Exercise 1.

(a) $$P(X|\theta) = \begin{cases} \theta & \text{if } x = \text{head} \\ 1-\theta & \text{if } x = \text{tail} \end{cases}$$

with the definition $P(X_k|\theta)$, We state the likelihood function

$$P(D|\theta) = \prod_{k=i}^{n} P(X_k|\theta) = \theta^5 \cdot (1-\theta)^2$$

Team = RCSKSK

Ramesha 387219

CHEN 387275

Shah 384926

Swamy 384418

Khalil 385402

Kao 387470

(b) $$f(x) = P(D|\theta) = \theta^5 \cdot (1-\theta)^2$$

$$L(D|\theta) = 5\log\theta + 2\log(1-\theta)$$

$$= \frac{5}{\theta} + \frac{2}{1-\theta}(-1)$$

$$= \frac{5}{\theta} - \frac{2}{1-\theta}$$

$$= \frac{5-5\theta-2\theta}{\theta(1-\theta)}$$

$$= \frac{5-7\theta}{\theta(1-\theta)}$$

$$5-7\theta=0$$

$$\hat{\theta} = \frac{5}{7}$$

With this fact, each sample x_k is generated independently

Compute $P(X_8 = \text{head}, X_9 = \text{head} \mid \hat{\theta})$

$$P(X_8 = \text{head}, X_9 = \text{head} \mid \hat{\theta}) = P(X_8 = \text{head} \mid \hat{\theta})\, P(X_9 = \text{head} \mid \hat{\theta}) = \hat{\theta}^2$$

$$= \frac{5}{7} \times \frac{5}{7}$$

$$= \frac{25}{49}$$

(c) $$P(\theta) = \begin{cases} 1 & \text{if } 0 \le \theta \le 1 \\ 0 & \text{else} \end{cases}$$

prior distribution

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{\int P(D|\theta)p(\theta)d\theta}$$

posterior distribution

$$= \alpha \prod_{k=1}^{7} P(X_k|\theta)\, P(\theta)$$

$$= \begin{cases} \alpha \cdot \theta^5 \cdot (1-\theta)^2 & \text{if } 0 \le \theta \le 1 \\ 0 & \text{else} \end{cases}$$

$$\alpha = \frac{1}{\int_0^1 \prod_{k=1}^{7} P(X_k|\theta)\, P(\theta)\, d\theta}$$

$$= \frac{1}{\int_0^1 \prod_{k=1}^{7} P(X_k|\theta)\, d\theta}$$

$$= \frac{1}{\int_0^1 \theta^7 - 2\theta^6 + \theta^5 d\theta}$$

$$= \frac{1}{\left[\frac{\theta^8}{8} - \frac{2\theta^7}{7} + \frac{\theta^6}{6}\right]_0^1} = \frac{1}{\frac{28}{168} - \frac{48}{168} + \frac{21}{168}} = \frac{1}{\frac{1}{168}} = 168$$

$$\int P(X_8=\text{head}, X_9=\text{head}\mid\theta)\,P(\theta\mid D)\,d\theta = \int P(X_8=\text{head}\mid\theta)\,P(X_9=\text{head}\mid\theta)\,P(\theta\mid D)\,d\theta$$

$$= \int_0^1 \theta^2\cdot\alpha\cdot\theta^5\cdot(1-\theta)^2\,d\theta$$

$$= 168\int_0^1 \theta^7(1-\theta)^2\,d\theta$$

$$= 168\int_0^1 \theta^9 - 2\theta^8 + \theta^7\,d\theta$$

$$= 168\left[\frac{\theta^{10}}{10} - \frac{2\theta^9}{9} + \frac{\theta^8}{8}\right]_0^1$$

$$= 168\cdot\frac{1}{360} = \frac{7}{15} \approx 0.4\bar{6}$$

Exercise 2.

(a) To show $\sigma_n^2 \le \min\left(\frac{\sigma^2}{n}, \sigma_0^2\right)$

$$\frac{1}{\sigma_n^2} = \frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}$$

$$\Leftrightarrow 1 = \frac{\sigma_0^2 n + \sigma^2}{\sigma^2\sigma_0^2}\cdot\sigma_n^2$$

$$\Leftrightarrow \sigma_n^2 = \frac{\sigma^2\sigma_0^2}{\sigma_0^2 n + \sigma^2}$$

We can show $\sigma_n^2 \le \min\left(\frac{\sigma^2}{n}, \sigma_0^2\right)$. Assme $\frac{\sigma^2}{n} \le \sigma_0^2$

$$\sigma_n^2 \le \min\left(\frac{\sigma^2}{n}, \sigma_0^2\right)$$

$$\Leftrightarrow \frac{\sigma^2\sigma_0^2}{\sigma_0^2 n + \sigma^2} \le \frac{\sigma^2}{n}$$

$$\Leftrightarrow \sigma^2\sigma_0^2 \le \sigma^2\sigma_0^2 + \frac{\sigma^4}{n}$$

$$\Leftrightarrow \sigma^2\sigma_0^2 \le \sigma^2\cdot\left(\sigma_0^2 + \frac{\sigma^2}{n}\right)$$

$$\Leftrightarrow \sigma_0^2 \le \sigma_0^2 + \frac{\sigma^2}{n}$$

$$\Leftrightarrow 0 \le \frac{\sigma^2}{n}$$

* This holds true since n is the number of features $\Rightarrow N(\mu, \sigma^2)$, thus $n, \sigma > 0$

