Q1

$$p(x|\theta) \sim U(0,\theta) = \begin{cases} 1/\theta & 0 \le x \le \theta \\ 0 & \text{otherwise} \end{cases}$$

(a) $D = \{x_1, \cdots, x_n\}$

$$I(x) = \begin{cases} 1 & \text{if logical value of } x - \text{True} \\ 0 & \text{else} \end{cases}$$

Basically $I(x)$ is indicator function.
So,

$$p(D|\theta) = \prod_{k=1}^{n} p(x_k|\theta)$$

$$= \prod_{k=1}^{n} \frac{1}{\theta} I(0 \le x_k \le \theta)$$

$$= \frac{1}{\theta^n} I\left(\theta \ge \max_k x_k\right) \cdot I\left(\min_k x_k \ge 0\right)$$

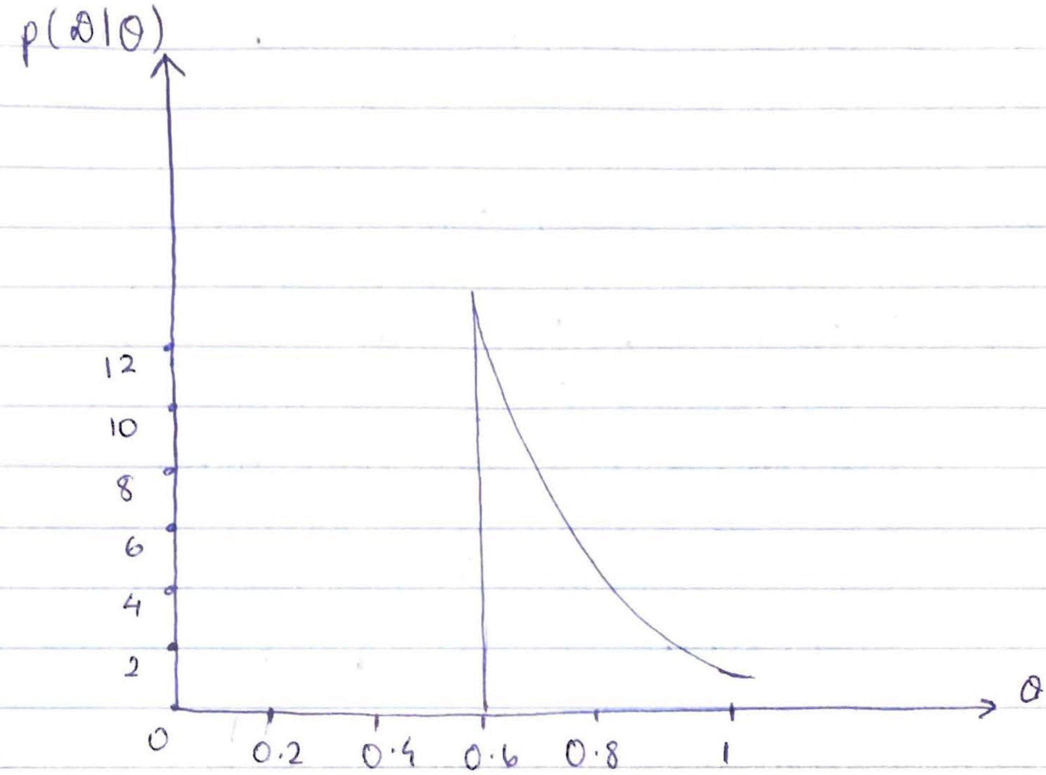
Here, as θ increases, $\frac{1}{\theta^n}$ decreases monotonically.

Also, $I\left(\theta \ge \max_k x_k\right) = 0$ if θ is less than the maximum value of x_k. So,

$$\boxed{\hat{\theta}_{ML} = \max_k x_k}$$

(b)

P.T.O



The maximum likelihood estimate for θ is max [D] i.e., the value of the maximum element in D as proved in part (a).

Here, $\max_k x_k = 0.6$

Also, $p(D|\theta) = \frac{1}{\theta^n} I()$ which is monotonically decreasing with increase in θ. Knowing the maximum value, means we know the starting point. Thus, without knowing the values of the other four points, we can plot a graph for likelihood $p(D|\theta)$ in range $0 \le \theta \le 1$

Q2.

$$p(x|\theta) = \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$$

$\mathcal{D} = [x_1, \dots, x_n]$ is a set of n samples independently drawn according to this probability density.

(a) $s = (s_1, \dots, s_d)^T$ is sum of n samples.

