

(科目:机器学习) 清华大学数学作业纸

编号: 班级: 姓名: 第 页

$$2.4. \text{二元分布 } P(X=x) = \mu^x (1-\mu)^{1-x} \quad P(X=1) = \mu \quad P(X=0) = 1-\mu$$

$$E(X) = \int x P(x) dx = 1 \cdot P(X=1) + 0 \cdot P(X=0) = \mu$$

$$\begin{aligned} \text{Var}(X) &= E[(X - EX)^2] = \int (x - \mu)^2 P(x) dx \\ &= \int [(1-\mu)^2 \cdot \mu + \mu^2 (1-\mu)] dx \\ &= \mu - \mu^2 \\ &= \mu(1-\mu) \end{aligned}$$

$$\text{二项分布 } P(Y=m) = \binom{N}{m} \mu^m (1-\mu)^{N-m}$$

$$Y = X_1 + X_2 + \dots + X_N, \quad X_i \sim B(x|\mu), \quad i=1, 2, \dots, N$$

$$\begin{aligned} EY &= E(X_1 + X_2 + \dots + X_N) = EX_1 + EX_2 + \dots + EX_N \\ &= NM \end{aligned}$$

$$\text{Var}(Y) = E[(Y - EY)^2] = E(Y^2) - (EY)^2$$

$$\begin{aligned} E(Y^2) &= \int Y^2 P(Y) dy \\ &= \int m^2 \binom{N}{m} \mu^m (1-\mu)^{N-m} dm \\ &= \int [m(m-1)+m] \frac{N!}{m!(N-m)!} \mu^m (1-\mu)^{N-m} dm \\ &= \int \frac{N!}{(m-2)!(N-m)!} \mu^m (1-\mu)^{N-m} dm + m \int \frac{N!}{m!(N-m)!} \mu^m (1-\mu)^{N-m} dm \\ &= N(N-1)\mu^2 \int \frac{(N-2)!}{(m-2)![N-2-(m-2)]!} \mu^{m-2} (1-\mu)^{[(N-2)-(m-2)]} dm \\ &\quad + \int m P(Y=m) dm \end{aligned}$$

$$= N(N-1)\mu^2 + NM$$

$$= N^2\mu^2 - NM^2 + NM$$

$$\text{Var}(Y) = E(Y^2) - (EY)^2$$

$$= N^2\mu^2 - NM^2 + NM - N^2\mu^2$$

$$= NM(1-\mu)$$

△

编号:

班级:

姓名:

第 页

$$2.7 \quad \{x_n\}_{n=1}^N, i.i.d., 1 \leq x_n \leq b$$

记6个面的概率分别为 $M_i$ ,  $1 \leq i \leq 6$ ,  $\sum_{i=1}^6 M_i = 1$

$\{x_n\}$ 符合二元分布, 概率函数为

$$P(x_n | M_i) = M_i^{x_n=i} (1-M_i)^{x_n \neq i}$$

$$L(M_i | x) = \prod_{n=1}^N M_i^{x_n=i} (1-M_i)^{x_n \neq i}$$

$$= M_i^{\sum_{n=1}^N x_n=i} (1-M_i)^{\sum_{n=1}^N x_n \neq i}$$

$$\frac{\partial \log L(M_i | x)}{\partial M_i} = \frac{\partial \left\{ \sum_{n=1}^N x_n=i \log M_i + \sum_{n=1}^N x_n \neq i \log (1-M_i) \right\}}{\partial M_i}$$

$$= \frac{1}{M_i} \sum_{n=1}^N x_n=i - \frac{1}{1-M_i} \sum_{n=1}^N x_n \neq i$$

$$= 0$$

$$(1-M_i) \sum_{n=1}^N x_n=i = M_i \sum_{n=1}^N x_n \neq i$$

$$M_i = \frac{1}{N} \sum_{n=1}^N x_n=i \quad (1 \leq i \leq 6)$$

$$2.9 \quad P(x_{1n}, x_{2n} | \rho) = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp \left( -\frac{x_{1n}^2 - 2\rho x_{1n} x_{2n} + x_{2n}^2}{2(1-\rho^2)} \right)$$

