Question 4

What is Parametric Bootstrap?

We can sample from the distribution F but instead we obtain a sample using an estimate $\hat{F}$. We can estimate the distribution function $F(x)$ using a parametric model $F(x;\theta)$ which is indexed by some parameters.

For a given sample S and parametric model $F(x;\theta)$ we obtain an estimate $\hat{\theta}$ based on the sample. We then generate B bootstrap samples $S_1^*, ..., S_B^*$ using $F(x;\hat{\theta})$. Here we generate samples from the model and not through sampling with replacement.

Parametric Bootstrap for Regression

We will now apply the parametric bootstrap in the context of regression.

- The assumed regression model is
$$Y_i = \alpha + \beta(x_i - \bar{x}) + R_i$$
where $R_i \approx_{iid} G(0,\sigma)$ (iid stands for independent and identically distributed random variables)
- We fit the model to obtain the estimates $\hat{\alpha}, \hat{\beta}$ and $\hat{\sigma}$
- After obtaining the above values $\hat{\alpha}, \hat{\beta}$ and $\hat{\sigma}$ we fit them in the above regression model

Now in order to obtain bootstrap samples, we will generate R_i^* from $G(0,\hat{\sigma})$ and set $y_i^* = \hat{\alpha} + \hat{\beta}(x_i - \bar{x}) + R_i^*$ where x_i are fixed.

We then obtain the bootstrap sample as:

$$S_b^* = \{(x_1, y_1^*), (x_2, y_2^*), ..., (x_n, y_n^*)\}$$

and as we saw x_i is fixed hence in this case $x_1, x_2, ..., x_n$ is fixed and $y_1^*, y_2^*, ..., y_n^*$ are different.

For each bootstrap sample S_b^* we estimate the parameters to get the bootstrap replicates $\hat{\alpha_b^*}, \hat{\beta_b^*}$ and $\hat{\sigma_b^*}$

- The parametric bootstrap motivates another way to re-sample data i.e. sampling the errors