Q1

a) Here $p(\theta) \sim N(25, 3^2)$ [from literature]

And when we take the instrument, ~~then~~ we get $y=28$ as output.

$$p(y|\theta) \sim N(\theta, 5^2)$$, 'θ' here is chosen from the prior gaussian.

Using Bayes theorem:

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

Here $p(y) = \int_{0}^{\infty} p(y|\theta) p(\theta) \, d\theta$, which will come out to be a constant

$\therefore$ we can say.

$$p(\theta|y) \propto p(\theta) \cdot p(y|\theta).$$

To find θ_{map}, we can find θ that maximizes $[p(\theta) \cdot p(y|\theta)]$

Plugging in the gaussians, we get

$$p(\theta|y) \propto \left[\frac{1}{\sqrt{2\pi}(3)} e^{-\frac{(\theta-25)^2}{2(3)^2}}\right] \cdot \left[\frac{1}{\sqrt{2\pi}(5)} e^{-\frac{(y-\theta)^2}{2(5)^2}}\right]$$

$$\therefore p(\theta|y) \propto K \cdot e^{\left(-\frac{(y-\theta)^2}{50} - \frac{(\theta-25)^2}{18}\right)} \quad \text{--- (1)}$$

To find which θ maximizes the expression (1), we will differentiate wrt θ.

$$\frac{d}{d\theta} p(\theta/y) = \frac{d}{d\theta}\left(e^{-\frac{(y-\theta)^2}{50} - \frac{(\theta-25)^2}{18}}\right)$$

$$= \left[e^{-\frac{(y-\theta)^2}{50} - \frac{(\theta-25)^2}{18}}\right]\left[-\frac{2(y-\theta)(-1)}{50} - \frac{2(\theta-25)(1)}{18}\right] = 0$$

Since e^{-x} is never 0, it can only be asymptotic.

We can say

$$-\frac{2(y-\theta)(-1)}{50} - \frac{2(\theta-25)}{18} = 0$$

$$\frac{28-\theta}{25} = \frac{\theta-25}{9}$$

$$34\theta = 625 + 28\times 9$$

$$\theta = 25.7941 \quad] \text{ Ans I.a}$$

b) Now, if we exchange the prior and the likelihood.

$p(\theta) \sim N(28, 5^2)$

& $p(y/\theta) \sim N(\theta, 3^2)$

then

$$\frac{d}{d\theta} p(\theta/y) = \frac{d}{d\theta}\left[e^{-\frac{(y-\theta)^2}{18} - \frac{(\theta-28)^2}{50}}\right]$$

$$= \left(e^{-\frac{(y-\theta)^2}{18} - \frac{(\theta-28)^2}{50}}\right) \cdot \left(\frac{-2(y-\theta)(-1)}{18} - \frac{2(\theta-28)}{50}\right)$$

Again,

$$\frac{2(y-\theta)}{18} - \frac{2(\theta-28)}{50} = 0$$

$$\frac{y-\theta}{9} = \frac{\theta-28}{25}$$

$$\frac{25(25-\theta)}{9} = 9\theta - 28\times 9$$

$$625 - 25\theta = 9\theta - 28\times 9$$

$$625 + 28\times 9 = 34\theta$$

$$\left.\theta = 25.79411\ \right]\ \text{Ans 1.b.}$$

'θ' did not change when we ~~reversed~~ exchanged the $p(\theta)$ (prior) and the $p(y|\theta)$ (likelihood), probably because

c) literature $\sim \mathcal{N}(\mu_1, \sigma_1)$.

Measurement $\sim \mathcal{N}(\theta, \sigma_2)$.

Using Bayes rule ~~p(y|θ)~~ $p(\theta|y) = \dfrac{p(y|\theta)\cdot p(\theta)}{p(y)}$.

Here we ignore $p(y)$ & it is just a normalizing factor, so.

For $\theta_{map} = \underset{\theta}{\operatorname{argmax}}\ p(y|\theta)\cdot p(\theta)$

$$= \underset{\theta}{\operatorname{argmax}}\ p(\theta)\cdot p(y|\theta)$$

$$= \underset{\theta}{\operatorname{argmax}}\ \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(\theta-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\theta)^2}{2\sigma_2^2}}$$

$$= \underset{\theta}{\operatorname{argmax}}\ \frac{1}{2\pi\sigma_1\sigma_2} \cdot e^{-\frac{(\theta-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\theta)^2}{2\sigma_2^2}}$$

Applying log

$$= \underset{\theta}{\operatorname{argmax}}\left[\log\frac{1}{2\pi\sigma_1\sigma_2} + \frac{-(\theta-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\theta)^2}{2\sigma_2^2}\right]$$

