## Problem 1

### 1.a.

Transition probability matrix:

[[ 0.5 0.5 0. ]

[ 0.25 0.5 0.25]

[ 0. 0.5 0.5 ]]

Transition probabilities after 2 steps

[[ 0.375 0.5 0.125]

[ 0.25 0.5 0.25 ]

[ 0.125 0.5 0.375]]

Transition probabilities after 5 steps

[[ 0.26562 0.5 0.23438]

[ 0.25 0.5 0.25 ]

[ 0.23438 0.5 0.26562]]

Transition probabilities after 10 steps

[[ 0.25049 0.5 0.24951]

[ 0.25 0.5 0.25 ]

[ 0.24951 0.5 0.25049]]

Transition probabilities after 25 steps

[[ 0.25 0.5 0.25]

[ 0.25 0.5 0.25]

[ 0.25 0.5 0.25]]

### 1.b.

Transition probability matrix with absorbing states:

[[ 1. 0. 0. ]

[ 0.25 0.5 0.25]

[ 0. 0. 1. ]]

Transition probabilities after 2 steps

[[ 1. 0. 0. ]

[ 0.375 0.25 0.375]

[ 0. 0. 1. ]]

Transition probabilities after 5 steps

[[ 1. 0. 0. ]

[ 0.48438 0.03125 0.48438]

[ 0. 0. 1. ]]

Transition probabilities after 10 steps

[[ 1. 0. 0. ]

[ 0.49951 0.00098 0.49951]

[ 0. 0. 1. ]]

Transition probabilities after 25 steps

[[ 1. 0. 0. ]

[ 0.5 0. 0.5]

[ 0. 0. 1. ]]

If a rat starts in room B it will end up in room A with probability .5 and room C with probability .5.

### 1.c.

After 84 iterations the rat has is in Room A with probability over 95% regardless of starting state. By 125th iteration that probability is over 99% regardless of starting state. In the limit, the rat will always end up in room A.

Transition probability matrix:

[[ 1. 0. 0. 0. 0. ]

[ 0.25 0.5 0.25 0. 0. ]

[ 0. 0.25 0.5 0.25 0. ]

[ 0. 0. 0.25 0.5 0.25]

[ 0. 0. 0. 0.5 0.5 ]]

Transition probabilities after 84 iterations:

[[ 1. 0. 0. 0. 0. ]

[ 0.98153 0.00281 0.0052 0.00679 0.00367]

[ 0.96587 0.0052 0.0096 0.01255 0.00679]

[ 0.9554 0.00679 0.01255 0.01639 0.00887]

[ 0.95173 0.00735 0.01358 0.01774 0.0096 ]]

Transition probabilities after 125 iterations:

[[ 1. 0. 0. 0. 0. ]

[ 0.99624 0.00057 0.00106 0.00138 0.00075]

[ 0.99305 0.00106 0.00196 0.00256 0.00138]

[ 0.99091 0.00138 0.00256 0.00334 0.00181]

[ 0.99017 0.0015 0.00277 0.00361 0.00196]]

## Problem 2.

### a.

Transition probability matrix:

0 1

0 0.876158 0.123842

1 0.428688 0.571312

### b.

Logit Regression Results ==============================================================================

Dep. Variable: union No. Observations: 21766

Model: Logit Df Residuals: 21758

Method: MLE Df Model: 7

Date: Sat, 25 Oct 2014 Pseudo R-squ.: 0.1847

Time: 17:27:04 Log-Likelihood: -9462.5

converged: True LL-Null: -11606. LLR p-value: 0.000 ===============================================================================

coef std err z P>|z| [95.0% Conf. Int.]

-------------------------------------------------------------------------------

Intercept -2.2960 0.385 -5.959 0.000 -3.051 -1.541 prior\_union 2.1210 0.038 56.117 0.000 2.047 2.195

year -0.0114 0.007 -1.700 0.089 -0.024 0.002

age 0.0237 0.006 3.846 0.000 0.012 0.036

grade 0.0464 0.008 5.911 0.000 0.031 0.062

south -0.6630 0.042 -15.734 0.000 -0.746 -0.580

black 0.5797 0.043 13.404 0.000 0.495 0.664

smsa 0.0476 0.043 1.099 0.272 -0.037 0.132 ===============================================================================

### c., d., e.

Parts c. and d. do not ask for output. Steps shown in code file.

