Я не смог придумать смешное название...

Байесовский подход

Чертоги разума

Модель: случайный параметр θ и выборка $\mathbf{X} = (X_1, \ldots, X_n)$. Плотность их совместного распределения можно расписать как

$$\rho_{\mathbf{X}|\theta}(\mathbf{x}, t) = \rho_\theta(t) \cdot \rho_{\mathbf{X}|\theta}(\mathbf{x}|t),$$

где:

- $\rho_\theta(t)$ — плотность *априорного распределения* на Θ;
- $\rho_{\mathbf{X}|\theta}(\mathbf{x}|t) = \prod_{i=1}^n \rho_{X_i|\theta}(x_i|t)$ — распределение выборки при фиксированном значении $\theta = t$.

Для оценки θ используется *апостериорное распределение*, то есть условное распределение при условии увиденной выборки:

$$\rho_{\theta|\mathbf{X}}(t|\mathbf{x}) = \frac{\rho_{\mathbf{X},\theta}(t|\mathbf{x})}{\rho_{\mathbf{X}}(\mathbf{x})} = \frac{\rho_{\mathbf{X},\theta}(\mathbf{x}, t)}{\int_\Theta \rho_{\mathbf{X},\theta}(\mathbf{x}, s)\, ds} = \frac{\rho_\theta(t) \cdot \rho_{\mathbf{X}|\theta}(\mathbf{x}|t)}{\int_\Theta \rho_\theta(s) \cdot \rho_{\mathbf{X}|\theta}(\mathbf{x}|s)\, ds}.$$

Если нужна точечная оценка, то обычно от апостериорного распределения берут матожидание:

$$\widehat{\theta}(\mathbf{x}) = \int_\Theta t \cdot \rho_{\theta|\mathbf{X}}(t|\mathbf{x})\, dt = \mathsf{E}(\theta|\mathbf{X} = \mathbf{x}).$$

Такую оценку называют *байесовской.*

Теорема. Байесовская оценка является наилучшей в байесовском подходе с квадратичной функции потерь, то есть

$$\widehat{\theta} = \arg\min_{\theta^*} \int_\Theta \mathsf{E}_t(\theta^*(\mathbf{X}) - t)^2 \cdot \rho_\theta(t)\, dt.$$

Какое априорное распределение выбрать?

- Придумать такое семейство распределений $\mathcal{Q}$, что если $\rho_\theta(t) \in \mathcal{Q}$, то и $\rho_{\theta|\mathbf{X}}(t|\mathbf{x}) \in \mathcal{Q}$ — в таком случае не придётся мучиться с интегралом в формуле апостериорной плотности. Такие семейства называют *сопряжёнными* к семейству $\{\rho(x|t)\}$.
- Взять *распределение Джеффриса*, чья плотность пропорциональна $\sqrt{\det I_{\mathbf{X}}(\theta)}$ — это своего рода равномерное распределение в пространстве $\mathcal{P}$.

  Замечание. Иногда $\sqrt{\det I_{\mathbf{X}}(\theta)}$ не интегрируемо на Θ, то есть эта плотность задаёт всего лишь меру, но при этом зачастую формула апострериорного распределения выше задаёт честную вероятностную меру на прямой.

Zadania

1. Найдите семейство распределений, сопряжённое
(а) $\{\mathrm{Bern}(p) \colon p \in (0,1)\}$;
(б) $\{\mathcal{N}(\theta, 1) \colon \theta \in \mathbb{R}\}$;
(в) $\{\mathcal{N}(0, 1/\theta) \colon \theta > 0\}$,
и найдите байесовскую оценку для параметра θ с найденным априорным сопряжённым распределением.

2. Посчитайте распределение Джеффриса в моделях
(а) $\{\mathrm{Bern}(p) \colon p \in (0,1)\}$;
(б) $\{\mathcal{N}(\theta, 1) \colon \theta \in \mathbb{R}\}$;
(в) $\{\mathcal{N}(0, \theta^2) \colon \theta > 0\}$
и найдите байесовскую оценку для параметров.

3. Пусть $T(\mathbf{X})$ — достаточная статистика. Докажите, что байесовская оценка является функцией от достаточной статистики, то есть

$$\mathsf{E}(\theta | \mathbf{X}) = \varphi(T(\mathbf{X})).$$

4. Докажите, что для любого априорного распределения на $\mathbb{R}$ в модели $X_i \sim \mathcal{N}(\theta, 1)$ выборочное среднее $\overline{X}$ не может быть байесовской оценкой.