Each question is worth 10 points. Explain your answers clearly.

1. You are given one sample that is either Uniform$(-1, 1)$ if $\Theta = 0$ or Uniform$(0, 4)$ if $\Theta = 1$. Your prior on Θ is equally likely ($\mathrm{P}(\Theta = 0) = \mathrm{P}(\Theta = 1) = 1/2$).

   (a) What is the MAP estimator for Θ from this sample?

   Solution: Let the sample be x. By Bayes' rule, $f_{\Theta|X}(\theta|x) \propto f_{X|\Theta}(x|\theta) f_\Theta(\theta)$. Since $\mathrm{P}(\Theta = 0) = \mathrm{P}(\Theta = 1) = 1/2$, the MAP estimator maximizes $f_{X|\Theta}(x|\theta)$, which gives $MAP = 0$ if $x < 1$ and $MAP = 1$ otherwise.

   (b) What is the error probability of the MAP estimator in part (a)?

   Solution: $\mathrm{P}(MAP \neq \Theta) = \mathrm{P}(\Theta = 0, X \geq 1) + \mathrm{P}(\Theta = 1, X < 1) = 0 + 1/2 \cdot 1/4 = 1/8$.

2. A fair n-sided die with equally likely face values $1, 2, \dots, n$ is tossed five times.

   (a) What is the maximum likelihood estimator $\hat{N}$ for n given two samples X_1, X_2?

   Solution: The likelihood is $\mathrm{P}(X_1 = x_2, X_2 = x_2 | n) = 1/n^2$ if $x_1, x_2 \leq n$ and zero otherwise. It is maximized when n is as small as possible, namely $n = \max\{x_1, x_2\}$. Thus $\hat{N} = \max\{X_1, X_2\}$.

   (b) Let X_3 be the next sample. What is the probability that the next sample X_3 takes one of the values $1, 2, \dots, \hat{N}$ in the limit as n tends to infinity?

   Solution: This is the probability of the event E = "$X_3 \leq \max\{X_1, X_2\}$." The complementary event E^c is the event $X_3 > \max\{X_1, X_2\}$, which is the same as "$X_3 > X_1$ and $X_3 > X_2$." Let D be the event that X_1, X_2, and X_3 are all distinct. Conditioned on D, all orderings of X_1, X_2, X_3 are equally likely so $\mathrm{P}(E^c|D) = 1/3$. As $\mathrm{P}(D) = \mathrm{P}(X_2 \neq X_1)\,\mathrm{P}(X_3 \neq X_1, X_2 | X_2 \neq X_1) = (1 - 1/n)(1 - 2/n)$, as n gets large $\mathrm{P}(D)$ tends to one. Therefore $\mathrm{P}(E^c)$ tends to $1/3$ and $\mathrm{P}(E)$ tends to $2/3$ in the limit $n \to \infty$.

3. A random variable has PMF $f(-1) = f(1) = \theta$, $f(0) = 1 - 2\theta$, where θ is unknown ($0 \leq \theta \leq \frac{1}{2}$).

   (a) What is the actual standard deviation σ of the random variable?

   Solution: The mean μ is zero, so $\sigma^2 = (-1)^2 \cdot \theta + 1^2 \cdot \theta + 0^2 \cdot (1 - 2\theta) = 2\theta$ and $\sigma = \sqrt{2\theta}$.

   (b) What is the PMF of the adjusted sample standard deviation S^2 for two samples?

   Solution: $S^2 = (X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 = (X_1 - X_2)^2/2$ takes value 0 when $X_1 = X_2$, value 2 when one of the samples is 1 and the other is -1, and value $1/2$ otherwise. Therefore $\mathrm{P}(S^2 = 0) = \mathrm{P}(X_1 = X_2) = 2\theta^2 + (1 - 2\theta)^2$, $\mathrm{P}(S^2 = 2) = \mathrm{P}(X_1 = 1, X_2 = -1) + \mathrm{P}(X_1 = -1, X_2 = 1) = 2\theta^2$, and $\mathrm{P}(S^2 = 1/2) = 1 - \mathrm{P}(S^2 = 0) - \mathrm{P}(S^2 = 2) = 4\theta(1 - 2\theta)$.

v	0	$1/2$	2
$\mathrm{P}(S^2 = v)$	$2\theta^2 + (1 - 2\theta)^2$	$4\theta(1 - 2\theta)$	$2\theta^2$

Each question is worth 10 points. Explain your answers clearly.

1. X is a Normal$(0, \Theta)$ random variable, where the prior PMF of the parameter Θ is $\mathrm{P}(\Theta = 1/2) = 1/2$, $\mathrm{P}(\Theta = 1) = 1/2$. You observe the following three independent samples of X: $1.0, 1.0, -1.0$.

