Третье задание по БМСО

Ярослав Аверьянов

Октябрь 2015

1 Первая задача

$$||P - Q||_{TV} = 2 \cdot \underset{B \in \mathbf{B}}{sup} |P(B) - Q(B)| \text{ - общая дисперсия } P \text{ и } Q.$$

Для непрерывных распределений P и Q с плотностями p и q обозначим

$$||P - Q||_d = \int |p - q| d\mu$$

Докажем эквивалентность $|| \cdot ||_{TV}$ и $|| \cdot ||_d$ для непрерывных распределений.

Пусть $A = [x | p(x) - q(x) \leq 0] \in \mathbf{B}$ - замкнутое и измеримое множество.

Пусть $B \in \mathbf{B}$:

$P(B) - Q(B) = \int_B (p(x) - q(x))dx = \int_A (p(x) - q(x))dx + \int_{B\ A}(p(x) - q(x))dx - \int_{A\ B}(p(x) - q(x))dx$

Где $\int_{B\ A}(p(x) - q(x))dx \leq 0, \int_{A\ B}(p(x) - q(x))dx \geq 0 \rightarrow P(B) - Q(B) \leq \int_A (p(x) - q(x))dx = P(A) - Q(A)$

$\overline{A} = x | p(x) - q(x) \geq 0$:

$Q(B) - P(B) \leq Q(\overline{A}) - P(\overline{A}) = P(A) - Q(A)$.

Поэтому $|P(B) - Q(B)| \leq P(A) - Q(A)$. $\rightarrow ||P - Q||_{TV} = 2 \cdot \underset{B \in \mathbf{B}}{sup} |P(B) - Q(B)| = 2 \cdot (P(A) - Q(A)) = P(A) - Q(A) - (1 - P(A)) + (1 - Q(A)) = \int_A (p - q) d\mu + \int_{\overline{A}} (p - q) d\mu = ||p - q||_d$.

2 Вторая задача

Пусть $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$. Используя то, что $d_{TV}(X, Y) = \frac{1}{2}\int_{-\infty}^{+\infty} |f(x) - g(x)|dx$, где f и g - плотности для X и Y. Далее предположим, что $\sigma_1 > \sigma_2$.

Введем следующие обозначения:

$$c = \frac{\sigma_1}{\sigma_2} \qquad \Delta = \frac{\mu_1 - \mu_2}{\sigma_2}$$
$$A = \frac{\sqrt{(c^2-1)2log(c)+\Delta^2} - c\Delta}{c^2-1}, \qquad B = -\frac{\sqrt{(c^2-1)2log(c)+\Delta^2} + c\Delta}{c^2-1}$$

Используя эти обозначения:

$$\int_{-\infty}^{+\infty} |f(x) - g(x)|dx = \int_{-\infty}^{+\infty} |\phi(z) - c\phi(\Delta + cz)|dz$$

И $\phi(z) \leq c\phi(\Delta + cz) <--> A \leq z \leq B$. Следовательно
$\int_{-\infty}^{+\infty} |f(x) - g(x)|dx = \int_A^B [c\phi(\Delta + cz) - \phi(z)]dz + \int_{-\infty}^A [\phi(z) - c\phi(\Delta + cz)]dz + \int_B^{+\infty} [\phi(z) - c\phi(\Delta + cz)]dz = 2[(\Phi(\Delta + cB) - \Phi(B)) - (\Phi(\Delta + cA) - \Phi(A))]$.
Где

$$\Delta + cB = \frac{c\sqrt{(c^2-1)2log(c)+\Delta^2} - \Delta}{c^2-1}$$
$$\Delta + cA = -\frac{c\sqrt{(c^2-1)2log(c)+\Delta^2} + \Delta}{c^2-1}$$

Если $\sigma_1 = \sigma_2$, формула сокращается до $\Phi(\frac{|\Delta|}{2}) - \Phi(-\frac{|\Delta|}{2})$

Следовательно, $d_{TV}(X, Y) = (\Phi(\Delta + cB) - \Phi(B)) - (\Phi(\Delta + cA) - \Phi(A))$

3 Третья задача

Показать, что если:

У1) множество параметров $\theta \in \Theta \subset \mathbf{R}$ - ограниченное открытое множество.

У2) $\exists \sigma > 0 : \forall \theta \in \Theta$ информация Фишера $I(\theta)$:

$$I(\theta) > \sigma > 0,$$

то для экспоненциального семейства распределений выполнено "предположение (5)"теоремы Бернштейна-фон-Мизеса.

