

CE 311S Lab 8

Using Hypothesis Tests to Validate a Model

Purpose

In this lab you will evaluate the value and statistical validity of a highway traffic model using hypothesis tests. This hypothesis tests will allow you to rigorously determine if the results of the model accurately reflect reality.

Introduction

Over the past few decades, societies have become more and more dependent on the use of road networks in order to function, resulting in greater congestion, crashes, and demand for infrastructure (road) maintenance. Finding ways to model road networks is an important step toward understanding their performance and impacts. However, this modeling process is extremely difficult, primarily because it involves human agents. An important test for a model is whether or not it “replicates” reality with strong success.

For this lab, an engineering firm has developed a simple model for simulating traffic flow, or the number of cars per time period, on a particular section of highway. This model generates traffic flow for a one-hour period and also calculates the average time it takes for each car to traverse the highway section. You have been assigned the task of bench-testing the model to determine whether or not it accurately simulates the traffic flow. In order to perform this evaluation, you will compare output from the model with real data (taken from the highway section on a single day).

It is important to note how the time (cost) it takes to traverse a freeway segment depends on traffic flow v ¹. A simple, though widely accepted form for the average travel time of a section of road is the classic Bureau of Public Roads (BPR) function:

$$t(v) = A \left(1 + B \left(\frac{v}{C} \right)^D \right)$$

where v represents flow and A , B , C , and D are constants determined by the properties of the road. What is important to note is that $t(v) \propto v^D$, where $D > 1$.

Procedure

- 1) Answer Questions 1 – 2.
- 2) Download and open the file “traffic.zip” from the lab Webpage.
- 3) Open the file “traffic.vi” and set the “Time of Day” setting to “6:00 – 7:00 a.m.” Click the “run” button (⇒) on the toolbar and record the model’s results in the Excel spreadsheet traffic.xls. Run 49 more trials and record the results.
- 4) Select another Time of Day and record 50 trials in the second sheet of the Excel spreadsheet.
- 5) Answer Questions 3 - 5.

Extra Credit:

Select a third Time of Day and record 50 trials on the third sheet of the spreadsheet, and add calculations for this data to your answers for Questions 3-5.

¹ Here, “flow” is to be interpreted as demand. Note that v is allowed to exceed the capacity C (as in question 1). Technically, flow can never exceed capacity while demand may.

Table 1: Observed Traffic Flows and Travel Times by Time of Day

Time of Day	Traffic Flow (vehicles/hour)	Average Travel Time (minutes)
6:00 – 7:00 a.m.	7500	1.40
7:00 – 8:00 a.m.	16500	3.66
8:00 – 9:00 a.m.	21000	5.75
9:00 – 10:00 a.m.	12000	2.24
10:00 – 11:00 a.m.	7800	1.44
11:00 – 12:00 noon	6600	1.30
12:00 – 1:00 p.m.	6225	1.26
1:00 – 2:00 p.m.	7350	1.38
2:00 – 3:00 p.m.	8175	1.49
3:00 – 4:00 p.m.	13125	2.54
4:00 – 5:00 p.m.	13875	2.76
5:00 – 6:00 p.m.	15375	3.25
6:00 – 7:00 p.m.	14475	2.94

Questions

- 1) Why is the constant D in the BPR function greater than 1? As a result of this, how will the distribution of the average travel time, $t(v)$, compare with that of the flow, v ? (For this question, assume that, for large values of v , $v > C$)

2) One of the effects of poor traffic modeling is that the interactions between the population and the traffic network (e.g. why they use it, how they drive when they are on it) are not correctly understood. What are some of the effects of not understanding these interactions correctly? That is, how does a lack of good foresight in the area of how people will utilize a transportation network affect an urban area, in both the short and long run? Are any of these effects apparent here in Austin?

3) With a 95% confidence interval, test the hypothesis that the 6:00 – 7:00 a.m. traffic flow predicted by the model is equal to the actual data given in Table 1. At the same level of confidence, test the same hypothesis for the Time-of-Day you selected in Part 4. Discuss your results (i.e., are they different or the same – and why you think this is happening).

4) With a 95% confidence interval, test the hypothesis that the cost predicted by the model is equal to the actual data given in Table 1 for both the 9:00 a.m. traffic flow and the Time-of-Day you selected in Part 4. How do these compare to your results from Question 3? Why do you think this is happening?

5) Based on your results from Question 3 and 4 (and possibly 2), make a decision as to whether or not you would recommend this traffic flow model for actual use. Back up your conclusions with both quantitative reasoning and practical reasoning, including considerations of the fact that this validation only took into account a single day's data for comparison. (use back if necessary).