Домашняя работа №1.

№1.

ОМП - оценка дисперсии:

$$\hat{\sigma}^2 = \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2}\right)^2, \quad \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Преобразуем:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = \frac{1}{n}\sum_{i=1}^{n}\Big((x_i - E\{x_1\}) - (\hat{\mu} - E\{x_1\})\Big)^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - E\{x_1\})^2 - \frac{2}{n}(\hat{\mu} - E\{x_1\})\sum_{i=1}^{n}(x_i - E\{x_1\}) +$$

$$+ (\hat{\mu} - E\{x_1\})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - E\{x_1\})^2 - (\hat{\mu} - E\{x_1\})^2$$

Найдём мат. ожидание:

$$E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - E\{x_1\})^2 - (\hat{\mu} - E\{x_1\})^2\right\} =$$

$$= E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - E\{x_1\})^2\right\} - E\left\{(\hat{\mu} - \overset{E\{\hat{\mu}\}}{E\{x_1\}})^2\right\} =$$

$$= E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i E\{x_1\} + (E\{x_1\})^2)\right\} - D\{\hat{\mu}\} =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(E\{x_i^2\} - (E\{x_1\})^2\right) - D\{\hat{\mu}\} = D\{x_1\} - D\{\hat{\mu}\} =$$

$$= D\{x_1\} - \frac{D\{x_1\}}{n} = \frac{n-1}{n}D\{x_1\} \neq D\{x_1\}$$

$\Rightarrow$ оценка я-ся смещённой.

№2.

1. Преобразуем:

$$\frac{1}{m}\sum_{i=1}^{m}\log\left(\frac{1}{P_\theta[x_i]}\right)+\frac{1}{m}\left(\log\left(\frac{1}{\theta}\right)+\log\left(\frac{1}{1-\theta}\right)\right)=$$

$$=-\frac{1}{m}\sum_{i=1}^{m}\log P_\theta(x_i)-\frac{1}{m}\log\theta-\frac{1}{m}\log(1-\theta)=$$

$$=-\frac{1}{m}\left(\sum_{i=1}^{m}\log P_\theta(x_i)+\log\theta+\log(1-\theta)\right)=$$

$$=\left[P_\theta(x_{m+1})=\theta,\ P_\theta(x_{m+2})=(1-\theta)\right]=-\frac{1}{m}\sum_{i=1}^{m+2}\log P_\theta(x_i)$$

Тогда если $x_{m+1}=1$, а $x_{m+2}=0$, то тогда ОМП-оценка представима в виде:

$$\hat\theta=\frac{1}{m+2}\sum_{i=1}^{m+2}x_i=\frac{1}{m+2}\left(\sum_{i=1}^{m}x_i+1\right)$$

$\cancel{\ }$

2. $|\hat\theta-\theta^*|=|\hat\theta-E\{\hat\theta\}+E\{\hat\theta\}-\theta^*|\leq\left[\text{неравенство треугольника}\right]\leq$

$$\leq|\hat\theta-E\{\hat\theta\}|+|E\{\hat\theta\}-\theta^*|$$

Из п. 1: $E\{\hat\theta\}=E\left\{\frac{1}{m+2}\left(\sum_{i=1}^{m}x_i+1\right)\right\}=$

$$=\frac{mE\{x_i\}}{m+2}+\frac{1}{m+2}=\frac{m\theta^*+1}{m+2}$$

Тогда получаем:

$$|\hat\theta-E\{\hat\theta\}|=\left|\frac{1}{m+2}\left(\sum_{i=1}^{m}x_i+1\right)-\frac{1+m\theta^*}{m+2}\right|=$$

$$=\frac{m}{m+2}\left|\frac{1}{m}\sum_{i=1}^{m}x_i-\theta^*\right|$$

$$|E\{\hat\theta\} - \theta^*| = \frac{1-2\theta^*}{m+2} \le \frac{1}{m+2}$$

Применим неравенство Хёфдинга к первому модулю:

$$P(|\hat\theta - E\{\hat\theta\}| \ge \varepsilon) \le 2e^{-2m\varepsilon^2}$$

$$P\left(|\hat\theta - \theta^*| \ge \frac{1}{m+2} + \varepsilon\right) \le 2e^{-2m\varepsilon^2}$$

Обозначим $\delta = P\left(|\hat\theta - \theta^*| \ge \frac{1}{m+2} + \varepsilon\right)$.

Тогда с вер-ю $1-\delta$

$$|\hat\theta - \theta^*| < \varepsilon + \frac{1}{m+2} = \sqrt{\frac{\log\left(\frac{2}{\delta}\right)}{2m}} + \frac{1}{m+2} = O\left(\sqrt{\frac{\log\left(\frac{1}{\delta}\right)}{m}}\right)$$

$\otimes$

№3. 1. $\max\limits_{c} \sum\limits_{y=1}^{k} \sigma_y \log c_y$, $c_y > 0$, $\sum\limits_y c_y = 1$ $\sigma \in \mathbb{R}^k_+$

Пусть $\sigma_y = \sum\limits_{i=1}^{m} P_{\theta^{(t)}}(Y=y \mid X=x_i)$, тогда на шаге M мы получаем задачу максимизации следующего вида:

$$\max_{c \in \mathbb{R}^k} \sum_{i=1}^{k} \sigma_y(\log(c_y) - C_i), \quad c_y \ge 0, \sum_y c_y = 1$$

где $C_i = \frac{1}{2}\|x_i - \mu_y\|^2$.

Получили формулу из условия как задачу максимизации на шаге M метода soft K-Means.

2. $c^* = \frac{1}{\sum\limits_y \sigma_y} \sigma$; т.к. $\sigma_y \ge 0$, то $c^* \ge 0$. Также рассмотрим $\sum\limits_y c_y^* = \frac{1}{\sum\limits_y \sigma_y} \sum\limits_y \sigma = 1$.

$\Rightarrow c^*$ — вектор вероятности