Hypothesis testing and estimation for proportions

[Name of the Writer]

[Name of the Institution]

**Answer 1:**

**Hypothesis:**

**H0**:  most people in B.C. does not think their local police are doing a good or very good job at treating people fairly

**H1**:  most people in B.C. think their local police are doing a good or very good job at treating people fairly

**Significance level:**

The significance level for this study is set at 0.05. This shows that we have a 5% chance of rejecting a null hypothesis when it is really true. This is also refered to as type 1 error. The type 2 error is the failure to reject the null hypothesis when it is really false. In this example the probability of committing type2 error is accounted for by selecting a reasonably large sample size for the analysis.

**Test Statistic:**

Z = (P’- p)/ (p.q)/n

**Where**

P’ = 1916/2400

P= 0.5

q = 1-p

 = 0.5

**Calculation:**

Putting values in the above formula

Z = (0.798-0.5)/ (0.5\*0.5)/2400

= 0.298/0.0102

= 29.21

**Tabulated value of Z:**

The tabulated value of Z is 1.645. This is because the test is one tailed. The tail of the test is decided by the way in which hypothesis are developed. If the hypothesis are developed in terms of = for H0 and ≠ for H1 then it will be considered a two tailed test. However in the case under consideration, the hypothesis are stated in terms of more than and less than. This way of stating the hypothesis leads us to one tailed test. In calculation, there is slight difference between the tabulated values of Z when we work with two tailed or one tailed tests.

**Conclusion:**

The calculated value of Z is greater than the tabulated value so we reject H0. This means that we could not gather enough evidence to accept the null hypothesis. It means that  most people in B.C. think their local police are doing a good or very good job at treating people fairly The hypothesis are developed in terms of the population proportion i.e. P. Then we have used the sample data for calculation purposes. In this analysis, we have stated the hypothesis theoretically. If we convert it to numeric form, we will have P≤ 0.5 and P> 0.5. It is easier to state the hypothesis in numerical terms and it is also easier to comprehend for the readers.

**Part 2**

**Confidence coefficient for proportions:**

A confidence interval is used to show that the value of a certain parameter lies within the specified limits. A parameter is a value calculated form a population. The figures used for the calculations undertaken are the statistics i.e. taken from the sample. I will choose a 99% confidence interval because the higher confidence interval would mean lesser error. The error term is 100-CI which comes to 1 % in this case. The other name of this figure is the level of significance.

**Calculating the proportion:**

This method is used when we need to know how many times does a certain attribute occur within a given set of data. In this case the attribute is the response of people with good or very good option. We have proportioned those people who think that the police is doing a good job at ensuring the safety of the people. The desired criteria in the given excel sheet shows that there are 2365 occurrences of good or very good. Thus the proportion would be 2365/2400 = 0.98. This is refered to as p^. This is also named as sample proportion. The symbol **P** is used to show the term population proportion.

 **Margin of error:**

The margin of error is calculated as Z\* P’(1-P’)

= 2.58\* (2365/2400)(1-2365/2400)

= 2.58\* (0.98)(0.02).

= 0.3612

**Formula for confidence interval:**

**P’-margin of error & P’+ margin of error**

**=** 0.98-0.3612, 0.98+0.3612

= 0.6188, 1.3412

**Part 3:**

In the first part we have enough evidence to reject H0. We could not have enough evidence to accept the null hypothesis. Thus we can say that the police is doing a good job in treating the people fairly. The statistical analysis done shows a clear tilt of the people towards the performance of the police department. We have accounted for both types of errors. Type 1 and type 2 errors.

In the second part we are 99% sure that a given value will lie between the given interval. There is only 1% chance that any given data point containing good or very good criterion will fall outside this limit. Also the limits are not very wide which shows that data are condensed. That is beacause we have used 99 % confidence interval for our analysis. If we represent it on a diagram, we will have a large majority of the values falling within a specified limit. We have estimated the limits within which we are 99% confident that the population proportion will fall.