   (a) What is the posterior PMF of Θ?

   Solution: By Bayes' rule

$$f_{\Theta|X_1X_2X_3}(\theta|1.0, 1.0, -1.0) \propto f_{X_1X_2X_3|\Theta}(1.0, 1.0, -1.0|\theta)\,\mathrm{P}(\Theta = \theta) \propto \frac{1}{\theta^3} e^{-3/2\theta^2}\,\mathrm{P}(\Theta = \theta).$$

   As Θ is equally likely to take values 1/2 and 1, the posterior PMF is

$$f_{\Theta|X_1X_2X_3}(1/2|1.0, 1.0, -1.0) = \frac{8e^{-6}}{8e^{-6} + e^{-3/2}} \quad f_{\Theta|X_1X_2X_3}(1|1.0, 1.0, -1.0) = \frac{e^{-3/2}}{8e^{-6} + e^{-3/2}}.$$

   (b) What is the MAP estimate of Θ?

   Solution: As $e^{-3/2} \approx 0.2231$ is larger than $8e^{-6} \approx 0.0198$, the MAP estimate is 1.

The true fraction of employees in some company that support longer lunch breaks is 80%. Ten employees are polled about their support for longer lunch breaks (randomly with repetition). Find the probability that at least 70% of the polled employees support longer lunch breaks

   (a) by direct calculation (you may use an online calculator if you provide a reference),

   Solution: The number of polled employees T supporting longer lunch breaks is a Binomial$(10, 0.8)$ random variable. We are looking for the probability that such a random variable takes value 7 or more. Using this online calculator we find $\mathrm{P}(\text{Binomial}(10, 0.8) \geq 7) \approx 0.879$.

   (b) using an approximation by a normal random variable.

   Solution: As T has mean $\mu = 8$ and standard deviation $\sigma = \sqrt{10 \cdot 0.8 \cdot 0.2} \approx 1.265$, $\mathrm{P}(T \geq 7) = \mathrm{P}(T \geq 8 - 0.791\sigma)$. A normal approximation would estimate this value by $\mathrm{P}(\text{Normal}(0, 1) \geq -0.791) \approx 0.785$.

   Because T is a discrete random variable, the event $T \geq 7$ is the same as $T > 6$, which can be expressed as $T > 8 - 1.581\sigma$. A normal approximation would then yield an estimate of $\mathrm{P}(\text{Normal}(0, 1) \geq -1.581) \approx 0.943$. In fact any estimate between 0.785 and 0.943 would be a valid answer to this question. The reason for this large range of valid estimates is that the error in applying the Central Limit Theorem for ten indicator samples is rather large.

2. A random variable X is Normal$(1,1)$ with probability p and Normal$(-1,1)$ with probability $1-p$, where the parameter p is unknown.

   (a) What is the maximum likelihood estimator of p from a single sample X?

   Solution: Let Θ be the indicator that X is Normal$(1,1)$. Then by total probability theorem,

$$\begin{aligned} f_X(x) &= f_{X|\Theta}(x|1)f_\Theta(1) + f_{X|\Theta}(x|0)f_\Theta(0) \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{(x-1)^2}{2}} \cdot p + \frac{1}{\sqrt{2\pi}}e^{-\frac{(x+1)^2}{2}} \cdot (1-p) \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2+1}{2}} \left(p(e^x - e^{-x}) + e^{-x}\right) \end{aligned}$$

   The maximum likelihood estimator finds the p between 0 and 1 for which $f_X(x)$ is maximized. Since $f_X(x)$ is a linear function in p with positive slope iff $e^x - e^{-x} > 0$, i.e. when $x > 0$, the ML estimate is

$$ML = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

   (b) Is the estimator in part (a) unbiased? Justify your answer.

   Solution: The expected value of ML (for a fixed value of the parameter p) is

$$\begin{aligned} \mathrm{E}[ML] = \mathrm{P}(X > 0) &= p\,\mathrm{P}(X > 0|\Theta = 1) + (1-p)\,\mathrm{P}(X > 0|\Theta = 0) \\ &= p\,\mathrm{P}(\mathrm{Normal}(1,1) > 0) + (1-p)\,\mathrm{P}(\mathrm{Normal}(-1,1) > 0) \\ &\approx p \cdot 0.841 + (1-p) \cdot 0.158. \end{aligned}$$

   This is not an unbiased estimator of p: For example when $p = 1$, $\mathrm{E}[ML] \approx 0.841$. In fact, you can conclude the answer is no without doing any calculation: $\mathrm{E}[ML] = \mathrm{P}(X > 0)$ is always strictly smaller than 1, so it cannot be an unbiased estimator when $p = 1$.

1. You are trying to estimate the fraction V of vegetarians in Hong Kong using Bayesian statistics. Your prior is that V is a Uniform$(0, 1/2)$ random variable.

