Political Methodology Minor Exam
Fall 2009

If you believe a question provides insufficient information to answer, explain why you think so, state some minimal additional assumptions necessary to answer, make those assumptions, and answer accordingly. There should be enough information, but this might earn partial credit. For questions that require calculations, please show your work.

The time for the exam is eight hours. This means that you have from 8 am to 5 pm on September 21.

This exam is open note and open book. You are not allowed to collaborate, however.

PART I Answer all questions. The answer to questions with mathematical derivations may be handwritten though it is preferable to type these parts if time allows.

1. Take the following joint pdf

$$f(x,y) = \begin{cases} 6xy^2 & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

   What is the expectation of XY? Are these variables independent?

2. Suppose we have data from a survey asking about approval of the way President Bush is handling his job as president. The survey has $n = 1,000$ respondents, contacted via a random sampling. The survey finds 55% of respondents approve of President Obama. Construct a 95% confidence interval around the estimated approval rating. Why can a confidence interval be formed for this estimated proportion without reference to the binomial distribution?

3. Below is the density for the geometric distribution. Write down the likelihood, log-likelihood, and derive the MLE for the parameter π. Prove that this estimate is unbiased. Explain what it means for this MLE to be MVUE. Sketch a proof that this MLE is MVUE.

$$f(y) = (1-\pi)^{y_i}\pi$$

4. Below are the results from two regression models. These regression models are based on data from Poe & Tate APSR (1994). Here a scale of human rights violations is the outcome variable. Higher scores on this variable indicate greater numbers of human rights violations. This variable is a function of the Polity III democracy scale, the natural log of population, percent change in population, GDP per capita in 1000's of dollars, a dummy for a leftist government, a dummy for a military regime, a dummy for whether the country was a British colony, a dummy for engagement in an international

war, and a dummy for whether the country was in a civil war. The descriptive statistics for these variables are below in that order:

Variable	Obs	Mean	Std. Dev.	Min	Max
democ3	1848	3.831169	4.379737	0	10
dem_sq	1848	33.84957	42.35903	0	100
lpop	2191	15.52331	1.932997	11.05	20.89
perchpop	2189	2.254591	3.7292	-48.45	86.03
pcgnp	2120	3.970184	6.16552	.053	36.67
left	2184	.1717033	.3772087	0	1
mil2	2198	.2711556	.4446574	0	1
brit	2202	.3723887	.483551	0	1
iwarcow2	2198	.0837125	.2770191	0	1
cwarcow	2182	.1035747	.3047779	0	1

