## Question 1:

a) MLE estimator for $\mu$ is: $\hat{\mu} = \bar{X}$

$$L_n(\bar{X}_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x_i - \bar{x}_n)^2}{2\sigma^2}\right)$$

$$L_n(\mu_0) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_i-\mu_0)^2}{2\sigma^2}\right)$$

$$\Rightarrow \frac{L_n(\bar{X}_n)}{L_n(\mu_0)} = \prod_{i=1}^{n} \exp\left(\frac{-(x_i-\bar{x}_n)^2}{2\sigma^2}\right) \cdot \exp\left(\frac{(x_i-\mu_0)^2}{2\sigma^2}\right) = \prod_{i=1}^{n} \exp$$

$$\Rightarrow 2\log\left(\frac{L_n(\bar{X}_n)}{L_n(\mu_0)}\right) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x}_n)^2\right)$$

$$= \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n}\left(2x_i(\bar{x}_n-\mu_0) + \mu_0^2 - \bar{x}_n^2\right) = 2\frac{\bar{x}_n-\mu_0}{\sigma^2} \cdot \sum_{i=1}^{n} x_i + \frac{n}{\sigma^2}(\mu_0^2 - \bar{x}_n^2)$$

$$= \frac{2}{\sigma^2} \cdot n \cdot \bar{x}_n^2 - \frac{2\mu_0}{\sigma^2} \sum_{i=1}^{n} x_i + \frac{n}{\sigma^2}\mu_0^2 - \frac{n}{\sigma^2}\bar{x}_n^2 = \underbrace{\frac{n}{\sigma^2} \cdot \bar{x}_n^2 + \frac{1}{\sigma^2}\left(n - 2\sum_{i=1}^{n} x_i\right)\mu_0}_{\chi_1^2}$$

$$= \frac{n}{\sigma^2}(\bar{x}_n^2 + \mu_0^2 - 2\bar{x}_n\mu_0) = \left(\frac{\bar{x}_n - \mu_0}{\sqrt{\sigma^2/n}}\right)^2 \sim \chi_1^2$$

Therefore we reject the null hypothesis if $\lambda > \chi^2_{1,\alpha}$.

b) MLE estimator for $\sigma^2$: $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$

$$L_n(\hat{\sigma}^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(\frac{-(x_i-\mu)^2}{2\hat{\sigma}^2}\right)$$

$$\frac{\hat{\sigma}^2}{\sigma_0^2} = \frac{1}{n} \cdot \frac{S^2}{\sigma^2} = \frac{1}{n^2} \cdot \chi_1^2$$

$$L_n(\sigma_0^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(\frac{-(x_i-\mu)^2}{2\sigma_0^2}\right)$$

$$2\log\left(\frac{L_n(\hat{\sigma}^2)}{L_n(\sigma_0^2)}\right) = 2\log\left(\left(\frac{\sigma_0^2}{\hat{\sigma}^2}\right)^{n/2}\right) \cdot \left(\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma_0^2} - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\hat{\sigma}^2}\right)$$

$$= n \cdot \log\left(\frac{\sigma_0^2}{\hat{\sigma}^2}\right) \cdot \left(\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{n}{2}\right) = n \cdot \log\left(\frac{\sigma_0^2}{\hat{\sigma}^2}\right) \cdot \left(\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{n}{2}\right) = \lambda$$

Therefore we reject the null hypothesis if $\lambda > \chi^2_{1,\alpha}$.

# Question 2:

a) $\mathbb{E}[AX] \overset{\text{by}}{=} \begin{pmatrix} \mathbb{E}(A^1 X) \\ \vdots \\ \mathbb{E}(A^m X) \end{pmatrix} = \begin{pmatrix} A^1 \mu \\ \vdots \\ A^m \mu \end{pmatrix} = A\mu$

matrix row independence

$\mathbb{E}[v^T X] = v^T \mu \quad \forall v \in \mathbb{R}^n$
by tutorial

b) $\Sigma = Var(X) = \mathbb{E}[(X-\mu_X)(X-\mu_X)^T]$

It is known that $\forall v \in \mathbb{R}^n$, $vv^T \succeq 0$ therefore $\mathbb{E}(vv^T) \succeq 0$
by non-negatively weighted sum of positive semidefinite matrices.

$\Rightarrow \mathbb{E}[(X-\mu_X)(X-\mu_X)^T] \succeq 0 \Rightarrow \Sigma \succeq 0.$

c) $Cov(AX, BY) = \mathbb{E}[A(X-\mu_X) \cdot B(Y-\mu_Y)] \overset{\text{by (a)}}{=} A \cdot Cov(X, BY)$

$\overset{\text{by (b)}}{=} A \cdot Cov(BY, X)^T = A \cdot \mathbb{E}[B(Y-\mu_Y)(X-\mu_X)]^T = A \cdot (B \cdot \mathbb{E}[(Y-\mu_Y)(X-\mu_X)])^T$

$= A \cdot Cov(Y, X)^T \cdot B^T = A \cdot Cov(X, Y) \cdot B^T$

# Question 3:

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y = (X^T X)^{-1} X^T (X\beta^* + \varepsilon) = \beta^* + (X^T X)^{-1} X^T \varepsilon$$

$$\hat{\beta}_{MLE} = \arg\max_{\beta} \prod_{i=1}^{n} P(Y_i = y_i) = \arg\max_{\beta} \prod_{i=1}^{n} P\left(\beta_0 + \sum_{j=1}^{k} \beta_j X_{ij} + \varepsilon_i = y_i\right)$$

$$= \arg\max_{\beta} \prod_{i=1}^{n} P\left(\varepsilon_i = y_i - \beta_0 - \sum_{j=1}^{k} \beta_j X_{ij}\right) = \arg\max_{\beta} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \cdot \frac{(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j X_{ij})^2}{\sigma^2}\right)$$

$$= \arg\max_{\beta} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \frac{(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j X_{ij})^2}{\sigma^2}\right) = \arg\max_{\beta} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{k} \beta_j X_{ij})^2\right)$$

$$\arg\min_{\beta} \left(\sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{k} \beta_j X_{ij})^2\right) = \arg\min_{\beta} S(\beta)$$

log monotone increasing + $e^x$ monotone increasing

$$\Longrightarrow \hat{\beta}_{MLE} = \arg\min_{\beta} S(\beta)$$

This is exactly the definition of $\hat{\beta}_{OLS}$ and therefore $\underline{\hat{\beta}_{MLE} = \hat{\beta}_{OLS}}$.