Плотность распределения: $p(x|\theta) = exp(-c(\theta) + \theta t(x))$

$1 = \int exp(-c(\theta) + \theta t(x))dx, c(\theta) = ln \int exp(\theta t(x))dx$ - нормировка

$$c^{'}(\theta) = \frac{\int t(x)exp(\theta t(x))dx}{\int exp(\theta t(x))dx} = \int exp(-c(\theta) + \theta t(x))t(x)dx = \mathbf{E}_\theta t(x)$$

Посчитаем информацию Фишера и логарифм правдоподобия:

$$I(\theta) = -\mathbf{E}[l^{''}(x|\theta)] = c^{''}(\theta) > \sigma > 0$$
$$L_n(\theta) = lnL(x_1, ..., x_n|\theta) = -nc(\theta) + \theta \sum_{i=1}^{n} t(x_i)$$
$$\frac{1}{n}(L_n(\theta) - L_n(\theta_0)) = c(\theta) - c(\theta_0) + \frac{1}{n}(\theta - \theta_0) \sum_{i=1}^{n} t(x_i).$$

Так как

$$c(\theta) - c(\theta_0) \geq (\theta - \theta_0)\mathbf{E}_\theta t(x) + \frac{\sigma}{2}(\theta - \theta_0)^2$$

Получаем:

$$\frac{1}{n}(L_n(\theta) - L_n(\theta_0)) \leq (\theta - \theta_0)(-\mathbf{E}_\theta t(x) + \frac{1}{n} \sum_{i=1}^{n} t(x_i)) - \frac{\sigma}{2}(\theta - \theta_0)^2$$

Т.к. $\frac{1}{n} \sum_{i=1}^{n} t(x_i) \xrightarrow{P} \mathbf{E}t(x), |\theta - \theta_0| < \gamma \rightarrow \forall \delta > 0 \exists \epsilon > 0 :$

$$P_{\theta_0}(\sup_{|\theta - \theta_0| < \delta} \frac{1}{n}(L_n(\theta) - L_n(\theta_0)) \leq -\epsilon) \geq P_{\theta_0}(\gamma| - \mathbf{E}t(x) + \frac{1}{n} \sum_{i=1}^{n} t(x_i)| - \frac{\sigma}{2}\delta^2 <$$
$$-\epsilon) = P_{\theta_0}(| - \mathbf{E}t(x) + \frac{1}{n} \sum_{i=1}^{n} t(x_i)| < \frac{1}{\gamma}(\frac{\sigma}{2}\delta^2 - \epsilon)) \rightarrow 1, \text{ где } \epsilon < \frac{\sigma}{2}\delta^2.$$

4 Четвертая задача

Получить оценку максимального правдоподобия $\hat{\theta_n}$ и байесовскую оценку $\theta_n^* = \int \theta p(\theta|x_1, ..., x_n)d\theta$ для среденего μ нормального распределения $N(\mu, \sigma^2)$ и априорного распределения $\theta \sim N(\mu_\Pi, \sigma_\Pi^2)$.

Посчитаем правдоподобие:

$$L(x_1, ..., x_n|\theta) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} exp(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2})$$

Оценка максимального правдоподобия:

$$\hat{\mu_n} = \underset{\mu}{argmax} L(x_1, ..., x_n|\mu) = \underset{\mu}{argmax} exp(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}) =$$
$$\underset{\mu}{argmax} \sum_{i=1}^{n}(x_i - \mu)^2 \rightarrow \hat{\mu_n} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Байесовская оценка $\mu_n^* = \int \mu \cdot p(\mu|x_1,...,x_n) d\mu$

Априорное распределение $\mu \sim N(\mu_\Pi, \sigma_\Pi^2)$

Апостериорное распределение:

$$p(\mu|x_1,...,x_n) \propto L(x_1,...,x_n|\mu) \cdot p(\mu) \propto exp(-\frac{\sum\limits_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} - \frac{(\mu^2-2\mu\mu_\Pi+\mu_\Pi^2)}{2\sigma_\Pi^2}) \propto$$
$$exp(-\frac{\mu^2}{2}(\frac{1}{\sigma^2}+\frac{n}{\sigma^2}) + \mu(\frac{\mu}{\sigma^2}+\frac{\sum\limits_i x_i^2}{2\sigma^2})) = exp(-\frac{(\mu^2-2\mu\mu_n^*+\mu_n^{*2})}{2\sigma_n^{*2}}) =$$
$$exp(-\frac{(\mu^2-2\mu\mu_n^*+\mu_n^{*2})}{2\sigma_n^{*2}}) = exp(-\frac{(\mu-\mu_n^*)^2}{2\sigma_n^{*2}})$$

Байесовская оценка μ:

$$\mu_n^* = \frac{\frac{\mu_\Pi}{\sigma_\Pi^2}+\frac{\sum\limits_i x_i}{\sigma^2}}{\frac{1}{\sigma_\Pi^2}+\frac{n}{\sigma^2}}$$

5 Пятая задача

Сравнить сходимость к нормальному распределению величин $\sqrt{n}(\theta_0 - \theta_n^*)$ и $\sqrt{n}(\theta_0 - \hat{\theta_n})$.

