MSBA7003 Quantitative Analysis Methods



03 Monte Carlo Simulation & Dynamic Decisions

Agenda

- Monte Carlo Simulation
- Risk Assessment
- Inventory Management
- Revenue Management
- Service Management



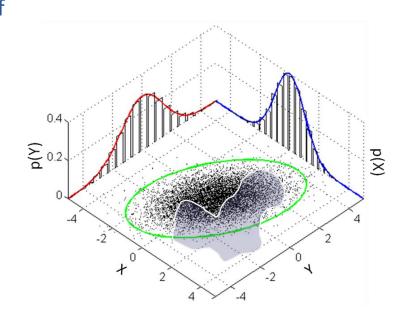
Monte Carlo Simulation

- In practice, we often make decisions under risk for very complicated problems.
 - Too many random variables involved
 - Too many or even infinitely many possible states
 - Objective functions are highly nonlinear
 - Cannot be explicitly solved
- Simulation is to use a computer to duplicate the features, appearance, and characteristics of a real system, and learn the outcome of a decision.

Monte Carlo Simulation

 Monte Carlo methods rely on repeated random sampling to obtain numerical results.





Simulation-Based Decision Making: Single Decision

• Mathematically, if the objective function V(X,D) depends on the value of a random variable (vector) X and a decision (vector) D, then the expected objective value given D can be approximated by randomly sample X independently N times and computing the average:

$$\frac{1}{N} \sum_{i=1}^{N} V(X_i, D)$$

ullet Then we can try different D to maximize or minimize the objective.

	Sampli				
Options	Sample (1)	Sample (2)	Sample (3)	 Sample (N)	Average Objective Value
Option (A)	VA1	VA2	VA3	 VAN	(VA1+VA2+VA3+ +VAN)/N
Option (B)	VB1	VB2	VB3	 VBN	(VB1+VB2+VB3+ +VBN)/N
Option (C)	VC1	VC2	VC3	 VCN	(VC1+VC2+VC3+ +VCN)/N

Simulation-Based Decision Making: Multiple Decisions

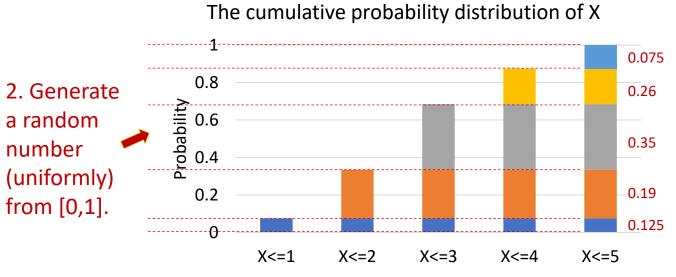
- When there are multiple stages, our objective is to maximize or minimize $\sum_t V_t(X_t, D_t)$ by choosing a set of decisions $\{D_t\}$. A feasible way to obtain an approximate solution is to find a policy $\Pi: (t, X_t) \to D_t$.
- We can evaluate Π through Monte Carlo simulation.

Fix ∏	Sample the	value of X_t at				
Simulation	Stage 1	Stage 2	Stage 3	•••	Stage k	Simulated Objective
Round 1	V11	V12	V13	•••	V1k	V1=V11+V12+V13+ +V1k
Round 2	V21	V22	V23	•••	V2k	V2=V21+V22+V23+ +V2k
•••	•••	•••	•••	•••	•••	•••
Round N	VN1	VN2	VN3	•••	VNk	VN=VN1+VN2+VN3+ +VNk
					Average	(V1+V2+ +VN)/N

• Then we can try different Π to maximize or minimize the objective.

Generating a Random Number

• Random sampling is a critical step. Suppose the model is ready (i.e., the distribution of X is known), how to sample the value of X?



3. Return the value of X corresponding to the segment that contains the random number.

1. Calculate the cumulative probability for each possible value of X and divide the [0,1] interval into segments corresponding to each possible value of X.

