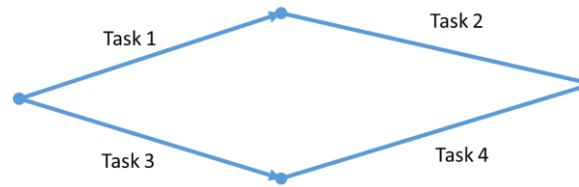


1. Consider the following workflow.



Task 1 follows a random uniform distribution between 5 and 9; the completion time for Task 2 follows a random normal distribution with mean=7 and sd =2; Task 3 follows an exponential distribution with rate = 8; Task 4 follows uniform distribution between 3 and 10. Compute the following.

- What are the mean and median times to complete all the tasks?
- What is the probability that all the tasks get completed in 12 hours?
- Create a plot of the density of the total completion time.

2. Suppose you just purchased 4 stocks.

Stock	Quantity Purchased	Purchase Price (per unit of stock)	Expected return
A	30	95	N(mean=10%,sd=5%)
B	15	100	U[-10%,15%]
C	25	25	U[-15%,25%]
D	50	50	N(mean=20%,sd=30%)

- What is the total expected profit on your investment?
- What is the median and standard deviation of your total expected profit?
- Suppose the transaction fee for purchasing these stocks is \$100. What is probability that you don't lose money? What is the probability you will make at least \$1500 profit?

3. Suppose you are testing two versions of a web homepage design with different features. Some visitors were randomly shown Version A and others were shown Version B. The table below shows the number of visits to each version and the number of 'completions'. A completion indicates someone completing the purchase process at the site. Our object of interest is 'conversion rate' – percentage of visits that result in completion. Which version of the homepage would you recommend rolling out? Explain your rationale and the process you followed.

		Homepage Version A		Homepage version B	
		Visits	Completion	Visits	Completion
	07-6-4	7,823	796	2,910	289
	07-6-5	5,611	541	3,049	262
	07-6-6	5,092	533	2,775	298
	07-6-7	16,407	1,001	3,266	191
	07-6-8	4,072	416	1,980	188
	07-6-9	2,802	268	1,512	129
	07-6-10	3,277	323	1,408	134
	07-6-11	8,159	808	2,709	258
	07-6-12	5,331	517	2,802	258
	07-6-13	5,217	542	2,720	272
	07-6-14	15,922	1,099	3,119	205
	07-6-15	4,360	415	2,091	182

4. Read the file Health.csv. There is a variable called ofp which indicates the number of office visits. Assuming this follows a poisson distribution, test the hypothesis that the mean number of office visits is 5.5.

Health\$ofp

```
[1] 5 1 13 16 3 17 9 3 1 0 0 44 2 1 19 19 0 3 2 12 2 3 1 1
5 1 12 6 2 4 2 2 1 0 1 4 26 3 1 1
[41] 1 3 0 17 5 5 9 2 1 9 11 3 1 1 0 3 2 4 3 5 5 1 4 0
8 8 1 5 0 4 3 11 13 0 7 5 3 2 22 6
```

5. Read the file Walmart.csv

head(walmart)

```
Store Dept Date weekly_Sales IsHoliday
1 1 1 2/5/2010 24924.50 FALSE
2 1 1 2/12/2010 46039.49 TRUE
3 1 1 2/19/2010 41595.55 FALSE
4 1 1 2/26/2010 19403.54 FALSE
5 1 1 3/5/2010 21827.90 FALSE
6 1 1 3/12/2010 21043.39 FALSE
```

Create 95% confidence intervals for the mean, median and standard deviation of the weekly sales of store 1 during non-holiday weeks. Create confidence intervals (a) assuming normality, and (b) using bootstrapping.

6. A customer complains to the owner of an independent fast-food restaurant that the restaurant is discriminating against the elderly. The customer claims that people 60 years old and older are given fewer french fries than people under 60. The owner responds by gathering data, collected without the knowledge of the employees so as not to affect their behavior. Here are data on the weight of french fries (grams) for the two groups of customers:

Age less than 60: 75 77 80 69 73 76 78 74 75 81 75 80

Age greater than 60: 68 74 77 71 73 75 80 77 78 72 69 71

Conduct a two sample test to evaluate the complaint.

**(A significant portion of your grade will depend on adequate commenting of the code. Make sure to include your name in the code as a comment).**

**(Please read: By submitting the assignment, you are confirming that this represents your own work, and that no part of the work submitted has been copied from others. Consequences for cheating can be quite serious. It can lead to not only failing the course but also dismissal from the program)**