Q3a.

~~We need~~ Let PDF be $P(X)$ & $P(X_1, X_2 - - X_n)$ be the joint PDF.

mean of the PDF.

Bias of $\bar{x}$

$$= E_{P(X_1, X_2 \cdots X_n)}\left[\frac{\sum x_i}{N}\right] - \mu$$

$$\boxed{\mu = \int_x x P(x)\, dx}$$

as x_n are drawn independently,

$$P(x_1, x_2, \cdots x_n) = P(x_1)P(x_2) \cdots P(x_n)$$

$$\Rightarrow E_{P(X_1 \cdots X_n)}\left[\frac{\sum x_i}{N}\right] = \frac{1}{N} \underbrace{\iiiint}_{n \text{ times}} P(x_1)P(x_2)\cdots P(x_n)\left(\sum x_i\right) dx_1 dx_2 \cdots dx_n$$

$$= \iint P(x_2)P(x_3)\cdots P(x_n)\left(\int_{x_1} P(x_1)\left(\sum x_i\right) dx_1\right) dx_2 \cdots dx_n$$

$$= \int_x P(x_1)(x_1 + x_2 - - x_n)\, dx_1 = \int_x x_1 P(x_1)\, dx_1 + (x_2 + \cdots x_n)\int_x P(x_1)\, dx_1$$

$$= \mu + (x_2 + \cdots x_n) \qquad \left(\because \int_x P(x)\, dx = 1\right)$$

So solving integral corresponding to x_1 turned x_1 to μ mathematically & didn't change other values. So similarly, we get

$$E_{P(X_1 \cdots X_n)}\left[\frac{\sum x_i}{N}\right] = \frac{1}{N}(\underbrace{\mu + \cdots \mu}_{N \text{ times}}) = \mu$$

Hence Bias $= \mu - \mu = 0$

Q3b.

let $P(X)$ be the PDF. & σ^2 be the variance of $P(X)$

$$\sigma^2 = E(X^2) - (E(X))^2 = E(X^2) - \mu^2 \longrightarrow (1)$$

$$E\left[\sum_{n=1}^{N} \frac{1}{N}(x_n - \bar{x})^2\right] = \frac{1}{N}\sum_{i=1}^{N} E[(x_i - \bar{x})^2] \longrightarrow (2)$$

$$E[(x_i - \bar{x})^2] = E[x_i^2 + \bar{x}^2 - 2x_i\bar{x}]$$

$$= E[x_i^2] + E[\bar{x}^2] - 2E[x_i\bar{x}]$$

Represent $\bar{x}$ as $\frac{x_i}{N} + \frac{\sum_{j \neq i} x_j}{N} = \frac{x_i}{N} + \frac{y}{N}$

Note that x_i & y are independent

$$E[x_i^2] = \sigma^2 + \mu^2 \quad (\text{from eq } (1))$$

$$= E[x_i^2] - \frac{2}{N}E[x_i(x_i + y)] + E[\bar{x}^2]$$

$$\cancel{= E[x_i^2] - \frac{2}{N}E[x_i^2]}$$

$$E(\bar{x}^2) = \frac{1}{N^2}E\left[\sum_k \left(x_k^2 + x_i\left(\sum_{j \neq k} x_j\right)\right)\right]$$

$$= \frac{1}{N}E[x_i(x_i + y)]$$ ($\because$ summing up all k is equivalent to $N \times$ (only case i))

$$= E[x_i^2] - \frac{2}{N}E[x_i(x_i + y)] + \frac{1}{N}E[x_i(x_i + y)]$$

$$= E(x_i^2) - \frac{1}{N}E[x_i(x_i + y)]$$

$$= E(x_i^2) - \frac{1}{N}\left[E(x_i^2) + E(x_i y)\right]$$

x_i & y are independent

$$= E(x_i^2) - \frac{1}{N}E(x_i^2) - \frac{1}{N}E(x_i)E(y)$$

$$E(y) = E(N\bar{x} - x_i) = E[N\bar{x}] - E(x_i)$$

$$= N\mu - \mu$$

$$= \left(\frac{N-1}{N}\right)(\sigma^2 + \mu^2) - \frac{\mu^2(N-1)}{N}$$

$$E[(x_i - \bar{x})^2] = \left(\frac{N-1}{N}\right)\sigma^2$$

from eq (2)