σ^2 is the variance of Gaussian distribution

Another case: $\left(\sigma_0^2 \le \frac{\sigma^2}{n}\right)$

$$\sigma_n^2 \le \min\left(\frac{\sigma^2}{n}, \sigma_0^2\right)$$

$$\Leftrightarrow \frac{\sigma^2\sigma_0^2}{\sigma_0^2 n + \sigma^2} \le \sigma_0^2$$

$$\Leftrightarrow \sigma^2\sigma_0^2 \le \sigma_0^2\cdot(\sigma_0^2 n + \sigma^2)$$

$$\Leftrightarrow \sigma^2 \le \sigma_0^2 n + \sigma^2$$

$$\Leftrightarrow 0 \le \sigma_0^2 n$$

This holds true since $n > 0$ and $\sigma_0^2 > 0$ since it's the variance of the Gaussian distribution

$N(\mu_0, \sigma_0^2)$

(b)

To show

$\min(\mu_0, \hat{\mu}_n) \leq \mu_n \leq \max(\mu_0, \hat{\mu}_n)$

solve μ_n, also using our σ_n^2 from above:

$$\frac{\mu_n}{\sigma_n^2} = \frac{n}{\sigma^2}\hat{\mu}_n + \frac{\mu_0}{\sigma_0^2}$$

$$\Leftrightarrow \mu_n = \left(\frac{n}{\sigma^2}\hat{\mu}_n + \frac{\mu_0}{\sigma_0^2}\right) \cdot \sigma_0^2$$

$$\Leftrightarrow \mu_n = \left(\frac{n}{\sigma^2}\hat{\mu}_n + \frac{\mu_0}{\sigma_0^2}\right) \cdot \frac{\sigma^2\sigma_0^2}{\sigma_0^2 n + \sigma^2}$$

$$\Leftrightarrow \mu_n = \frac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\hat{\mu}_n + \frac{\sigma^2}{n\sigma_0^2+\sigma^2}\mu_0$$

Assume $\mu_0 < \hat{\mu}_n$ (1) and show first part of the inequality

$\min(\mu_0, \hat{\mu}_n) \leq \mu_n$

$$\Leftrightarrow \mu_0 \leq \mu_n$$

$$\Leftrightarrow \mu_0 \leq \frac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\hat{\mu}_n + \frac{\sigma^2}{n\sigma_0^2+\sigma^2}\mu_0$$

$$\Leftrightarrow \mu_0 \cdot (n\sigma_0^2 + \sigma^2) \leq n\sigma_0^2\hat{\mu}_n + \sigma^2\mu_0$$

$$\Leftrightarrow \mu_0 n\sigma_0^2 \leq n\sigma_0^2\hat{\mu}_n$$

$$\Leftrightarrow \mu_0 \leq \hat{\mu}_n$$

This is true due to assumption (1) above. Show the 2nd part of the inequality

$\mu_n \leq \max(\mu_0, \hat{\mu}_n)$

$$\Leftrightarrow \mu_n \leq \hat{\mu}_n$$

$$\Leftrightarrow \frac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\hat{\mu}_n + \frac{\sigma^2}{n\sigma_0^2+\sigma^2}\mu_0 \leq \hat{\mu}_n$$

$$\Leftrightarrow n\sigma_0^2\hat{\mu}_n + \sigma^2\mu_0 \leq \hat{\mu}_n \cdot (n\sigma_0^2 + \sigma^2)$$

$$\Leftrightarrow \sigma^2\mu_0 \leq \hat{\mu}_n\sigma^2$$

$$\Leftrightarrow \mu_0 \leq \hat{\mu}_n$$

This is true due to assumption (1) above also

Now Assume $\hat{\mu}_n < \mu_0$ (2) and show that our inequality still holds

We'll start with the first Part again:

$$\min(\mu_0, \hat{\mu}_n) \leq \mu_n$$

$$\Leftrightarrow \hat{\mu}_n \leq \mu_n$$

$$\Leftrightarrow \hat{\mu}_n \leq \frac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\hat{\mu}_n + \frac{\sigma^2}{n\sigma_0^2+\sigma^2}\mu_0$$

$$\Leftrightarrow \hat{\mu}_n(n\sigma_0^2+\sigma^2) \leq n\sigma_0^2\hat{\mu}_n + \sigma^2\mu_0$$

$$\Leftrightarrow \hat{\mu}_n\sigma^2 \leq \sigma^2\mu_0$$

$$\Leftrightarrow \hat{\mu}_n \leq \mu_0$$

This hold true due to assumption (2) above

Let's show the second part of the inequality

$$\mu_n \leq \max(\mu_0, \hat{\mu}_n)$$

$$\Leftrightarrow \mu_n \leq \mu_0$$

$$\Leftrightarrow \frac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\hat{\mu}_n + \frac{\sigma^2}{n\sigma_0^2+\sigma^2}\mu_0 \leq \mu_0$$

$$\Leftrightarrow n\sigma_0^2\hat{\mu}_n + \sigma^2\mu_0 \leq \mu_0(n\sigma_0^2+\sigma^2)$$

$$\Leftrightarrow n\sigma_0^2\hat{\mu}_n \leq \mu_0 n\sigma_0^2$$

$$\Leftrightarrow \hat{\mu}_n \leq \mu_0$$

This holds the true due to assumption (2) above