Let

$$x_k = (x_{k1}, \dots, x_{kd})^T \quad \text{for } k \in 1, \dots, n$$

Then,

$$s_i = \sum_{k=1}^{n} x_{ki} \quad i = 1, \dots, d$$

Also,

$$p(\mathcal{D}|\theta) = P(x_1, \dots, x_n | \theta)$$

$$= \prod_{k=1}^{n} (x_k|\theta) \longrightarrow x_k \text{ are independent}$$

$$= \prod_{k=1}^{n} \prod_{i=1}^{d} \theta_i^{x_{ki}} (1-\theta_i)^{1-x_{ki}}$$

$$= \prod_{i=1}^{d} \theta_i^{\sum_{k=1}^{n} x_{ki}} (1-\theta_i)^{\sum_{k=1}^{n} (1-x_{ki})}$$

$$= \prod_{i=1}^{d} \theta_i^{s_i} (1-\theta_i)^{n-s_i}$$

Hence Proved.

(b) Assume uniform prior for θ

$$\therefore \; p(\theta) = 1 \quad \text{for} \quad 0 \le \theta_i \le 1 \quad i = 1, \dots, d$$

Now,

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta) \cdot p(\theta)}{p(\mathcal{D})}$$

From (a),

$$p(\mathcal{D}|\theta) = \prod_{i=1}^{d} \theta_i^{s_i} (1-\theta)^{n-s_i}$$

So,

$$p(\mathcal{D}) = \int p(\mathcal{D}|\theta)\, p(\theta)\, d\theta$$

$$= \int \prod_{i=1}^{d} \theta_i^{s_i} (1-\theta_i)^{n-s_i} d\theta$$

$$= \int_0^1 \cdots \int_0^1 \prod_{i=1}^{d} \theta_i^{s_i} (1-\theta_i)^{n-s_i} d\theta_1 d\theta_2 \cdots d\theta_d$$

$$= \prod_{i=1}^{d} \int_0^1 \theta_i^{s_i} (1-\theta_i)^{n-s_i} d\theta_i$$

Now, $s_i = \sum_{k=1}^{n} x_{ki}$ takes values in set $\{0, 1, \cdots, n\}$ for $i = 1, \cdots, d$ and if we use the identity

$$\int_0^1 \theta^m (1-\theta)^n d\theta = \frac{m!\, n!}{(m+n+1)!}$$

Substituting above,

$$p(\mathcal{D}) = \prod_{i=1}^{d} \int_0^1 \theta_i^{s_i} (1-\theta_i)^{n-s_i} d\theta_i = \prod_{i=1}^{d} \frac{s_i!\, (n-s_i)!}{(n+1)!}$$

So, $p(\theta|\mathcal{D}) = \dfrac{p(\mathcal{D}|\theta) \cdot p(\theta)}{p(\mathcal{D})}$

$$= \frac{\prod_{i=1}^{d} \theta_i^{s_i} (1-\theta_i)^{n-s_i}}{\prod_{i=1}^{d} \frac{s_i! \, (n-s_i)!}{(n+1)!}}$$

$$= \prod_{i=1}^{d} \frac{(n+1)!}{s_i! \, (n-s_i)!} \cdot \theta_i^{s_i} (1-\theta_i)^{n-s_i}$$

c) $d = 1, n = 1$.

$$p(\theta_1|\mathcal{D}) = \frac{2!}{s_1! \, (n-s_1)!} \theta_1^{s_1} (1-\theta_1)^{n-s_1}$$

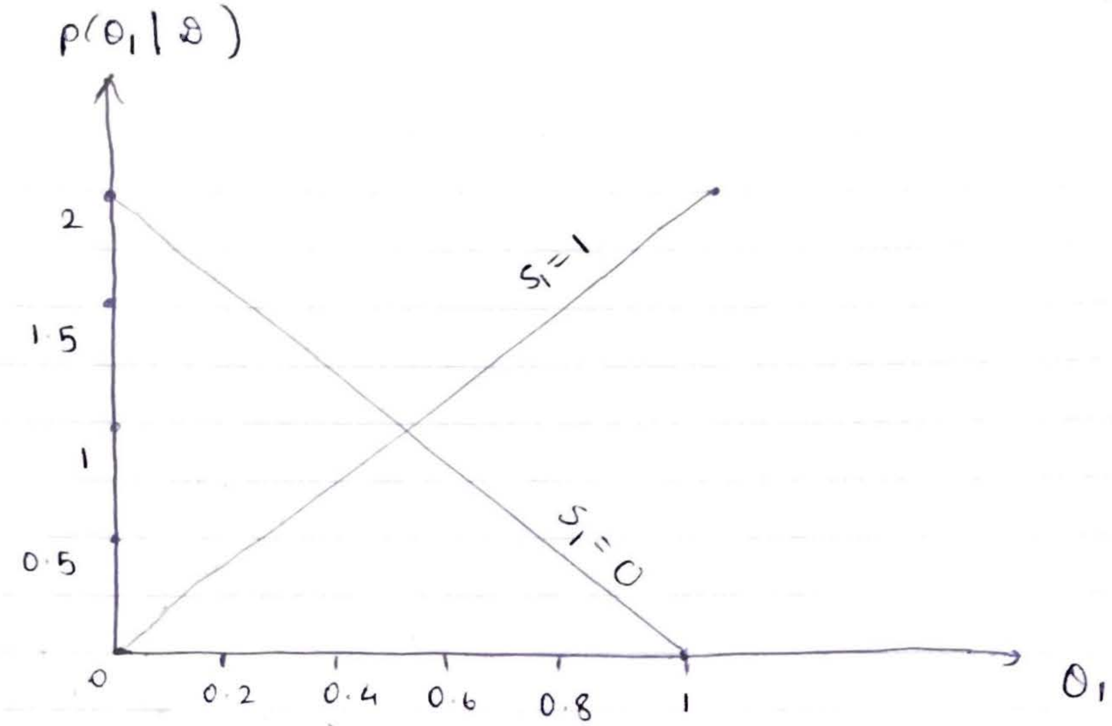
$$= \frac{2}{s_1! \, (1-s_1)!} \theta_1^{s_1} (1-\theta_1)^{1-s_1}$$