$$E\left[\sum_{n=1}^{N}\frac{1}{N}(x_n - \bar{x})^2\right] = \frac{1}{N} \times N \times \left(\frac{N-1}{N}\right)\sigma^2$$

$$= \left(\frac{N-1}{N}\right)\sigma^2$$

$$\text{Bias} = \left(\frac{N-1}{N}\right)\sigma^2 - \sigma^2$$

$$= \frac{-\sigma^2}{N}$$

Q3C.

$$E\left[\sum_{n=1}^{N} \frac{1}{(N-1)} (x_n - \bar{x})^2\right] = \frac{1}{N-1} \sum_{i=1}^{N} E[(x_i - \bar{x})^2]$$

from prevQuestion,

$$E[(x_i - \bar{x})^2] = \left(\frac{N-1}{N}\right)\sigma^2$$

$$= \left(\frac{1}{N-1}\right) N \times \left(\frac{N-1}{N}\right)\sigma^2$$

$$= \sigma^2$$

$$\therefore \text{Bias} = \sigma^2 - \sigma^2 = 0$$

Q4a.

$$p(x;\mu,b) = \frac{1}{2b} e^{-\left(\frac{|x-\mu|}{b}\right)}$$

let $L(\mu,b)$ be likelihood fuction.

$$L(\mu,b) = \prod_i p(x_i;\mu,b) \equiv \frac{1}{2b} e^{-\left(\frac{\sum_i |x_i-\mu|}{b}\right)}$$ ($\because$ x_n are drawn independently)

$$\log(L) = -\left(\frac{\sum_i |x_i-\mu|}{b}\right) - \log(2b)$$

We can ignore $-\log(2)$ as we are only interested in terms containing μ & b.

$$\log(L) = -\left(\frac{\sum_i |x_i-\mu|}{b}\right) - \log b$$

μ_0 ~~μ_{MLE}~~ $\equiv$ Maximum ~~to~~ likelihood estimate of μ

b_0 ~~b_{MLE}~~ $\equiv$ ~~Maxim~~ " " " of b

$\mu_0 =$ ~~arg~~ μ at which $\log(L)$ attains it's maximum value

$b_0 =$ b at which $\log(L)$ attains it's maximum value.

Q4b. $\boxed{\mu \text{ Estimation}}$: (treat b as a constant i.e independent of μ)

$$\log(L) = -\left(\frac{\sum_i |x_i - \mu|}{b}\right) - \log b$$

$\sum_i |x_i - \mu|$ should be min for $\log(L)$ to be maximum.

We know that $\sum_i |x_i - k|$ is minimum when

k is the median of the data $\{x_n\}$.

$$\boxed{\begin{array}{l} \therefore \mu_0 = \text{Sample median} \\ (\text{MLE}) \end{array}}$$

$\boxed{\text{b estimation}}$:

for this put $\mu = \mu_0$ (MLE)

$$\log(L) = -\left(\frac{\sum_i |x_i - \mu_0|}{b}\right) - \log b$$

let $\sum_i |x_i - \mu_0| = K$

$$\log(L) = -\frac{K}{b} - \log b$$

$$\frac{d}{db}(\log L) = \frac{K}{b^2} - \frac{1}{b} = 0$$

$$\boxed{b = K}$$

$$\left.\frac{d^2}{d^2 b}(\log L)\right|_{\text{at } b=K} = \left.-\frac{2K}{b^3} + \frac{1}{b^2}\right|_{b=K} = -\frac{1}{b^2} < 0$$

