1. Various topics.

(a) Consider OLS regression in the linear regression model $y = X\beta + \epsilon$, where $X$ is $N \times k$ and $\epsilon \sim N(0, \sigma^2 I)$. Suppose we wish to predict the actual value of $y$ given $x = x^*$. Give a 95% confidence interval for this prediction.

(b) Suppose we regress $y$ on $X_1$ only yielding OLS estimator $\hat{\beta}_1 = (X_0'X_1)^{-1}X_1'y$. In fact the true model is $y = X_1\beta_1 + X_2\beta_2 + \epsilon$, where $\epsilon \sim N(0, \sigma^2 I)$. Find $E[\hat{\beta}_1]$. What do you conclude?

(c) Consider the usual multiple regression model with intercept, which can be written as $y = \beta_1 l + X_2\beta_2 + \epsilon$, where $l$ is an $N \times 1$ vector of ones, $X_2$ is $N \times (k-1)$. Assume $\epsilon \sim N(0, \sigma^2 I)$. By subtracting means of regressors, an equivalent transformed model is

$$ y = \alpha_1 l + X_2^*\beta_2 + \epsilon, $$

where $X_2^* = X_2 - l\bar{X}_2'$, with $\bar{X}_2' = \frac{1}{N}l'X_2$ a $1 \times (k-1)$ vector of means of the regressors. Write this transformed model as $y = Z\gamma + \epsilon$, for appropriate $Z$ and $\gamma$, use the result that $X_2^*l = 0$, and the general result that $\hat{\gamma} = (Z'Z)^{-1}Z'y$ to show that OLS regression in the transformed model yields, after some algebra, gives estimates

$$ \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \bar{y} \\ (X_2^*X_2^*)^{-1}X_2^*y \end{bmatrix}. $$

2. Consider OLS regression in the linear regression model $y = X\beta + \epsilon$, where $X$ is $N \times k$ and $\epsilon \sim N(0, \sigma^2 I)$. Consider test of $H_0 : R\beta = r$ against $H_a : R\beta \neq r$, where $R$ is $q \times k$.

(a) Suppose $\sigma^2$ is known. Given knowledge of the distribution of the OLS estimator in this setting, derive a test statistic for $H_0$ and gives its distribution under $H_0$.

(b) Use this test statistic to test $H_0 : \beta_1 = 2\beta_2$ against $H_a : \beta_1 \neq 2\beta_2$ at level 0.05 when $\hat{\beta}_1 = 5$, $\hat{\beta}_2 = 2$, $\sigma^2 = 0.1$ and $X'X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. What do you conclude?

(c) Suppose now that $\sigma^2$ is unknown. Propose a variation to your test statistic in part (a) that can be used if $N \to \infty$. 
3. Various topics

(a) Give a mathematical definition for convergence in probability.

(b) Give a mathematical definition for a multivariate central limit theorem.

(c) Give the formula for the IV estimator and state any desirable properties of this estimator.

(d) Give the formula for the GLS estimator and state any desirable properties of this estimator.

(e) For regression with heteroskedasticity give the formula that permits valid inference based on the OLS estimator.

(f) Find the estimator that minimizes \((y - X\beta)'ZZ'(y - X\beta)\), where \(X\) is \(N \times k\), \(Z\) is \(N \times m\), \(m > k\) and both matrices are of full column rank.

4. Consider the estimator \(\hat{\beta} = (X'AX)^{-1}X'Ay\), where \(A\) is a symmetric \(N \times N\) matrix of constants, in the multiple regression model

\[ y = X\beta + u, \]

where \(y\) is an \(N \times 1\) vector, \(X\) is an \(N \times k\) matrix, \(\beta\) is a \(k \times 1\) vector and \(u\) is an \(N \times 1\) vector. We assume that \(X\) is fixed, and \(u \sim [0, \Sigma]\) where \(\Sigma \neq \sigma^2 I\).

For asymptotic theory there is no need to use laws of large numbers and central limit theorems. Instead just use \(\text{plim} \frac{1}{N} X'X\) exists, \(\text{plim} \frac{1}{N} X'Au = 0\) and \(\frac{1}{\sqrt{N}} X'Au \overset{d}{\to} N[0, B]\).

(a) Obtain \(E[\hat{\beta}]\) and \(V[\hat{\beta}]\).

(b) Show whether or not \(\hat{\beta}\) is consistent.

(c) Obtain the limit distribution of \(\sqrt{N}(\hat{\beta} - \beta)\).

