Research Methods in Political Science I, Homework Assignment 11

Due: 9pm on 12 January 2016 (**Tuesday**)

How to submit: Send as an email attachment

Email subject: Research Methods 1, Assignment 11

File names: `hw11-YourName.pdf`

Part I

Q1: Experiment 1. You have a coin with the probability of heads θ. You flipped it 10 times and got 7 heads. Answer the following questions.

1. Visualize the likelihood function of θ given the data.
2. Visualize the log-likelihood function of θ given the data.
3. Obtain the maximum likelihood estimate (MLE) of θ.
4. Calculate the observed Fisher information of θ.
5. Obtain the likelihood intervals that are equivalent to the 90 percent and 95 percent confidence intervals. Then, visually display the intervals with the normalized likelihood function. You may display two intervals in a single figure or create a figure for each.
6. Calculate the Wald statistic where the null hypothesis is $\theta = 0.5$. Conduct the statistical test using the 5 percent significance level.
7. Compare the inference based on the likelihood intervals with that on the Wald statistic (Wald intervals).

Q2: Experiment 2. You have a coin with the probability of heads θ. You flipped it 100 times and got 70 heads. Answer the following questions.

1. Visualize the likelihood function of θ given the data.
2. Visualize the log-likelihood function of θ given the data.
3. Obtain the MLE of θ.
4. Calculate the observed Fisher information of θ.
5. Obtain the likelihood intervals that are equivalent to the 90 percent and 95 percent confidence intervals. Then, visually display the intervals with the normalized likelihood function. You may display two intervals in a single figure or create a figure for each.
6. Calculate the Wald statistic where the null hypothesis is $\theta = 0.5$. Conduct the statistical test using the 5 percent significance level.
7. Compare the inference based on the likelihood intervals with that on the Wald statistic (Wald intervals).

Q3. Compare the experiments above. What are the differences between them? What caused the differences?

Part II

Q1. A random variable Y_i is independently distributed as follows.

$$\begin{aligned} Y_i &\sim \text{Bin}(n, \pi_i) \\ \pi_i &= \text{logit}^{-1}(\beta_1 + \beta_2 x_i) \end{aligned}$$

Using the data presented in Table 1, run logistic regression to estimate $\boldsymbol{\beta}$ and answer the following questions. You are allowed to use the function `glm()` only for No. 4. Up to No. 3, use the (log-)likelihood function you specified. You may use `maxLik()` (or other optimization functions) to find maxima.

Table 1: Observed Data

x_i	n_i	y_i
5	8	0
10	5	0
15	8	4
20	10	2
25	10	7
30	6	3
35	3	1
40	5	5
45	4	4
50	4	4

1. Visualize the log-likelihood function of θ given the data. Fixing the value of β_1, plot the log-likelihood versus β_2.
2. Obtain the MLE.
3. Calculate AIC.
4. Get the MLE and AIC using `glm()` and compare the results. (Hint: you can either (1) deconstruct groups into Bernoulli trials or (2) fit logistic regression to grouped data.)

Q2. Using the dataset you gathered for Homework Assignment 9, conduct logistic regression analysis. Compare multiple models if necessary.

1. Without `glm()`, define the (log-)likelihood function and obtain the MLE and AIC.
2. With `glm()`, obtain the MEL and AIC and compare the results to those you got for No.1.
3. Assuming that the response were 1 for observations with the predicted probability equal to or larger than 0.5, calculate the rate of correct prediction.
4. Draw an ROC curve and evaluate the fit.