Generating a Random Number

- For discrete random variables
 - Step 1: Build a cumulative distribution table with $X_i < X_{i+1}$.
 - Step 2: Assign the interval $(0, F(X_1)]$ to X_1 and $(F(X_{i-1}), F(X_i)]$ to X_i from i=2.
 - ullet Step 3: generate a random probability and find the corresponding X value.
- Generating a discrete random number in Excel:
 - Binomial distribution: BINOM.INV(n, p_s, probability)
 - An arbitrary distribution: VLOOKUP(probability, distribution table, column index for the return value)

Harry's Auto Tire

• Harry's Auto Tire sells a popular radial tire and replenish its inventory every 10 days. If the daily demand independently follows the distribution shown in the table below and the initial inventory level is 40 units, what is the expected sales in a 10-day period?

Daily Demand	0	1	2	3	4	5
Probability	0.05	0.10	0.20	0.30	0.20	0.15

• Note that 10-day sales = min(10-day demand, total inventory).

Harry's Auto Tire

First, build a cumulative distribution table.

Next, suppose we decompose the simulation of daily demand into two steps: random number & vlookup

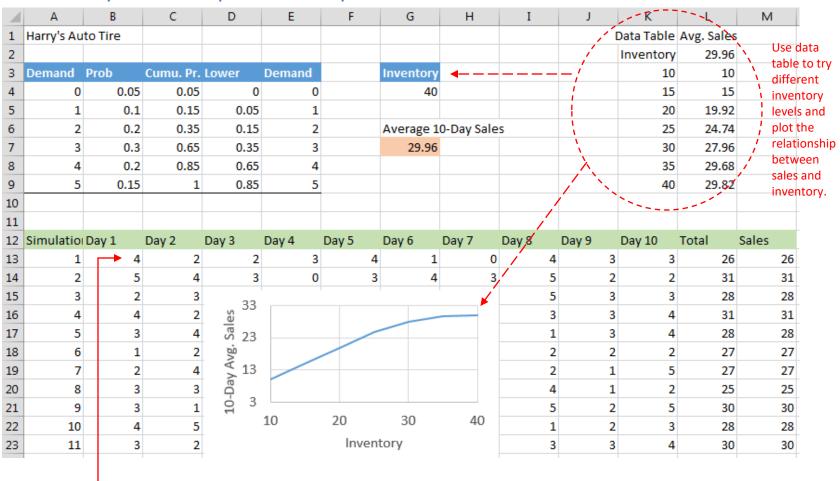
А	В	С	D	Е
Harry's Auto Ti	re			
		г		
Daily Demand	Probability	Cumulative Pr	Lower	Daily Demand
0	0.05	0.05	0	0
1	0.1	0.15	0.05	1
2	0.2	⟨.0.35	0.15	> _ 2
3	0.3	0.65	0.35	3
4	0.2	0.85	0.65	4
5	0.15	ز	0.85	5
Simulation				
Days	Random Numb	Demand 🖊		
Day 1	0.283855951	2	=vlookup(B13,	\$D\$4:\$E\$9,2)
Day 2	0.351300493	3		
	Daily Demand 0 1 2 3 4 5 Simulation Days Day 1	Harry's Auto Tire Daily Demand Probability 0 0.05 1 0.1 2 0.2 3 0.3 4 0.2 5 0.15 Simulation Random Number Day 1 0.283855951	Daily Demand Probability Cumulative Probability 0 0.05 0.05 1 0.1 0.15 2 0.2 0.35 3 0.3 0.65 4 0.2 0.85 5 0.15 1 Simulation Days Random Numb Demand ✓ Day 1 0.283855951 2	Daily Demand Probability Cumulative Process 0

The VLOOKUP function looks up the random number in the leftmost column of the defined lookup table. It moves downward through this column until it finds a cell that is equal to or bigger than the random number. It then goes to the previous row and gets the value from the specified column.

