# Exam

**Problem 1:** (15 points)
Suppose that the classical regression model applies but that the true value of the constant is zero. In order to answer the following questions assume just one independent variable.

1. Give the formulae for the two least squares slope estimators (the one with and the one without the constant).
2. Calculate their variances.
3. Compare the variance of the least squares slope estimator computed without a constant term with that of the estimator computed with an unnecessary constant term.

**Solution:**

1.

$$\begin{aligned} y &= \beta_1 + \beta_2 x + \varepsilon \rightarrow \beta_2 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} \\ y &= \tilde{\beta}_2 x + \varepsilon \rightarrow \tilde{\beta}_2 = \frac{\sum_i x_i y_i}{\sum_i x_i^2} \end{aligned}$$

2.

$$\begin{aligned} Var(\beta_2) &= \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2} \\ Var(\tilde{\beta}_2) &= \frac{\sigma^2}{\sum_i x_i^2} \end{aligned}$$

3. The ratio of these two variances is

$$\begin{aligned}
\frac{Var(\tilde{\beta}_2)}{Var(\beta_2)} &= \frac{\frac{\sigma^2}{\sum_i x_i^2}}{\frac{\sigma^2}{\sum_i (x_i-\bar{x})^2}} \\
&= \frac{\sum_i (x_i - \bar{x})^2}{\sum_i x_i^2} \\
& \sum (x_i - \bar{x})^2 = \sum (x_i^2 - 2x_i\bar{x} + \bar{x}) = \sum x_i^2 - 2n\bar{x}\bar{x} + n\bar{x}^2 = \sum x_i^2 - n\bar{x}^2 \\
&= \frac{\sum_i x_i^2 - n\bar{x}^2}{\sum_i x_i^2} \\
&= 1 - \frac{n\bar{x}^2}{\sum_i x_i^2} \leq 1
\end{aligned}$$

It follows that fitting the constant term when it is unnecessary inflates the variance of the least squares estimator if the mean of the regressor is not zero.

**Problem 2:** (15 points)
Suppose that $y$ has the pdf $f(y|\mathbf{x}) = \frac{1}{\boldsymbol{\beta}'\mathbf{x}} e^{-y/(\boldsymbol{\beta}'\mathbf{x})}, y > 0$. Then $E[y|\mathbf{x}] = \boldsymbol{\beta}'\mathbf{x}$ and $Var[y|\mathbf{x}] = (\boldsymbol{\beta}'\mathbf{x})^2$. For this model, prove that the GLS and MLE estimators are the same, even though this distribution involves the same parameters in the conditional mean function and the disturbance variance.

**Solution:**
First the GLS estimator:

$$\hat{\boldsymbol{\beta}}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y = \left(\sum_i \frac{\mathbf{x_i}\mathbf{x_i}'}{(\boldsymbol{\beta}'\mathbf{x_i})^2}\right)^{-1}\left(\sum_i \frac{\mathbf{x_i}y_i}{(\boldsymbol{\beta}'\mathbf{x_i})^2}\right)$$

Next the MLE estimator:

$$\begin{aligned}
L &= \prod_i \frac{1}{\boldsymbol{\beta}'\mathbf{x_i}} e^{-y_i/(\boldsymbol{\beta}'\mathbf{x_i})} \\
\ln L &= -\sum_i ln(\boldsymbol{\beta}'\mathbf{x_i}) - \sum_i y_i/(\boldsymbol{\beta}'\mathbf{x_i}) \\
\frac{\partial \ln L}{\partial \boldsymbol{\beta}} &= -\sum_i \frac{\mathbf{x_i}}{\boldsymbol{\beta}'\mathbf{x_i}} + \sum_i y_i/(\boldsymbol{\beta}'\mathbf{x_i})^2\mathbf{x_i} = 0 \\
\sum_i \frac{y_i\mathbf{x_i}}{(\boldsymbol{\beta}'\mathbf{x_i})^2} &= \sum_i \frac{\mathbf{x_i}}{\boldsymbol{\beta}'\mathbf{x_i}} \qquad \text{Now write } \sum_i \frac{\mathbf{x_i}}{\boldsymbol{\beta}'\mathbf{x_i}} = \sum_i \frac{\mathbf{x_i}\mathbf{x_i}'\boldsymbol{\beta}}{(\boldsymbol{\beta}'\mathbf{x_i})^2} \\
\sum_i \frac{y_i\mathbf{x_i}}{(\boldsymbol{\beta}'\mathbf{x_i})^2} &= \sum_i \frac{\mathbf{x_i}\mathbf{x_i}'\boldsymbol{\beta}}{(\boldsymbol{\beta}'\mathbf{x_i})^2} \\
\hat{\boldsymbol{\beta}}_{MLE} &= \left(\sum_i \frac{\mathbf{x_i}\mathbf{x_i}'}{(\boldsymbol{\beta}'\mathbf{x_i})^2}\right)^{-1} \sum_i \frac{y_i\mathbf{x_i}}{(\boldsymbol{\beta}'\mathbf{x_i})^2}
\end{aligned}$$

**Problem 3:** (15 points)
The following model is estimated using a balanced panel of five firms over 20 years: $I_{it} = \beta_1 F_{it} + \beta_2 C_{it} + \varepsilon_{it}$, where the regressors are market value ($F$) and capital ($C$) and the dependent variable is investment ($I$). Suppose that the true error structure of the model is $\varepsilon_{it} = \alpha_i + \eta_{it}$, where $\alpha$ is uncorrelated with the regressors.

