宿題 1

微分と積分の順序が交換できるとき，不偏推定量 $\hat{\theta}$ について，次式を示す。

$$\mathbb{E}_{\boldsymbol{x}_1,\ \ldots,\ \boldsymbol{x}_n \sim q(\cdot;\ \theta)} \left[\hat{\theta} \sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log q(\boldsymbol{x}_i;\ \theta) \Big|_{\theta=\theta^*} \right] = 1 \tag{1}$$

i.i.d. な訓練標本 $\{\boldsymbol{x}_i\}_{i=1}^{n}$ がモデル $q(\boldsymbol{x};\ \theta)$ から生起するとき，その確率は，

$$p(\boldsymbol{x}_1,\ \ldots,\ \boldsymbol{x}_n) = \prod_{i=1}^{n} q(\boldsymbol{x}_i;\ \theta) \tag{2}$$

となる。このとき，対数尤度とその θ による微分は，

$$L(\theta) = \log p = \sum_{i=1}^{n} \log q(\boldsymbol{x}_i;\ \theta) \tag{3}$$

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta}(\log p) = \sum_{i=1}^{n} \frac{\partial}{\partial \theta}(\log q(\boldsymbol{x}_i;\ \theta)) \tag{4}$$

である。

いま，式 (1) の左辺は次のように変形できる。

$$\mathbb{E}_{\boldsymbol{x}_1,\ \ldots,\ \boldsymbol{x}_n \sim q(\cdot;\ \theta)} \left[\hat{\theta} \sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log q(\boldsymbol{x}_i;\ \theta) \Big|_{\theta=\theta^*} \right] = \mathbb{E}_{\boldsymbol{x} \sim p(\cdot;\ \theta)} \left[\hat{\theta}\ \frac{\partial}{\partial \theta}(\log p) \Big|_{\theta=\theta^*} \right] \tag{5}$$

$$= \int p \hat{\theta}\ \frac{\partial}{\partial \theta}(\log p) \Big|_{\theta=\theta^*} \mathrm{d}\boldsymbol{x}$$

$$= \int \hat{\theta} \left\{ p \frac{\partial}{\partial \theta}(\log p) \right\} \Big|_{\theta=\theta^*} \mathrm{d}\boldsymbol{x} \tag{6}$$

また，$\mathbb{E}\left[\hat{\theta}\right] = \theta^*$ の両辺を θ^* で偏微分することを考える。微分と積分の順序が交換できるとき，

$$1 = \frac{\partial}{\partial \theta^*} \mathbb{E}\left[\hat{\theta}\right] \tag{7}$$

$$= \frac{\partial}{\partial \theta^*} \int p \hat{\theta}\ \mathrm{d}\boldsymbol{x} \tag{8}$$

$$= \int \frac{\partial}{\partial \theta^*} \left(p \hat{\theta}\right) \mathrm{d}\boldsymbol{x} \tag{9}$$

$$= \int \hat{\theta} \frac{\partial p}{\partial \theta^*}\ \mathrm{d}\boldsymbol{x} \tag{10}$$

となる。いま，

$$p(\boldsymbol{x};\ \theta) \left(\frac{\partial}{\partial \theta} \log p(\boldsymbol{x};\ \theta) \right) = \frac{\partial}{\partial \theta} p(\boldsymbol{x}; \theta) \tag{11}$$

であるから，結局，

$$\int \hat{\theta} \left\{ p \frac{\partial}{\partial \theta}(\log p) \right\} \Big|_{\theta=\theta^*} \mathrm{d}\boldsymbol{x} = 1 \tag{12}$$

となる。

式 (6)，(12) から，

$$\mathbb{E}_{\boldsymbol{x}_1,\,\ldots,\,\boldsymbol{x}_n\sim q(\cdot;\,\theta)}\left[\hat{\theta}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log q(\boldsymbol{x}_i;\,\theta)\bigg|_{\theta=\theta^*}\right]=\int\hat{\theta}\left\{p\frac{\partial}{\partial\theta}(\log p)\right\}\bigg|_{\theta=\theta^*}\,\mathrm{d}\boldsymbol{x}=1 \tag{13}$$

であることが示された。