Harry's Auto Tire

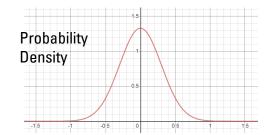
A more compact & comprehensive spreadsheet model:

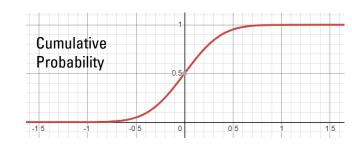
=VLOOKUP(RAND(),\$D\$4:\$E\$9,2,1)



Generating a Random Number

- For continuous random variables, the idea is the same
 - Step 1: derive the inverse function of Y = F(X)
 - E.g., for exponential distribution with parameter λ , the cumulative distribution function (CDF) is $Y = F(X) = 1 e^{-\lambda X}$ and the inverse CDF is $X = F^{-1}(Y) = -\lambda^{-1} \cdot \ln(1 Y)$.
 - Step 2: generate a random probability (i.e., a random number uniformly from [0,1]) and assign it to Y to get X.





- Generating a continuous random number in Excel:
 - Normal distribution: NORM.INV(probability, mean, st_dev)
 - Standard Normal distribution: NORM.S.INV(probability)
 - Uniform distribution on [0, 1]: RAND()

Generating a Random Number in Python

- In Python, we can use numpy to generate various kinds of random numbers.
- from numpy import random
- Normal distribution: random.normal(mean, std, size=None), wherein mean and std could be vectors.
- [0,1] Uniform distribution: random.rand(size=None).
- Random integers: random.randint(low, high, size=None).
- Binomial distribution: random.binomial(n, p, size=None).
- Exponential distribution: random.exponential(mean, size=None).

Generating a Random Number in Python

For an arbitrary discrete distribution:

```
from numpy import random
dist_table = \{0:[4,0.05], 1:[5,0.2], 2:[6,0.45], 3:[7,0.65], 4:[8,0.8], 5:[9,0.9], 6:[10,1.0]\}
def randnum(d_table, size=None):
  if size==None or size==1:
     temp=random.rand()
     for j in range(len(dist_table)):
        if dist_table[j][1]>temp:
          pointer=j
          break
     return dist_table[pointer][0]
  else:
     numbers=[]
     temp=random.rand(size)
     for i in range(size):
        for j in range(len(dist_table)):
          if dist_table[j][1]>temp[i]:
             pointer=i
             break
        numbers.append(dist_table[pointer][0])
     return numbers
```

Potential Problems in the Data

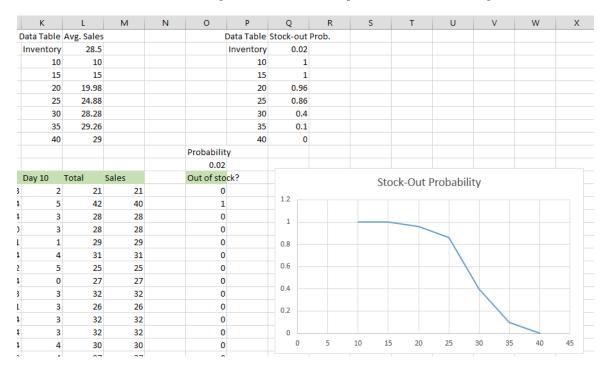
- We may encounter two types of problems when we use historical data to build simulation models.
- I. Data truncation.
 - The bias caused by incomplete observation. For example, the observed daily sales quantity of an ice-cream may be less than the true demand due to stockouts. For another example, the observed arrival rate of hourly parking customers may be less than the true demand rate due to the carpark closure when it is full.
- II. Endogeneity.
 - The bias caused by non-random decisions. For example, apparel shops tend to mark down their prices during high traffic periods, which could create an illusion of high price sensitivity among the customers. For another example, automakers and consumer electronics manufacturers would like to cyclically introduce new products, which creates a cyclic pattern of demand.

