hw_markevich_21

środa, 27 listopada 2024 13:02

Задание 1

1. Апостериорное распределение для параметра распределения Пуассона.

Количество срабатываний счетчика Гейгера за минуту n подчиняется распределению Пуассона:

$$P_\lambda(n) = \frac{\lambda^n}{n!}e^{-\lambda}.$$

1. В ходе эксперимента счетчик Гейгера сработал за минуту m раз. С помощью теоремы Байеса определите апостериорное распределение на λ. Указание: априорную плотность вероятности λ можно считать постоянной (так как мы изначально ничего не знаем про λ) .
2. Эксперимент повторили еще раз, в этот раз счетчик Гейгера сработал за минуту m' раз. Как обновилось апостериорное распределение на λ ?

Примечание.

Такая плотность вероятности не будет нормируема. Чтобы сделать рассуждение более строгим, можно ввести обрезку на очень больших λ (так как это нереалистичные значения). Другими словами, можно считать, что априорная плотность вероятности $p_0(\lambda)-$ это какая-то очень медленно меняющаяся функция и как-то убывающая на бесконечности. Тогда в числителе и знаменателе формулы Байеса она будет домножаться на гораздо более быструю функцию и поэтому можно заменить $p_0(\lambda) \to p_0(0)$. Константа $p_0(0)$ должна сократиться в ходе вычислений.

$$P_\lambda(n) = \frac{\lambda^n}{n!}\exp(-\lambda); \quad p(\lambda(m)) = \frac{p_\lambda(m)\cdot p(\lambda)}{\int d\lambda\, p_\lambda(m)\, p(\lambda)} = \frac{p(m|\lambda)\cdot p_0(0)}{\int d\lambda\, p(m|\lambda)\cdot p_0(0)} =$$

$$= \frac{\frac{\lambda^m}{m!}e^{-\lambda}}{\int d\lambda\, p_\lambda(m)} = \frac{\lambda^m}{m!}e^{-\lambda}$$

$$P'(\lambda) = P(\lambda|m) \Longrightarrow P'(\lambda|m') = \frac{P'(m'|\lambda)P'(\lambda)}{P'(m')} = \frac{P'(m'|\lambda)\,P'(\lambda)}{\int d\lambda\, P(m'|\lambda)P'(\lambda)} =$$

$$= \frac{\frac{\lambda^{m'+m}}{m'!\,m!}e^{-2\lambda}}{\int d\lambda\, \frac{\lambda^{m'+m}}{m'!\,m!}e^{-2\lambda}} = \frac{\lambda^{m'+m}e^{-2\lambda}}{2^{-m-m'-1}(m'+m)!}$$

Задание 2

2. Апостериорное распределение для аргумента нормального распределения.

Пусть имеется априорное распределение на вектор $\boldsymbol{x}$, задаваемое симметричной положительно определенной матрицей A :

$$p_0(\boldsymbol{x}) = \frac{1}{Z}e^{-\frac{\boldsymbol{x}^T A \boldsymbol{x}}{2}}.$$

Было произведено измерение величин $\boldsymbol{x}$, которое дало значение $\boldsymbol{x}_1$. Найдете апостериорное распределение на $\boldsymbol{x}$.

Hint: Мы изначально ничего не знаем об A - надо ввести на него какое-то априорное распределение. Предлагается брать равномерное.

$$p_0(\bar{x}) = \frac{1}{Z}\exp\left(-\frac{1}{2}\bar{x}^T A\bar{x}\right)$$

$$P(A|X_1) = \frac{P(X_1|A)\,P(A)}{P(X_1)} = \frac{P(X_1|A)\,P(A)}{\int dA\, p(X_1|A)\,P(A)}$$

равномерное распределение на A: $P(A) = \alpha \in \mathbb{R}$

тогда $$P(A|X_1) = \frac{\exp\left(-\frac{1}{2}\bar{X}_1^T A\bar{X}_1\right)}{\int dA \exp\left(-\frac{1}{2}\bar{X}_1^T A\bar{X}_1\right)} = \frac{\exp\left(-\frac{1}{2}\lambda^i(x_1^i)^2\right)}{\int d\lambda^1 \ldots d\lambda^n \exp\left(-\frac{1}{2}\lambda^i(x_1^i)^2\right)} =$$

$$= // \int_0^\infty d\lambda_i \exp\left(-\frac{1}{2}\lambda^i(x_1^i)^2\right) = \frac{1}{(X_1^i)^2} // = \frac{\exp\left(-\frac{1}{2}\lambda^i(x_1^i)^2\right)}{(X_1^1 \cdot \ldots \cdot X_1^n)^2} =$$

$$= \frac{\exp\left(-\frac{1}{2}\bar{X}_1^T A\bar{X}_1\right)}{(X_1^1 \cdot \ldots \cdot X_1^n)^2}$$

$P(x|X_1) = \int [illegible]$

$$= \frac{\exp(\ldots)}{(x_1^1 \cdot \ldots \cdot x_1^n)^2}$$

$$P(x|x_1) = \int dA$$

Задание 5

5. Лассо Тибширани: связь L1 и L2 между собой.

Покажите, что задача минимизации квадратичной функции потерь с дополнительным ограничением (лассо Тибширани):

$$\mathcal{L} = \|Xw - y\|^2 \to \min_w, \quad \sum_\alpha |w_\alpha| < C$$

эквивалентна L1-регуляризации. Указание: можно воспользоваться условиями Каруша - Куна — Таккера (обобщение метода Лагранжа). Link .

$L = \|Xw - y\|^2 \to \min\limits_w ; \quad \sum\limits_\alpha |w_\alpha| < C$

$\exists \hat{x}$ – искомый $\Rightarrow$ из необходимого условия

Каруша–Куна–Таккера: (для $\tilde{L}(x) = \lambda_0 L + \sum \lambda_i g_i(x)$)

$\min \tilde{L}(x) = \tilde{L}(\hat{x})$ – это буквально L_1 – регуляризация