$$L(\rho | x_1, x_2) = \prod_{n=1}^N P(x_{1n}, x_{2n} | \rho)$$

$$= \left( \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \right)^N \exp \left( -\sum_{n=1}^N \frac{x_{1n}^2 - 2\rho x_{1n} x_{2n} + x_{2n}^2}{2(1-\rho^2)} \right)$$

$$\frac{\partial \log L(\rho | x_1, x_2)}{\partial \rho} = \frac{\partial \left\{ -N \log(2\pi(1-\rho^2)^{\frac{1}{2}}) - \sum_{n=1}^N \frac{x_{1n}^2 - 2\rho x_{1n} x_{2n} + x_{2n}^2}{2(1-\rho^2)} \right\}}{\partial \rho}$$

$$\begin{aligned} \frac{\partial}{\partial \rho} \left\{ -N \log(2\pi(1-\rho^2)^{\frac{1}{2}}) \right\} &= -N \cdot \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \cdot \frac{1}{2}(1-\rho^2)^{-\frac{1}{2}} \cdot -2\rho \\ &= \frac{N\rho}{2\pi(1-\rho^2)} \end{aligned}$$

(科目: ) 清华大学数学作业纸

编号:

班级:

姓名:

第 页

$$\begin{aligned}
 \frac{\partial}{\partial p} \left\{ - \sum_{n=1}^N \frac{x_{1n}^2 - 2px_{1n}x_{2n} + x_{2n}^2}{2(1-p^2)} \right\} &= - \sum_{n=1}^N \frac{\partial}{\partial p} \frac{x_{1n}^2 - 2px_{1n}x_{2n} + x_{2n}^2}{2(1-p^2)} \\
 &= - \sum_{n=1}^N \frac{-2x_{1n}x_{2n} \cdot 2(1-p^2) - (x_{1n}^2 + x_{2n}^2 - 2px_{1n}x_{2n}) \cdot (-4p)}{4(1-p^2)^2} \\
 &= - \sum_{n=1}^N \frac{-x_{1n}x_{2n}(1-p^2) + p(x_{1n}^2 + x_{2n}^2 - 2px_{1n}x_{2n})}{(1-p^2)^2} \\
 \frac{\partial \log L(p | x_1, x_2)}{\partial p} &= \frac{NP}{2(1-p^2)} - \sum_{n=1}^N \frac{-x_{1n}x_{2n}(1-p^2) + p(x_{1n}^2 + x_{2n}^2 - 2px_{1n}x_{2n})}{(1-p^2)^2} = 0 \\
 NP^3 - \sum_{n=1}^N x_{1n}x_{2n}p^2 + (\sum_{n=1}^N (x_1^2 + x_2^2) - N)p - \sum_{n=1}^N x_{1n}x_{2n} &= 0
 \end{aligned}$$

解即为  $p$  的估计

(科目: ) 清华大学数学作业纸

编号:

班级:

姓名:

第 页

$$2.11 P(\mu | x) = \frac{P(x|\mu)P(\mu)}{P(x)}$$

$$\begin{aligned} &= \frac{1}{P(x)} \prod_{n=1}^N \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp(-\frac{1}{2}(x_n - \mu)^T \Sigma^{-1} (x_n - \mu)) \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_0|^{\frac{1}{2}}} \exp(-\frac{1}{2}(\mu - \mu_0)^T \Sigma_0^{-1} (\mu - \mu_0)) \\ &= \frac{1}{P(x)} \left( \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \right)^N \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_0|^{\frac{1}{2}}} \exp(-\frac{1}{2}(x_n - \mu)^T \Sigma^{-1} (x_n - \mu) - \frac{1}{2}(\mu - \mu_0)^T \Sigma_0^{-1} (\mu - \mu_0)) \\ &= C \cdot \exp\left(-\frac{1}{2} \left\{ \sum_{n=1}^N (x_n^T \Sigma^{-1} x_n - x_n^T \Sigma^{-1} \mu - \mu^T \Sigma^{-1} x_n + \mu^T \Sigma^{-1} \mu) + \mu^T \Sigma_0^{-1} \mu \right. \right. \\ &\quad \left. \left. - \mu^T \Sigma_0^{-1} \mu_0 - \mu_0^T \Sigma_0^{-1} \mu + \mu_0^T \Sigma_0^{-1} \mu_0 \right\} \right) \\ &= C \cdot \exp\left(-\frac{1}{2} \left\{ \mu^T (N \Sigma^{-1} + \Sigma_0^{-1}) \mu - \mu^T \left( \Sigma^{-1} \sum_{n=1}^N x_n - \Sigma_0^{-1} \mu_0 \right) - \left( \sum_{n=1}^N x_n^T \Sigma^{-1} - \mu_0^T \Sigma_0^{-1} \right) \mu \right. \right. \\ &\quad \left. \left. + \sum_{n=1}^N x_n^T \Sigma^{-1} x_n - \mu_0^T \Sigma_0^{-1} \mu_0 \right\} \right) \end{aligned}$$

