Economics 522A, Homework 7 Suggested Solutions

1. Ruud, 13.8:

   (a) In class, we showed that

$$
\begin{aligned}
s^2 &= \frac{n}{n-k}\left[\frac{1}{n}\sum_{i=1}^n y_i^2 - \hat{\beta}'\left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)\hat{\beta}\right] \\
&= \frac{n}{n-k}\left[\frac{1}{n}\left(y'y - \hat{\beta}'X'X\hat{\beta}\right)\right].
\end{aligned}
$$

First, consider the term $y'y$. Write $y = y - X\beta + X\beta$. Then we can write:

$$
\begin{aligned}
y'y &= (y - X\beta + X\beta)'(y - X\beta + X\beta) \\
&= (y - X\beta)'(y - X\beta) + \beta'X(y - X\beta) + (y - X\beta)'X\beta + \beta'X'X\beta \\
&= (y - X\beta)'(y - X\beta) + \beta'Xy + y'X\beta - \beta'X'X\beta
\end{aligned}
$$

But

$$
\begin{aligned}
\beta'X'y &= \beta'(X'X)(X'X)^{-1}X'y \\
&= \beta'X'X\hat{\beta},
\end{aligned}
$$

and

$$
\begin{aligned}
y'X\beta &= y'X(X'X)^{-1}X'X\beta \\
&= \hat{\beta}'X'X\beta
\end{aligned}
$$

So we can write

$$
y'y = (y - X\beta)'(y - X\beta) + \hat{\beta}'X'X\beta + \beta'X'X\hat{\beta} - \beta'X'X\beta
$$

and

$$
\begin{aligned}
y'y - \hat{\beta}'X'X\hat{\beta} &= (y - X\beta)'(y - X\beta) - \left[\hat{\beta}'X'X\hat{\beta} + \beta'X'X\beta - \hat{\beta}'X'X\beta - \beta'X'X\hat{\beta}\right] \\
&= (y - X\beta)'(y - X\beta) - (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
s^2 &= \frac{n}{n-k}\left[\frac{1}{n}(y - X\beta)'(y - X\beta) - \frac{1}{n}(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta)\right] \\
&= \frac{n}{n-k}\left[\frac{1}{n}\sum_{i=1}^n (y_i - x_i'\beta)^2 - (\hat{\beta} - \beta)'\left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)(\hat{\beta} - \beta)\right].
\end{aligned}
$$

Next, write

$$
\begin{aligned}
s^2 - \sigma^2 &= s^2 - \left[\frac{n}{n-k}\sigma^2 - \frac{k}{n-k}\sigma^2\right] \\
&= s^2 - \left[\frac{n}{n-k}\frac{1}{n}\sum_{i=1}^n \sigma^2\right] + \frac{k}{n-k}\sigma^2 \\
&= \frac{k}{n-k}\sigma^2 + \frac{n}{n-k}\left[\frac{1}{n}\sum_{i=1}^n \left\{(y_i - x_i'\beta)^2 - \sigma^2\right\} - (\hat{\beta} - \beta)'\left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)(\hat{\beta} - \beta)\right].
\end{aligned}
$$

So

$$\sqrt{n}(s^2-\sigma^2) = \frac{k\sqrt{n}}{n-k}\sigma^2+\frac{n}{n-k}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{(y_i-x_i'\beta)^2-\sigma^2\right\}-\sqrt{n}(\hat{\beta}-\beta)'\left(\frac{1}{n}\sum_{i=1}^{n}x_ix_i'\right)(\hat{\beta}-\beta)\right].$$

(b) Consider

$$\sqrt{n}(s^2-\sigma^2)-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{(y_i-x_i'\beta)^2-\sigma^2\right\}.$$

This has three terms:

$$\frac{k\sqrt{n}}{n-k}\sigma^2-\frac{k}{n-k}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{(y_i-x_i'\beta)^2-\sigma^2\right\}\right]-\frac{n}{n-k}\left[\sqrt{n}(\hat{\beta}-\beta)'\left(\frac{1}{n}\sum_{i=1}^{n}x_ix_i'\right)(\hat{\beta}-\beta)\right].$$

