## Practice questions

1. Alice is mailing letters to solicit donations from CUHK alums. From past experience she knows that 30% of the alums make a 500 dollar donation, and 10% of the alums make a 1,000 dollar donation. Use the central limit theorem to estimate the number of letters Alice should mail to meet a 50,000 dollar donation target with probability 90%.

   **Solution:** The alumni donations $X_1, \dots, X_n$ are independent random variables with mean $\mu = \mathrm{E}[X_i] = 500 \cdot 0.3 + 1000 \cdot 0.1 = 250$ dollars and variance $\sigma^2 = \mathrm{Var}[X_i] = 500^2 \cdot 0.3 + 1000^2 \cdot 0.1 - 250^2 = 112500$ dollars. By the central limit theorem we approximate the sum of their donations $X_1 + \cdots + X_n$ as a normal random variable with mean $\mu n$ and standard deviation $\sigma\sqrt{n}$. A Normal$(0, 1)$ random variable takes values $-1.2816$ or higher 90% of the time. To achieve the desired probability of success, we need to pick $n$ so that $\mu n - 1.2816\sigma\sqrt{n}$ is at least $50,000$ dollars. Plugging in the values of $\mu$ and $\sigma$ we obtain the inequality

   $$250n - 1.2816 \cdot \sqrt{112,500} \cdot \sqrt{n} \geq 50,000.$$

   By the quadratic formula, it is enough to choose $n = 226$ for this. Hence, Alice needs to mail about 226 alums.

2. After much dating experience, Romeo concludes that girls show up Exponential$(\Theta)$ hours late to a date, where $\Theta$ is a Uniform$(0, 1)$ random variable whose value is encoded deep in a girl's heart. On the first two dates Juliet shows up 10 minutes late and 30 minutes late. What is Romeo's posterior PDF for Juliet's $\Theta$?

   **Solution:** The prior PDF is $f_\Theta(\theta) = 1$ if $0 \leq \theta \leq 1$. Let $X_1, X_2$ be the number of hours Juliet is late to the first two dates. The PDF of $X$ conditioned on $\Theta$ is $f_{X|\Theta}(x|\theta) = \theta e^{-\theta x}$ if $x \geq 0$. Therefore, the posterior PDF for Juliet's $\Theta$ is

   $$\begin{aligned} f_{\Theta|X_1,X_2}(\theta|1/6, 1/2) &\propto f_\Theta(\theta) f_{X|\Theta}(1/6|\theta) f_{X|\Theta}(1/2|\theta) \\ f_{\Theta|X_1,X_2}(\theta|1/6, 1/2) &= C \cdot \theta^2 e^{-2\theta/3} \quad \text{if } 0 \leq \theta \leq 1 \end{aligned}$$

   where $C$ is the proportionality constant. Probabilities add up to one so $C^{-1} = \int_0^1 \theta^2 e^{-2\theta/3} d\theta$. Using an online integrator (or integration by parts) this evaluates to $(27 - 51e^{-2/3})/4$. Therefore, the posterior PDF is $4/(27 - 51e^{-2/3}) \cdot \theta^2 e^{-2\theta/3} \approx 4.9036\theta^2 e^{-2\theta/3}$ for $0 \leq \theta \leq 1$.

3. The ratio of A, B, and C students at CUHK is 1:2:2. The probability that A, B, and C students answer a true/false exam question correctly are 90%, 70%, and 55%, respectively. Bob takes a 4 question exam and answers 3 of them correctly. The teacher determines Bob's grade using a MAP estimate, assuming his answers are independent. What should Bob's grade be?

