**Capstone Project: Bias in Promotion Exams**

**Introduction:**

**The Capstone project is a semester long, critical thinking project utilizing a single data set. You will be examining this data from the point of view of concepts in each chapter of the statistics course. The goal is to give you experience in manipulating and analyzing data and to help you understand how decisions are made and conclusions reached using data.**

**Intermediate assessments will have deadlines scattered throughout the semester. The assignments range from answering questions and making graphs to running hypothesis tests and drawing conclusions regarding the meaning of the data set. For all written responses you are expected to use proper English grammar and to write in complete sentences and, where needed, paragraphs. Statistical graphs should be properly labeled and in a format large and clear enough to be easily understood.**

**Links to the original raw data will be provided.**

***Adapted from data and article in “Journal of Statistics Education” by Weiwen Miao, Haverford College.***

**Part 1: Chapter 1: Data**

**Open the data provided. The data includes all of the results for a promotion exam cycle at the New Haven Fire Department in 2003. The columns represent the race of the examinee coded as W for white applicants, B for black applicants and H for Hispanic applicants. The next columns are the scores for the oral and then the written parts of the exam. Under an agreement between the firefighter’s union and the city government, the written exam was weighted as 60% of total score and the oral exam was weighted as 40% of total. The final column of data will be the aggregate score for each examinee. Only applicants with an aggregate score of 70% or above were eligible for promotion. In addition, the department was required by the city charter of New Haven to fill the “*x”* available positions from a list of the top “*x+2”* candidate scores. This “eligibility list” could be used for a two year period of time to fill any additional vacancies that occurred.**

|  |  |
| --- | --- |
| **Variable Label** | **Measured In** |
| **Race** | **W(white), B(black),H(Hispanic)** |
| **Position** | **Lieutenant, Captain** |
| **Oral exam** | **Numerical, 0-100** |
| **Written exam** | **Numerical, 0-100** |

**Based on the description of the data, answer the following questions.**

**Task 1:**

* **Does the data represent a sample or a population? Explain your answer.**
* **Would this be considered an observational or experimental project? Explain your choice.**

**Task 2: Determine the level of measurement for each of the variable.**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Level of Measurement** | **Explanation** |
| **Race** |  |  |
| **Position** |  |  |
| **Oral exam** |  |  |
| **Written exam** |  |  |

**Part II: Chapter 2: Summarizing and Graphing Data**

**Task 1: Prepare a frequency distribution for the written exam data for the 118 firefighters. Use a starting point of 40 and determine what class width will give you 6 classes. This class width is\_\_\_\_. Now fill out the frequency table below:**

|  |  |
| --- | --- |
| **Class** | **Frequency** |
| **40 -** |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Convert this table into a relative frequency table:**

|  |  |
| --- | --- |
| **Class** | **Relative Frequency** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Task 2: Use Statdisk to construct a histogram of the written exam score data. Be sure to use a starting point of 40 and the same class width as you used in the frequency tables. Be sure to print out the histogram and correct the labels as needed to properly reflect the data.**

* **Describe the shape of this distribution using proper statistical terminology.**
* **Which classes are the largest?**
* **Estimate the average score of the participants based on the histogram.**

**Task 3: Use Statdisk to sort the written exam data for individual race categories. Now make a stem and leaf plot for each of the three race categories below:**

* **How many test scores from each racial category are between 40 and 59?**
* **Which race had the highest test scores? Lowest test scores?**

**Part III: Chapter 3: Descriptive Statistics**

**Task 1: Descriptive Statistics:**

* **Use the sample transformer to find the combined score, which is a weighted mean of 40% oral and 60% written scores. Copy these scores into column 5 and label the column as “combined score”. Save this data set to your flash drive or H drive. Print out the data list and attach to this worksheet.**
* **Use Statdisk to get the descriptive statistics for oral exam, written exam and combined score and insert these results into the table below.**

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Task 2: Fill in the following chart concerning the measures of center for the variables:**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Best Measure of Center** | **Explain Why** |
| **Race** |  |  |
| **Position** |  |  |
| **Written exam** |  |  |
| **Oral Exam** |  |  |
| **Combined score** |  |  |

**Task 3: Sort the data according to position and create a modified, side-by-side boxplot of the combined score data for the positions. Copy and insert the boxplots below being sure to properly label the plots. Outliers are shown with an asterisk.**

**Are there any outliers in this data?**

* **To determine the highest and lowest boundary data values, you must first calculate the Interquartile Range (IQR): IQR = Q3 – Q1**
* **Lowest boundary value = Q1 -1.5 (IQR)**
* **Highest boundary value = Q3 + 1.5 (IQR)**