We've already shown that:

$$\begin{aligned}\hat{\beta}-\beta &\xrightarrow{\text{p}} 0\\ \frac{1}{n}\sum_{i=1}^{n}x_ix_i' &\xrightarrow{\text{p}} E[x_ix_i']\\ \sqrt{n}(\hat{\beta}-\beta) &\xrightarrow{\text{d}} N(0,\sigma^2E[x_ix_i']^{-1}).\end{aligned}$$

Also, we know that $n/(n-k)\to 1$, $k/(n-k)\to 0$, and $(k\sqrt{n})/(n-k)\to 0$. So, by Slutsky's Lemma,

$$\frac{n}{n-k}\left[\sqrt{n}(\hat{\beta}-\beta)'\left(\frac{1}{n}\sum_{i=1}^{n}x_ix_i'\right)(\hat{\beta}-\beta)\right]\xrightarrow{\text{p}} 0.$$

Also, since ordinary convergence implies convergence in probability,

$$\frac{k\sqrt{n}}{n-k}\sigma^2\xrightarrow{\text{p}} 0.$$

Finally, the term

$$\frac{k}{n-k}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{(y_i-x_i'\beta)^2-\sigma^2\right\}\right]\xrightarrow{\text{p}} 0,$$

since the term inside the brackets converges in distribution to a normal distribution, by the central limit theorem (see the next section). Therefore,

$$\sqrt{n}(s^2-\sigma^2)-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{(y_i-x_i'\beta)^2-\sigma^2\right\}\xrightarrow{\text{p}} 0.$$

(c) Consider $\epsilon_i^2=(y_i-x_i'\beta)^2$.

$$E[\epsilon_i^2]=E[E[\epsilon_i^2|x_i]]=\sigma^2.$$

$$
\begin{aligned}
V[\epsilon_i^2] &= E\left[(\epsilon_i^2)^2\right] - \left(E[\epsilon_i^2]\right)^2 \\
&= E[\epsilon_i^4] - (\sigma^2)^2 \\
&= \mu_4 - \sigma^4.
\end{aligned}
$$

So by the central limit theorem, assuming the moments exist,

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{(y_i - x_i'\beta)^2 - \sigma^2\right\} \xrightarrow{\text{d}} N(0, \mu_4 - \sigma^4),
$$

and the result follows by the continuous mapping theorem.

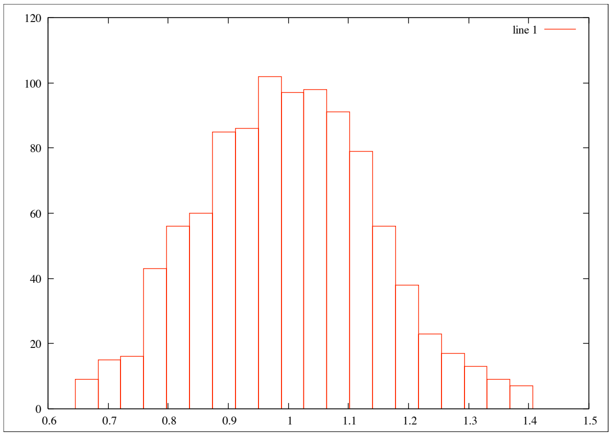
(d) We already have a consistent estimator of σ^2, s^2. We can estimate μ_4 by

$$
\hat{\mu_4} = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i'\hat{\beta})^4.
$$