Derivating | Calculating the gradient

$$= \frac{d}{d\theta}\left[\log\frac{1}{2\pi\sigma_1\sigma_2} + \frac{-(\theta-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\theta)^2}{2\sigma_2^2}\right]$$

$$= 0 - \frac{\cancel{2}(\theta-\mu_1)(1)}{\cancel{2}\sigma_1^2} + \frac{\cancel{2}(y-\theta)(1)}{\cancel{2}\sigma_2^2}$$

$$\nabla = -\frac{(\theta-\mu_1)}{\sigma_1^2} + \frac{y-\theta}{\sigma_2^2}$$

Equating this to 0.

$$-\frac{(\theta-\mu_1)}{\sigma_1^2} = \frac{\theta - y}{\sigma_2^2}$$

$$-\frac{\theta}{\sigma_1^2} + \frac{\mu_1}{\sigma_1^2} = \frac{\theta}{\sigma_2^2} - \frac{y}{\sigma_2^2}$$

$$\cancel{\frac{\theta}{\sigma_1^2}} \; \frac{\mu_1}{\sigma_1^2} + \frac{y}{\sigma_2^2} = \frac{\theta}{\sigma_2^2} + \frac{\theta}{\sigma_1^2}$$

$$\theta\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) = \frac{\mu_1}{\sigma_1^2} + \frac{y}{\sigma_2^2}$$

$$\theta = \frac{\mu_1\left(\frac{1}{\sigma_1^2}\right)}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} + \frac{y\left(1/\sigma_2^2\right)}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

$$\therefore \theta_{map} = \frac{1/\sigma_1^2}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}\mu_1 + \frac{\frac{1}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}y$$

Q2

$$\hat{p}_v = \frac{N_v + m}{N + |v| m}$$

Here $\hat{p}$ is the estimator which is the smoothed probability for the frequency estimator.

$p_v = P(X = v)$

$$Y_v = \begin{cases} 1 & \text{if } x_t = v \\ 0 & \text{otherwise} \end{cases}$$

We define N_v to be a random variable which takes values from Y

$\therefore \; N_v = Y_1 + Y_2 + Y_3 + \cdots + Y_N \quad \forall x^t \in X$

x^t here are the samples drawn from 'D'.

Now, $\hat{p}_v$ is the estimator and the original value is $p_v = \frac{N_v}{N}$

Bias $= E(d) - \theta$

Here $d = \hat{p}_v$
$\theta = p_v$

$$\text{Bias} = E\left[\frac{N_v + m}{N + |v| m}\right] - \frac{N_v}{N}$$

Since '$N + |v|m$' is a constant, we take it out

$$= \frac{1}{N + |v| m}\left[E[N_v] + E[m]\right] - \frac{N_v}{N}$$

Calculating $E[N_v]$

$= E[Y_1 + Y_2 + Y_3 + \cdots Y_N]$

$= E[Y_1] + E[Y_2] + \cdots + E[Y_N]$

$= \sum 0 \cdot (1-p(0)) + 1 \cdot p(1)) + \cdots + \sum 0 \cdot (1-p(0)) + 1(p(1))$

$= \underbrace{p_v + p_v + p_v + \cdots + p_v}_{N \text{ terms}}$

$= Np_v$

Bias $= E[d] - \theta$ Here $\theta = p_v$.

$= E\left[\frac{N_v + m}{N + |V| m}\right] - p_v$

$= \frac{1}{N + 0.3}(Np_v + m) - p_v$

$= \frac{Np_v + m - (Np_v + p_v(0.3))}{N + 0.3}$

$= \boxed{\frac{m - 0.3 p_v}{N + 0.3}}$

$= \cancel{0.1 - 0.3 p_v}$ $\boxed{\frac{0.1 - 0.3 p_v}{N + 0.3}}$ Ans

Q3 let us build a cost matrix for the company.

$x = \$350$

	Stolen	Not Stolen
Approved	350	-7
Not Approved	0	1750

Given $P(\text{Stolen}|x) = 0.26$

$\therefore P(\overline{\text{Stolen}}|x) = 0.74$

a) If purchase is approved & the card is stolen, then the cost would be \$350.

b) The cost to the company will be (\$-7), i.e the company will gain money frome the transaction.

