

Project 3: Population Proportions

Zack Garza

Math 142: Elementary Statistics
Effective Date of Report: July 19, 2014

CONTENTS

I	Introduction	1
II	Data	2

I. INTRODUCTION

THe objective of this project is to test a claim about two independent population proportions. Due to the ever-increasing cost of fuel and the intricate link between fossil fuels and international conflicts, the overall fuel efficiency in miles-per-gallon (MPG) has become increasingly important to both consumers and manufacturers. The question addressed in this project is whether or not spikes in global oil prices drive innovation by increasing the average miles per gallon rating in newly manufactured cars.

II. DATA

To analyze this hypothesis, we turn to the UCI Machine Learning Repository, which provides a data set of matching vehicles made between 1970 and 1982 and their efficiencies in MPG. The time period between 1978 and 1982 will be examined, as this time interval includes the 1979 Energy Crisis, during which the US price per barrel of crude oil spiked and purportedly led to increased fuel economy in the 1980s.

The data set includes samples of individual car models from each year, so the first population will be a sample of cars automobiles made in 1978, and the second population will be those made in 1982. The 10 year average MPG between 1970 and 1980 will be used as a baseline – anything above that average will be considered a "success", and anything below it will be considered a "failure". We will take the null hypothesis to be that the proportion of 1978 cars that were above the ten year average is equal to the the same proportion of 1982 cars. The alternative hypothesis is that they were not in fact equal, which is summarized as follows:

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 \neq p_2$$

where p_1 is the proportion of 1978 cars for which the MPG is above the 10 year average, and p_2 is the proportion of 1982 cars with MPGs above the 10 year average.

First, the baseline must be computed, so the data set is read in, and the MPGs for all cars made between 1970 and 1980 are summed and divided by the number of cars sampled.

Ten Year Average: 23.515

The statistical characteristics of these populations are given below.

Population	Number of Samples (n)	Average MPG (μ)	Proportion Above 10-Year Average (\bar{p})
1 (1978)	$n_1 = 36$	$\mu_1 = 24.061$	$p_1 = 0.389$
2 (1982)	$n_2 = 31$	$\mu_2 = 31.71$	$p_2 = 0.935$