240A Winter 2007: Solutions to Midterm Exam

1.(a) We have

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 0 & 2 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 1 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix} \text{ and } \mathbf{X}'\,\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\,\mathbf{y} = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix}^{-1} \begin{bmatrix} 8 \\ 12 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 8 \\ 12 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

(b) We have $\mathbf{y}' = \begin{bmatrix} 4 & 0 & 2 & 2 \end{bmatrix}$ and $\widehat{\mathbf{y}} = \begin{bmatrix} 3 & 1 & 3 & 1 \end{bmatrix}$, so

$$s^2 = \frac{1}{2}(\mathbf{y} - \widehat{\mathbf{y}})'(\mathbf{y} - \widehat{\mathbf{y}}) = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2} \times 4 = 2.$$

$$\Rightarrow \mathrm{V}\left[\widehat{\boldsymbol{\beta}}\right] = s^2(\mathbf{X}'\mathbf{X})^{-1} = 2\begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

So standard error of $\widehat{\beta}_2 = \sqrt{1/2} = 0.707$.

(c) Easiest to use $R^2 = 1 - \sum_i \widehat{u}_i^2 / \sum_i (y_i - \bar{y})^2 = 1 - 4/(2^2 + 2^2 + 0^2 + 0^2) = 0.5$.
(where use $\sum_i \widehat{u}_i^2 = (\mathbf{y} - \widehat{\mathbf{y}})'(\mathbf{y} - \widehat{\mathbf{y}})$ computed in part (c).

(d)(i)(ii) The Stata commands are

```
. input x y
  1. 2 4
  2. 0 0
  3. 2 2
  4. 0 2
  5. end
. regress y x
```

2.(a) Differentiate

$$\frac{\partial Q(\gamma)}{\partial \gamma} = \frac{\partial \mathbf{v}'\mathbf{A}\mathbf{v}}{\partial \gamma} = \frac{\partial}{\partial \gamma}\{\mathbf{y}'\mathbf{A}\mathbf{y} - 2\mathbf{y}'\mathbf{A}\mathbf{Z}\gamma + \gamma'\mathbf{Z}'\mathbf{A}\mathbf{Z}\gamma\} = -2\mathbf{Z}'\mathbf{A}\mathbf{y} + 2\mathbf{Z}'\mathbf{A}\mathbf{Z}\gamma = -2\mathbf{Z}'\mathbf{A}(\mathbf{y} - \mathbf{Z}\gamma).$$

Alternatively use chain rule for matrix differentiation

$$\frac{\partial Q(\gamma)}{\partial \gamma} = \frac{\partial \mathbf{v}'\mathbf{A}\mathbf{v}}{\partial \gamma} = \frac{\partial \mathbf{v}'}{\partial \gamma} \times \frac{\partial Q(\gamma)}{\partial \mathbf{v}} = -\mathbf{Z}' \times 2\mathbf{A}\mathbf{v} = -2\mathbf{Z}'\mathbf{A}(\mathbf{y} - \mathbf{Z}\gamma).$$

Set to zero

$$\mathbf{Z}'\mathbf{A}(\mathbf{y} - \mathbf{Z}\gamma) = \mathbf{0} \quad \Rightarrow \quad \mathbf{Z}'\mathbf{A}\mathbf{y} = (\mathbf{Z}'\mathbf{A}\mathbf{Z})\gamma \quad \Rightarrow \widetilde{\gamma} = (\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{y}.$$

(b) We have since $\mathbf{y} = \mathbf{Z}\gamma + \mathbf{v}$,

$$\begin{aligned} \widetilde{\gamma} &= (\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}(\mathbf{Z}\gamma + \mathbf{v}) \\ &= \gamma + (\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{v}. \end{aligned}$$

So $\mathrm{E}[\widetilde{\gamma}] = \boldsymbol{\gamma}$ as $\mathrm{E}[\mathbf{v}] = \mathbf{0}$ and

$$\begin{aligned}
\mathrm{V}[\widetilde{\gamma}] &= \mathrm{E}[(\widetilde{\gamma}-\gamma)(\widetilde{\gamma}-\gamma)'] \\
&= \mathrm{E}[(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{v})((\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{v})'] \\
&= (\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\mathrm{E}[\mathbf{v}\mathbf{v}']\mathbf{A}\mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1} \\
&= (\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\sigma^2\mathbf{I}\mathbf{A}\mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1} \\
&= \sigma^2(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{A}\mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}.
\end{aligned}$$

