(Q1.) λ_1 : Alice

λ_2 : Bob

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

$$f_a = \frac{1}{11} \quad , \quad f_b = \frac{10}{11}$$

Now $P(A/D) = \frac{1}{11}\left(\frac{\lambda_1^{12} e^{-\lambda_1}}{12!} \cdot \frac{\lambda_1^{10} e^{-\lambda_1}}{10!} \cdot \frac{\lambda_1^{11} e^{-\lambda_1}}{11!} \cdot \frac{\lambda_1^{4} e^{-\lambda_1}}{4!} \cdot \frac{e^{-\lambda_1} \cdot \lambda_1^{11}}{11!}\right)$

$$P(B/D) = \frac{10}{11}\left(\frac{\lambda_2^{12+10+11+4+11} \; e^{-5\lambda_2}}{12! \cdot 10! \cdot 11! \cdot 4! \cdot 11!}\right)$$

$$\therefore \frac{P(\text{data with Alice})}{P(\text{data with bob})} = \frac{1}{10}\left(\left(\frac{10}{15}\right)^{48} e^{25}\right)$$

$$= 25.39$$

(Q2.) $\hat{\theta} = \theta + \varepsilon$ $\qquad f(\varepsilon) \sim N(0, \sigma^2 = 4)$

$\therefore f(\hat{\theta}) \sim N(\theta, 4)$

$f(\theta) \sim N(5, 9)$

$$f(\theta) = N(5, 9) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(\theta-5)^2}{2\sigma_1^2}}$$

$\hat{\theta}$ = likelihood estimate

$\therefore f(x/\theta) \sim N(\theta, 4)$

$$f(x, \theta) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2\sigma_2^2}}$$

Date: / /

Now we need
posterior pdf for $\theta = f\left(\frac{\theta}{x}\right)$

$$\therefore f\left(\frac{\theta}{x}\right) = f\left(\frac{x}{\theta}\right) \cdot f(\theta)$$

$$= \frac{1}{2\pi.(6)} e^{-\frac{(\theta-5)^2}{18} - \frac{(x-\theta)^2}{8}}$$

given $x = 6$

$$f\left(\frac{\theta}{x}\right) = \frac{1}{12\pi} e^{-\frac{1}{2}\left(\frac{(\theta-5)^2}{9} + \frac{(\theta-6)^2}{4}\right)}$$

$$f\left(\frac{\theta}{x}\right) = \frac{1}{12\pi} e^{-\frac{1}{2}\left(\frac{13\theta^2+100+324-40\theta-108\theta}{36}\right)}$$

$$f\left(\frac{\theta}{x}\right) = \frac{1}{12\pi} e^{-\frac{1}{2}\left(\frac{13\theta^2-148\theta+424}{36}\right)}$$

$$f\left(\frac{\theta}{x}\right) = \frac{1}{12\pi} e^{-\frac{1}{2}\left(\frac{\theta^2-\frac{148}{13}\theta+\frac{424}{13}}{36/13}\right)}$$

$$f\left(\frac{\theta}{x}\right) = \frac{1}{12\pi} e^{-\frac{1}{2}\left(\frac{\theta^2-2\left(\frac{74}{13}\right)\theta+\left(\frac{74}{13}\right)^2+c}{36/13}\right)}$$

$$f\left(\frac{\theta}{x}\right) = \frac{1}{12\pi} e^{-\frac{1}{2}\left(\frac{\left(\theta-\frac{74}{13}\right)^2+c}{36/13}\right)}$$

$$f\left(\frac{\theta}{x}\right) \sim N\left(\frac{74}{13}, \frac{36}{13}\right)$$

Date: / /

(b) Noise = $n_1, n_2, n_3 \ldots n_n$
and $\bar{x}$ = sample mean.

Now $n=4$, $\bar{x}=6$
$f(\theta)$= prior $\sim N(5, 9)$
$\mu_{prior} = 5$, $\sigma_{prior} = 3$

$\sigma^2 = 4$ as $\hat{\theta} = \theta + \varepsilon$ ∴ $f(\hat{\theta}) \sim N(\theta, 4)$
as $f(\varepsilon) \sim N(0, 4)$

Now $a = \frac{1}{9}$, $b = \frac{4}{4} = 1$

$$\mu_{post} = \frac{(\frac{1}{9})(5) + (1)(6)}{\frac{1}{9}+1} = \frac{59}{10} = 5.9$$