**Any data value outside these bounds is an outlier. List any outliers.**

**Task 4: Analyze the side-by-side modified boxplots from Task 3 and summarize the differences you see between the two groups. Explain why these results are reasonable.**

**Task 5:**

* **Determine the z score for the identified outlier in the Captain’s combined scores data.**
* **Complete the following statement: The z-score of \_\_\_\_ means that the individual data value \_\_\_\_\_\_ is \_\_\_\_\_ standard deviation units above/below (circle one) the mean value of \_\_\_\_\_\_\_.**
* **Is this an unusual z score? Why or why not?**

**Part IV: Chapter 4: Relative Frequency Probability Estimates**

**Task 1: You will need to manipulate the data and construct some contingency tables in order to find the probabilities for Task 2. There is a prepared handout on this topic which can be obtained from your instructor.**

|  |  |  |
| --- | --- | --- |
|  | **Lieutenant candidate** | **Captain candidate** |
| **White** |  |  |
| **Black** |  |  |
| **Hispanic** |  |  |

**Task 2: Use the contingency table to estimate the following probabilities.**

* **What is the probability of selecting a black candidate if you choose one name at random from the data?**

**P(black candidate) =**

* **What is the probability of selecting a captain candidate if you choose one name at random from the data?**

**P(captain candidate) =**

* **What is the probability of selecting a Hispanic lieutenant candidate if you choose one name at random from the data?**

**P(Hispanic lieutenant candidate) =**

* **What is the probability of selecting a lieutenant candidate or a white candidate?**

**P(lieutenant OR white) =**

* **If you selected 2 participants without replacement what is the probability of selecting 2 consecutive white candidates?**

**P(white and then white) =**

**Part V: Chapter 5: Probabilities for Binomial Distributions**

**Task 1: For the entire set of 118 test takers, calculate the probability of scoring above 70 which was the promotion cutoff score. Open the full data set and sort the combined scores to get the number of successful candidates.**

**P(success) =**

**Task 2: Suppose that in an adjacent city with a comparable population and general demographics ten firefighters are randomly selected to take the same national promotional exam. Determine if this group can be treated as a binomial distribution. State each requirement.**

**Task 3:**

* **Calculate the probability that none of the test takers in the group of ten has obtained a score high enough to qualify for promotion.**
* **Calculate the probability that at least one of the test takers in the group of ten has obtained a score high enough to qualify for promotion.**
* **Describe the relationship between these results.**

**Task 4: Would it be considered statistically unusual if more than 8 out of the 10 randomly selected candidates scored high enough to be eligible for promotion? Compute the probability and explain why.**

**Task 5: Would it be considered statistically unusual if exactly 7 out of the 10 randomly selected candidates scored high enough to be eligible for promotion? Compute the probability and explain why.**

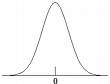
**Part VI: Chapter 6: Probabilities for Normal Distributions**

**Task 1: Construct a histogram and boxplot of the combined scores of the all test takers using the “explore data” function in Statdisk.**

**Describe how this histogram and boxplot fit the definition of approximately normal. Be sure to paste the results screen below. Identify the mean and standard deviation.**

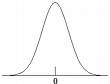
**Task 2: What is the probability that a participant will get a score less than 65?**

**Z = P(x<65) =**



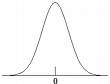
**Task 3: What is the probability that a selected participant scores between 80 and 100 on the exam?**

**Z = P(80<x<100) =**

****

**Task 4: If a group of 9 participants are randomly selected find the probability that the mean score for the group is over 72?**

**Z = P(x>72) =**



**Part VII: Chapter 7: Confidence Intervals**

**For the purposes of confidence intervals, we will consider these 118 candidates to be a sample of the larger population of firefighters nation-wide who took this promotion exam.**

**Task 1: Use Statdisk to complete the contingency tables for passing status versus race for both the Lieutenants and Captains exams.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Lieutenants:** | **Black** | **Hispanic** | **White** |
| **Pass** |  |  |  |
| **Fail** |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Captains:** | **Black** | **Hispanic** | **White** |
| **Pass** |  |  |  |
| **Fail** |  |  |  |

**Task 2: Use the data set to find for each of the following:**

* **Among the black candidates for lieutenant exam, what was the percentage who passed?**
* **Among the Hispanic candidates for lieutenant exam, what was the percentage who passed?**
* **Among the white candidates for lieutenant exam, what was the percentage who passed?**
* **Among the black candidates for captain exam, what was the percentage who passed?**
* **Among the Hispanic candidates for captain exam, what was the percentage who passed?**
* **Among the white candidates for captain exam, what was the percentage who passed?**