Рассмотрим оценку максимального правдоподобия:
$\hat{\theta_n} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i, \hat{\theta_n} \sim N(\theta_0, \frac{\sigma^2}{n}), \sqrt{n}(\theta_0 - \hat{\theta_n}) \sim N(0, \sigma^2)$

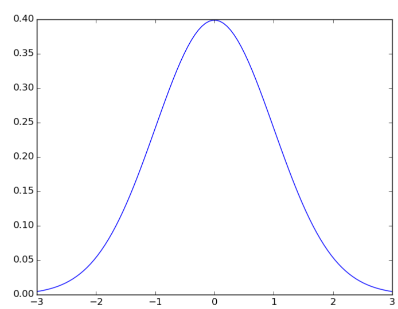

Рис. 1: Стандартное нормальное распределение N(0,1)

Перейдем к байесовской оценке:

Пусть $\theta \sim N(\theta_\Pi, \sigma_\Pi^2) \to \theta_n^* = \frac{\frac{\theta_\Pi}{\sigma_\Pi^2}+\frac{\sum\limits_i x_i}{\sigma^2}}{\frac{1}{\sigma_\Pi^2}+\frac{n}{\sigma^2}}, \theta_n^* \sim N(\frac{\frac{\theta_\Pi}{\sigma_\Pi^2}+\frac{n\theta_0}{\sigma^2}}{\frac{1}{\sigma_\Pi^2}+\frac{n}{\sigma^2}}, \frac{n\sigma^2}{(\frac{\sigma^2}{\sigma_\Pi^2}+n)^2})$

$(\theta_0 - \hat{\theta_n^*}) \sim N(\frac{\frac{1}{\sigma_\Pi^2}(\theta_\Pi-\theta_0)}{\frac{1}{\sigma_\Pi^2}+\frac{n}{\sigma^2}}, \frac{n\sigma^2}{(\frac{\sigma^2}{\sigma_\Pi^2}+n)^2}) \sim N(\frac{(\theta_\Pi-\theta_0)}{1+\frac{n\sigma_\Pi^2}{\sigma^2}}, \frac{n\sigma^2}{(\frac{\sigma^2}{\sigma_\Pi^2}+n)^2})$.
Т.к. $\theta_0 \sim N(\theta_\Pi, \sigma_\Pi^2) \rightarrow$
$(\theta_0 - \hat{\theta_n^*}) \sim N(0, \frac{n\sigma^2}{(\frac{\sigma^2}{\sigma_\Pi^2}+n)^2} + \frac{\sigma_\Pi^2}{(1+\frac{n\sigma_\Pi^2}{\sigma^2})^2})$
Следовательно,
$\sqrt{n}(\theta_0 - \hat{\theta_n^*}) \sim N(0, \frac{n^2\sigma^2}{(\frac{\sigma^2}{\sigma_\Pi^2}+n)^2} + \frac{n\sigma_\Pi^2}{(1+\frac{n\sigma_\Pi^2}{\sigma^2})^2}) \sim N(0, \frac{n^2\sigma^2+\frac{n\sigma^4}{\sigma_\Pi^2}}{(\frac{\sigma^2}{\sigma_\Pi^2}+n)^2}) \sim N(0, \frac{n\sigma^2}{(\frac{\sigma^2}{\sigma_\Pi^2}+n)}) \sim N(0, \frac{\sigma^2}{\frac{\sigma^2}{n\sigma^2}+1})$.
Следовательно, для байесовской оценки мы получаем нормально распределенную случайную величину, с мат.ожиданием, равным нулю и дисперсией $\frac{\sigma^2}{\frac{\sigma^2}{n\sigma_\Pi^2}+1}$;
Возьмем $\sigma = 1, \sigma_\Pi = \frac{1}{2}$:

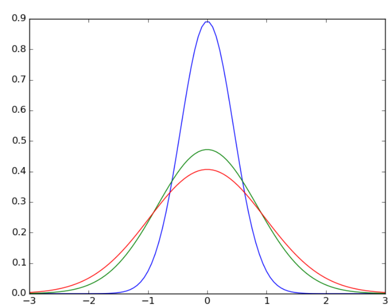

Рис. 2: N$(0, \frac{1}{1+\frac{4}{n}})$ при $n = 1, 10, 100$