta.$$

3.1.1 Почему дисперсия уменьшается?

По формуле полной дисперсии

$$\text{Var}_\theta(\theta^*) = \text{Var}_\theta(\mathbb{E}_\theta[\theta^* | S]) + \mathbb{E}_\theta[\text{Var}_\theta(\theta^* | S)].$$

Следовательно,

$$\text{Var}_\theta(T) \leq \text{Var}_\theta(\theta^*).$$

Таким образом, условное усреднение по достаточной статистике удаляет из оценки вариацию, не содержащую дополнительной информации о параметре.

3.2 Следствия

💡 Утверждение

Если $\tau(\theta)$ допускает оптимальное оценивание, то оптимальная оценка является функцией от достаточной статистики.

💡 Утверждение

Если $\hat{\theta}(X)$ и $\theta^*(X)$ — две оптимальные оценки одной и той же функции параметра, то

$$\hat{\theta}(X) = \theta^*(X) \quad P_\theta\text{-п.н.}$$

для любого $\theta \in \Theta$.

Таким образом, оптимальная несмещённая оценка, если она существует, по существу единственна.

4 Полные статистики

📖 Определение

Статистика $S(X)$ называется **полной** для семейства распределений $\mathcal{P}_\Theta$, если для любой борелевской функции f из

$$\mathbb{E}_\theta f(S(X)) = 0 \quad \text{для всех } \theta \in \Theta$$

следует

$$f(S(X)) = 0 \quad P_\theta\text{-п.н. для всех } \theta \in \Theta.$$

Иначе говоря,

$$[\forall \theta \in \Theta : \mathbb{E}_\theta f(S(X)) = 0] \implies [\forall \theta \in \Theta : f(S(X)) = 0 \quad (P_\theta\text{-п.н.})].$$

Полнота означает, что среди функций достаточной статистики не существует нетривиальной статистики, математическое ожидание которой равно нулю при всех значениях параметра.

5 Теорема Лемана–Шеффе

! Теорема (Лемана–Шеффе)

Пусть $S(X)$ — **полная достаточная статистика** для семейства $\mathcal{P}_\Theta$.
Если

$$\mathbb{E}_\theta \varphi(S(X)) = \tau(\theta),$$

то

$$\hat{\theta}(X) = \varphi(S(X))$$

является оптимальной несмещённой оценкой для $\tau(\theta)$.

Таким образом, для построения оптимальной несмещённой оценки достаточно:

1. найти полную достаточную статистику $S(X)$;
2. найти функцию φ , для которой

$$\mathbb{E}_\theta \varphi(S(X)) = \tau(\theta).$$

Тогда $\varphi(S(X))$ автоматически является единственной (с точностью до равенства почти наверное) оптимальной несмещённой оценкой.

6 Экспоненциальные семейства

Распределения экспоненциального семейства естественным образом порождают достаточные статистики.

В общем виде плотность можно записать как

$$p_\theta(x) = h(x) \exp \left\{ \sum_{j=1}^k a_j(\theta) T_j(x) - A(\theta) \right\}.$$

Для независимой выборки $X_1, \dots, X_n$ совместная плотность имеет вид

$$p_\theta(x_1, \dots, x_n) = \left(\prod_{i=1}^n h(x_i) \right) \exp \left\{ \sum_{j=1}^k a_j(\theta) \sum_{i=1}^n T_j(x_i) - nA(\theta) \right\}.$$

По критерию факторизации естественной достаточной статистикой является

$$T(X) = \left(\sum_{i=1}^n T_1(X_i), \dots, \sum_{i=1}^n T_k(X_i) \right).$$

6.1 Полнота в экспоненциальных семействах

! Теорема об экспоненциальных семействах

Пусть

$$\theta \in \Theta \subset \mathbb{R}^k$$

и P_θ — семейство распределений экспоненциального типа.

Пусть, кроме того, множество значений

$$(a_1(\theta), \dots, a_k(\theta)), \quad \theta \in \Theta,$$

содержит k -мерный параллелепипед.

Тогда

$$T(X) = (T_1(X), \dots, T_k(X))$$

является **полной достаточной статистикой**.

Это особенно полезно в сочетании с теоремой Лемана–Шеффе: если удаётся представить несмещённую оценку как функцию такой статистики, оптимальность следует автоматически.

7 Анциллярные статистики и теорема Басу

i Определение

Статистика $A(X)$ называется **анциллярной** (*ancillary statistic*), если её распределение не зависит от параметра θ .

! Теорема Басу

Пусть $S(X)$ — полная достаточная статистика, а $A(X)$ — анциллярная статистика.

Тогда для любого $\theta \in \Theta$ статистики

$$A(X) \quad \text{и} \quad S(X)$$

независимы.

Теорема Басу связывает два противоположных типа статистик:

- $S(X)$ содержит всю доступную информацию о параметре θ ;
- распределение $A(X)$ вообще не зависит от θ .

Если $S(X)$ при этом полна, то эти статистики оказываются независимыми.

8 Основная схема

Связь между основными понятиями можно представить следующим образом:

$$\boxed{\text{Достаточная статистика}} + \boxed{\text{Полнота}} \implies \boxed{\text{Полная достаточная статистика}}.$$

Если существует функция φ такая, что

$$\mathbb{E}_\theta \varphi(S) = \tau(\theta),$$

то по теореме Лемана–Шеффе

$$\boxed{\varphi(S) \text{ — оптимальная несмещённая оценка } \tau(\theta) .}$$

Кроме того, если A — анциллярная статистика, то по теореме Басу

$$\boxed{S \perp A} .$$