Exercises Chapter 5

1. To answer the following questions, you can refer to Assumptions 1-4 of Chapter 3 or Assumptions 1-4* of Chapter 5 (both is possible).
   (a) Which assumptions are needed for the unbiasedness of the Ordinary Least Squares (OLS) estimator in the linear regression model?
   (b) Which additional assumptions are needed to make OLS the best linear unbiased estimator?
   (c) Correct or false: The phrase "linear" in part (b) refers to the fact that we are estimating a linear model.

2. (a) Explain, why the elasticity

$$El_x f(x) = \frac{x}{f(x)} \frac{\partial f(x)}{\partial x}$$

   of a deterministic and differentiable function f, with $y = f(x)$, can be interpreted as the approximate percentage change in y per 1% change in x.

   Moreover, show that the elasticity with respect to x can be written as

$$El_x f(x) = \frac{\partial \log(f(x))}{\partial \log(x)}.$$

   (b) Consider the following log-log regression model (our regularity Assumptions 1-4 are assumed to be true):

$$\log(Y_i) = \gamma_1 + \gamma_2 \log(X_i) + \varepsilon_i, \quad i = 1, \ldots, n. \tag{1}$$

   The parameter γ_2 in this regression model is often interpreted as the elasticity of $f(x) = E(Y|X = x)$ with respect to $x > 0$, i.e. as the approximate percentage change in $E(Y|X = x)$ per 1% change in x. When is this interpretation true?

3. Consider the following multiple linear regression model

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i, \quad i = 1, \ldots, n.$$

   Assumptions 1-4* of Chapter 5 are assumed to hold. Consider a smallish sample size of $n = 25$ and let

$$\begin{aligned} \beta &= (\beta_1, \beta_2, \beta_3)' = (2, 3, 4)' \\ X_{i2} &\sim U[1, 4] \\ V_i &\sim N(0, 1) \\ X_{i3} &= 2X_{i2} + V_i \\ \varepsilon_i\,|X &\sim N(0, 2/3) \end{aligned}$$

   Hint: Observe that X_2 and X_3 are constructed to correlate with each other.

   This is how you can draw $n = 25$ realizations of the regressors and the error term. Observe the strong correlation between X_2 and X_3.

```
> set.seed(1234)
> n   <- 25
> X_2 <- runif(n, 1, 4)
> V   <- rnorm(n)
> X_3 <- 2 * X_2 + V
> eps <- rnorm(n, sd=sqrt(2/3))
> ## Sample correlation between X_2 and X_3:
> cor(X_2, X_3)
```

```
[1] 0.8547267
```

Take the following X matrix as the observed X to condition on.

```
> X   <- cbind(rep(1, n), X_2, X_3)
```

(a) Write a Monte Carlo (MC) simulation with 500 MC replications to produce 500 realizations of the OLS estimator $\hat{\beta}_2$ for β_2 conditional on X. Approximate the conditional bias of $\hat{\beta}_2$ given X, i.e. $\text{Bias}(\hat{\beta}_2|X)$, using the 500 realizations of $\hat{\beta}_2|X$. What do you observe?

(b) Repeat the MC simulation in (a), but, when computing the OLS estimates, omit X_{i3} from the estimation formula for all $i = 1, \dots, n$. Approximate again conditional bias of $\hat{\beta}_2$ given X, i.e. $\text{Bias}(\hat{\beta}_2|X)$, using the 500 realizations of $\hat{\beta}_2|X$. What do you observe?

(c) Does omitting X_{i3}, $i = 1, \dots, n$, violate the exogeneity assumption (Assumption 2)? Explain your answer using mathematical derivations.

4. Correct or false? Justify your answers.

   a) A regression of the OLS residuals on the regressors included in the model yields, by construction, an R^2 of zero.

   b) If an estimate $\hat{\beta}_k$ is significantly different from zero at the 10% level, it is also significantly different from zero at the 5% level.

5. Consider the following multiple linear regression model:

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i \quad i = 1, \dots, n,$$

where n is a small sample size (e.g. $n = 15$). Assume the model assumptions of Chapter 5 are fulfilled.

(a) How can one test the hypothesis that $\beta_3 = 1$ against a two-sided alternative. State H_0, H_A, the test-statistic, its distribution, and explain how the test decision is conduced.

(b) Consider the null hypothesis $\beta_2 + \beta_3 = 0$ against a two-sided alternative. Use mathematical derivations to show that in this case an F-test simplifies to a t-test.

(c) How can one test the hypothesis that $\beta_2 = \beta_3 = 0$. State H_0, H_A, the test-statistic, its distribution, and explain how the test decision is conduced.