

**STAT40720 Intro. to Data Analytics**

**Assignment 1**



**Name** Kevin O’Connell

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

**Submission Date** 5thth October 2015

**Question 1**

Chebyshev’s Theorem provides the proportion of a dataset lying within **k** standard deviations of the mean is given by:

So, for 75% of the population:

75% of the population will lie between 2 standard deviations of the mean, yielding the range:

75% of the houses will sell for a price between €414,000 and €642,000.

**Question 2**

**(a) At most 2 lines are in use**

**(b) At least 3 lines are in use**

**(c) Between 2 and 3 line, inclusive, are not in use**

**(d) At least 3 lines are not in use**

**Question 3**

**(a) What is the value of *i*?**

**(b) What is the probability that at most 2 documents are required?**

**Question 4**

**(a) Determine P(X ≤ 2)?**

**(b) Determine P(1 ≤ X ≤ 4)?**

**(b) Determine P(0)?**

**Question 5**

**(a) Determine P(x < 30km/h)?**

But the Normal distribution is symmetric about zero, yielding:

Note, the value for above was taken by linear interpolation between probabilities for 0.31 and 0.32 in the NCST tables.

**(b) Determine P(x > 45km/h)?**

**(c) Determine P(40km/h < x < 60km/h)?**