$$\begin{aligned} \mu^T (N \Sigma^{-1} + \Sigma_0^{-1}) \mu &= \mu^T N \Sigma^{-1} (\Sigma_0 + \frac{1}{N} \Sigma) \Sigma_0^{-1} \mu \\ &= \mu^T (\Sigma_0 (\Sigma_0 + \frac{1}{N} \Sigma) \frac{1}{N} \Sigma)^{-1} \mu \\ &= \mu^T \Sigma_N^{-1} \mu \end{aligned}$$

$$\begin{aligned} -\mu^T \left( \Sigma^{-1} \sum_{n=1}^N x_n - \Sigma_0^{-1} \mu_0 \right) &= -\mu^T (N \Sigma^{-1} (\Sigma_0 + \frac{1}{N} \Sigma) \Sigma_0^{-1} \Sigma_0 (\Sigma_0 + \frac{1}{N} \Sigma)^{-1} \frac{1}{N} \sum_{n=1}^N x_n \\ &\quad + \Sigma^{-1} (\Sigma_0 + \frac{1}{N} \Sigma) \Sigma_0^{-1} \Sigma (\Sigma_0 + \frac{1}{N} \Sigma)^{-1} \mu_0) \\ &= -\mu^T (N \Sigma^{-1} (\Sigma_0 + \frac{1}{N} \Sigma) \Sigma_0^{-1}) (\Sigma_0 (\Sigma_0 + \frac{1}{N} \Sigma)^{-1} \frac{1}{N} \sum_{n=1}^N x_n \\ &\quad + \frac{1}{N} \Sigma (\Sigma_0 + \frac{1}{N} \Sigma)^{-1} \mu_0) \\ &= -\mu^T \Sigma_N^{-1} \mu_N \end{aligned}$$

$$\text{同理可得 } -(\sum_{n=1}^N x_n^T \Sigma^{-1} - \mu_0^T \Sigma_0^{-1}) \mu = -\mu_N^T \Sigma_N^{-1} \mu$$

$$\begin{aligned} \therefore P(\mu | x) &= C' \cdot \exp\left(-\frac{1}{2} (\mu^T \Sigma_N^{-1} \mu - \mu^T \Sigma_N^{-1} \mu_N - \mu_N^T \Sigma_N^{-1} \mu + \mu^T \Sigma_N^{-1} \mu)\right) \\ &= C' \exp\left(-\frac{1}{2} (\mu - \mu_N)^T \Sigma_N^{-1} (\mu - \mu_N)\right) \end{aligned}$$

$$\therefore P(\mu | x) = N(\mu | \mu_N, \Sigma_N)$$

$$\begin{aligned}
 2.12 \quad P(x) &= \frac{1}{\sqrt{2\pi} b} \exp\left(-\frac{(x-\mu)^2}{2b^2}\right) \\
 &= \frac{1}{\sqrt{2\pi} b} \exp\left(-\frac{\mu^2}{2b^2}\right) \exp\left(-\frac{x^2}{2b^2}\right) \exp\left(\frac{\mu x}{b^2}\right) \\
 &= \frac{1}{\sqrt{2\pi} b} \exp\left(-\frac{b^2}{2}\left(\frac{\mu}{b^2}\right)^2\right) \exp\left(-\frac{x^2}{2b^2}\right) \exp\left(\frac{\mu}{b^2} x\right) \\
 &= g\left(\frac{\mu}{b^2}\right) h(x) \exp\left(\frac{\mu}{b^2} x\right)
 \end{aligned}$$

△