University of Technology Sydney
School of Mathematical and Physical Sciences

Mathematical Statistics (37262) –
Class 8 Preparation Work
SOLUTIONS

1.

i) $\bar{y} = \frac{1}{10}\sum_{i=1}^{10} y_i = \frac{36.9}{10} = 3.69$

ii) The sample mean is an unbiased estimator of the population mean hence $\hat{\mu} = 3.69$.

iii) The 95% confidence interval for μ is given by $\bar{y} \pm 1.96\sqrt{\frac{\sigma^2}{n}}$ (since, for $Z \sim N(0,1)$, $P(-1.96 < Z < 1.96) \approx 0.95$)

This gives an interval of $3.69 \pm \frac{3.92}{\sqrt{10}}$ i.e. $(2.45,\ 4.93)$.

iv) Given that we know the variance is $\sigma^2 = 2^2$, our new observation $Y_{new} \sim N(\mu, 2^2)$. The distribution of the sample mean is $\bar{Y} \sim N\left(\mu, \frac{2^2}{10}\right)$

hence we have that $Y_{new} - \bar{Y} \sim N\left(0,\ \left(2^2 + \frac{2^2}{10}\right)\right)$

Our 95% prediction interval for Y_{new} is then given by $\bar{y} \pm 1.96\sqrt{\sigma^2 + \frac{\sigma^2}{10}}$.

This is approximately $(-0.42,\ 7.80)$

v) $s^2 = \frac{1}{10}\sum_{i=1}^{10}(y_i - \bar{y})^2 = 6.7729$.

vi)

$$E(s^2) = E\left(\frac{1}{10}\sum_{i=1}^{10}(y_i - \bar{y})^2\right)$$

$$= \frac{E\left(\sum_{i=1}^{10}(y_i-\mu)^2\right)}{10} + \frac{E\left(\sum_{i=1}^{10}(\bar{y}-\mu)^2\right)}{10} - \frac{2E\left(\sum_{i=1}^{10}(y_i-\mu)(\bar{y}-\mu)\right)}{10}$$

$$\frac{E\left(\sum_{i=1}^{10}(y_i-\mu)^2\right)}{10} - E\left((\bar{y}-\mu)^2\right)$$

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

$E(s^2) = \sigma^2 - \dfrac{\sigma^2}{n}$ hence $bias(s^2,\sigma^2) = \sigma^2 - \dfrac{\sigma^2}{n} - \sigma^2 = -\dfrac{\sigma^2}{n} \neq 0.$

vii) The confidence intervals for the population mean μ would be wider when assuming $Y \sim N(\mu,\sigma^2)$ rather than $Y \sim N(\mu,2^2)$ since our uncertainty is greater when the variance is not known.

vii) The 95% confidence interval for μ is given by $\bar{y} \pm 2.62\sqrt{\dfrac{\hat{\sigma}^2}{n}}$ where

$\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{10}(y_i - \bar{y})}{10-1}$, since for $T \sim t_9$, $P(-2.262 < Z < 2.262) \approx 0.95$.

This gives an interval of $3.69 \pm \dfrac{2.262}{\sqrt{10}}\sqrt{\dfrac{67.729}{9}}$ i.e. $(1.99,\ 5.39)$.

2. $$f(x)=\begin{cases}(2\theta)^{-1} & x\in[0,2\theta]\\ 0 & \text{otherwise}\end{cases}.$$

i) Calculate $E(X)=\int_0^{2\theta} x(2\theta)^{-1}dx=\left[\frac{x^2}{4\theta}\right]_0^{\theta}=\theta$.

ii) $L(\{x_1,x_2,...x_n\}\mid\theta)=(2\theta)^{=n}$ .hence $\ell(\{x_1,x_2,...x_n\}\mid\theta)=-n\ln(2\theta)$.

$\frac{\partial}{\partial\theta}\ell(\{x_1,x_2,...x_n\}\mid\theta)=-\frac{n}{\theta}$. This is never zero. The likelihood is maximised when θ is as small as possible.

$x\in[0,2\theta]$ so we set $2\hat{\theta}_{MLE}=\max\{x_1,x_2,...,x_n\}$

i.e. $\hat{\theta}_{MLE}=\frac{\max\{x_1,x_2,...,x_n\}}{2}$.

iii) $E(X)=\theta$ so, by matching first moments, we obtain $\hat{\theta}_{MM}=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$.

iv) Consider $\{x_1,x_2,...x_5\}=\{1,\ 2,\ 3,\ 4,\ 90\}$.

The sample mean is $\bar{x}=\frac{1}{5}(1+2+3+4+90)=20$.

This would give $\hat{\theta}_{MM}=20$ implying that $x\in[0,40]$ which we know is not true, as we have an observation $x_5=90$.