   (a) You poll a random person and they are not a vegetarian. What is the posterior PDF of V?

   Solution: By Bayes's rule, $f_{V|X}(v|0) \propto \mathrm{P}(X = 0|V = v)f_V(v) \propto 1 - v$ for $v \in [0, 1/2]$. As $\int_0^{1/2}(1-v)dv = \frac{3}{8}$, the posterior PDF is $\frac{8}{3}(1-v)$ for $v \in [0, 1/2]$.

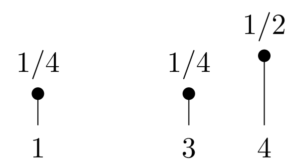
   (b) What is the expected posterior probability that the next polled person will be a vegetarian?

   Solution: It is $\mathrm{E}[V|X = 0] = \int_0^{1/2} v \cdot \frac{8}{3}(1-v)dv = \frac{8}{3}(\frac{1}{8} - \frac{1}{24}) = \frac{2}{9} \approx 0.222$.

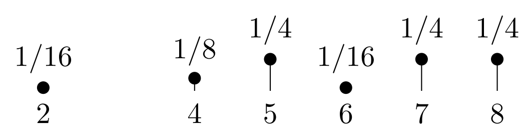
2. A food company produced 1000 boxes of biscuits, 500 of which contain 4 biscuits, 250 contain 3 biscuits and 250 contain one biscuit. You sample two boxes (with repetition) and record the sample mean $\overline{X}$ of the number of biscuits.

   (a) What is the PMF of $\overline{X}$?

   Solution: The marginal PMF of each sample is $f(4) = \frac{1}{2}$, $f(3) = \frac{1}{4}$, $f(1) = \frac{1}{4}$:



   We can calculate the PMF of $X_1 + X_2$ using the convolution formula to get:



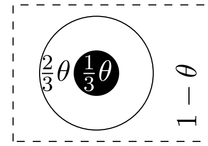
   The PMF of $\overline{X} = (X_1 + X_2)/2$ is then obtained by scaling the value by $\frac{1}{2}$:

x	1	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4
$\mathrm{P}(\overline{X} = x)$	1/16	1/8	1/4	1/16	1/4	1/4

   (b) What is the probability that the sample mean equals the actual mean?

   Solution: The actual mean is $\mu = 1 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{2} = 3$, so $\mathrm{P}(\overline{X} = \mu) = 1/16$.

3. An archer hits the bull's eye with probability $\frac{1}{3}\theta$, the rest of the target with probability $\frac{2}{3}\theta$, and misses the target with probability $1-\theta$, where $\theta \in [0,1]$ is a parameter that models the archer's skill.



(a) The archer hits the bull's eye twice and misses the board once. What is the maximum likelihood estimate of their skill θ (assuming their shots are independent)?

Solution: The probability of this outcome is proportional to $(\frac{1}{3}\theta)^2 \cdot (1-\theta) \propto \theta^2 - \theta^3$. The critical points are those θ for which $(d/d\theta)(\theta^2-\theta^3) = 2\theta - 3\theta^2$ equals zero, namely 0 and $2/3$, together with the other endpoint $\theta = 1$. The maximum occurs at $\theta = 2/3$. This is the maximum likelihood estimate.

(b) Describe an unbiased estimator (3 points) of minimum variance (+2 points) for the player's skill θ from a single attempt. (**Hint:** The estimator assigns a "score" to each outcome.)

Solution: One unbiased estimator assigns a score S of 1 for hitting the board (bull's eye or not) and 0 for missing it. Then $\mathrm{E}[S] = 1 \cdot \frac{2}{3}\theta + 1 \cdot \frac{1}{3}\theta + 0 \cdot (1-\theta) = \theta$.

A general estimator S will assign scores a, b, c for the bull's eye, the rest of the board, and a miss. For the estimator to be unbiased we need that $\frac{1}{3}\theta a + \frac{2}{3}\theta b + (1-\theta)c = \theta$ for all possible skill levels θ, from where we must choose $c = 0$ and $a + 2b = 3$. The variance of the estimator is $\mathrm{Var}[S] = \mathrm{E}[S^2] - \mathrm{E}[S]^2 = \frac{1}{3}\theta a^2 + \frac{2}{3}\theta b^2 - \theta^2$, so we need to minimize $\frac{1}{3}a^2 + \frac{2}{3}b^2 = \frac{1}{3}(3-2b)^2 + \frac{2}{3}b^2$. This is an increasing quadratic function of b so it is minimized at the critical point where its first derivative $\frac{4}{3}(2b-3) + \frac{4}{3}b$ vanishes, namely $b = 1$ and $a = 1$. So the proposed unbiased estimator happens to be the one of minimum variance.