No.
Date

MLE MAP

M 回顾 Likelihood：有 N 个 i.i.d 样本：$D=\{x^{(1)},\cdots,x^{(N)}\}$ of r.v. X
L 若 X 离散，PMF 为 $p(x|\theta)$，则 Likelihood of D is：
E
$$L(\theta)=\prod_{n=1}^{N}p(x^{(n)}|\theta)$$
若 X 连续，PDF 为 $f(X|\theta)$，则：
$$L(\theta)=\prod_{n=1}^{N}f(x^{(n)}|\theta)$$
考虑 toss coin：$P(\text{Head})=\theta$　$P(\text{Tail})=1-\theta$
$$\hat\theta=\arg\max_\theta P(D|\theta)=\arg\max_\theta\prod_{i=1}^{n}P(x_i|\theta)$$
$$=\arg\max_\theta \underbrace{\theta^{\alpha_H}(1-\theta)^{\alpha_T}}\quad J(\theta)$$

$\partial J(\theta)/\partial\theta=0\Rightarrow\hat\theta_{MLE}=\dfrac{\alpha_H}{\alpha_H+\alpha_T}$

考虑 $X\sim$ Gaussian (μ,σ)，欲 $\arg\max_{\mu,\sigma^2}:\underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2}}_{J(\theta)}$

$\frac{\partial J}{\partial\mu}=\sum_{i=1}^{n}(x_i-\mu)/\sigma^2=0$；$\frac{\partial J}{\partial\sigma^2}=-n+\alpha*\frac{1}{\sigma^2}=0$

$\therefore\hat\mu_{MLE}=\frac{1}{n}\sum_{i=1}^{n}x_i$　$\hat\sigma^2_{MLE}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat\mu)^2$

△ 附：MLE 推出的 σ^2 有 bias！　$\hat\sigma^2_{unbiased}=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\hat\mu)^2$

why？因为 $\hat\sigma^2_{MLE}$ 中用的是 sample mean 而非 true mean

*：若 $p(X)\sim Beta(\alpha,\beta)$，
则 $p(X)$ max：$x^*=\frac{\alpha-1}{\alpha+\beta-2}$
$E(X)=\frac{\alpha}{\alpha+\beta}$

M ：The Bayesian Way：
A
P
$$P(\theta|D)=\frac{P(D|\theta)\,P(\theta)}{\underbrace{P(D)}_{\text{常数}}}\propto\underbrace{P(D|\theta)}_{\text{Likelihood}}\underbrace{P(\theta)}_{\text{prior}}$$
or：后验

欲 $\arg\max_\theta\, p(D|\theta)p(\theta)\Leftrightarrow\arg\max_\theta[\log P(D|\theta)+\log p(\theta)]$

考虑 flip coin：$P(D|\theta)=\binom{n}{\alpha_H}\theta^{\alpha_H}(1-\theta)^{\alpha_T}$，if $P(\theta)\sim Beta(\beta_H,\beta_T)$
则 $p(\theta|D)\sim Beta(\beta_H+\alpha_H,\beta_T+\alpha_T)$，$\hat\theta_{MAP}\overset{*}{=}\dfrac{\alpha_H+\beta_H-1}{\alpha_H+\beta_H+\alpha_T+\beta_T-2}$
$(\hat\theta_{MAP}=\arg\max_\theta p(\theta|D))$

总结：Frequentist：Sample 少时表现不好；Bayesians：不同 prior 不同 answer

KOKUYO

扫描全能王 创建