   **Solution:** Let $\Theta$ be Bob's grade, where $\Theta = 2$ means he gets an A and so on. Its prior PMF is

   $$p_\Theta(\theta) = \begin{cases} 0.4 & \text{if } \theta = 0 \text{ or } 1 \\ 0.2 & \text{if } \theta = 2 \end{cases}$$

   Let $X$ be the number of exam questions Bob answers correct, which is a Binomial$(4, p)$ random variable with $p = 0.55, 0.7, 0.9$ when $\theta = 0, 1, 2$, respectively. Therefore the conditional PMF is

$$p_{X|\Theta}(x|\theta) = \begin{cases} \binom{4}{x} 0.55^x 0.45^{4-x} & \text{if } \theta = 0 \text{ and } 0 \le x \le 4 \\ \binom{4}{x} 0.7^x 0.3^{4-x} & \text{if } \theta = 1 \text{ and } 0 \le x \le 4 \\ \binom{4}{x} 0.9^x 0.1^{4-x} & \text{if } \theta = 2 \text{ and } 0 \le x \le 4 \end{cases}$$

The MAP rule finds $\theta$ such that the posterior PDF

$$p_{\Theta|X}(\theta|3) \propto p_\Theta(\theta) p_{X|\Theta}(3|\theta)$$

is maximized. Since the denominator $p_X(3)$ is independent of $\theta$, we can disregard it in our MAP estimate calculation. The numerator for $\theta = 0, 1, 2$ and $x = 3$ are approximately

| $\theta$ | 0 | 1 | 2 |
|---|---|---|---|
| $p_\Theta(\theta) p_{X|\Theta}(3|\theta)$ | 0.120 | 0.165 | 0.0583 |

As it is the largest when $\theta = 1$, Bob's grade is B.

4. The lifetimes of a normal and a defective light bulb are exponential random variables with averages of 5 years and 2 years, respectively. On average, one of three light bulbs comes out defective.

   (a) For which values of $y$ is a light bulb that burns out in $y$ years more likely to be defective than not under a MAP estimate?

   (b) What is the estimation error (the probability of an incorrect conclusion) in part (a)?

**Solution:**

(a) Let $\Theta$ indicates whether the light bulb is defective. The prior PMF of $\Theta$ is

$$p_\Theta(\theta) = \begin{cases} 1/3 & \text{if } \theta = 1 \\ 2/3 & \text{if } \theta = 0 \end{cases}$$

Let $T$ be the life of a light bulb, then $T|\Theta = 1$ is Exponential(1/2) and $T|\Theta = 0$ is Exponential(1/5). The conditional PDF of $T$ given $\Theta$ is

$$f_{T|\Theta}(t|\theta) = \begin{cases} \frac{1}{2} e^{-y/2} & \text{if } \theta = 1 \\ \frac{1}{5} e^{-y/5} & \text{if } \theta = 0 \end{cases}$$

The MAP estimator $\hat{\Theta}$ maximizes $p_\Theta(\theta) f_{T|\Theta}(y|\theta)$ over $\theta$. It estimates $\hat{\Theta} = 1$ when

$$\begin{aligned} p_\Theta(1) f_{T|\Theta}(y|1) &> p_\Theta(0) f_{T|\Theta}(y|0) \\ \frac{1}{3} \cdot \frac{1}{2} e^{-y/2} &> \frac{2}{3} \cdot \frac{1}{5} e^{-y/5} \\ y &< \frac{10}{3} \ln \frac{5}{4} = 0.7438... \end{aligned}$$

and $\hat{\Theta} = 0$ otherwise.

(b) The estimation error is the probability that the MAP estimate $\hat{\Theta}$ is inconsistent with the real $\Theta$. Let $t = \frac{10}{3} \ln \frac{5}{4}$ be the threshold for estimation. Then

$$\begin{aligned} \mathrm{P}(error) &= \mathrm{P}(\hat{\Theta} \neq \Theta) \\ &= \mathrm{P}(\hat{\Theta} = 1, \Theta = 0) + \mathrm{P}(\hat{\Theta} = 0, \Theta = 1) \\ &= \mathrm{P}(\hat{\Theta} = 1 \mid \Theta = 0) P(\Theta = 0) + \mathrm{P}(\hat{\Theta} = 0 \mid \Theta = 1) P(\Theta = 1) \\ &= 2/3 \cdot \mathrm{P}(T < t \mid \Theta = 0) + 1/3 \cdot \mathrm{P}(T \ge t \mid \Theta = 1) \\ &= 2/3 \cdot F_{T|\Theta}(t \mid 0) + 1/3 \cdot (1 - F_{T|\Theta}(t \mid 1)) \\ &\approx 0.322 \end{aligned}$$

which is only slightly better than trivially always guessing $\theta = 0$.