1. If the model is estimated as a fixed effects model, what will be the statistical properties, in terms of efficiency and consistency, of the estimates?

2. The estimates for pooled OLS, fixed effects (using dummies) and random effects models are given in the table below. Use the statistics shown to decide whether the data support a fixed effects or random effects specification. Carefully explain your reasoning.

Dependent Variable is Investment

| Estimation | Constant | Market Value | Capital |
|---|---|---|---|
| (a) OLS | -48.030 (-2.236) | 0.10509 (9.236) | 0.30537 (7.019) |
| (b) Fixed Effects | - | 0.10598 (6.669) | 0.34666 (14.348) |
| (c) Random Effects | -61.575 (-0.775) | 0.10549 (6.859) | 0.34641 (14.350) |

(t-ratios are shown in brackets)
Breush-Pagan LM test for random effects (1 df): 453.82
Hausman test of fixed vs random effects (2 df): 1.27

**Solution:**

1. If the individual effects are strictly uncorrelated with the regressors then a random effects model is the appropriate model. However, if a fixed effect model is estimated the estimates will be consistent but not efficient.

2. Breush-Pagan LM test: Test statistic is 453.82, the critical value from the chi-squared table is 3.84, so the null hypothesis that random effects are not needed can be rejected.
   Hausman Test: Test statistic is 1.27, the critical value from the chi-squared table is 5.99, so the null hypothesis of the random effects model cannot be rejected.

**Problem 4:** (15 points)
Consider the stochastic processes given below. For each process determine what the effects of first differencing the process, i.e. computing $y_t - y_{t-1}$, on autocorrelation are, e.g. reduction of the autocorrelation.

1. $y_t = y_{t-1} + \varepsilon_t$, where $\varepsilon_t$ is normally distributed white noise.
2. $y_t = \beta_0 + \beta_1 t + \varepsilon_t$, where $\varepsilon_t$ is normally distributed white noise.
3. $y_t = \boldsymbol{\beta}'\mathbf{x}_t + \varepsilon_t$, where $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ and $u_t$ is normally distributed white noise.
[Hint: Compare the autocorrelation of $\varepsilon_t$ and the autocorrelation of $(\varepsilon_t - \varepsilon_{t-1})$.]

**Solution:**

1. $\Delta y_t = y_t - y_{t-1} = \varepsilon_t$, white noise, no more autocorrelation
2. $\Delta y_t = y_t - y_{t-1} = \beta_1 + \varepsilon_t - \varepsilon_{t-1}$. This is an MA(1) process with autocorrelation $\frac{\theta}{1+\theta^2} = \frac{1}{1+1} = \frac{1}{2}$.
3. $\Delta y_t = y_t - y_{t-1} = \beta'(x_t - x_{t-1}) + v_t$, where $v_t = \varepsilon_t - \varepsilon_{t-1}$.

$$
\begin{aligned}
Var(\varepsilon_t) &= \frac{\sigma_u^2}{1-\rho^2} \\
Var(v_t) &= Var(\varepsilon_t - \varepsilon_{t-1}) = Var(\rho\varepsilon_{t-1} - \varepsilon_t + u_t) \\
&= Var[(\rho-1)\varepsilon_{t-1} + u_t] = (\rho-1)^2\frac{\sigma_u^2}{1-\rho^2} + \sigma_u^2 = \frac{2\sigma_u^2}{1+\rho} \\
Cov[v_t, v_{t-1}] &= Cov[\varepsilon_t - \varepsilon_{t-1}, \varepsilon_{t-1} - \varepsilon_{t-2}] \\
&= E[\varepsilon_t\varepsilon_{t-1} - \varepsilon_{t-1}^2 - \varepsilon_t\varepsilon_{t-2} + \varepsilon_{t-1}\varepsilon_{t-2}] \\
&= \rho\frac{\sigma_u^2}{1-\rho^2} - \frac{\sigma_u^2}{1-\rho^2} - \rho^2\frac{\sigma_u^2}{1-\rho^2} + \rho\frac{\sigma_u^2}{1-\rho^2} = \frac{\sigma_u^2(2\rho - 1 - \rho^2)}{1-\rho^2} \\
&= \frac{\sigma_u^2(\rho-1)^2}{(\rho-1)(\rho+1)} = \frac{\sigma_u^2(\rho-1)}{\rho+1} \\
\frac{Cov[v_t, v_{t-1}]}{Var[v_t]} &= \frac{\rho-1}{2}
\end{aligned}
$$

Compare the two autocorrelations:

$$\begin{aligned} |\rho| &> \left|\frac{\rho-1}{2}\right| & \text{Assume } \rho > 0 \text{ and } |\rho| < 1 \\ \rho &> -\frac{\rho-1}{2} \rightarrow \rho > \frac{1}{3} \end{aligned}$$

If the original autocorrelation is greater than 1/3 (For economic data, this is likely to be fairly common.) the differenced process has a smaller autocorrelation.