Problem 1 $\longrightarrow$

(A.) given $P(\theta) = \text{Beta}(\theta|1,1) \propto x^{1-1}(1-x)^{1-1}$

hence $P(\theta)$ is a uniform distribution

$\cancel{P(X<3) = P(X=0) + P(X=1) + P(X=2)}$

let's define the random variable X to denote the number of heads.

$$P(\theta|X<3) = \frac{P(X<3|\theta)\,P(\theta)}{P(X<3)}$$

as it is the normalisation constant we take it to be some $K \in \mathbb{R}$

$$P(\theta|X<3) \propto P(X<3|\theta)\,P(\theta)$$

$$\propto \big(P(X=0|\theta) + P(X=1|\theta) + P(X=2|\theta)\big)\,P(\theta)$$

$$= \left\{\text{Bin}(X=0|\theta,5) + \text{Bin}(X=1|\theta,5) + \text{Bin}(X=2|\theta,5)\right\} P(\theta)$$

$$= \left\{\binom{5}{0}\theta^0(1-\theta)^5 + \binom{5}{1}\theta^1(1-\theta)^4 + \binom{5}{2}\theta^2(1-\theta)^3\right\}$$

$\text{Beta}(\theta|1,1)$ $P(\theta)$ $\Downarrow$ $\mathcal{U}(0,1)$

hence, $P(\theta|X<3) \propto \binom{5}{0}\theta^0(1-\theta)^5 + \binom{5}{1}\theta^1(1-\theta)^4 + \binom{5}{2}\theta^2(1-\theta)^3$

(B.) $D = \{x_1, x_2 \cdots x_n\}$

(1) to write the likelihood we assume that x_i's are independent hence $P(D|\alpha) = \prod_{i=1}^{n} p(x_i|\alpha)$ where we use the product rule.

$$P(D|\alpha) = \prod_{i=1}^{n} \frac{1}{2\alpha} \mathbb{I}(x_i \in [\alpha, \alpha]) = \frac{1}{(2\alpha)^n} \prod_{i=1}^{n} I(x_i \in [-\alpha, \alpha])$$

note :-

(a) $f(\alpha) = \frac{1}{(2\alpha)^n} \Rightarrow f(\alpha) > 0 \quad f'(\alpha) = \frac{-n}{2^n \alpha^{n+1}} < 0$

$$\frac{d}{d\alpha} \ln(f(\alpha)) = \frac{f'(\alpha)}{f(\alpha)} > 0$$

hence it is monotonically decreasing

(b) $$\prod_{i=1}^{n} I(x_i \in [-\alpha, \alpha]) = \begin{cases} 0 & \text{if any } x_i \notin [-\alpha, \alpha] \\ 1 & \text{if all } x_i \in [-\alpha, \alpha] \end{cases}$$

hence to find the MLE we need to find α^* such that D is a subset of $[-\alpha^* \ \alpha^*]$. let's define $|D| = \{|x_1|, |x_2| \cdots |x_n|\}$ then we can say that the MLE estimate will be

$$\alpha^* = \sup \text{ or } \max(|D|) \quad \text{or} \quad \max(|x_i|)$$

(ii) $$p(x_{n+1}) = \frac{1}{2a^*} \mathbb{I}\left(x_{n+1} \in [-a^*, a]\right) = \begin{cases} 1/2a^* & \text{if } a^* \geq |x_{n+1}| \\ 0 & \text{else.} \end{cases}$$

(iii) if we see carefully then if $x_{n+1} > \sup(|D|)$ then we will get $p(x_{n+1}) = 0$ no matter what using the MLE estimate. so if our dataset is small then accuracy will drop.

a better approach will be to use a MAP estimate and Bayesian analysis.

Problem 2

we aim to solve this problem by converting each of the standard representations into the form of an exponential family.

