ECONOMETRICS I: PROBLEM SET 3

EXERCISE: MLE

a) Show that, under $e \sim N(0, \sigma^2)$ MLE estimators of β and σ^2 are equivalent to OLS.

b) Execute the following R code, which serves as an example to obtain robust standard errors for heteroscedasticity and the corresponding hypothesis t-tests.

Example:

```
#Environment:
library(AER)              # install.packages("AER")
data("CPSSWEducation")
attach(CPSSWEducation)    # ?CPSSWEducation

# Model
reg <- lm(earnings ~ education)
summary(reg)

# Plot observations and add the regression line
plot(education, earnings, ylim = c(0, 150))
abline(labor_model, col = "steelblue", lwd = 2)

# Compute homoskedastic-robust standard errors.
t <- linearHypothesis(reg, "X = 0")$'Pr(>F)'[2] < 0.05

# Compute heteroskedasticity-robust (HC1) standard errors
t.rob <- linearHypothesis(reg,
                          "X = 0", white.adjust = "hc1")$'Pr(>F)'[2] < 0.05

# Show both t-tests, where 1="true" meaning Ho is "true" at the 95% level.
round(cbind(t = mean(t), t.rob = mean(t.rob)), 3)

# Same fot the varcovar matrix
(vcov <- vcovHC(labor_model, type = "HC1"))

# Compute the square root of the diagonal elements in vcov
(robust_se <- sqrt(diag(vcov)))
```

```r
# We use `coeftest()` on our robust model and show the assumed-homoskedastic
# model:
coeftest(labor_model, vcov. = vcov)
summary(labor_model)
```

- Generate heteroscedastic data for two variables (X) and (Y) (as in 1.c), which satisfy the population equation $Y = \alpha + \beta X + e$, where $\beta = 1$.
- Create a scatterplot of these variables with a fitted line showing their relationship.
- Perform 10,000 regressions, each with new random samples drawn from the same distribution established in (1.c).
- For each $i = 1, \ldots, 10,000$, perform the "HC0" and "HC2" t-tests and store them in vectors `t` and `t.rob`.
- Compute the percentage of rejections of the null hypothesis $\beta = 1$.
- What is the conclusion of this result?

c) MLE is the technique that helps us determine the parameters of the distribution that best describe the given data. Imagine that we have a sample that was drawn from a normal distribution with mean $\mu = 5$ and variance $\sigma^2 = 100$. The objective is to estimate these parameters with MLE.

The normal log-likelihood function is given by: $l = -\frac{1}{2}nln(2\pi) - \frac{1}{2}nln(\sigma^2) - \frac{1}{2\sigma^2}\sum(yi - \mu)^2$

*Note that minimizing a negative likelihood function is the same as maximizing the likelihood function.

```r
# We define:
X <- rnorm(n = 1000, mean = 5, sd = 10)
df <- data.frame(X)

# We program the log-likelihood function in R:
normal.lik1 <- function(theta,y){
  mu<-theta[1]
  sigma2<-theta[2]
  n<-nrow(y)
  logl<- -.5*n*log(2*pi)-.5*n*log(sigma2)-(1/(2*sigma2))*sum((y-mu)^2)
  return(-logl)
}

# Here theta is a vector containing the two parameters of interest
# (i.e. theta[1] is equal to mu). The remainder sets n, and the log-likelihood function.

# we use optim(starting values, log-likelihood, data) with starting values 0 and 1.
optim(c(0,1), normal.lik1, y=df)

#We can ask for the method-of-moments-mean directly:
mean(X)
var(X)
```

- Now, estimate the MLE parameters β and σ^2 with Y = 5 + 2X + e.

d) Show that plugging in the estimators β^2_{mle} and $\hat{\sigma}^2_{mle}$ into:

$$l_n(\beta, \sigma^2) = -\frac{n}{2} log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - X_i'\beta)^2$$

We get the maximized log likelihood function. How does it work as a measure o fit?

e) Let $x_1, x_2, \ldots, x_n$ be a random sample from each of the following probability density function (pdf):

(a) Bernoulli distribution

$$f(x;\theta) = \theta^x(1-\theta)^{1-x}, \quad 0 \leq \theta \leq 1, \quad x = 0, 1$$

(b) Poisson distribution

$$f(x;\theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, 2, \ldots, \quad 0 \leq \theta < \infty$$

(c) Exponential distribution

$$f(x;\theta) = \theta e^{-\theta x}, \quad 0 < x < \infty, \quad 0 < \theta < \infty$$

Derive the maximum likelihood estimator (MLE) of θ for a single case. Verify that the second derivative of the log-likelihood is negative, ensuring a maximum of the likelihood function.