In-Class Exercises

• A jewellery store sells a necklace at the price of \$1,100. The number of visits to the store per day follows a binomial distribution with a mean of 300 and maximum value of 1,000. Among the arriving customers, one out ten will show interests in this necklace. Customer willingness-to-pay for the product follows a normal distribution with a mean of \$1,000 and a standard deviation of \$300. What is the average daily sales quantity of this necklace?

Probability Estimation

- Monte Carlo method can not only assess the average performance of a decision but can also assess the probability of an event.
- For Harry's Auto Tire, what is the probability of running out of stock?



- The owner of Simkin's hardware store wants to find a good, low cost inventory policy for an electric drill (Ace Model). Because all the sales and inventory are recorded in the computer and the supplier accepts orders at any time, so continuous inventory monitoring is adopted.
- How to optimize the order quantity (Q) and reorder point (r), given the distributions of daily demand and reorder lead time below?

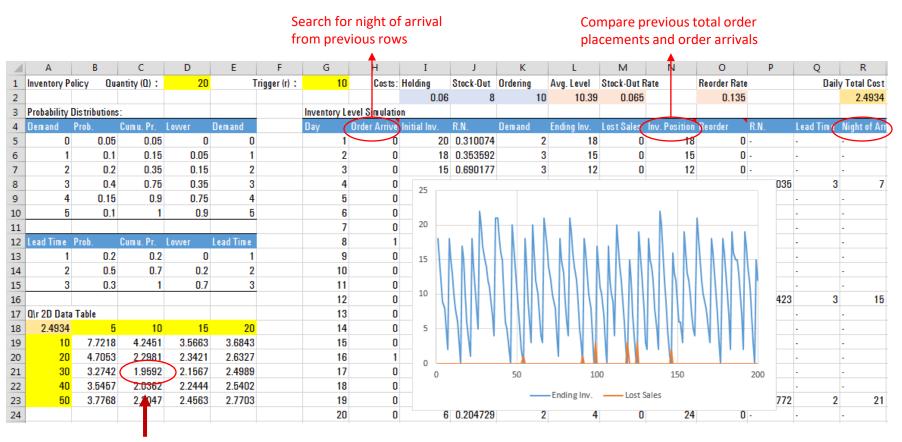
Daily Demand	Frequency (Days)	Prob.	Cumu. Prob.
0	15	0.05	0.05
1	30	0.10	0.15
2	60	0.20	0.35
3	120	0.40	0.75
4	45	0.15	0.90
5	30	0.10	1.00
	300	1.00	

Lead Time (Days)	Frequency (Orders)	Prob.	Cumu. Prob.
1	10	0.20	0.20
2	25	0.50	0.70
3	15	0.30	1.00
	50	1.00	

• Suppose Simkin's wants to test the policy of **Q=10 and r=5**. The process is simulated manually for a 10-day period.

DAY	UNITS RECEIVED	BEGINNING INVENTORY	RANDOM NUMBER	DEMAND	ENDING INVENTORY	LOST SALES	ORDER	RANDOM NUMBER	LEAD TIME
1	0	10	0.06	1	9	0	No		
2	0	9	0.63	3	6	0	No		
3	0	6	0.57	3	3	0	Yes	0.02	1
4	0	3	0.94	5	0	2	No		
5	10	10	0.52	3	7	0	No		
6	0	7	0.69	3	4	0	Yes	0.33	2
7	0	4	0.32	2	2	0	No		
8	0	2	0.30	2	0	0	No		
9	10	10	0.48	3	7	0	No		
10	0	7	0.88	4	3	0	Yes	0.14	1
				-	Total 41	2			

- The objective is to find a low-cost solution so Simkin's must determine the costs associated with carrying inventory, lost sales, and ordering cost.
- Simkin's store is open 200 days a year, and the holding cost is \$12 per drill per year. The estimated ordering cost is \$10 per order. Lost sales cost \$8 per unit.
- The average (ending) inventory level = 41/10 = 4.1 units per day
- The rate of lost sales = 2/10 = 0.2 units per day
- The rate of orders = 3/10 = 0.3 times per day
- The average daily cost = 4.1*12/200 + 0.2*8 + 0.3*10 = 4.85