(c) $\widetilde{\mathbf{v}} = \mathbf{y} - \mathbf{Z}\widetilde{\gamma} = \mathbf{Z}\gamma + \mathbf{v} - \mathbf{Z}\widetilde{\gamma} = \mathbf{v} - \mathbf{Z}(\widetilde{\gamma}-\gamma) = \mathbf{v} - \mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{v}$.
So $\widetilde{\mathbf{v}} = (\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A})\mathbf{v}$, and

$$\begin{aligned}
\mathrm{E}[\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}'] &= \mathrm{E}[(\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A})\mathbf{v}\mathbf{v}'(\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A})] \\
&= \sigma^2(\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A})(\mathbf{I} - \mathbf{A}\mathbf{Z}(\mathbf{Z}'\mathbf{A}\mathbf{Z})^{-1}\mathbf{Z}') \\
&\neq \sigma^2\mathbf{I} = \mathrm{E}[\mathbf{v}\mathbf{v}'].
\end{aligned}$$

3.(a) We have

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'(\mathbf{Z}\boldsymbol{\beta}+\mathbf{u}) = \boldsymbol{\beta} + (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{u}. \\
\mathrm{E}[\widehat{\boldsymbol{\beta}}] &= \boldsymbol{\beta} + \mathrm{E}_{\mathbf{Z},\mathbf{X},\mathbf{u}}[(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{u}] = \boldsymbol{\beta} + \mathrm{E}_{\mathbf{Z},\mathbf{X}}[(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}']\mathrm{E}_{\mathbf{u}}[\mathbf{u}] = \boldsymbol{\beta},
\end{aligned}$$

where we use independence of $\mathbf{u}$ from $\mathbf{X}$ and $\mathbf{Z}$ and $\mathrm{E}[\mathbf{u}] = \mathbf{0}$.

(b) We have

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= \boldsymbol{\beta} + \left(N^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1} N^{-1}\mathbf{Z}'\mathbf{u} \\
&\xrightarrow{p} \boldsymbol{\beta} + \left(\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1} \operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{u} \\
&\xrightarrow{p} \boldsymbol{\beta}
\end{aligned}$$

if we assume $\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{X}$ exists and is finite nonzero
and $\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{u} = \mathbf{0}$, which is likely as $N^{-1}\mathbf{Z}'\mathbf{u} = \frac{1}{N}\sum \mathbf{z}_i u_i$, and $\mathbf{z}_i u_i$ are iid with mean 0.

(c) We have

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= \left(N^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1} \frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{u} \\
&\xrightarrow{d} \left(\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1} \times \mathcal{N}[\mathbf{0},\ \sigma^2 \operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{Z}] \\
&\xrightarrow{d} \mathcal{N}\left[\mathbf{0}, \sigma^2 \left(\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1} \left(\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{Z}\right) \left(\operatorname{plim} N^{-1}\mathbf{X}'\mathbf{Z}\right)^{-1}\right].
\end{aligned}$$

Where we assume a central limit theorem applies so that $\frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{u} \xrightarrow{d} \mathcal{N}[\mathbf{0},\ \sigma^2 \operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{Z}]$, and the limit variance is $\sigma^2 \operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{Z}$ since $\mathrm{V}[\frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{u}] = \mathrm{E}_{\mathbf{Z},\mathbf{u}}[N^{-1}\mathbf{Z}'\mathbf{u}\mathbf{u}'\mathbf{Z}] = \sigma^2 N^{-1}\mathrm{E}_{\mathbf{Z}}[\mathbf{Z}'\mathbf{Z}]$.

