ML-HW3 ۸۱۰۱۹۷۰۳۹ امیرحسین عباسکوهی

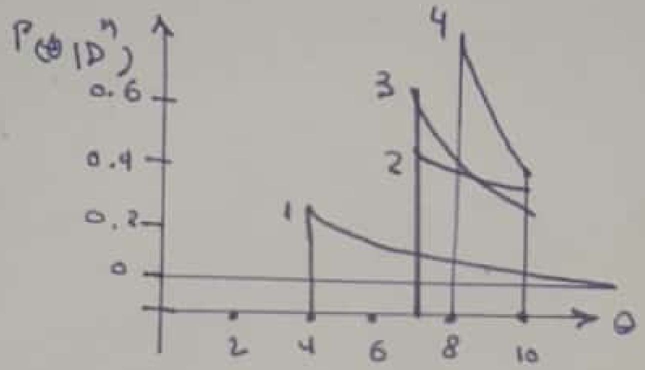
(۱

$$x_1 = 4 \rightarrow P(\theta | D^1) \propto p(x|\theta)\, p(\theta | D^0) = \begin{cases} \frac{1}{\theta} & \text{for } 4 \leq \theta \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

next data point : $x_2 = 7 \rightarrow P(\theta | D^2) \propto p(x|\theta)\, p(\theta | D^1) = \begin{cases} \frac{1}{\theta^2} & 7 \leq \theta \leq 10 \\ 0 & \text{otherwise} \end{cases}$

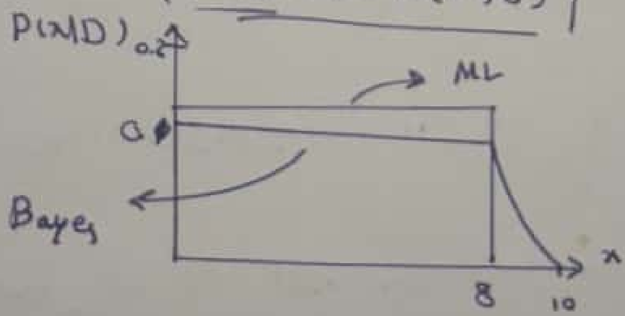
⋮

$$P(\theta | D^n) \propto \frac{1}{\theta^n} \text{ for } \max_x |D^n| \leq \theta \leq 10 :$$



حال براساس نمودار به این نتیجه می‌توان رسید که Maximum Likelihood برابر با $\hat{\theta} = 8$ خواهد بود. این به ما نتیجه می‌دهد که:

$$\boxed{p(x|D) \sim U(0,8)}$$



براساس دیتاست داده‌شده، توزیع $p(x|D)$ براساس maximum likelihood به ما

$p(x|\hat{\theta}) \sim U(0,8)$ می‌دهد.

2

(2

$$P(x|\theta) = \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$$

$$P(x_1 \dots x_n|\theta) = \prod_{k=1}^{n} \prod_{i=1}^{d} \theta_i^{x_{ki}} (1-\theta_i)^{1-x_{ki}}$$

$$\ell(\theta) = \sum_{k=1}^{n} \sum_{i=1}^{d} x_{ki} \ln \theta_i + (1-x_{ki}) \ln(1-\theta_i)$$

$$|\nabla_\theta L(\theta)|_i = \nabla_{\theta_i} L(\theta) = \frac{1}{\theta_i} \sum_{k=1}^{n} \frac{1}{1-\theta_i} \sum_{k=1}^{n} (1-x_{ki}) = 0$$

برای هر i

$$\Rightarrow \frac{1}{\hat{\theta}_i} \sum_{k=1}^{n} x_{ki} = \frac{1}{1-\hat{\theta}_i} \sum_{k=1}^{n} (1-x_{ki}) \rightarrow (1-\hat{\theta}_i) \sum_{k=1}^{n} x_{ki} = \hat{\theta}_i (n - \sum_{k=1}^{n} x_{ki})$$

$$\hat{\theta}_i = \frac{1}{n} \sum_{k=1}^{n} x_{ki}$$

$$\sigma_c \rightarrow \hat{\theta} = \frac{1}{n} \sum_{k=1}^{n} x_k$$

(۴

a

$$Q(\theta, \theta^0) = \mathcal{E}_{x_{32}} [\ln p(x_g, x_b ; \theta) / \theta^0, D_g]$$

$$= \int_{-\infty}^{\infty} (\ln p(x_1|\theta) + \ln p(x_2|\theta) + \ln p(x_3|\theta)\, p(x_{32}|\theta^0, x_{31}=2)\, dx_{32}$$