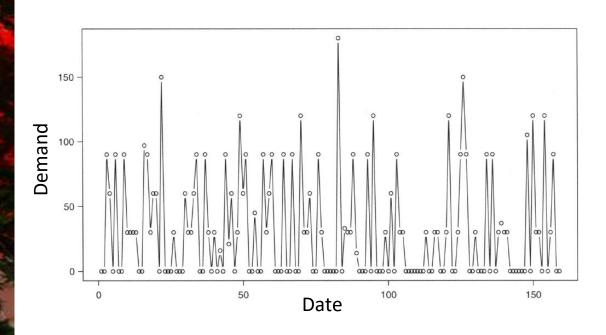
Use 2D data table to try different (r,0) combination and find the optimal policy.

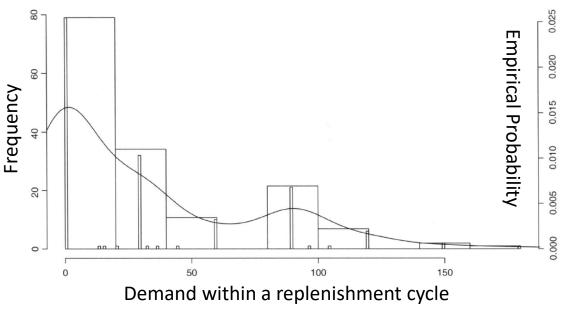
- Company Overview:
- 1,950 in-store pharmacies
- Each stocks about 2 to 3 thousand drugs
- Periodic inventory review: (s,S) policy
- Each pharmacy faces unique demand

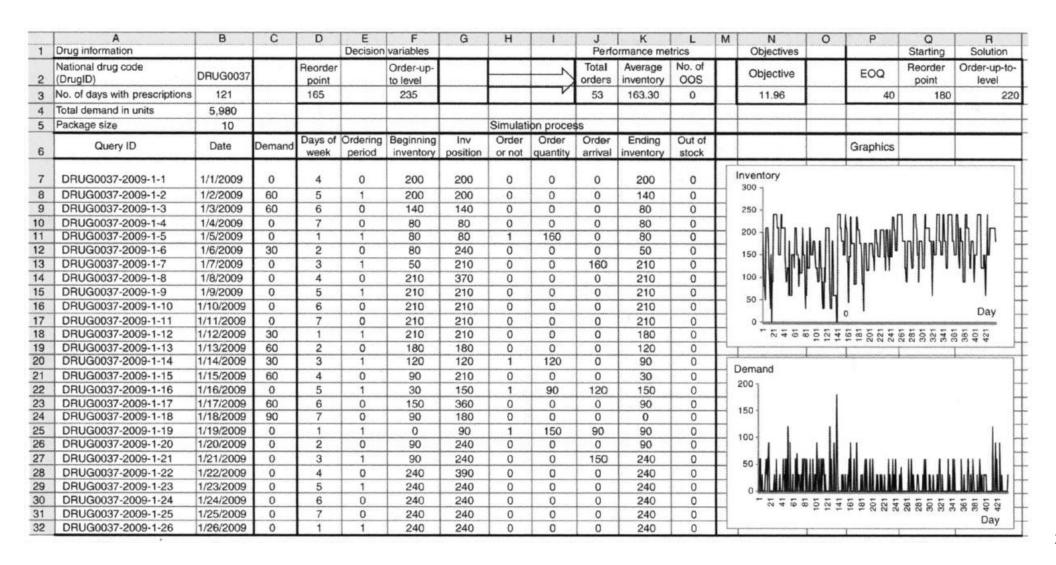


- The Challenge of Inventory Management:
- The demand distribution is highly irregular and cannot be captured by standard distributions; complicated models are not acceptable

- The historical demand data of a medicine sold at a certain pharmacy store.
- The basic dose is one pill per day and the prescriptions are often written for 30 days.