5. In this question you will investigate sampling using Bayesian statistics. You have a coin of an unknown probability of heads $P$. Your prior is that $P$ is a Uniform$(0,1)$ random variable.

   (a) The coin is flipped 10 times and 9 of the 10 flips are heads. What is the posterior probability that $P > 80\%$?

   (b) (**Optional**) Let $P$ indicate the probability that a triangle whose vertex coordinates are independent Uniform$(0,1)$ random variables is acute. Estimate $P$ by running a computer simulation 500 times. What is the probability that $P$ is within 5% of your estimate?

   (c) A second coin, whose prior is also uniform and independent of the first coin, is flipped 10 times and all flips are heads. What is the probability that the second coin is more biased towards heads than the first one?

**Solution:**

(a) The number of heads $X$ is a Binomial$(10, P)$ random variable. The posterior PDF of $P$ is a Beta$(10,2)$ random variable so

$$f_{P|X}(p|9) = \frac{11!}{9! \cdot 1!} p^9(1-p) = 110p^9(1-p).$$

The posterior probability that $P > 0.8$ is then

$$\Pr(P > 0.8|X = 9) = \int_{0.8}^{1} f_{P|X}(p|9)dp = 110\left(\frac{p^{10}}{10} - \frac{p^{11}}{11}\right)\Bigg|_{0.8}^{1} \approx 0.6779.$$

(b) Running the simulation code on Slide 34 from Lecture 11, I obtained 124 positives and 379 negatives in 500 tries. We want to estimate the probability that $P$ is within 5% of $124/500 = 0.248$ given that $X = 124$, where $X$ is Binomial$(500, P)$. When the prior on $P$ is Uniform$(0,1)$ this is the probability that a Beta$(125, 377)$ random variable is either less than $0.248 - 0.05 = 0.198$ or greater than $0.248 + 0.05 = 0.298$. Using a Beta random variable calculator we find that $\Pr(\text{Beta}(125, 377) < 0.198) \approx 0.283\%$ and $\Pr(\text{Beta}(125, 377) > 0.298) \approx 0.699\%$. So the Bayesian error estimate is less about 0.982%.

Recall that the Central Limit Theorem gave us a comparable error estimate of 1% for 664 samples. Unlike the Bayesian estimate, that one did not assume a prior on $P$.

(c) Let $Q$ and $Y$ be the bias and number of heads of the second coin. We assume the prior of $Q$ is Uniform$(0,1)$ and independent of $P$. We are interested in the probability that $Q > P$ given that $X = 9, Y = 10$. By the total probability theorem,

$$\text{P}(Q > P|X = 9, Y = 10) = \int_0^1 \Pr(Q > p|X = 9, Y = 10, P = p) \cdot f_{P|X,Y}(p|9, 10)dp.$$

Since the coins are independent and $P$ given $X = 9$ is a Beta$(10,2)$ random variable,

$$f_{P|X,Y}(p|9, 10) = f_{P|X}(p|9) = 110p^9(1-p).$$

By the same reasoning, $Q$ given $Y = 10$ is a Beta$(11,1)$ random variable with PDF $f_{Q|Y}(q|10) = 11q^{10}$ and

$$\Pr(Q > p|X = 9, Y = 10, P = p) = \Pr(Q > p|Y = 10) = \int_p^1 q^{10}dq = 1 - p^{11}.$$

Therefore

$$\text{P}(Q > P \mid X = 9, Y = 10) = \int_0^1 (1-p^{11}) \cdot 110p^9(1-p) \cdot dp = \frac{16}{21} \approx 0.7619.$$