Q5a. $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ → Modelling 3 coordinates (independent)

$$p(x; \mu, C) = \frac{1}{(2\pi)^{3/2}\sqrt{|C|}} e^{-\frac{1}{2}(x-\mu)^T \frac{I}{\alpha}(x-\mu)}$$

* $|C| = \left|\frac{I}{\alpha}\right| = \frac{1}{\alpha^3}$

$$\boxed{p(x; \mu, C) = \left(\frac{\alpha}{2\pi}\right)^{3/2} e^{\frac{(x-\mu)^T(x-\mu)}{2\alpha}}}$$

$$Q(x; \mu, \alpha) = 0 \quad \text{if } \|x\|_2 \neq 1$$

else:

$$\cancel{Q(x; \mu, \alpha)} \cancel{\propto}$$

$$P(X \mid \|x\|_2 = 1) \propto e^{\frac{\sum_i (x_i - \mu_i)^2}{2\alpha}}$$

$$\propto e^{\frac{1}{2\alpha}\left(\sum x_i^2 + \sum \mu_i^2 - 2\sum x_i \mu_i\right)}$$

$$\propto e^{\frac{1}{2\alpha}\left(1 + 1 - 2x^T\mu\right)} \quad (\because \mu \text{ is of unit norm})$$

$$\propto e^{\frac{1 - x^T\mu}{\alpha}}$$

$$\therefore Q(x; \mu, \alpha) \propto e^{\frac{1 - x^T\mu}{\alpha}}$$

$$Q(x; \mu, \alpha) = \begin{cases} 0 & \text{if } \|x\|_2 \neq 1 \\ K e^{\frac{1 - x^T\mu}{\alpha}} & \text{if } \|x\|_2 = 1 \end{cases}$$

where K is normalization constant.

$\therefore$ at $b=k$ is ~~a~~ ~~maximum~~ $\log(L)$ is maximum.

$$\boxed{\therefore\ b_0 = \left(\sum_i |x_i - \mu_0|\right)}$$

Q5b We need to find x's such that $\cancel{\frac{d}{dx} Q(x;\mu,\alpha) = 0}$

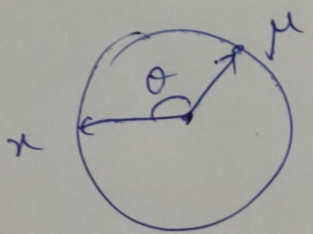
$Q(x;\mu,\alpha)$ has local maxima

Notice that μ is also on the sphere.

~~tate~~. We need to find local minimas of $x^T\mu$

this just just dot product of μ with another vector on the sphere.

Consider a plane along origin & μ. ~~a~~ intersection with the sphere will be circle



$x^T\mu = |\mu||x|\cos\theta = \cos\theta \quad \theta \in [0,\pi]$

$\cos\theta$ has only one local minima

i.e at $\theta = \pi$.

$x^T\mu = \cos\pi = -1$

So $\cancel{x^T = -1}$ $\boxed{x = -\mu}$

$\therefore$ only one mode is possible

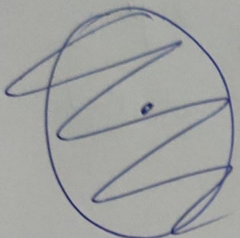
& that is $-\mu$

Q5C.

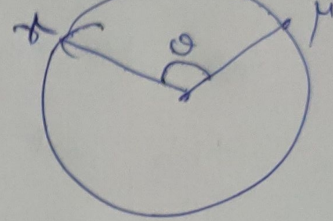
$$Q(x; \mu, \alpha) = k e^{\frac{1 - x^T \mu}{\alpha}}$$

as I wrote in the previous question,

$$x^T \mu = \cos\theta$$



$$(1 - \cos\theta)_{min} = 0$$



& it occurs when $\theta = 0$

Hence when $x = \mu$, $Q(x; \mu, \alpha)$ takes its min value.