- Important notes in practice:
 - Check whether lost sales (or the true demand) were recorded.
 - Check whether demand is stationary across different replenishment cycles.
 - Historical sales were affected by various factors, which may be under different conditions currently.
- Possible Solutions:
 - Check whether stock-outs were solely caused by low inventory levels; if so, data points with stock-outs can be dropped.
 - Adopt different ordering policies for different replenishment cycles.
 - Use multiple linear regression to model the demand and perform simulation.

K-Fashion Revisited

- There are two prices to make. One is the original tag price, and the other is the markdown price for the third month. The store traffic is smaller in the first two months. According to data, the monthly visit numbers follow a normal distribution with a mean of 4,000 and a standard deviation of 800. In the third month, the number of visit also follows a normal distribution with a mean of 8,000 and a standard deviation of 1,500.
- According to experience, 1/30 of the arriving customers will show interests in this style (e.g., asking about the price and/or trying on). These customers only like one of the three SKUs, and the chance is 1/3 for each SKU. The customer willingness-to-pay for this style in the first two months follows a uniform distribution on [\$0, \$1,000], and in the last month the range becomes [\$0, \$600].

K-Fashion Revisited

• Consider the tag price \$869 and the markdown policy:

$$p = 599 - 8\sum_{i=A,B,C} n_i - 24/\left(1 + \sum_{i=A,B,C} l_i\right)$$

- n_i is the beginning inventory level of SKU i in the third month;
- $I_i = 0$ or 1, which indicates whether the SKU is available in the third month.

Month	Beginning Inventory A	Beginning Inventory B	Beginning Inventory C	Price	Customer Arrival	Interested	Target A	Target B	Target C	Buy A	Buy B	Buy C	Sales Value
1	10	10	10	869	3059	129	36	52	41	9	5	3	14773
2	1	5	7	869	4313	159	59	51	49	1	4	2	6083
3	0	1	5	543	8386	268	91	84	93	0	1	5	3258
1	10	10	10	869	5263	179	60	57	62	6	8	9	19987

Service Management

- When customers obtain utility by occupying and using firm resources, services take place.
- A critical decision in service management is to optimize the service capacity, which determines the cost and the level of congestion.



Port of New Orleans



- Fully loaded barges arrive at night for unloading
- The number of barges each night varies from 0 5
- The number of barges vary from day to day
- Barges are unloaded first-in, first-out
- The unloading rate also varies from day to day
- Barges must wait for unloading, which is expensive
- The dock manager wants to do a simulation study to enable him to make better staffing decisions







Port of New Orleans

• Historical data suggests the following distributions:

NUMBER OF ARRIVALS	PROBABILITY	CUMULATIVE PROBABILITY
0	0.13	0.13
1	0.17	0.30
2	0.15	0.45
3	0.25	0.70
4	0.20	0.90
5	0.10	1.00

DAILY UNLOADING RATE	PROBABILITY	CUMULATIVE PROBABILITY
1	0.05	0.05
2	0.15	0.20
3	0.50	0.70
4	0.20	0.90
5	0.10	1.00

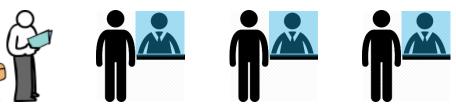
Port of New Orleans

DAY	NUMBER DELAYED FROM PREVIOUS DAY	RANDOM NUMBER	NUMBER OF NIGHTLY ARRIVALS	TOTAL TO BE UNLOADED	RANDOM NUMBER	Unloading Rate	NUMBER UNLOADED
1	_	0.52	3	3	0.37	3	3
2	0	0.06	0	0	0.63	3	0
3	0	0.50	3	3	0.28	3	3
4	0	0.88	4	4	0.02	1	1
5	3	0.53	3	6	0.74	4	4
6	2	0.30	1	3	0.35	3	3
7	0	0.10	0	0	0.24	3	0
8	0	0.47	3	3	0.03	1	1
9	2	0.99	5	7	0.29	3	3
10	4	0.37	2	6	0.60	3	3
11	3	0.66	3	6	0.74	4	4
12	2	0.91	5	7	0.85	4	4
13	3	0.35	2	5	0.90	4	4
14	1	0.32	2	3	0.73	4	3
15	0	0.00	5	5	0.59	3	3
	20		41				39
	Total delays		Total arrivals				Total unloaded