$\therefore$ Cost to company = \$ (-7)

i.e profit to company = \$7.

c) Expected cost to the company if purchase is approved is

$\cancel{R(A|x) = 350 \times P(A|S, x) + (-7) P(}$

$$R(A|x) = 350 \times P(S|x) + (-7) P(\bar{S}|x)$$

$$= 350 \times (.26) + (-7)(.74)$$

$$= \$85.75$$

(d) Expected cost to the company if the purchase is denied.

$$R(\text{Denied}|x) = 0 \cdot P(S|x) + 1750 \cdot P(\bar{S}|x)$$

$$= 0 + 1750 \times 0.74$$

$$= \$1295$$

(e) Minimum risky decision among the two is to approve the purchase

Q_4

a) $D = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \\ 1 & 5 & 4 \\ 1 & 8 & 7 \end{bmatrix}$ $\quad r = \begin{bmatrix} 3 \\ 7 \\ 2 \\ 1 \end{bmatrix}$

$D^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \\ 3 & 2 & 4 & 7 \end{bmatrix}$

To compute w's, we use the formula.

$$w = (D^T D)^{-1} D^T r$$

$$D^T D = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \\ 1 & 5 & 4 \\ 1 & 8 & 7 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \\ 3 & 2 & 4 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 18 & 16 \\ 18 & 106 & 87 \\ 16 & 87 & 78 \end{bmatrix}$$

Now,

$$(D^T D)^{-1} = \frac{1}{|D^T D|} \times \text{Adjugate}(D^T D)$$

$|D^T D|$ = Determinant of $D^T D$, & Solving for $|D^T D|$,

we get $|D^T D| = 500$.

$$\text{Adjugate}(D^TD) = \begin{bmatrix} 699 & -12 & -130 \\ -12 & 56 & -60 \\ -130 & -60 & 100 \end{bmatrix}$$

$$(D^TD)^{-1} = \frac{1}{500}\begin{bmatrix} 699 & -12 & -130 \\ -12 & 56 & -60 \\ -130 & 60 & 100 \end{bmatrix}$$

Now,

$$(D^TD)^{-1}D^T = \frac{1}{500}\begin{bmatrix} 699 & -12 & -130 \\ -12 & 56 & -60 \\ -130 & 60 & 100 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \\ 3 & 2 & 4 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} 0.594, & 0.782, & 0.238, & -0.614 \\ -0.272, & 0.184, & 0.056, & 0.032 \\ 0.22, & -0.34, & -0.06, & 0.18 \end{bmatrix}$$

Now, last step

$$W = (D^TD)^{-1}D^T r$$

$$= (D^TD^{-1})D^T\begin{bmatrix} 3 \\ 7 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 7.118 \\ 0.616 \\ -1.66 \end{bmatrix}$$

(b) Using only the first two examples from the dataset, we get,

$$D = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \end{bmatrix} \quad y = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

$$D^T = \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 3 & 2 \end{bmatrix}$$

$$D^T D = \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \end{bmatrix}$$

$$= \cancel{\begin{bmatrix} 2 & 5 & 5 \\ 5 & 17 & 11 \\ 5 & 11 & 13 \end{bmatrix}} \quad \begin{bmatrix} 2 & 5 & 5 \\ 5 & 17 & 11 \\ 5 & 11 & 13 \end{bmatrix}$$

$$(D^T D)^{-1} = \frac{1}{|D^T D|} \times \text{Adjugate}(D^T D)$$

$$\cancel{|D^T D| = 2(17 \times 13 - 11 \times 5) - 5(5 \times 13 - 5 \times 5) + 5(5 \times 11 - 17 \times 5)}$$

$$\cancel{= 332 - 200 + 5 - 150}$$

$$\cancel{= 332 - 350}$$

$$\cancel{= -18}$$

$$D^T D = 2(17 \times 13 - 11^2) - 5(13 \times 5 - 11 \times 5) + 5(11 \times 5 - 17 \times 5)$$

$$= 0$$

$$(D^TD)^{-1}D^T = \begin{bmatrix} -1 & 0 \\ -0.1875 & 0.375 \\ 0.625 & 0.25 \end{bmatrix}$$

$$(D^TD)^{-1}D^T y = \begin{bmatrix} -1 & 0 \\ -0.1875 & 0.375 \\ 0.625 & 0.25 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

$$= \begin{bmatrix} -3 \\ 2.0625 \\ 3.625 \end{bmatrix}$$

Here $(D^TD)^{-1}$ cannot be calculated as its determinant ~~inverse~~ comes out to be 0.

$\therefore$ D^TD is a singular matrix

Reason for why it happens :

We can see that in the matrix,

$$\begin{bmatrix} 2 & 5 & 5 \\ 5 & 17 & 11 \\ 5 & 11 & 13 \end{bmatrix}$$

if we observe closely, there is linear dependency present in the matrix.