$f(x|\theta)$ be a PDF where "θ" are the parameters then it is called to be a exponential family if it is in the following form →

$$f(x|\theta) = h(x) \cdot \exp(\eta(\theta) \cdot T(x) - A(\theta))$$

h, η, T and A are known postive functions

$T \equiv$ sufficient statistics, A = log-partition

① normal disribution with known S.D. →

$$P(x|\mu, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{1}{2\sigma_0^2}(x-\mu)^2\right)$$

$$= \underbrace{\left\{\frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(\frac{-x^2}{2\sigma_0^2}\right)\right\}}_{h(x)} * \left\{\exp\left(\underbrace{\frac{\mu}{\sigma_0^2}}_{\eta(\theta)} \underbrace{x}_{T(x)} - \underbrace{\frac{\mu}{\sigma_0^2}}_{A(\theta)}\right)\right\}$$

same proof for the case with a known mean.

(ii) binomial distribution → $P(x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$ $\quad x = 0, 1, \dots n$

$$= \binom{n}{x} \exp(x\ln(\theta) + (n-x)(\ln(1-\theta)))$$

$$= \underbrace{\binom{n}{x}}_{h(x)} \exp\left(\underbrace{\ln\left(\frac{\theta}{1-\theta}\right)}_{\eta(\theta)} \underbrace{x}_{T(x)} + \underbrace{n\ln(1-\theta)}_{-A(\theta)}\right)$$

(iii) poisson's distribution → $P(x|\theta) = \frac{\theta^x}{x!}e^{-\theta}$ $\quad x = 0, 1 \dots$

$$= \underbrace{\frac{1}{\lfloor x+1}}_{h(x)} \exp\left(\underbrace{-\theta}_{A(\theta)} + \underbrace{x}_{T(x)} \underbrace{\ln(\theta)}_{\eta(\theta)}\right)$$

(iv) geometric distribution → $P(x|\theta) = (1-\theta)^{K-1}\theta$

$$= \exp(\ln(\theta) + (K-1)\ln(1-\theta))$$

$$= \exp\left(\ln\left(\frac{\theta}{1-\theta}\right) + K\ln(1-\theta)\right)$$

$h(x) = 1$, $\eta(\theta) = \ln\left(\frac{\theta}{1-\theta}\right)$, $T(x) = 1$, $A(\theta) = -K\ln(1-\theta)$

exponential distribution $\rightarrow$ $P(x|\theta) = \begin{cases} \theta \exp(-\theta x) & x > 0 \\ 0 & x \leq 0 \end{cases}$

$$P(x|\theta) = \begin{cases} \exp(\ln(\theta) - \theta x) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

alternate form for an exponential family is $\rightarrow$ $f(x|\theta) = \underset{=1}{h(x)}\ \underset{=\theta}{g(\theta)}\ \exp(\underset{-\theta}{\eta(\theta)}\ \underset{x}{T(x)})$

hence we can conclude that all the given distributions are exponential family forms.

Problem 3 $\rightarrow$

as $f_\theta(x)$ is a PDF $\Rightarrow \int_{-\infty}^{\infty} f_\theta(x)\,dx = 1$ $\quad\cdots$ ①

define a function $l_\theta(x) = \ln(f_\theta(x)) \Rightarrow l' = \left(\frac{f'}{f}\right)$ {differentiating with respect to θ}

$$E[l_\theta'] = \int_{-\infty}^{\infty} l' f \, dx = \int_{-\infty}^{\infty} \left(\frac{f'}{\cancel{f}}\right)\cancel{f}\,dx = \int_{-\infty}^{\infty} f'\,dx = 0$$

note that from ① we have

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{\infty} f\,dx = \int_{-\infty}^{\infty} \frac{\partial f}{\partial\theta}\,dx = 0 \quad \text{or} \quad \int_{-\infty}^{\infty} f'\,dx = 0$$

if f is an exponential family then we can consider the following $\rightarrow \ln(f_\theta(x)) = \ln(h(x)) + \eta T(x) - A(\theta)$

$\eta \rightarrow$ or θ (considering the definition in the notes)

$$\rightarrow \frac{d}{d\theta}\left(\ln(f_\theta(x))\right) = 0 + T(x) - A'(\theta)$$

$$E[l_\theta'] = 0 \Rightarrow E[T(x) - A'(\theta)] = 0$$

$$\Rightarrow E[T(x)] = A'(\theta) = \frac{\partial A}{\partial\theta}$$