$$\sigma^2_{post} = \frac{9}{10} = 0.9$$

∴ $f\left(\frac{\theta}{x}\right) \sim N(5.9, 0.9) \longrightarrow$ posterior,

and $f(\theta) \sim N(5, 9) \rightarrow$ prior

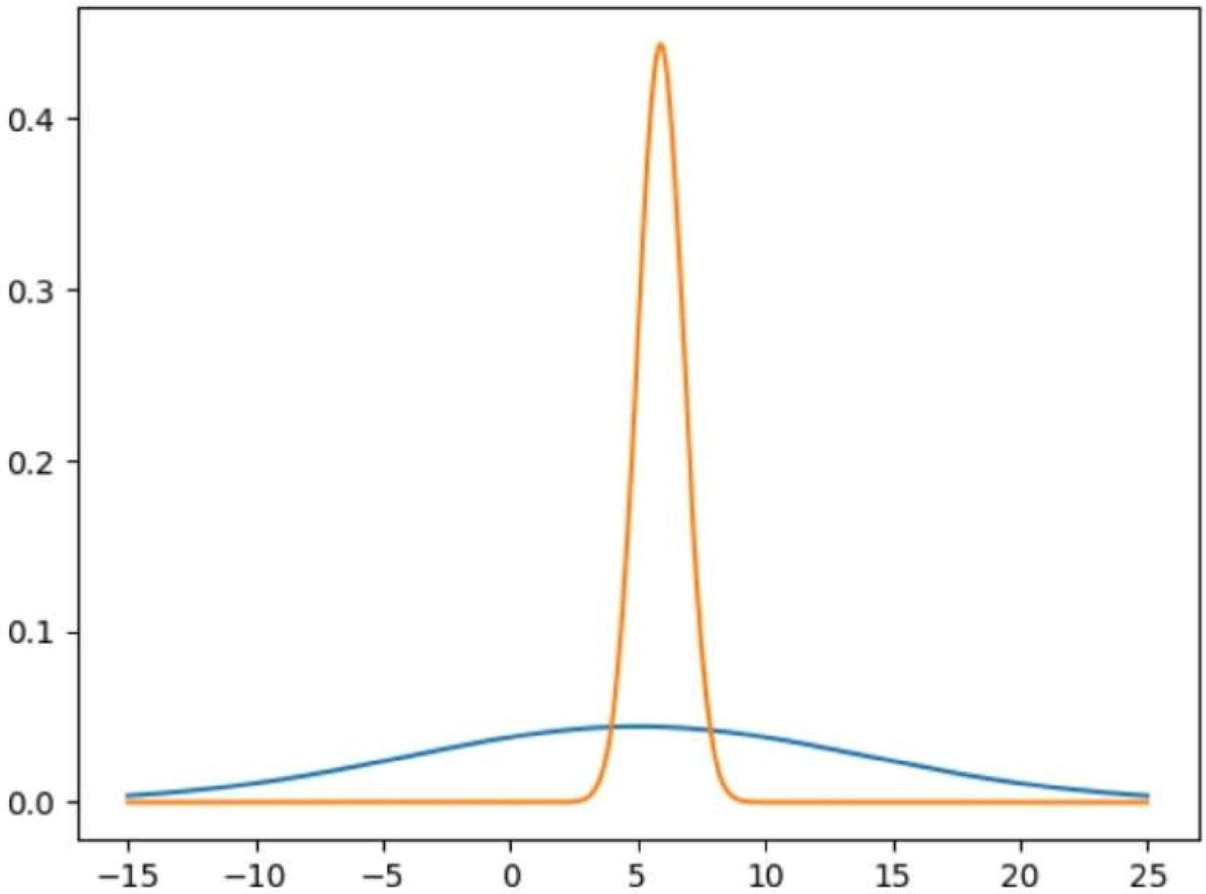
∴ after receiving the data we can see there is not much change in the mean ~~the~~ but the variance changes much ∴ we get more accurate value of θ

In [5]:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

x_axis = np.arange(-15, 25, 0.01)

plt.plot(x_axis, norm.pdf(x_axis, 5, 9))
plt.plot(x_axis, norm.pdf(x_axis, 5.9, 0.9))
plt.show()
```



Date: / /

(C) As more data is given the mean becomes more accurate and variance decreases as in (b) so its closer to true value of θ.
as $n\uparrow$ $b\uparrow$ and $\sigma^2_{post}\downarrow$

(d) $IQ \sim N(100, 15^2)$

$f(\theta) \sim N(100, 15^2)$

$\hat{\theta} = \theta + \varepsilon$, $f(\varepsilon) \sim N(0, 10^2)$

$f(\hat{\theta}) = N(\theta, 10^2)$

According to formulas given above
if $n=1$ (tested once)

$a = \frac{1}{15^2}$, $b = \frac{1}{10^2}$

$$\mu_{post} = \frac{\frac{1}{15^2}\cdot 100 + \frac{1}{10^2}\cdot 80}{\frac{1}{15^2} + \frac{1}{10^2}}$$

$\mu_{post} = 88.031$

$\therefore$ as $E(N(\mu_{post}, \sigma^2_{post})) = \mu_{post}$

$\Rightarrow E(\text{true IQ}) = 88.031$

(ii) Using same formulas :

Date: / /

$$\mu_{post} = \frac{\frac{1}{15^2} \cdot 100 + \frac{1}{10^2} \cdot 150}{\frac{1}{15^2} + \frac{1}{10^2}}$$

$$\mu_{post} = 129.92$$

Q3)

$$L(\theta) = \prod_{i=0}^{n} f(x, \theta_i)$$

$$L(\theta) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\sum_{i=0}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}}$$

$$\log L(\theta) = -n(\log \sigma + \log(\sqrt{2\pi})) - \sum_{i=0}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial(\log L(\theta))}{\partial \mu} = -\sum \frac{2(x_i-\mu)}{2\sigma^2} \cdot (-1)$$

$$0 = \frac{\sum_{i=0}^{n} (x_i-\mu)}{2\sigma^2}$$

$$0 = \sum x_i - n\mu$$

$$\hat{\mu} = \frac{\sum x_i}{n}$$

~~(ii)~~ ~~without figure~~
maximum 2 points ~~with binary classification~~

(Q5) (a) Constant function: It can shatter maximum 2 points, for any set of 3 points there is a combination which doesn't get shattered by the constant function

(b) Linear function in d dimension: It can shatter max (d+1) points in d dimensions as in 2d it goes max 3 in 3D it goes max 4 and even in 1 D it goes max 2 $\therefore$ (d+1) in d dimension

(c) Axis aligned rectangle: can shatter max 4 points

(d) Interval: We can have max 2 points shattered as any combination of (−, +, −) would not be separated.