1. ( 25 pts / 5 pts each ) Simulation Concepts: Provide short answer (a few sentences or a brief paragraph) addressing 1-5.
   1. Briefly discuss the distinction between a real system, a model, and a simulation.

A real system is a collection or composition of real-world processes that work in conjunction with or parallel to one another to achieve a goal or complete a task. A model is a simplified representation of a real system. With the aim to distill key features of a real system that will be approximated as closely as possible while balancing simplicity and realism. While a simulation is the application of a model to study the behavior and performance of a real system while manipulating variables that cannot be controlled in real life.

* 1. Describe how the concepts of a system state, state transition function, state variables and descriptive variables are related to a model-simulation.

A system state is a snapshot or moment of a simulation’s modeled representation of a real system, a state transition function determines how the simulation state changes in response to an event, the state variables are a collection of variables which describe the system at a particular time.

* 1. Describe the distinction between model validation and verification of a simulation.

Validation for a simulation would mean that the model being simulated was representative of the real system with respect to all assumptions being made, while verification would mean that the model being simulated is outputting accurate data with respect to all assumptions being made of the real system.

* 1. Discuss the levels of validity that can exist between a real system and a model.
  2. Discuss the difference between a continuous state/time system simulation and a discrete state/time simulation.

The major difference between a continuous state simulation and a discrete state simulation would be how the model used represents the states of the system. A discrete state model would represent the system in states that change instantaneously at separate points in time, similar to our drive-thru simulation. While with a continuous model the state variables change continuously with respect to time, like the position and speed a moving car.

1. ( 45 pts ) Suppose we want to simulate the behavior of a small gas station that has one island containing two gasoline pumps as shown below:
   1. (5 pts) List the import statements

Text

Description automatically generated

* 1. (15 pts) Provide code to create appropriate processes and resource. Also provide code to initialize the simulation, launch processes and start the execution of the simulation.

Text

Description automatically generated

* 1. (10 pts) Provide code to implement a customer generator process.

Text

Description automatically generated

* 1. (15 pts) Provide code to implement a customer simulation process.

Text

Description automatically generated

1. ( 30 pts ) Input Distributions
   1. (10 pts) Suppose you have been asked to model a system’s input behavior. After observing the system, you form the histogram shown below. You have decided to fit a standard distribution to the data. Select the “best” distribution to describe this input and sketch a plot of the corresponding probability density function. What is the mean of your chosen distribution. Make sure to clearly state why you selected the specific distribution and rejected the others.

A left skewed Weibull distribution would describe this distribution “best”, my mean would be around 30.

* 1. (10 pts) Suppose a simulation study was conducted of a grocery store that had several checkout stations. In the report of the simulation results, it was noted to save cost, the analyst decided to collected data on only one checkout station and recorded the number of stations open. To estimate the true arrival rate, they decided to multiply the observed arrival times by the number of checkout stations. For example, if the analyst observed an arrival time 10 and three checkouts were open, they assume there were three arrivals at time 10. This results in the following type of input data: observed: 10, 12, 17, 25, 32, ... that was simply tripled to form and augmented data set: (10, 10, 10, 12, 12, 12, 17, 17, 17, 25, 25, 25, 32, 32, 32, ...). Finally, the augmented data set was used to make a histogram, generated parameters (e.g. sample mean and variance) and select a theoretical distribution.

What is the flaw in this approach to selecting an input distribution? Do you suspect that this approach overestimates or underestimates the arrival rate? What would you do to improve the quality of this simulation?

The flaw in this approach to selecting an input distribution is that all the data collected was tripled in occurrence to accommodate for a lack of data collection which no consideration for how this affected the validity of the model. I would suspect that this drastically overestimates the arrival rate. To improve the quality of the simulation I might look for a longer period of data collection, changing which lane is being observed if only one can be at a time, increase the number of lanes being observed at one time and, simplify the model down to simulate a single check out station and then increase the checkout stations being simulated.

* 1. (10 pts) Suppose you are given the following histogram of interarrival times for a small shop. Hypothesize a theoretical distribution for this data and estimate the parameter(s) for the chosen distribution. Use the midpoint of the interval defining each bin to represent the typical value. Explain why you believe your chosen distribution is appropriate. Also describe how you would perform a Chi-Square test for the quality of the fit of your chosen distribution to the sample data

I would likely use a log normal distribution with a sigma near 2 pushing the curve to the far left. I would use this over other distribution as the highest level of frequency is near the shortest interval of time, and this frequency decreases as time increases. To perform a Chi-Square test we would