(d) We have

$$\widehat{\boldsymbol{\beta}} \overset{a}{\sim} \mathcal{N}[\boldsymbol{\beta},\ \sigma^2 N^{-1} \left(\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1} \left(\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{Z}\right) \left(\operatorname{plim} N^{-1}\mathbf{X}'\mathbf{Z}\right)^{-1}].$$

[or $\widehat{\boldsymbol{\beta}} \overset{a}{\sim} \mathcal{N}[\boldsymbol{\beta},\ \sigma^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}]$ will also do].
(e) Here $\frac{1}{N}\sum_i z_i u_i$ is the average of $z_i u_i$ which is iid since z_i and u_i are each iid and independent of each other. So can apply Khinchines theorem

$$\operatorname{plim} \frac{1}{N}\sum\nolimits_i z_i u_i = \mathrm{E}[z_i u_i] = \mathrm{E}[z_i]\mathrm{E}[u_i] = 0 \text{ as } \mathrm{E}[u_i] = 0.$$

(f) Again $\frac{1}{N}\sum_i z_i u_i$ is the average of $z_i u_i$ which is iid since z_i and u_i are each iid and independent of each other. $\mathrm{E}[z_i u_i]$ and $\mathrm{V}[z_i u_i] = \mathrm{E}[z_i^2 u_i^2] = \sigma^2\mathrm{E}[z_i^2] = \sigma^2\mathrm{E}[z^2]$. Can apply the Lindberg Levy CLT with

$$\begin{aligned}
&\frac{\frac{1}{N}\sum_i z_i u_i - 0}{\sqrt{\sigma^2\mathrm{E}[z^2]/N}} \xrightarrow{d} \mathcal{N}\left[\mathbf{0}, 1\right] \\
&\frac{\frac{1}{\sqrt{N}}\sum_i z_i u_i}{\sigma^2\mathrm{E}[z^2]} \xrightarrow{d} \mathcal{N}\left[\mathbf{0}, 1\right] \\
&\frac{\frac{1}{\sqrt{N}}\sum_i z_i u_i}{\sigma^2\mathrm{E}[z^2]} \xrightarrow{d} \mathcal{N}\left[\mathbf{0}, \sigma^2\mathrm{E}[z^2]\right] \text{ or } \mathcal{N}\left[\mathbf{0}, \sigma^2 \operatorname{plim} N^{-1}\sum_i z_i^2\right].
\end{aligned}$$

4.(a)(i) $\beta_2 - \beta_3 - 2 = 0$ is $\mathbf{R} = \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix}$ and $\mathbf{r} = 2$.

(ii) $\beta_2 = 0, \beta_3 = 0, \beta_4 = 0$ is $\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ and $\mathbf{r} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

(b) We have

$$\begin{aligned}
& \widetilde{\boldsymbol{\beta}} \overset{a}{\sim} \mathcal{N}[\boldsymbol{\beta},\ \mathbf{V}] \\
\Rightarrow\quad & \mathbf{R}\widetilde{\boldsymbol{\beta}} - r \overset{a}{\sim} \mathcal{N}[\mathbf{R}\boldsymbol{\beta} - \mathbf{r},\ \mathbf{RVR}'] \\
\Rightarrow\quad & \mathbf{R}\widetilde{\boldsymbol{\beta}} - r \overset{a}{\sim} \mathcal{N}[\mathbf{0},\ \mathbf{RVR}'] \text{ under } H_0 \\
\Rightarrow\quad & \mathrm{W} = (\mathbf{R}\widetilde{\boldsymbol{\beta}} - r)'[\mathbf{RVR}']^{-1}(\mathbf{R}\widetilde{\boldsymbol{\beta}} - r) \overset{a}{\sim} \chi^2(q) \text{ under } H_0.
\end{aligned}$$

5.(a) The dgp is $N = 5$ independent observations with

$$\begin{aligned}
y &= 1 + 2x + u \\
u &\sim \chi^2(1) - 1 \text{ with mean } 0 \\
x &\sim \mathcal{N}[1,\ 2^2]
\end{aligned}$$

(b) Stochastic regressors (actually random sampling of both y and x), since in different simulations we will have different regressors. [NOTE: It is not enough to say that x is drawn from the normal. The big thing is that different draws of x are used in each simulation].