$$= \ln p(x_1|\theta) + \ln p(x_2|\theta) + \int_{-\infty}^{\infty} \ln p(x_3|\theta)\, p(x_{32}|\theta^0, x_{31}=2)\, dx_{32}$$

$$= \ln p(x_1|\theta) + \ln p(x_2|\theta) + \int_{-\infty}^{\infty} p\left(\binom{2}{x_{32}}\Big|\theta\right) \cdot \frac{p\left(\binom{2}{x_{32}}\Big|\theta^0\right)}{\int_{-\infty}^{\infty} p\left(\binom{2}{x'_{32}}\Big|\theta^0\right) dx'_{32}}\, dx_{32}$$

$$= \ln p(x_1|\theta) + \ln p(x_2|\theta) + k$$

Scanned with CamScanner

B

برای k سه حالت داریم :

(1) $3 \leq \theta_2 \leq 4 \rightarrow K = \frac{1}{4} \int_0^{\theta_2} \ln\left(\frac{1}{\theta_1} e^{-2\theta_1} \frac{1}{\theta_2}\right) dx_{32}$

$$= \frac{1}{4} \theta_2 \ln\left(\frac{1}{\theta_1} e^{-2\theta_1} \frac{1}{\theta_2}\right)$$

(2) $\theta_2 \geq 4 \rightarrow K = \frac{1}{4} \int_0^{4} \ln\left(\frac{1}{\theta_1} e^{-2\theta_1} \frac{1}{\theta_2}\right) dx_{32}$

$$= \frac{1}{4} \times 4 \ln\left(\frac{1}{\theta_1} e^{-2\theta_1} \frac{1}{\theta_2}\right)$$

$$= \ln\left(\frac{1}{\theta_1} e^{-2\theta_1} \frac{1}{\theta_2}\right)$$

(3) Otherwise $\rightarrow K = 0$

$$Q(\theta, \theta^{\circ}) = \ln p(x_1 | \theta) + \ln p(x_2 | \theta) + k$$

$$= \ln\left(\frac{1}{\theta_1} e^{-\theta_1} \frac{1}{\theta_2}\right) + \ln\left(\frac{1}{\theta_1} e^{-3\theta_1} \frac{1}{\theta_2}\right) + k$$

$$= -\theta_1 - \ln(\theta_1 \theta_2) - 3\theta_1 - \ln(\theta_1 \theta_2) + k$$

$$= -4\theta_1 - 2\ln(\theta_1 \theta_2) + k$$

b

There are 2 cases:

(1) $3 \leq \theta_2 \leq 4$

$$Q(\theta, \theta^{\circ}) = -4 - \left(2\ln\theta_2 + \frac{1}{4}\theta_2 (2 + \ln\theta_2)\right)$$

$Q(\theta, \theta^{\circ}) = 3 \rightarrow$ leads to max : $Q(\theta, \theta^{\circ}) = -8.5212$

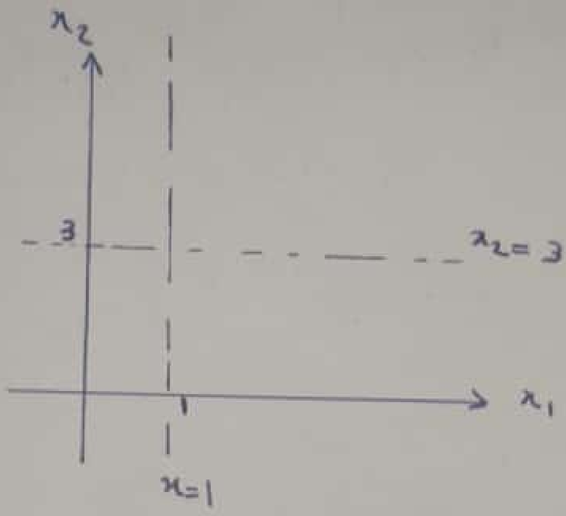
(2) $\theta_2 \geq 4 \longrightarrow Q(\theta, \theta^0) = -6 - 3\ln\theta_2$

argmax$_{\theta_2}$ монотоник یعنی ـ6

argmax$_{\theta_2}$ $Q(\theta, \theta^0) = 4 \xrightarrow{\text{leads}} Q(\theta, \theta^0) \approx -10.159$