(d) Your answer in part (c) will depend in the matrix \(B\). Give the likely form of \(B\) in terms of \(X\), \(A\) and \(\Sigma\). Hence give the asymptotic distribution of \(\hat{\beta}\) in terms of \(X\), \(A\) and \(\Sigma\).
5. The following analyzes data on Median Housing Rents and related data for 58 California Counties from the 2000 Census. The variables are:
   - rent = Monthly rent (median in the county) for Occupied Housing in 2000 in $
   - value = Value (median in the county) of an Occupied House in the county in 2000 in $
   - vacrate = Percentage of Rental Housing Units in the county unoccupied in 2000
   - hhsize = number of persons per household in Occupied Housing in the county
   - percrent = Percentage of Occupied Housing in the county that is rental (rather than nonrental)

Key Stata output is provided on the last page. All three OLS regressions report heteroskedastic-robust standard errors.

(a) Consider the first regression. Show that the predicted value of rent when value equals its sample average is equal to the sample average of rent. Are you surprised? Explain.

(b) Suppose we wish to do a one-sided test of the claim that rent increases with household size. Provide the p-value for this test and state whether or not we reject the claim at five percent.

(c) Does the second model appear to be an improvement on the first model with just value as a regressor? Give two explanations, aside from a formal F-test.

(d)(i) Suppose you wanted to do the F-test of the second model versus the first. What Stata command would you use: give as much detail as possible.
(ii) Give the complete Stata command that produces the estimates for the first regression output.

(e) Test at level 0.05 the hypothesis that the elasticity of house rent with respect to house value equals 1.
   [Hint: Do this the easiest possible way].

(f) For these data provide two distinct methods that you would use, aside from a formal hypothesis test of heteroskedasticity, to determine whether the errors are heteroskedastic in OLS regression of level of rent on level of value? [Note: that the data provided does not answer this question: you need to propose two methods].

(g) Suppose heteroskedasticity is determined to be present and we assume that $V[\text{rent} | \text{value}] = \sigma^2 \times \text{value}^2$. Provide a simple method to obtain parameter estimates in the model $\text{rent} = \beta_1 + \beta_2 \text{value}$ that are more efficient than the OLS estimates.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>county</td>
<td>0</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>rent</td>
<td>58</td>
<td>668.2414</td>
<td>180.9138</td>
<td>429</td>
<td>1185</td>
</tr>
<tr>
<td>value</td>
<td>58</td>
<td>179420.7</td>
<td>96920.5</td>
<td>72900</td>
<td>493300</td>
</tr>
<tr>
<td>vacrate</td>
<td>58</td>
<td>5.62931</td>
<td>3.427086</td>
<td>1.8</td>
<td>20.9</td>
</tr>
<tr>
<td>hhsize</td>
<td>58</td>
<td>2.720172</td>
<td>.3093428</td>
<td>2.29</td>
<td>3.33</td>
</tr>
<tr>
<td>percrent</td>
<td>58</td>
<td>36.57241</td>
<td>7.425839</td>
<td>21.3</td>
<td>65</td>
</tr>
<tr>
<td>logrent</td>
<td>58</td>
<td>6.472257</td>
<td>.2509759</td>
<td>6.061457</td>
<td>7.077498</td>
</tr>
<tr>
<td>logvalue</td>
<td>58</td>
<td>11.982</td>
<td>.464486</td>
<td>11.19684</td>
<td>13.10887</td>
</tr>
</tbody>
</table>

Linear regression

|            | Coef. | Std. Err. | t     | P>|t| | [95% Conf. Interval] |
|------------|-------|-----------|-------|------|----------------------|
| value      | .0017691 | .0001068  | 16.56 | 0.000 | .0015551 .0019831 |
| _cons      | 350.8251 | 17.23109  | 20.36 | 0.000 | 316.3071 385.3431 |

Linear regression

|            | Coef. | Std. Err. | t     | P>|t| | [95% Conf. Interval] |
|------------|-------|-----------|-------|------|----------------------|
| value      | .0018117 | .0001071  | 16.92 | 0.000 | .0015969 .0020265 |
| vacrate    | -2.003118 | 2.654439  | -0.75 | 0.454 | -7.327249 3.321014 |
| hhsize     | 69.26762  | 35.20702  | 1.97  | 0.054 | -1.34872 139.884 |
| percrent   | -2.904163 | 1.310289  | -2.22 | 0.031 | -5.532269 -.2760572 |
| _cons      | 272.2524  | 90.60301  | 3.00  | 0.004 | 90.52576 453.979 |

Linear regression

|            | Coef. | Std. Err. | t     | P>|t| | [95% Conf. Interval] |
|------------|-------|-----------|-------|------|----------------------|
| logrent    | .5149396 | .0208438  | 24.70 | 0.000 | .4731844 .5566947 |
| _cons      | .3022529 | .2487881  | 1.21  | 0.230 | -.1961295 .8006354 |

Number of obs = 58
F( 1, 56) = 274.18
Prob > F = 0.0000
R-squared = 0.8983
Root MSE = 58.219

Number of obs = 58
F( 4, 53) = 136.15
Prob > F = 0.0000
R-squared = 0.9194
Root MSE = 53.276

Number of obs = 58
F( 1, 56) = 610.32
Prob > F = 0.0000
R-squared = 0.9082
Root MSE = 0.07671