Markov transition matrix using the logistic model:

0.87142 0.12858

0.46298 0.53702

### f.

Steady state distribution:

0.78264 0.21736

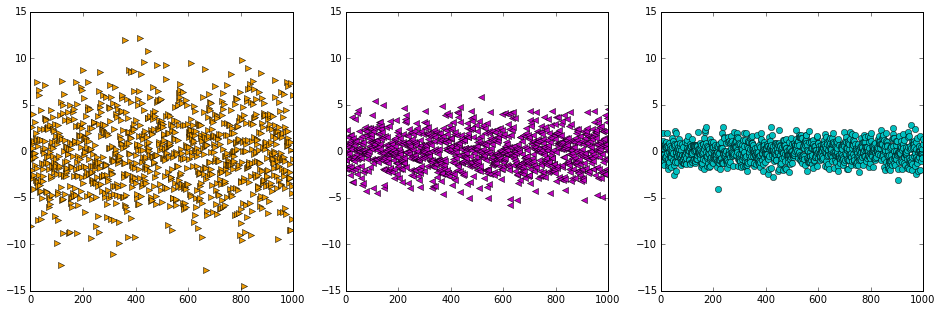
0.78264 0.21736

Data are probably available to compare the model’s prediction to real outcomes. During a short search I could not find a summary of NLSW data that came from the same or similar source as the data we are given. Unsummarized NLSW data are available up to year 2003, but I wasn’t confident I could aggregate and meaningfully summarize them to check prediction error.

## Problem 3.

### 3.a.

The three plots below are of random normal vectors with variances 4, 2, and 1.



### 3.b.

For this exercise I thought it would be fun to get a very low P-value, so I did not set seed and ran the code until I got the OLS with a P-value below 0.05. Naturally, it occurs about 5% of the time, so it didn’t take long.

OLS Regression Results

==============================================================================

Dep. Variable: y R-squared: 0.008

Model: OLS Adj. R-squared: 0.007

Method: Least Squares F-statistic: 8.435

Date: Sat, 25 Oct 2014 Prob (F-statistic): 0.00376

Time: 18:33:53 Log-Likelihood: -1460.1

No. Observations: 1000 AIC: 2924.

Df Residuals: 998 BIC: 2934.

Df Model: 1

==============================================================================

coef std err t P>|t| [95.0% Conf. Int.]

------------------------------------------------------------------------------

const -0.0014 0.033 -0.042 0.967 -0.066 0.063

x1 0.0960 0.033 2.904 0.004 0.031 0.161

==============================================================================

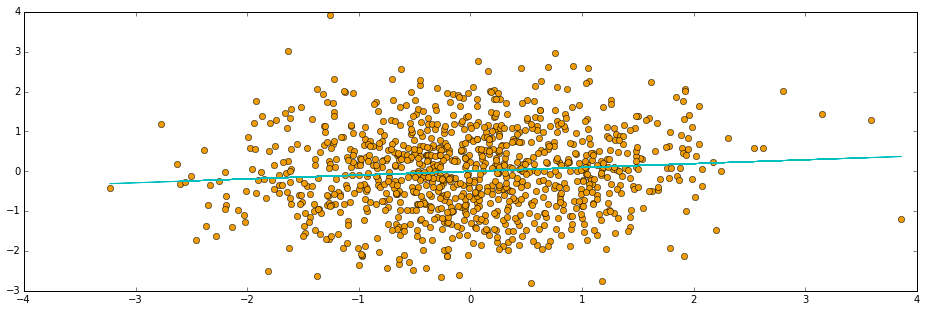
Omnibus: 2.243 Durbin-Watson: 2.051

Prob(Omnibus): 0.326 Jarque-Bera (JB): 2.315

Skew: 0.105 Prob(JB): 0.314

Kurtosis: 2.894 Cond. No. 1.02

==============================================================================



### 3.c.

Since all of the OLS conditions are met, the slope coefficients have a t-distribution. The histogram is consistent with a t-distribution.