To prove the linear dependency, we can follow the following transformations to the matrix.

$$\begin{array}{ccc} A & B & C \end{array}$$
$$\begin{bmatrix} 2 & 5 & 5 \\ 5 & 17 & 11 \\ 5 & 11 & 13 \end{bmatrix}$$

let $C = 2 \times B + C$

$$\begin{bmatrix} 2 & 5 & 15 \\ 5 & 17 & 45 \\ 5 & 11 & 35 \end{bmatrix},$$

Now taking out 5 common from third column, we get

$$\frac{1}{5}\,\cancel{\begin{bmatrix} 2 & 5 & 3 \\ 5 & 17 & 9 \\ 5 & 11 & 7 \end{bmatrix}}$$

$$\frac{1}{5}\begin{bmatrix} \frac{2}{5} & 1 & 3 \\ 1 & 17/5 & 9 \\ 1 & 11/5 & 7 \end{bmatrix}$$

Now, $C = 5(A+B)$

Since this linear tranformation exists, the determinant is zero and the matrix is singular.

c)

$$E_{2,r} = \frac{1}{2}\left(\sum_t \frac{1}{2}\left(g(x^t) - r^t\right)^2 + \lambda \sum_{i=1}^{2} w_i^2\right)$$

$$= \frac{1}{2}\left(\frac{1}{2}\left[(3w_2+w_1+w_0-3)^2 + (2w_2+4w_1+w_0-7)^2 + (4w_2+5w_1+w_0-2)^2 + (7w_2+8w_1+w_0-1)^2\right] + \lambda(w_2^2+w_1^2+w_0^2)\right]$$

d) Taking partial derivatives after letting $\lambda = 2$

$$\frac{\partial E_{2,r}}{\partial w_2} = \frac{1}{2}\left(\frac{1}{2}\left[2(3w_2+w_1+w_0-3)(3) + 2(2w_2+4w_1+w_0-7)(2) + 2(4w_2+5w_1+w_0-2)(4) + 2(7w_2+8w_1+w_0-1)(7)\right] + 2\left(2w_2(1)+0+0\right)\right)$$

$$= \frac{1}{2}\left(\frac{1}{2}\left[12w_2+6w_1+6w_0-18 + 8w_2+16w_1+4w_0-34 + 98w_2+112w_1+14w_0-14 + 32w_2+40w_1+8w_0-16\right] + 4w_2\right)$$

Now, Summing up all these and dividing by 2

$$= \frac{0.060516 + 0.544644 + 2.4273 + 0.329}{2}$$

$$= 1.68073$$

$$= \frac{1}{2}\left(\frac{1}{2}\left[150 w_2 + 174 w_1 + 32 w_0 - 82\right] + 4 w_2\right)$$

$$= \frac{1}{4}\left(150 w_2 + 174 w_1 + 32 w_0 - 82\right] + 2 w_2$$

$$= 37.5 w_2 + 43.5 w_1 + 8 w_0 - 20.5 + 2 w_2$$

$$= 39.5 w_2 + 43.5 w_1 + 8 w_0 - 20.5 = 0 \quad \text{—— (1)}$$

Now.

$$\frac{\partial E_{2,r}}{\partial w_1} = \frac{1}{2}\Bigg(\frac{1}{2}\Big(2(3w_2 + w_1 + w_0 - 3)(1) + 2(2w_2 + 4w_1 + w_0 - 7)(4)$$
$$+ 2(4w_2 + 5w_1 + w_0 - 2)(5)$$
$$+ 2(7w_2 + 8w_1 + w_0 - 1)(8)$$
$$+ 2(2w_1)(1) + 0 + 0\Big)\Bigg)$$

Rearranging we get

$$\frac{87}{2} w_2 + 55 w_1 + 9 w_0 - \frac{49}{2} \quad \text{—— (2)}$$

Q4

d) Contd

Now, Similarly

$$\frac{\partial E_{2,8}}{\partial w_0} = 8w_2 + 9w_1 + 2w_0 - 13/2 \quad (3)$$

(e) Solving the equations

$$2w_0 + 9w_1 + 8w_2 - \frac{13}{2} = 0$$

$$9w_0 + 55w_1 + \frac{87}{2}w_2 - \frac{49}{2} = 0$$

$$8w_0 + \frac{87}{2}w_1 + 41w_2 - 19 = 0$$

By using np.linalg.solve, we get

$w_0 = 6.2079124 6$
$w_1 = 0.13131313$
$w_2 = -0.88720539$

$\therefore$ The equation of the linear function

is

$$\boxed{g(x) = -0.8872x_2 + 0.1313x_1 + 6.2079}$$

Ans