

<b>Course website</b>	<a href="http://www.stat.wisc.edu/courses/st572-larget/">http://www.stat.wisc.edu/courses/st572-larget/</a>
<b>TA</b>	Xu He(Mario)
<b>Email</b>	hexu@stat.wisc.edu
<b>Website of TA</b>	<a href="http://www.stat.wisc.edu/~hexu">www.stat.wisc.edu/~hexu</a>
<b>Office</b>	1275F, MSC
<b>Office hour</b>	Th 11:00-12:00, 4:00-6:00
<b>Phone</b>	(608)334-9792

## 1 A Simulation problem for a game

It's a sort of gambling game. The players should pay \$2 first, and then pull 5 balls out of a bag. The bag contains 10 balls, all identical except that 5 are red, 5 are green. The player are not allowed to see the bag while picking balls, i.e., balls will be picked randomly without replacement. The player will be rewarded some money right after the game if certain number of red balls were picked.

0 or 5 red balls: \$ 100 prize  
1 or 4 red balls: \$ 5 prize(actually \$3 gain)  
2 to 3 red balls: no prize

- Write a function to generate 5 balls automatically. Use that function to pick balls for 3 times. Also, write a function to generate the dollars gain from a single game.
- If we play the game for sufficiently many times. Will we win in the last? Answer this question by simulate the game for 10000 times, and calculate the mean of gain or lose.
- If we only play once, what's the possibility we win? What if playing 10 times? Answer this question by simulate the game for 10000 times and record how many times we shall win.
- Continue with (c), what's the 5%, 15%, ... , 95% quantile of dollars gain if we play 10 times? Draw a count histogram to see the distribution.
- Extra problem: What is the possibility we get 4 red balls in a single game?

## 2 Simulation for coefficients from a linear model

The following data is collected to study the relationship between frequency of chirps made by a ground cricket and the corresponding ground temperature. The frequency is measured as the number of wing vibrations per second.

chirps (/second)	20.0	16.0	19.8	18.4	17.1	15.5	14.7	17.1	15.4	16.2	15.0	17.2	16.0	17.0	14.1
temperature (F)	88.6	71.6	93.3	84.3	80.6	75.2	69.7	82.0	69.4	83.3	78.6	82.6	80.6	83.5	76.3

- Fit the data by a simple linear model. Another model would be  $chirps \sim 1$ . Which model is better? Do an ANOVA test and report the null and alternative hypotheses and p-value.
- Simulate 1000 reasonable values of  $\beta$  and  $\sigma$  from the model  $chirps \sim temp$ . Find the 95% confidence interval for the slope of the line. Compare this interval with one you can find from `display()`.
- Do 4 groups of simulations. Each group contains 10 simulations. Each simulation is of size 10, 100, 1000 or 10000. I.e., take ten simulations of size 10, ten for size 100, ten for size 1000, ten for size 10000. For each group of simulations, calculate the mean and standard error of the standard error of those 10, 100, 1000, or 10000 simulations for coefficient of the slope. What's the proper simulation size so that we get a good simulation of coefficient of the slope?
- Find 0.01 quantile of the distribution for coefficient corresponding to temp. What simulation size do we need to have a standard smaller than 0.002?
- Have a 95% prediction interval for the chirp of temperature 70F. compare this interval with the one you can find using `predict()`.