Q5d. We need spread ~~around the~~
around the mode.

~~So~~ ~~Assuming~~

- only α ~~decides the~~ models the spread of ~~PMF~~ PDF
around mode.
Because it is the only variable in family of
such PDFS.

~~Q 6d.~~

Q 5~~e~~.

For a 1D guassian,

when μ changes, the PDF (~~og~~ graph) is just translated (w.r.t when $\mu = 0$). ~~The~~ Everything else remains same. So does the Normalization Constant & ~~Similarly~~ ~~This do~~ The change will be done in the exponent term only.

This doen't depend on dimension.

Hence true for 3D also.

~~Q5d.~~

5f

~~Like~~ Let L be the likelihood function.

$$L = \prod_i p(x_i; \mu, C)$$

(assuming ~~data~~ x_i, x_j are independ when $i \neq j$)

$$= \left(\frac{\alpha}{2\pi}\right)^{\frac{3N}{2}} e^{\frac{\sum_i (x_i-\mu)^T (x_i-\mu)}{2\alpha}}$$

Let μ_0 = Maximum likelihood fuction for μ.

$\mu_0 = \mu$ at which L is maximized

$= \mu$ at which $\sum_i (x_i-\mu)^T (x_i-\mu)$ is maximized.

5h.

Let ~~$P(\mu_i;\theta) =$~~ $\cancel{e^{-\mu A^T}}$ ~~where A=~~

Let $P(\mu_i;\theta) = k e^{-A^T \mu}$ where $A = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

where k is normalizing constant.

Using Bayesian, ~~(assume, we are given Sample of Size N)~~

$$\text{posterior} \propto e^{-(\sum_i x_i^T)\mu} \cdot e^{-A^T \mu}$$

$$\propto e^{-(\sum x_i^T + A^T)\mu}$$

$$\propto e^{-(\sum x_i + A)^T \mu}$$

$\therefore$ ~~posterior & prior belongs to the Same family~~

$\therefore$ ~~proposed prior is Conjugate prior.~~

As proved in previous Questions,

$-(\sum x_i + A)^T \mu$ is like dot product between two vectors. For it to be minimum (for the mode) one should lie opposite to the other on sphere.

Hence mode of posterior PDF $= (\sum x_i + A)$
(μ_{mode})

~~Q5b.~~

Q 6a.

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Let Jeffreys prior be ~~Q(k)~~ $Q(\lambda)$

$Q(\lambda)$ ~~Q(k)~~ should be ~~$\propto \sqrt{I(k)}$~~ $\propto \sqrt{I(\lambda)}$

~~I(k) =~~

$$I(\lambda) = E_{P_L(k|\lambda)}\left[\left(\frac{\partial}{\partial \lambda} \log P_L(k|\lambda)\right)^2\right]$$

where $P_L(k|\lambda)$ = likelihood function of $P(k)$

~~this k is a dataset {k}~~

$$= E_{P(k|\lambda)}\left[\left(\frac{\partial}{\partial \lambda} \log P(k|\lambda)\right)^2\right]$$ (taking dataset of size 1)

$$= \sum_k \frac{\lambda^k e^{-\lambda}}{k!}\left(\frac{k}{\lambda} - 1\right)^2$$

~~w.k.t $\sum_k \frac{\lambda^k e^{-\lambda}}{k!} = 1$~~

~~$= \frac{1}{\lambda} \sum_k \frac{\lambda^k e^{-\lambda}}{k!} \cdot k - 1$~~

~~$= \left(\sum_k \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!}\right) - 1$~~

w.k.T $\sum_{k} \frac{\lambda^k e^{-\lambda}}{k!} = 1$

$$= \sum_{k=0} \frac{\lambda^k e^{-\lambda}}{k!} \left(\frac{k^2}{\lambda^2} + 1 - \frac{2k}{\lambda} \right)$$