Below are two regression models. In the first, the model includes an interaction between the democracy variable and GDP per capita. Below that model is the variance-covariance matrix. In the second model, the specification includes a democracy-squared term but no interaction.

```
. reg ai democ3 lpop perchpop pcgnp left mil2 brit iwarcow2 cwarcow dem_gdp

      Source |       SS       df       MS              Number of obs =    1511
-------------+------------------------------           F( 10,  1500) =  129.91
       Model |  779.654192    10  77.9654192           Prob > F      =  0.0000
    Residual |  900.197562  1500  .600131708           R-squared     =  0.4641
-------------+------------------------------           Adj R-squared =  0.4605
       Total |  1679.85175  1510  1.11248461           Root MSE      =  .77468

------------------------------------------------------------------------------
          ai |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      democ3 |  -.0237621   .0072756    -3.27   0.001    -.0380335   -.0094907
        lpop |   .1592597    .013965    11.40   0.000     .1318667    .1866526
    perchpop |   .0107945   .0059476     1.81   0.070    -.0008721     .022461
       pcgnp |  -.0158005   .0086908    -1.82   0.069    -.0328479    .0012469
        left |  -.3545474   .0588517    -6.02   0.000    -.4699877   -.2391071
        mil2 |   .1492423   .0494041     3.02   0.003     .0523339    .2461507
        brit |  -.3008154   .0446049    -6.74   0.000    -.3883099   -.2133209
    iwarcow2 |   .4150772   .0676111     6.14   0.000      .282455    .5476995
     cwarcow |   1.202353   .0633387    18.98   0.000     1.078111    1.326595
     dem_gdp |  -.0057364   .0009988    -5.74   0.000    -.0076956   -.0037771
       _cons |   .3523365   .2308823     1.53   0.127    -.1005499    .8052229
------------------------------------------------------------------------------

        e(V) |     democ3        lpop    perchpop       pcgnp        left        mil2
-------------+------------------------------------------------------------------------
      democ3 |  .00005293
        lpop | -5.370e-06   .00019502
    perchpop |  3.846e-06   4.927e-06   .00003537
```

```
       pcgnp |  .00001721   .00001608  -5.718e-06   .00007553
        left |  .00013823    -.000026   .00003863   .00002056   .00346352
        mil2 |   .0001427  -6.569e-07  -.00001637   .00010333   .00023354   .00244076
       brit | -5.648e-06   .00005476  -.00001466  -.00001727    .0006143   .00016537
    iwarcow2 |  .00002614  -.00018699  -9.036e-06  -.00005474  -.00071064   .00016944
     cwarcow | -3.855e-06  -.00011036  -.00003928   .00006199  -.00053406   .00018491
     dem_gdp | -3.484e-06  -2.054e-06   7.150e-07  -7.689e-06  -2.668e-06  -.00001045
       _cons | -.00015064  -.00314358   -.0001623  -.00042004  -.00093393  -.00146642

       e(V) |       brit    iwarcow2     cwarcow     dem_gdp       _cons
------------+------------------------------------------------------------
       brit |  .00198959
    iwarcow2 | -.00022766   .00457125
     cwarcow |  2.485e-06   .00030619   .00401179
     dem_gdp |  2.992e-06   3.429e-06  -2.724e-06   9.977e-07
       _cons | -.00164562   .00269952   .00125884   .00004958   .05330663
```

```
. reg ai democ3 dem_sq lpop perchpop pcgnp left mil2 brit iwarcow2 cwarcow

      Source |       SS       df       MS              Number of obs =    1511
-------------+------------------------------           F( 10,  1500) =  132.86
       Model |  789.031723    10  78.9031723           Prob > F      =  0.0000
    Residual |  890.820031  1500  .593880021           R-squared     =  0.4697
-------------+------------------------------           Adj R-squared =  0.4662
       Total |  1679.85175  1510  1.11248461           Root MSE      =  .77064

------------------------------------------------------------------------------
          ai |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      democ3 |   .1441662   .0275604     5.23   0.000     .0901051    .1982273
      dem_sq |   -.020609   .0029405    -7.01   0.000     -.026377    -.014841
        lpop |   .1403654   .0137777    10.19   0.000     .1133398    .1673909
    perchpop |   .0143653   .0058741     2.45   0.015      .002843    .0258876
       pcgnp |   -.047843   .0043716   -10.94   0.000     -.056418   -.0392679
        left |  -.2951271   .0594488    -4.96   0.000    -.4117388   -.1785155
        mil2 |   .1411464   .0486006     2.90   0.004      .045814    .2364788
        brit |  -.2493438   .0445409    -5.60   0.000     -.336713   -.1619747
    iwarcow2 |   .4220745   .0671957     6.28   0.000      .290267    .5538821
     cwarcow |   1.170763   .0629905    18.59   0.000     1.047204    1.294322
       _cons |   .6319825   .2243066     2.82   0.005     .1919945     1.07197
------------------------------------------------------------------------------

       e(V) |     democ3      dem_sq        lpop    perchpop       pcgnp        left
------------+------------------------------------------------------------------------
      democ3 |  .00075958
      dem_sq | -.00007886   8.647e-06
        lpop | -.00003952   2.972e-06   .00018982
    perchpop |  4.210e-06   2.266e-07   6.411e-06    .0000345
       pcgnp |  .00003703  -5.106e-06  -1.511e-06  -3.391e-07   .00001911
        left |  .00041364  -.00003137  -.00004194    .0000393   .00001852   .00353416
        mil2 |  .00030409  -.00002182  -.00002945  -9.356e-06   .00003541    .0002826
        brit |  .00013588  -.00001438   .00005534  -.00001701   .00001422   .00066797
```

```
    iwarcow2 | -.00001095   5.336e-06  -.00017622  -.00001123  -.00003117  -.00071352
     cwarcow | -.00007418   6.684e-06  -.00011246  -.00003677   .00003662  -.00055995
       _cons |  1.450e-06   2.283e-06  -.00300902  -.00019571  -.00003885  -.00080128

        e(V) |       mil2        brit    iwarcow2     cwarcow       _cons
-------------+------------------------------------------------------------
        mil2 |  .00236202
        brit |  .00023095    .0019839
    iwarcow2 |  .00018976  -.00024433   .00451527
     cwarcow |  .00013788  -5.715e-07   .00031639    .0039678
       _cons | -.00094284  -.00177941   .00250417   .00138144   .05031347
```

(a) What is the expected value for Y assuming a non-leftist military regime that was a British colony, engaged in a civil war but not an international war holding the rest of the variables constant? Use the first model.

(b) Write down the marginal effect for the nonlinear democracy term at the mean. Calculate the standard error for this marginal effect and test whether it is statistically significant at the mean.

(c) Assume that democracy is the modifier variable in the interaction. Calculate the marginal effect of GDP at the mean, minimum, and maximum of democracy. Test whether each of these marginal effects is statistically significant.

(d) Assume that Model 1 is misspecified because of an omitted variable. Write down the proof for how this will affect the coefficient on the `left` variable. Under what conditions is it possible that an omitted variable will not affect the coefficient for leftist governments? What direction is the bias for the coefficient on `left`? Elaborate all possibilities.

