AMS 550.400 HW SET 1 Due Date: Oct 8

Last Compiled on October 29, 2012

General Instruction: To complete the homework set, you are required to do the followings. Your solutions must be typed in LaTeX using the course homework template. The progression of your homework solution is to be "recorded" by making a git folder specifically for this homework set. The burden of proof is on you, and if your git commit history is sparse, then you may be liable for a penalty. A paper copy of the PDF output of your LaTeX file is to be submitted to your instructor in class on the due date. *After* submitting the paper copy, but *before* the end of the due date, you will upload your work to your github by making a remote repository specifically for the homework, and post the link to the repository at the designated *Discussion* forum in Blackboard by making a thread just for you. The repository name in your github should be `550400.homeworkset.1` and the discussion forum thread should be named `YourFirstNameMiddleInitialLastName`, e.g., `BaracHObama` and `WillardMRommey`. You have till the end of the due date to finalize your github repository. However, any commit made after the class time of the due date will be inadmissible. *Your attention to details in following this instruction will be critical, and if not followed exactly at the time of collection, the homework set may be graded at* 90% *of the full score.*

Problem 4 (aka. Fair Play, 40 pts): Answer the following question:

> Is the tennis game fair?

Note that unlike Problem 3, this question is vaguely stated. This is intensional, whence to begin, you will first need to clarify what exactly your question is. You may use the class discussion on this particular problem, but you *may not* directly refer to our discussion. Instead, formulate the model carefully but concisely in your own words.

Problem 4 (aka. Fair Play, Solution):

To answer this question we need to restate the problem more clearly. Question asked here is whether the game is fair or not. In this question, one can ask what is fair. In the tennis game, fair game means the rule of the game should not favor toward any player. In mathematical term, each player has equal chance of winning.

We begin to analysis the problem by studying the rules of tennis. One player severe first, then opponents serve. The first to win 6 games, by two, wins the set. The first to win 2 sets wins the match. If the score is 6-6, a tie-breaker is played.The first team to score 7 points winning by two wins the set. The tiebreaker continues until one side wins by two.

In the beginning of the game, a toss is used to decide who serve first and after each games, server and receiver switch. Obviously, the server has the advantage. It seems fair because server and receiver are switched after each game. Both players have equal amount of serve and receiver times. To check whether tennis is fair games is equal to testing whether if the player who choose to server first has a better chance of winning. So formulate the problem statement: In the tennis games, if a player serve first, does he has a better chance of winning or his probability of winning is increased.

In this model, we want to investigate the correlation between the winning of the game and order of the player to server in the game. There are many other elements that can affect the results of the game. For example: weather, skills of the player, physical strength...etc. Since both players are subject to the same condition, these variables are exogenous variables. The endogenous variables here are the order of the player to server in the game and the results of the game.Variable results of the game will substitute the variable probability of the win. Because the probability of the win is difficulty to obtain and results of the game is a direct measure and a reasonable estimator.

If the probability of winning is 0.5, then we will conclude the tennis is a fair game. To conduct the experiments, we need to collect enough tennis results data including who server first in the game. To exclude the effect from the exogenous variables, the game data collected needs certain restriction. For example: Both player need to have same level of skills, same level of physical strength. Also the number of the game data required is determined by the sample size calculation. If the percentage of win equal to the percentage of lose in the game which the player served first, then we will conclude the tennis is a fair game.

Since the restriction on the datasets made it impossible to obtain real life data. In order to test the model, we will test the model in the computer simulation program. In the simulation program, the physical strength, tennis skills level can be controlled accurately. Since cost of obtain a virtual tennis game results is cheap, we can have very large sample size. Because from the large sample, we can have better confidence interval for our hypothesis.

As mentioned in the above paragraph, computer simulation is the only method to obtain the data that fits the requirement of excluding the exogenous variables. Without the real life data, the reliability of the model is still questionable. The solution will be combine this model with other model that deals the exogenous variables. For example: there's one model that is able to predict the probability of the win from the physical strength of the player and the skills level of the player. We will combine our model into this model. The new combined model has less restriction on the datasets. Then we can test our model's reliability with real life data.