$$\cancel{= \sum_{k=0}^{\infty} \left(\frac{\lambda^{k-2}}{(k-1)!} k e^{-\lambda} + \underbrace{\frac{\lambda^k e^{-\lambda}}{k!}}_{1} - 2 \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} \right)}$$

$$= \sum_{k=0}^{\infty} \left(\underset{①}{\frac{k^2 \lambda^k e^{-\lambda}}{\lambda^2 k!}} + \underset{②}{\frac{\lambda^k e^{-\lambda}}{k!}} - \underset{③}{\frac{2k\lambda^k e^{-\lambda}}{k! \lambda}} \right)$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \cdots$$

$$xe^x = \sum_{k=0}^{\infty} \frac{x^{k+1}}{k!}$$

$$\frac{d}{dx}(xe^x) = \sum_{k=0}^{\infty} \frac{(k+1)x^k}{k!} = e^x + xe^x$$

② $= 1$

$$① = \frac{1}{\lambda} \left(\sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!} e^{-\lambda} \right) = \frac{e^{-\lambda}}{\lambda} \sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!}$$

$$= \frac{e^{-\lambda}}{\lambda} (e^{\lambda})(1+\lambda) = \frac{1+\lambda}{\lambda}$$

$$③ = -2 e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = -2$$

$$I(\lambda) = \frac{1+\lambda}{\lambda} + 1 - 2 = \frac{1}{\lambda}$$

$$\therefore \; Q(\lambda) \propto \frac{1}{\sqrt{\lambda}}$$

Q6b let $S = \{k_i\}_{i=1,N}$

$$P(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

L = likelihood fuction

$$L = \frac{\lambda^{\Sigma k_i} e^{-N\lambda}}{\Pi(k_i!)}$$

Take terms that are dependent on λ.

$$L \approx \lambda^{\Sigma k_i} e^{-N\lambda} \quad \text{let } \Sigma k_i = S$$

$$\cancel{\frac{dL}{d\lambda} = S\lambda^{S-1} e^{-N\lambda} - N\lambda^{S} e^{-N\lambda-1} = 0}$$

$$\cancel{\Rightarrow e^{\lambda} S = N\lambda^2}$$

~~let for a large N, λ is small, ($e^{\lambda} \approx 1$)~~

$$\cancel{\Rightarrow \lambda = \sqrt{\frac{S}{N}}}$$

$$\cancel{\therefore \hat{\lambda}(S) = \sqrt{\frac{\Sigma k_i}{N}}} \quad \text{for large N.}$$

~~Q6c~~

$$\frac{dL}{d\lambda} = S\lambda^{S-1}e^{-N\lambda} - N\lambda^{S}e^{-N\lambda} = 0$$

$$\Rightarrow S = N\lambda$$

$$\Rightarrow \boxed{\lambda = \frac{S}{N}}$$

$$\therefore T(S) = \left(\frac{\Sigma K_i}{N}\right) = \text{sample mean of } S$$

Q6c

Let the distribution be $Q(x)$

$$\cancel{\text{Sm}}\, Q(x) = \sum_{S_{mean}=x} P(S) \xrightarrow{} \text{joint probability.}$$

$$= \sum_{S_{mean}=x} P(k_1, k_2 \cdots k_N)$$

$$= \sum_{S_{mean}=x} \frac{\lambda^{\Sigma k_i} e^{-N\lambda}}{\Pi(k_i!)} \quad \cdots (\Sigma k_i = Nx) \quad (\text{put } \lambda = x)$$

$$\cancel{= \lambda \; \frac{Nx}{\lambda} e^{-N}}$$

$$= \sum_{S_{mea}=x} \frac{x^{Nx} e^{-Nx}}{\Pi(k_i!)}$$

$$= x^{Nx} e^{-Nx} \left(\sum_{S_{mea}=x} \frac{1}{\Pi(k_i)!} \right)$$

No. This distribution is not poisson.