s_1 takes discrete values 0 and 1. Thus, densities are:

$s_1 = 0$: $p(\theta_1|\mathcal{D}) = 2(1-\theta_1)$
$s_1 = 1$: $p(\theta_1|\mathcal{D}) = 2\theta_1$

So,

for $0 \leq \theta_1 \leq 1$, we have,



d) $$p(x|\mathcal{D}) = \int p(x|\theta)\, p(\theta|\mathcal{D})\, d\theta$$

$$= \int \prod_{i=1}^{d} \theta_i^{x_i+s_i} (1-\theta_i)^{1-x_i+n-s_i} \frac{(n+1)!}{s_i!\,(n-s_i)!}\, d\theta$$

$$= \prod_{i=1}^{d} \frac{(n+1)!}{s_i!\,(n-s_i)!} \int \theta_i^{x_i+s_i} (1-\theta_i)^{1-x_i+n-s_i}\, d\theta$$

$$= \prod_{i=1}^{d} \frac{(n+1)!}{s_i!\,(n-s_i)!} \frac{(x_i+s_i)!\,(1-x_i+n-s_i)!}{(n+2)!}$$

$$= \prod_{i=1}^{d} \frac{(x_i+s_i)!\,(1-x_i+n-s_i)!}{(n+2)\, s_i!\,(n-s_i)!}$$

Now,

for $x_i = 0$;

$$\frac{(x_i + s_i)!\,(1 - x_i + n - s_i)!}{(n+2)\, s_i!\,(n - s_i)!}$$

$$= 1 - \frac{s_i + 1}{n + 2}$$

For $x_1 = 1$,

$$\frac{(x_i + s_i)!\,(1 - x_i + n - s_i)!}{(n+2)\, s_i!\,(n - s_i)!}$$

$$= \frac{s_i + 1}{n + 2}$$

Hence, $$P(x \mid \mathcal{D}) = \left(\frac{s_i + 1}{n + 2}\right)^{x_i} \left(1 - \frac{s_i + 1}{n+2}\right)^{1 - x_i}$$

e) From part (d), effective Bayesian estimate for θ is

$$\hat{\theta}_i = \frac{s_i + 1}{n + 2} \qquad i = 1, 2, 3, \ldots d$$

This is because,

$$P(\underline{x} \mid s_i, \underline{z}_i) \approx P(\underline{x} \mid \hat{\underline{\theta}}_i, \mathcal{S}_i)$$

Thus, we can compare

$$P(x \mid \mathcal{D}) \text{ and } P(x \mid \theta)$$

f) Baye's Minimum Error rule states that

PTO.

if $p(x|S_k)\, p(S_k) > p(x|S_j)\, p(S_j)$
Then $x \in S_k$

Here, $p(x|S_k) \approx p(x|\theta; S_k)$

$$= \prod_{i=1}^{d} \left(\frac{s_i^{(k)} + 1}{n+2} \right)^{x_i} \left(1 - \frac{s_i^{(k)} + 1}{n+2} \right)^{(1-x_i)}$$

and $P(S_k) = \dfrac{n_k}{\sum_{i=1}^{c} n_i}$

Hence, Baye's minimum error decision rule is $\Rightarrow$

$$\prod_{i=1}^{d} \left[\frac{s_i^{(k)} + 1}{n+2} \right]^{x_i} \left[1 - \frac{s_i^{(k)} + 1}{n+2} \right]^{1-x_i} \frac{n_k}{\sum_{j=1}^{c} n_j}$$

$$> \prod_{i=1}^{d} \left[\frac{s_i^{(j)} + 1}{n+2} \right]^{x_i} \left[1 - \frac{s_i^{(j)} + 1}{n+2} \right]^{1-x_i} \frac{n_j}{\sum_{j=1}^{c} n_j}$$

Then $\underline{\underline{x \in S_k}}$