TP5 Zhe HUANG 22208901 M1AI

Problem 1

(1) $P(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

The log-likelihood function is:

$$\log P(x|\mu,\sigma^2) = -\frac{\log(2\pi)}{2} - \frac{\log(\sigma^2)}{2} - \frac{(x-\mu)^2}{2\sigma^2}$$

Taking the partial derivatives with respect to μ and σ^2, we get:

$$\frac{\partial}{\partial\mu}\log p(x|\mu,\sigma^2) = \frac{(x-\mu)}{\sigma^2}, \quad \frac{\partial}{\partial\sigma^2}\log P(x|\mu,\sigma^2) = -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2\sigma^4}$$

Taking the second partial derivatives with respect to μ and σ^2, we get:

$$\frac{\partial^2}{\partial\mu^2}\log p(x|\mu,\sigma^2) = -\frac{1}{\sigma^2}, \quad \frac{\partial^2}{\partial(\sigma^2)^2}\log p(x|\mu,\sigma^2) = \frac{1}{2\sigma^4} - \frac{(x-\mu)^2}{\sigma^6}$$

$$\frac{\partial^2}{\partial\mu\partial\sigma^2}\log P(x|\mu,\sigma^2) = \frac{(x-\mu)}{\sigma^4}$$

Now we need to take the expectations of these partial derivatives under the distribution.

$$E[(x-\mu)^2] = \sigma^2, \quad E[(x-\mu)^4] = 3\sigma^4, \quad E[(x-\mu)] = 0$$

Using these expeactations, we can compute the elements of the Fisher information matrix

$$I_{\mu,\mu} = -E\left[\frac{\partial^2}{\partial\mu^2}\log P(x|\mu,\sigma^2)\right] = \frac{1}{\sigma^2}$$

$$I_{\sigma^2,\sigma^2} = -E\left[\frac{\partial^2}{\partial(\sigma^2)^2}\log P(x|\mu,\sigma^2)\right] = \frac{1}{2\sigma^4}$$

$$I_{\mu,\sigma^2} = -E\left[\frac{\partial^2}{\partial\mu\partial(\sigma^2)}\log P(x|\mu,\sigma^2)\right] = 0$$

Therefore, the fisher information matrix for a Gaussian distribution with μ and σ^2 is

$$\begin{bmatrix} I_{\mu,\mu} & I_{\mu,\sigma^2} \\ I_{\mu,\sigma^2} & I_{\sigma^2,\sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix}$$

(2) We have already computed the Fisher information matrix for this model:

$$I_{\mu,\mu} = \frac{N}{\sigma^2}, \quad I_{\mu,\sigma^2} = 0, \quad I_{\sigma^2,\sigma^2} = \frac{N}{2\sigma^4}$$ where N is the sample size

We can use the formula for the Cramér-Rao bound to find a lower bound on the variance of any unbiased estimator for 'μ'. In particular, we can use the estimator.

$\hat{\mu} = \frac{1}{N}\sum_{k=1}^{N} X_k$ which is an unbiased estimator of 'μ'.

We can compute its variance as:
$$\begin{aligned} Var(\hat{\mu}) &= Var\left(\frac{1}{N}\sum_{k=1}^{N} X_k\right) \\ &= \frac{1}{N^2} Var\left(\sum_{k=1}^{N} X_k\right) \\ &= \frac{1}{N^2}\sum_{k=1}^{N} Var(X_k) \\ &= \frac{1}{N^2}\cdot N \cdot \sigma^2 \quad \text{// } X_k \text{ are iid sampled.} \\ &= \frac{\sigma^2}{N} \end{aligned}$$

Thus:
$$\begin{aligned} Var(\hat{\mu}) &\geq \frac{1}{I_{\mu,\mu}} \\ &= \frac{\sigma^2}{\frac{N}{\sigma^2}} \\ &= \frac{\sigma^4}{N} \end{aligned}$$

Therefore, the Cramer-Rao bound for the parameter μ is $\frac{\sigma^2}{N}$, which is achieved by $\hat{\mu} = \frac{1}{N}\sum_{k=1}^{N} X_k$

(3) To find the variance of σ^2, we can use the fact that:

$$\frac{(N-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^{2(N-1)}$$

Therefore, we have:
$$\begin{aligned} Var(\hat{\sigma}^2) &= Var\left(\frac{(N-1)\hat{\sigma}^2}{\sigma^2}\cdot\frac{1}{N-1}\right) \\ &= \frac{1}{(N-1)^2} Var\left((N-1)\frac{\hat{\sigma}^2}{\sigma^2}\right) \\ &= \frac{2(N-1)}{N-3}\sigma^4 \end{aligned}$$

And we know $I_{\sigma^2,\sigma^2} = \frac{N}{2\sigma^4}$

So, the Cramér-Rao bound for σ^2 is:
$$\begin{aligned} Var(\hat{\sigma}^2) &\geq \frac{1}{I_{\sigma^2,\sigma^2}} \\ &= \frac{2\sigma^4}{N} \end{aligned}$$

ie. $\frac{2(N-1)}{N-3}\sigma^4 \geq \frac{2\sigma^4}{N}$ $(N-1)(N-3) \geq N \cdot N^2 - 4N + 3 \geq 0$ $(N-1)(N-3) \geq 0$

Problem 2

(1) For the Bernoulli distribution with parameter θ, the log-likelihood function is:

$$l(\theta) = \sum_{i=1}^{N} \log P(X_i|\theta) = \sum_{i=1}^{N} [X_i \log\theta + (1-X_i)\log(1-\theta)]$$

Taking the first derivative of the log-likelihood function with respect to θ, we get.

$$\frac{\partial l(\theta)}{\partial\theta} = \sum_{i=1}^{N}\left[\frac{X_i}{\theta} - \frac{1-X_i}{1-\theta}\right]$$

Taking the second derivative, we get.

$$\frac{\partial l(\theta)}{\partial\theta^2} = -\sum_{i=1}^{N}\left[\frac{X_i}{\theta^2} + \frac{1-X_i}{(1-\theta)^2}\right]$$

Therefore, the fisher information for the parameter θ is:

$$I(\theta) = -\mathbb{E}\left[\frac{\partial^2 l(\theta)}{\partial\theta^2}\right] = \mathbb{E}\left[\sum_{i=1}^{N}\left(\frac{X_i}{\theta^2} + \frac{1-X_i}{(1-\theta)^2}\right)\right] = \sum_{i=1}^{N}\left[\frac{\theta}{\theta^2} + \frac{1-\theta}{(1-\theta)^2}\right] = \frac{N}{\theta(1-\theta)}$$

(2) The unbiased estimator for θ is $\hat{\theta} = \frac{1}{N}\sum_{k=1}^{N} X_k$

The variance of X wrt Bernoulli distribution is $Var(X) = \theta(1-\theta)$

Then, the variance of $\hat{\theta}$ is:

$$Var(\hat{\theta}) = Var\left(\frac{1}{N}\sum_{k=1}^{N} X_k\right) = \frac{1}{N^2}\sum_{k=1}^{N} Var(X_k) = \frac{1}{N^2}\cdot N\cdot\theta(1-\theta) = \frac{\theta(1-\theta)}{N}$$

Therefore, the Cramér-Rao low bound for the variance of $\hat{\theta}$ is

$$Var(\hat{\theta}) \geqslant \frac{1}{I(\theta)} = \frac{\theta(1-\theta)}{N}$$