Ans 1

$$P(x_1, x_2, \dots x_n \mid \mu) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(\sum_{i=1}^{n} - \left(0.5 \frac{(x_i-\mu)^2}{\sigma^2}\right)\right)$$

To maximize the likelihood f^{x^n} we have to minimize $\sum_{i=1}^{n}(x_i-\mu)^2$ w.r.t μ.

$$\frac{d}{d\mu}\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}\frac{d((x_i-\mu)^2)}{d\mu} = \sum_{i=1}^{n} -(2(x_i-\mu)) = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i - n\mu = 0$$

$$\Rightarrow \mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

$\frac{d^2}{d\mu^2}\left(\sum_{i=1}^{n}(x_i-\mu)^2\right) > 0$

Thus it is a minima.

Thus $$\boxed{\hat{\mu}_{ML} = \frac{\sum_{i=1}^{n} x_i}{n}}$$

For map1 the posterior will also be a gaussian with mean

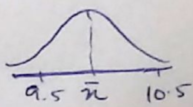
Thus. $$\hat{\mu}_{map1} = \frac{\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)\sigma_{prior}^2 + \mu_{prior}\frac{\sigma_{true}^2}{n}}{\sigma_{prior}^2 + \sigma_{true}^2/n}$$

For $\hat{\mu}_{map2}$
$$P(\theta) = \begin{cases} \frac{1}{2} & \text{if } \theta \in (9.5, 11.5) \\ 0 & \text{otherwise} \end{cases}$$

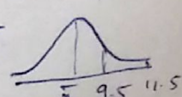
$$\text{Posterior} = P(\theta \mid x_1, x_2 \dots, x_n) = \frac{P(x_1, x_2, \dots x_n \mid \theta)\, P(\theta)}{P(x_1, x_2, \dots x_n)}$$

$$\begin{cases} \propto \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-0.5\sum_{i=1}^{n}\frac{(x_i-\theta)^2}{\sigma^2}\right) * \frac{1}{2} & \text{if } \theta \in (9.5, 11.5) \\ = 0 & \text{otherwise} \end{cases}$$

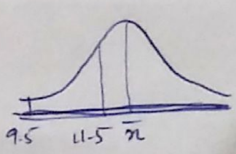
Thus if $\frac{\sum_{i=1}^{n} x_i}{n} \in (9.5, 11.5)$ then $\hat{\mu}_{map2} = \frac{\sum_{i=1}^{n} x_i}{n}$



Otherwise if $\frac{\sum_{i=1}^{n} x_i}{n} < 9.5$ then $\hat{\mu}_{map2} = 9.5$



if $\frac{\sum_{i=1}^{n} x_i}{n} > 11.5$ then $\hat{\mu}_{map2} = 11.5$



(i) In general as N increases the error decreases for maximum likelihood estimate because sample mean approaches true means as n becomes larger & larger.

For ~~the~~ $\hat{\mu}_{map1}$, $\hat{\mu}_{map2}$, for smaller values of n. prior dominates but as n becomes larger the data starts dominating the posterior distribution due to which there might be some fluctuations in errors as n increases but in general as n increases, errors decrease.

(ii) We will prefer the $\hat{\mu}_{map1}$ estimate ~~prior~~ because it has lesser relative error than the other two estimates.