Pinevalley Bank

 The X branch of Pinevalley Bank has four service counters, but sometimes it is not necessary to open all four counters. Historical data shows that customers arrive every 10 minutes on average between 2 to 3 p.m. on a Thursday afternoon. The inter-arrival time follows exponential distribution. The service time for each customer also follows exponential distribution with a mean of 5 minutes. The question is, in order to make sure the average waiting time for each customer is less than 3 minutes, what is the minimum number of counters needed during this period of time?











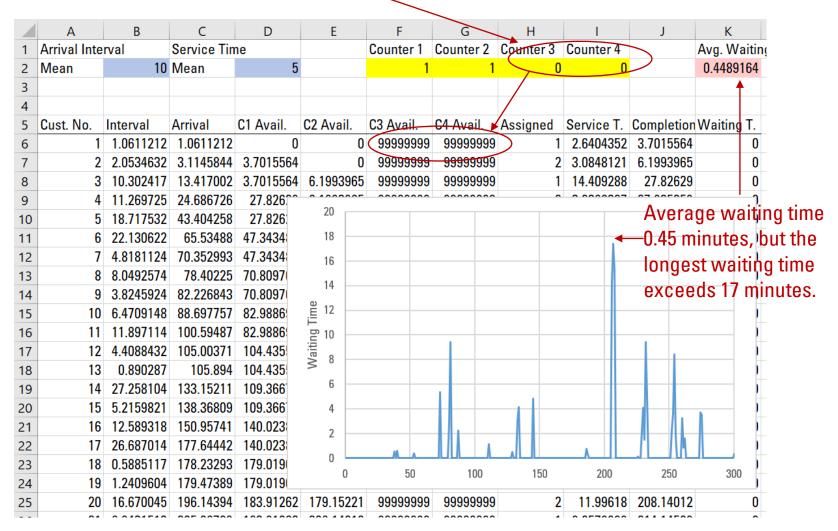
Pinevalley Bank

• We first manually simulate the process of customer arrival and service (assuming that two counters are open):

Customer No.	Arrival Interval (Minute)	Arrival Time (Minute)	Counter 1 Available From (Minute)	Counter 2 Available From (Minute)	Assigned Counter	Service Time (Minute)	Service Complete (Minute)	Waiting Time (Minute)
1	6	6	0	0	1	8	14	0
2	9	15	14	0	2	3	18	0
3	1	16	14	18	1	7	23	0
4	1	17	23	18	2	4	22	1
5	7	24	23	22	2	3	27	0
6	12	36	23	27	1	4	40	0

Pinevalley Bank

Counters not available will be available at an infinitely far time point.



Verification and Validation

- It is important that a simulation model be checked to see that it is working properly and providing good representation of the real-world situation
- <u>Verification</u> process involves determining that the computer model is internally consistent and following the logic of the conceptual model
 - Verification answers the question "Did we build the model right?"
- <u>Validation</u> is the process of comparing a simulation model to the real system it represents to make sure it is accurate
 - Validation answers the question "Did we build the right model?"

After-Class Exercises

True or False?

- Ken is good at shooting three-pointers. For each shot, the chance of success is 0.92. To simulate the total score after 10 shots, in total 10 random numbers should be generated.
- We want to simulate the operations of a restaurant. Assume customer dining time per table follows an exponential distribution with a mean of 1 hour. Currently, there are 20 tables of dining customers. Among them, the number of customers still eating after half an hour can be simulated by a binomial distribution.
- For Pinevally Bank, suppose only one window is always open and the second window is open only when the queue length is equal or longer than 3 people. Will the average waiting time for the customers satisfy the requirement?