(e) Provide an interpretation for the logged population variable from the first model.

5. Say you have a bivariate regression

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Assume X is measured with normal random measurement error. In other words, $X = X^* + U_i$, where $U_i \sim N(0, 1)$. Sketch a proof for the effect of the measurement error on the estimate of β_1. Does this proof generalize to the multivariate case?

6. Delineate the difference between leverage and influence. What combination of leverage and influence is required to bias OLS estimates? How is each measured? If these measures are problematic, what are some alternatives to OLS that are more robust to outliers?

7. Under heteroskedasticity, the least squares regression model is inefficient. An alternative is the use of generalized least squares (GLS). Derive the GLS estimator. What is the chief drawback of GLS?

8. Table 1 contains a logistic regression model of voter turnout as a function of education in years (1-15, mean 12), age (18-85, mean 47), weekly earnings (0-2884, mean 158), a dummy for females, and a dummy for blacks.

   - Are logistic regression estimates unbiased?
   - Calculate the change in predicted probabilities for a ten year increase in age from the mean for a black female with average education and earnings.
   - Why is it possible to calculate statistical significance here in the exact same way as for a linear regression model?

Table 1: Logit Model of Voter Turnout

Education	0.27
	(0.05)
Age	0.035
	(0.007)
Earnings	0.019
	(0.0002)
Female	0.02
	(0.22)
Black	0.15
	(0.43)
Constant	−12.03
	(2.12)

- Describe one method for calculating a confidence interval for a predicted probability. Why is such a method necessary?

9. Provide a formal description of overdispersion in the poisson regression model (PRM). If you estimate a PRM with overdispersed data, how will this affect the model estimates? How does one test for overdispersion? What other pathologies of count data can manifest as a positive test for overdispersion? What are the two major methods of interpretation for the PRM? Under what conditions are these two methods more or less useful?

10. What is the Independence of Irrelevant Alternatives assumption? Use conditional probabilities to sketch how it might be violated. What does it imply for models of unordered choices? What models might one use to model data when IIA is violated?

11. Models of civil war have become common in international relations. Table 2 contains the results from a Cox model with the outcome being the length of the civil war in years. This is modeled as a function of a series of dummy variables. The first is whether the war started from a popular revolution, whether the country is in Sub-Saharan Africa, whether the war is between an ethnic minority and majority, and finally whether the rebels receive financing from the drug trade. Use hazard ratios to interpret the popular revolutions and the drugs variables. What is the key diagnostic for Cox models and how is it remedied?

PART II. Answer three of the questions below.

1. Explain the logic behind selection models. Derive the Heckman selection model. What are some of the criticisms of this model?

2. For data with a multilevel structure explain the consequences of estimating a pooled, unpooled, and partially pooled (the random effects model) with this data. What is the best model?

3. Derive the simple random effects estimator for the mean. Explain the logic behind this estimator. How does its properties differ from the usual estimator for the mean?

Table 2: Cox Model of Civil War Duration

Popular Rev.	1.38
	(0.27)
SS Africa	1.22
	(0.34)
Ethnic War	−1.33
	(0.40)
Drugs	−1.06
	(0.43)

4. Outline the biprobit and seemingly unrelated regression models. Under what conditions would these models be necessary? How are these models related to Heckman models?

5. Demonstrate that models for panel data are identical to multi-level models. What other additional pathologies tend to be present with panel data models that are absent with cross-sectional multi-level models?

6. A classic game used to demonstrate a mixed strategy Nash equilibrium is called Rock-Paper-Scissors. In this game, two players simultaneously select one of the three moves. If both players select the same move, both receive utility of zero. If one selects Rock and the other plays Paper, the player selecting Rock receives utility -1 and the one playing Paper receives +1 (Paper covers Rock). If one selects Rock and the other plays Scissors, the one playing Rock receives +1 and the one playing Scissors receives -1 (Rock crushes Scissors). Finally, if one plays Paper and the other plays Scissors, the one playing Paper receives -1 and the one playing Scissors receives +1 (Scissors cuts Paper). The unique Nash equilibrium to this game is that both players play the mixed strategy 1/3 Rock, 1/3 Paper, 1/3 Scissors.

   In this problem, consider a variant on this classic game. In this variant, if one player plays Rock and the other plays Scissors, the Rock player receives +2 and the Scissors player receives -2 (Rock REALLY crushes Scissors!). The utilities from all other pairs of moves played against one another remain as before.

   (a) Draw the extensive form of this game.

   (b) Construct the normal form of this game.

   (c) Find all Nash equilibria of this game (consider both pure and mixed strategies).

   (d) Show that, for Player 1 playing the mixed strategy equilibrium derived in part (c), Player 2 receives the same expected utility no matter what strategy she plays. What is that expected utility?

   (e) In the mixed strategy Nash equilibrium to this game, is Rock or Paper or Scissors played with the greatest frequency? Why?