

**STAT40720 Intro. to Data Analytics**

**Assignment 3**



**Name** Kevin O’Connell

**Student ID** ########

**Submission Date** 2ndth November 2015

**Question 1**

**(a) Mean Bulb Lifetime longer than 1,000 hours**

H0: μ = 1000 hours  
HA: μ > 1000 hours

This is an upper-tailed test. If the test statistic lies in the rejection region, i.e. z > zCRITICAL value, the null hypothesis should be rejected. There is then sufficient evidence to state the mean lifetime of the bulbs is greater than 1000 hours.

**(b) Mean Male/Female Population Proportions of Gym Attendees equal**

H0: MALE - FEMALE = 0  
HA: MALE - FEMALE ≠ 0

…where MALE and FEMALE are the proportions of male and female   
students respectively who attend the gym 4 or more times per week.

This is a two-tailed test. If the test statistic lies in the rejection regions, i.e. z > zCRITICAL or z < - zCRITICAL values, the null hypothesis should be rejected. There is then sufficient evidence to state the proportion of students that attend the gym 4 or more times per week is different from males to females.

**Question 2**

1. **Hypothesis Construction**

If the restaurant want to test if the entire batch is spoiled, they should conduct a hypothesis test with the following Null and Alternate hypotheses:

H0: SPOILED < 1  
HA: SPOILED = 1

…where SPOILED is the proportion of chicken packages that are spoiled.

However it is more likely the chicken is mostly unspoiled, and the restaurant wants to randomly sample for spoiled chicken. In this case, the following hypotheses should be adopted:

H0: SPOILED > 0  
HA: SPOILED = 0

…where SPOILED is the proportion of chicken packages that are spoiled.

Here, the Null Hypothesis (H0) assumes the true proportion of spoiled chicken is greater than zero. After randomly sampling the chicken for spoiled packages, if there is not sufficient evidence that the true spoiled proportion is greater than zero, the Null Hypothesis should be rejected, and accept the Alternate Hypothesis, i.e. the entire batch is safe.

1. **Type I & Type II Error**

Proceeding with the more likely scenario above, i.e.:

H0: SPOILED > 0  
HA: SPOILED = 0

A Type I error refers to the incorrect rejection of the Null Hypothesis, and acceptance of the Alternative Hypothesis, when the Null Hypothesis is, in fact, true. In this scenario, it refers to the assumption that the entire batch of chicken is safe, when in fact, there is a spoiled package.

A Type II error refers to the failure to reject a false Null Hypothesis when the Alternative Hypothesis is true. In the context of this question, it is a failure to reject the assumption that there is a proportion of the population that is spoiled, when in fact, the entire population is not spoiled, and safe to eat.

**Question 3**

1. **95% Confidence Interval**

The zCRITICAL value for a 95% confidence interval is 1.96.

1. **Hypothesis Test**

The following hypotheses should be used:

H0: μ = 40 packets  
HA: μ > 40 packets

This is an upper-tailed test.

The zCRITICAL value for a significance *α = 0.05* is 1.6449. The z-statistic is greater than the zCRITICAL value, therefore the Null Hypothesis should be rejected. There is sufficient evidence to state the mean quantity of packets sold each week is greater than 40.

1. **Profit Assurance**

The P-value (obtained from Table 9, NCST) associated with a zCRITICAL value of 5.1649, with 30 degrees of freedom is greater than 0.9999. The tables stop listing values above zCRITICAL = 4.4. Therefore, the pharmaceutical company can be more than 99.99% certain the mean sales in the Kilkenny pharmacy are greater than 40 per week, and equally that average profits will exceed €400.