Введение в байесовскую статистику

Семинар был посвящен обсуждению ключевых отличий и преимуществ байесовской статистики по отношению к классической частотной.

Мы рассмотрели класический статистический эксперимент Фишера «Леди, дегустирующая чай» в разных постановках и убедидись в том, что порой априорная информация о структуре эксперимента прямо влияет на нашу «меру удивления» от его результатов. Байесовская постановка задач математической статистики максимально учитывает это априорную информацию.

Затем, мы проанализировали байесовский подход к решению задач проверки гипотез и оценивания неизвестных параметров на примере одной модельной задачи.

Пусть $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ — выборка, причем X_1 имеет распределение с плотностью f_θ, где $\theta \in \Theta \subseteq \mathbb{R}^d, d \geq 1$.

Байесовская постановка задачи проверки гипотез. Задача состоит в том, чтобы различить простую основную гипотезу

$$H_0 = \{\theta = \theta_0\}$$

и простую альтернативу

$$H_1 = \{\theta = \theta_1\}.$$

Предположим, что нам известна некоторая дополнительная информация об этих гипотезах, а имеено, их априорное распределение (т.е., вероятность $\mathbf{P}(H_0) = 1 - \mathbf{P}(H_1)$). Байесовский критерий в этом случае имеет вид:

$$\delta(\boldsymbol{X}) = \begin{cases} H_0, \text{ если } \mathbf{P}(H_0|\boldsymbol{X})/\mathbf{P}(H_1|\boldsymbol{X}) < c; \\ H_1, \text{ если } \mathbf{P}(H_0|\boldsymbol{X})/\mathbf{P}(H_1|\boldsymbol{X}) \geq c. \end{cases}$$

Классические «бенчмарки» для константы c таковы: при $c = 3$ имеются существенные основания отвергнуть основную гипотезу, а при $c = 10$ — сильные основания.

Байесовская постановка задачи оценивания параметров. Задача состоит в том, чтобы по заданной выборке и некоторой априорной информации построить оценку параметра θ. В основе построения байесовских оценок лежит формула Байеса:

$$\mathbf{P}(\theta|\boldsymbol{X}) = \frac{L(\boldsymbol{X};\theta)\pi(\theta)}{\int_\Theta L(\boldsymbol{X};\theta)\pi(\theta)d\theta},$$

где $L(\boldsymbol{X};\theta) = \prod_{j=1}^{n} f_\theta(X_j)$ — функция правдоподобия, а $\pi(\theta)$ — некоторое априорное распределение параметра θ. Сами оценки при этом могут иметь следующий вид:

$$\theta^* = \mathbf{E}\left(\theta|\boldsymbol{X}\right) \qquad \text{или } \theta^{**} = \arg\max_{\theta\in\Theta} \mathbf{P}\left(\theta|\boldsymbol{X}\right).$$

Наконец, мы кратко обсудили три способа получения априорной информации: субъективный, объективный и частотный.