National Taiwan University
Institute of Industrial Engineering
Jakey BLUE

Data Analytics

109-2 Homework #02

Due at 23h59, March 7, 2021; files uploaded to NTU-COOL

1. (15%) Simulate the averages of $[2, 3, 4, 5]$ dices for 1000 times. Draw the four histograms for the sample averages of $[2, 3, 4, 5]$ dice, respectively.
   *Reproduce the CLT results on p. 22 of DA01 slides.

2. To validate the Kruskal's count on p. 6 of DA01 slides, we play the game with one deck of cards, i.e., 52 cards, for 10000 times. Each time, the 52 cards are randomly shuffled. We then start from the first 10 cards, and the face cards (J/Q/K) are counted as 5 steps.
   a. (15%) What is the probability that all the first 10 cards reach the same end?
   b. (15%) Vary the simulation settings:
      the # of cards: $[52, 104]$;
      the # of steps for face cards $= [1, 3, 5, 7, 9]$.
      What are the $2 \times 5 = 10$ probabilities? Discuss your observation?