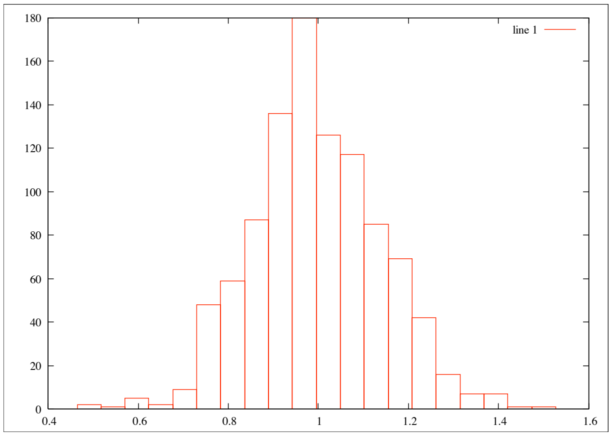
So we can use

$$
\hat{\mu_4} - s^4.
$$

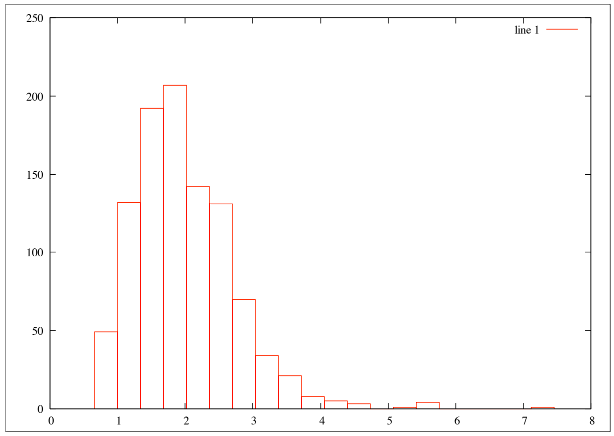
2. The Matlab program `hw72.m` carries out the simulation study. Here is the histogram of draws for $\hat{\beta}_1$:



And here is the histogram of draws for $\hat{\beta}_2$:



We see that both appear to be approximately normally distributed. However, s^2 has a slightly skewed distribution:



So the normal approximation for s^2 would miss the asymmetry in the distribution, although it is not terrible.

Next, we'll consider whether the mean and variance of $\hat{\beta}$ are close to the asymptotic approximations. According to our asymptotic theory,

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\text{d}} N(0, \sigma^2 E[x_i x_i']^{-1}).$$

So the approximate distribution of $\hat{\beta}$ is:

$$\hat{\beta} \overset{\text{a}}{\sim} N(\beta, \frac{\sigma^2}{n} E[x_i x_i']^{-1}).$$

We can be a bit more specific here:

$$\sigma^2 = V[y_i | x_i] = V[v_i | x_i] = 2.$$

Also,

$$E[x_i x_i'] = E \begin{bmatrix} 1 & x_{i2} \\ x_{i2} & x_{i2}^2 \end{bmatrix}.$$

But since $E[x_{i2}] = 0$ and $V[x_{i2}] = 1$,

$$E[x_i x_i'] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So for our study, where $n = 100$,

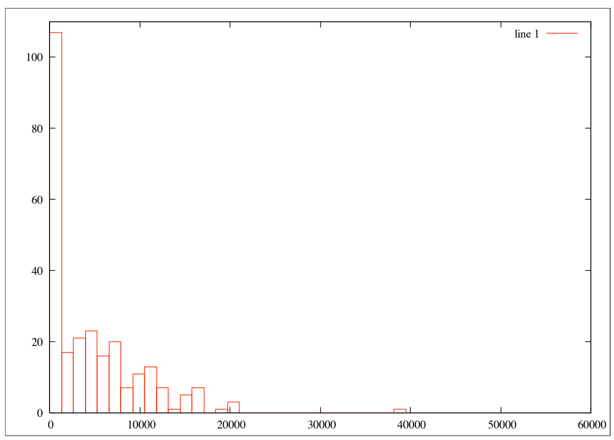
$$V(\hat{\beta}) \approx \begin{pmatrix} \frac{2}{100} & 0 \\ 0 & \frac{2}{100} \end{pmatrix}, \quad E[\hat{\beta}] \approx (1, 1)'.$$

Here are the numerical results:

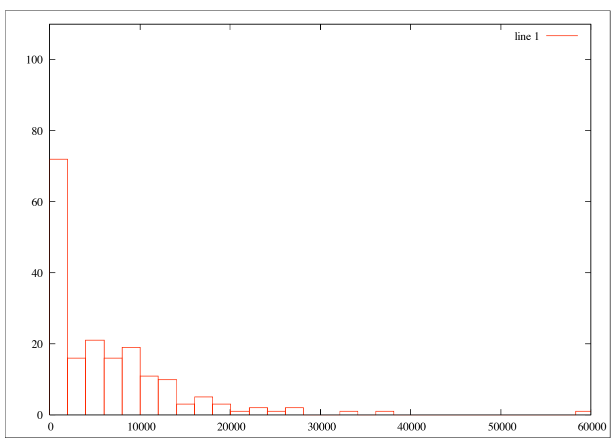
```
average of betahat:
  1.00019  0.99717

covariance matrix of betahat:
    2.0989e-02   -2.8885e-06
   -2.8885e-06    2.0432e-02

average of s2:
  1.9999
```

We see that the results line up very closely with the asymptotic approximations, with the exception of the asymmetry of the distribution of s^2.

3. (a) Here is the histogram of the distribution of y_i given treatment 0:



And here is the distribution given treatment 1:



Clearly, these conditional distributions do not appear to be normal: they are very asymmetric, with considerable mass at 0.

(b) Interpretation of this model: it may help to think of the four different categories that individuals can fall into, as defined by x_{i2} and x_{i3}. Individuals can be treated

and high education, treated and low education, untreated and high education, or untreated and low education.
So

$$\begin{aligned} E[y_i \mid \text{untreated, low ed}] &= \beta_1 \\ E[y_i \mid \text{treated, low ed}] &= \beta_1 + \beta_2 \\ E[y_i \mid \text{untreated, high ed}] &= \beta_1 + \beta_3 \\ E[y_i \mid \text{treated, high ed}] &= \beta_1 + \beta_2 + \beta_3. \end{aligned}$$

So we can interpret β_2 as the difference in average earnings between treated and untreated groups, holding fixed education level. The model makes an assumption that this difference is the same for the low education subgroup, as for the high education subgroup. Similarly, β_3 is the difference in average earnings between high and low education groups, holding fixed treatment status, and again the model assumes that this difference is the same for treated and untreated subgroups.
The Matlab program `hw73.m` carries out the calculations for this problem. Here are the results:

```
betahat =

  4443.3
  1570.1
  4141.2

SE =

   406.58
   632.52
  1438.02

Confidence Intervals

  3646.42  5240.20
   330.32  2809.80
  1322.70  6959.74

W = 16.465
```

We see that the coefficient on the treatment is 1570.1, indicating that treated individuals have about \$1570 higher annual earnings than untreated individuals, controlling for education. This is fairly similar to the result from HW6, where we found that unconditionally, treated individuals had \$1794 higher earnings.
To test that $\beta_2 = \beta_3 = 0$, we can use the restriction matrix:

$$R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The restriction becomes:

$$R\beta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The test statistic is 16.5, so we reject the null hypothesis that $\beta_2 = \beta_3 = 0$ at the 5% level, using the table for the χ^2_2 distribution.

(c) To understand this model, note that

$$\begin{aligned}
E[y_i \mid \text{untreated, low ed}] &= \beta_1 \\
E[y_i \mid \text{treated, low ed}] &= \beta_1 + \beta_2 \\
E[y_i \mid \text{untreated, high ed}] &= \beta_1 + \beta_3 \\
E[y_i \mid \text{treated, high ed}] &= \beta_1 + \beta_2 + \beta_3 + \beta_4.
\end{aligned}$$

So, now, among low educated individuals, treated individuals have a mean that is β_2 higher than untreated individuals. Among high education individuals, the difference in means is $\beta_2 + \beta_4$. So the treated-untreated mean difference can depend on education level.

The Matlab program `hw74.m` does the calculations, and we get:

```
betahat =

  4548.57
  1308.15
   231.49
  5841.80

SE =

   409.05
   645.25
  2492.97
  3047.31

Confidence Intervals

    3.7468e+03    5.3503e+03
    4.3465e+01    2.5728e+03
   -4.6547e+03    5.1177e+03
   -1.3093e+02    1.1815e+04

W = 12.018
```

We reject the null hypothesis that $\beta_3 = \beta_4 = 0$.