$\Rightarrow \theta = (1 \; 3)^t$

ج



$$\mathcal{L}(N_{ij})(\lambda) = \prod_{i=1}^{n} \prod_{j=1}^{m} e^{a_{ij}\lambda_j} \frac{(\lambda_j a_{ij})^{N_{ij}}}{N_{ij}!}$$

$$\xrightarrow[\text{likelihood}]{\log} L_{(N_{ij})_{ij}}(\lambda) = \sum_{i=1}^{n} \sum_{j=1}^{m} (-\lambda_j a_{ij} + N_{ij} \log(\lambda_j a_{ij}) - \log(N_{ij}!))$$

$$\frac{d}{d\lambda_j} E[L(N_{ij} \mid (Y_i)_i] = \sum_{i=1}^{n} -a_{ij} + \frac{1}{\lambda_j} E[N_{ij} \mid (Y_i)_i] \quad \forall j = 1, \dots, m$$

distribution of $N_{ij} \mid (Y_i)_i$ = distribution of $N_i \mid Y_i$

($k = i$ بجز) مستقل است Y_k ها از N_{ij} چون

Lemma: Let X_1, X_2 be independent Poisson distribution with $X_1 \sim P(\lambda_1)$, $X_2 \sim P(\lambda_2)$

Then $\rightarrow X_1 \mid (X_1 + X_2) \sim Bin\left(X_1 + X_2, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$

$X_1 = N_{ij}$ and $X_2 = Y_i - N_{ij} \rightarrow N_{ij} \mid Y_i \sim Bin\left(Y_i, \frac{a_{ij}\lambda_j}{\sum_{k=1}^{m} a_{ik}\lambda_k}\right)$

Expectation of Bionomial with parameters n and p $\rightarrow np \Rightarrow$

$$E[N_{ij} \mid Y_i] = \frac{Y_i a_{ij} \lambda_j}{\sum_{k=1}^{m} a_{ik}\lambda_k}$$

Therefore:

$$\frac{d}{d\lambda_j} E[l_{(N_{ij})_{ij}} \mid (Y_i)_i] = \sum_{i=1}^{n} -a_{ij} + \frac{1}{\lambda_j} \frac{Y_i a_{ij} \lambda_j^{old}}{\sum_{k=1}^{m} a_{ik}\lambda_k^{old}} \quad \forall j = 1, \dots, m$$

$\xrightarrow{\text{derivative}=0}$ $$0 = \sum_{i=1}^{n} -a_{ij} + \frac{1}{\lambda_j} \frac{Y_i a_{ij} \lambda_j^{old}}{\sum_{j=1}^{m} a_{ik}\lambda_k^{old}}$$

$$\downarrow$$

$$\lambda_j = \frac{\lambda_j^{old}}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \frac{Y_i a_{ij}}{\sum_{k=1}^{m} a_{ik}\lambda_k^{old}}$$

6

(4

a

$$\mu_n = \Big[\underbrace{\frac{1}{\frac{n}{\delta^2}+\frac{1}{\delta_0^2}}}_{\delta_n^2}\Big]\frac{n}{\delta^2}\hat{\mu}_n + \Big[\underbrace{\frac{1}{\frac{n}{\delta^2}+\frac{1}{\delta_0^2}}}_{\delta_n^2}\Big]\frac{\mu_0}{\delta_0^2}$$

$$\xrightarrow{n_0=\frac{\delta^2}{\delta_0^2}} \qquad \mu_0 = n_0^{-1}\sum_{k=-n_0+1}^{0} x_k$$

$$\mu_n = \Big[\frac{1}{1+\frac{\delta^2}{n\delta_0^2}}\Big]\hat{\mu}_n + \Big[\frac{1}{1+\frac{n\delta_0^2}{\delta^2}}\Big]\mu_0$$

$$= \Big[\frac{1}{1+\frac{n_0}{n}}\Big]\frac{1}{n}\sum_{k=1}^{n} x_k + \Big[\frac{1}{1+\frac{n}{n_0}}\Big]\frac{1}{n}\sum_{k=-n_0+1}^{0} x_k$$

$$= \frac{1}{n+n_0}\sum_{k=1}^{n} x_k + \frac{1}{n+n_0}\sum_{k=-n_0+1}^{0} x_k$$

$$= \frac{1}{n+n_0}\sum_{k=-n_0+1}^{n} x_k$$

b

μ_n ← یک میانگین وزن دار از نمونه های real و fiction است

چون:

$$\mu_n = \frac{n}{n+n_0}\Big[\frac{1}{n}\sum_{k=0}^{n} n\Big] + \frac{n_0}{n+n_0}\Big[\frac{1}{n_0}\sum_{k=-n_0+1}^{0} x_k\Big] = \frac{1}{n+n_0}\sum_{k=-n_0+1}^{n} x_k$$

δ_n^2 ← اگر در نظر بگیریم reciprocal است، δ_n^{-2} precision است

$$\delta_n^{-2} = n\delta^{-2} + n_0\delta^{-2}$$

پس تعداد نمونه های fictious یک تخمین اطمینان برای precision

می دهند، هرچه قدر نمونه real اضافه شود به صورت خطی زیاد می شود