MATH2033 Mathematical Statistics
Assignment 4 Suggested Solutions

1. (a) Since $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$, we have:

$$\begin{aligned}
(n-1)s^2 &= \sum_{i=1}^{n}\left(X_i^2 + \bar{X}^2 - 2X_i\bar{X}\right)\\
&= \sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n}\bar{X}^2 - 2\bar{X}\sum_{i=1}^{n} X_i\\
&= \sum_{i=1}^{n} X_i^2 + n\bar{X}^2 - 2\bar{X}n\bar{X}\\
&= \sum_{i=1}^{n} X_i^2 - n\bar{X}^2
\end{aligned}$$

Then,

$$\begin{aligned}
(n-1)\mathrm{E}\left(s^2\right) &= \sum_{i=1}^{n}\mathrm{E}\left(X_i^2\right) - n\mathrm{E}\left(X^2\right)\\
&= \sum_{i=1}^{n}\Big(\mathrm{Var}\left(X_i\right) + (\mathrm{E}X_i)^2\Big) - n\mathrm{E}\left(\bar{X}^2\right) && (\mathrm{Var}\, X_i = \mathrm{E}X_i^2 - (\mathrm{E}X_i)^2)\\
&= n\,\mathrm{Var}\, X_1 + \sum_{i=1}^{n}(\mathrm{E}X_i)^2 - n\mathrm{E}\left(\bar{X}^2\right) && (\mathrm{Var}(X_i) = \mathrm{Var}(X_1))\\
&= n\sigma^2 + \sum_{i=1}^{n}\Big((\mathrm{E}X_i)^2 - \mathrm{E}\left(X^2\right)\Big)\\
&= n\sigma^2 - \sum_{i=1}^{n}\Big(\mathrm{E}\left(X^2\right) - (\mathrm{E}X_i)^2\Big)\\
&= n\sigma^2 - \sum_{i=1}^{n}\left(\mathrm{E}\left(\bar{X}^2\right) - (\mathrm{E}\bar{X})^2\right) && (\mathrm{E}X_i = \mathrm{E}\bar{X} = \mu)\\
&= n\sigma^2 - \sum_{i=1}^{n}\mathrm{Var}(\bar{X})\\
&= n\sigma^2 - n\frac{\sigma^2}{n}\\
&= (n-1)\sigma^2
\end{aligned}$$

Cancelling $(n-1)$ from both sides, it follows $\mathrm{E}\left(s^2\right) = \sigma^2$.

(b) No. By Jensen's inequality,

$$\mathrm{E}(s) = \mathrm{E}(\sqrt{s^2}) \leq \sqrt{\mathrm{E}(s^2)} = \sqrt{\sigma^2} = \sigma.$$

Thus $\mathrm{E}(s) \leq \sigma$. It follows that $\mathrm{E}(s)$ is not always equal to σ, or s is not an unbiased estimate of σ.

(c) Based on the result of part (a),

$$\mathrm{E}(\frac{s^2}{n}) = \frac{1}{n}\mathrm{E}(s^2) = \frac{1}{n}\sigma^2 = \sigma^2_{\bar{X}}.$$

2. (a) False. The histogram of the sample will resemble the histogram of the population. It is the sampling distribution of the sample mean that is approximately normal.
   (b) True.
   (c) False. A 95% confidence interval for μ must contain the sample mean as its center.
   (d) False.
   (e) False. Not necessarily to be exactly 95 out of 100.

3. The first two moments of the gamma distribution are:

$$\mu_1 = \frac{\alpha}{\beta},\ \mu_2 = \frac{\alpha(\alpha+1)}{\beta^2}.$$

To use the method of moments, we have to represent the α and β with μ_1 and μ_2. From

$$\mu_2 = \mu_1^2 + \frac{\mu_1}{\beta},$$

we get

$$\beta = \frac{\mu_1}{\mu_2 - \mu_1^2}.$$

Then

$$\alpha = \beta\mu_1 = \frac{\mu_1^2}{\mu_2 - \mu_1^2}.$$

Since $\hat{\sigma}^2 = \hat{\mu}_2 - \hat{\mu}_1^2$, the formulas for estimating the parameters α and β are

$$\alpha = \frac{\bar{X}^2}{\hat{\sigma}^2},\ \hat{\beta} = \frac{\bar{X}}{\hat{\sigma}^2}.$$

4. $E(Y) = \int_0^1 y \cdot \theta y^{\theta-1} dy = \theta/(\theta+1)$, so $\hat{\theta} = \frac{\bar{Y}}{1-\bar{Y}} = 0.54$.