This problem sheet has several questions which all require the use of R. The questions are all related, and it would be best if all the questions were done together in an R Markdown report. The template for the coursework would be suitable, or simply open R Studio, click on the "New File" icon in the top-left corner, and select either "R Notebook" or "R Markdown". To compile the notebook, "Knit to HTML".

Question 1

Use `dnorm` to plot the probability density function of the standard normal random distribution on the interval $[-4, 4]$.

Hint: Use the `seq` function to generate 1000 evenly spaced points on the interval $[-4, 4]$.

Question 2

Use `dgamma` to plot the probability density function of a $\Gamma(2, 0.5)$ random variable on the interval $[0, 20]$. Note that we are using the shape/rate parametrisation here, i.e. $\alpha = 2$ is the shape and $\beta = 0.5$ is the rate.

Question 3

(a) For $X_1, X_2, \ldots, X_n \sim \Gamma(\alpha, \beta)$, use R to sample observations $x_1, x_2, \ldots, x_n$, where $n = 1000$ and $\alpha = 2$ and $\beta = 0.5$.

(b) From these $x_1, x_2, \ldots, x_n$ values, compute the the standardised $z_1, z_2, \ldots, z_n$, where

$$z_i = \frac{x_i - \mathrm{E}[X_i]}{\sqrt{\mathrm{Var}[X_i]}}.$$

(c) Compute the weighted sum

$$S = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_i.$$

(Note the square root in the fraction $1/\sqrt{n}$; this is **not** the sample mean.)

(d) Repeat steps (a) to (c) t times (using a loop), and save the resulting sums $S_1, S_2, \ldots, S_t$ to a vector $\mathbf{S}$. It is suggested that t is set to $t = 10,000$.

(e) Plot a histogram of the values $S_1, S_2, \ldots, S_t$. In the `hist` function, set the parameters `freq=FALSE` and `breaks=30`.

(f) Does this histogram look familiar? Use the `lines` function in R to plot the probability density function of an appropriate distribution over the histogram.

Hint: For $X \sim \Gamma(\alpha, \beta)$, $\mathrm{E}[X] = \frac{\alpha}{\beta}$ and $\mathrm{Var}[X] = \frac{\alpha}{\beta^2}$.