(c) Here θ is the coefficient of x, with $\theta = 2$ in the dgp.
We expect $\widehat{\theta}$ to be unbiased as the error is inedependent with zero mean.
We expect $\widehat{\theta}$ to be nonormal distributed since this is a small sample with nonnormal errors.
[Also even if error was normal, still get $\widehat{\theta}$ nonnormal for small sample as regressors are stochastic].
We find the average $\widehat{\theta} = 2.016591$ which is close to 2 as expected.
The distribution of $\widehat{\theta}$ is very nonormal, since symmetry statistic is a long way from zero and kurtosis statistic is a long way from three.

6.(a) Yes. The overall F statistic is 134.68 with a p-value of 0.0.

(b) Yes. The t-test statistic of $H_0 : \beta_{\text{incomesq}} = 0$ against $H_a : \beta_{\text{incomesq}} \neq 0$ is -1.12 with a p-value of 0.262 which is > 0.05. So do not reject H_0. The squared income term is not needed.
And the linear term income is needed since it is statistically significant with p-value of 0.000.

(c) $H_0 : \beta_{\text{famsize}} = 700$ against $H_a : \beta_{\text{famsize}} \neq 700$.
$t = (805.701 - 700)/61.51642 = 1.718 < t_{.025;700} \simeq x_{.025} = 1.96$.
Do not reject H_0. Conclude that it is associated with a \$700 increase.
[ASIDE: I had meant to ask this as a one-sided test question.]

	Exam / 50			Exam / 50		
$75th$ percentile	41	(80%)			B+	28 and above
Median	36	(72%)	A	40 and above	B	22 and above
$25th$ percentile	31	(62%)	A-	34 and above		

Exam Discussion

The exam was graded very carefully. Generally if you made a mistake in details you lost a point even if final answer was correct. For example, multiplying noncomformable matrices due to misplaced transposes. The goal here is to inform you about details you are missing.

1.(b) Many missed this.

2.(a) Need to have all details correct along the way.
(b) The variance is formally $\mathrm{V}[\widetilde{\boldsymbol{\gamma}}] = \mathrm{E}[(\widetilde{\boldsymbol{\gamma}} - \mathrm{E}[\widetilde{\boldsymbol{\gamma}}])(\widetilde{\boldsymbol{\gamma}} - \mathrm{E}[\widetilde{\boldsymbol{\gamma}}])']$.
So you to show $\mathrm{E}[\widetilde{\boldsymbol{\gamma}}] = \boldsymbol{\gamma}$ before using $\mathrm{E}[(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma})(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma})']$.
(c) The question asked for $\mathrm{E}[\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}']$, an $N \times N$ matrix.
Many instead considered the scalar $\mathrm{E}[\widetilde{\mathbf{v}}'\widetilde{\mathbf{v}}]$ which is more difficult to obtain.

3. This question had stochastic regressors, not fixed regressors.
Also we are dealing with matrices not scalars.
(a) Several people said $\mathrm{E}[(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{u}] = \mathrm{E}[(\mathbf{Z}'\mathbf{X})^{-1}]\mathrm{E}[\mathbf{Z}'\mathbf{u}]$ which cannot be the case if $\mathbf{Z}$ is stochastic (e.g. $\mathrm{E}[\mathbf{Z}'\mathbf{Z}] \neq \mathrm{E}[\mathbf{Z}']\mathrm{E}[\mathbf{Z}]$.)
(b) This was done poorly.
(c) Since dealing with matrices we do not have e.g. $\xrightarrow{d} \mathcal{N}\left[\mathbf{0}, \sigma^2 \frac{\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{Z}}{(\operatorname{plim} N^{-1}\mathbf{Z}'\mathbf{X})^2}\right]$.
(f) This is harder with a lot of details along the way.

4.(b) The question asked for a derivation, not just the formula.

5.(a) You need to write down the data generating process (and mathematically).
Also, note that this is not the dgp is not the same as the simulation design.
(b) This was (deliberately) a tricky question. See the solution.
(c) To get the full 4 points credit you need to have all details correct on both the expected distribution and the observed distribution from the simulation.

6. This was done fine. The last part was easy as I mistakenly asked for a two-sided test when I had meant to ask for a one-sided test of claim that it exceeds $700.