Practical – Session 2

1. If X and Y are independent and follow normal distribution with $X \sim N(0,2)$ and $Y \sim N(2,3)$, then $X+Y \sim N(2,5)$. Simulate 10,000 random variables from each of the distribution of X and Y and verify this result by plotting a 2×2 panel of graphs, with proper axis labels and titles, showing

   (a) The distribution of X.
   (b) The distribution of Y.
   (c) A scatter plot between X and Y.
   (d) The distribution of $X + Y$.

   (e) Calculate the mean and variance in your simulated samples of X, Y and $X + Y$.
   (f) Plot a scatterplot between X and $X+Y$.

2. Read in the data "mvc" to R.

   (a) Transform the continuous variable height to a categorical variable with 3 height categories: 155-167, 168-172, 173-180cm.
   (b) Fit a linear regression model with MVC as dependent variable and age, height (category) as explanatory variables.
   (c) Set the middle height group as the reference level and refit the regression model. Is the result consistent with (b)?
   (d) Summarize the results of (c) in the following table:

Variable	Estimate	95% CI	p-values
Age			
Height			
R^2=	adj. R^2=		

   (e) Refit the regression model in (d) by limiting to those aged 40y or below. Is the result the same as (d) qualitatively?