**Task 3: Use the data set to find for the percent of the sample that are applying for the captain’s position.**

* **Describe what requirements must be met for this interval to be valid and whether you think that this data set meets these requirements.**
* **Based on this data, construct a 95% confidence interval for the percentage of firefighters that will apply for a Captain’s rank.**
* **Interpret the meaning of the confidence interval in words.**
* **Can we confidently state that less than half of the candidates are taking the test for Captain? Why or why not?**

**Task 4: Use the mean score for the oral exam of all candidates to construct and complete the confidence interval table below.**

|  |  |  |
| --- | --- | --- |
| **Confidence Level** | **Confidence Interval** | **Margin of Error** |
| **80%** |  |  |
| **90%** |  |  |
| **95%** |  |  |
| **98%** |  |  |
| **99%** |  |  |

**Be examining this table, answer the following questions.**

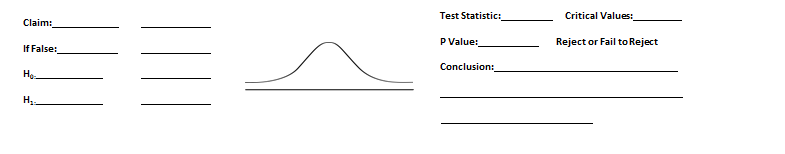
* **As the confidence level increases, what happens to the width of the interval?**
* **As the confidence level increases, what happens to the margin of error?**
* **If the confidence level were 92%, would the confidence interval estimate be more or less precise than for 95%?**

**Part VIII: Chapter 8: Hypothesis Tests**

**Task 1: Sort according to position and delete the lieutenant data, then sort according to race and select and move the data for white candidates to new columns on the right. Construct a side-by-side boxplot for mean combines scores and insert below.**

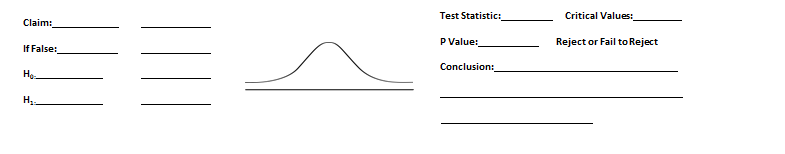
**What does examination of this boxplot suggest about the difference in exam scores between white and minority candidates for promotion?**

**Task 2: Examine the claim that the mean score on this exam is lower than 74.11 for the minority candidates for the captain position.**

****

* **If the mean score for white candidates in New Haven is 74.11, what conclusions can we draw from our hypothesis test?**
* **Which procedure, a visual examination of a boxplot or a formal hypothesis test, is more valid proof of the idea that white candidates have a higher mean score?**

**Task 3: Use this data and a .05 significance level to test the claim that less than 58% of the minority candidates (Black and Hispanic candidates grouped together) pass the lieutenants exam.**

****

**Part IX: Chapter 10: Linear Regression Relationships**

**Task 1:**

* **Create a scatterplot for oral exam versus written exam scores. Insert that scatterplot here then perform a linear regression test on the data.**
* **Report the correlation coefficient, Pearson critical values, and write the equation for the linear relationship.**
* **Determine if there is a valid linear relationship.**

**Task 2:**

* **Create a scatterplot for written exam versus oral exam scores. (This is simply a reversal of x, y values from Task 1.) Insert that scatterplot here then perform a linear regression test on the data.**
* **Report the correlation coefficient, Pearson critical values, and write the equation for the linear relationship.**
* **Determine if there is a valid linear relationship.**

**Task 3: Describe the differences between the results from Task 1 and Task 2 by completing the statement below.**

**When the x, y values are reversed, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ remain the same while the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ changes.**

**Part X: Chapter 11: Independence Tests**

**One way to examine discrimination data is to analyze the relationship between variables. In the Ricci case, we want to look at the possible relationship between whether or not a candidate passed their exam and their race. To do this we can make contingency tables for each exam and then do Chi-squared tests of independence between passing and race.**

**Task 1: Use Statdisk to create two contingency tables. Fill in the tables below:**

**Lieutenant Exam:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Black** | **Hispanic** | **White** |
| **Pass** |  |  |  |
| **Fail** |  |  |  |

**Captain Exam:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Black** | **Hispanic** | **White** |
| **Pass** |  |  |  |
| **Fail** |  |  |  |

**Task 2: Perform the chi-squared independence test for the lieutenant’s position. Insert the results below.**

* **Can we conclude that the lieutenant exam results are independent of race?**
* **What does this imply about fairness in the promotion exam?**

**Task 3: Perform the chi-squared independence test for the captain’s position. Insert the results below.**

* **Can we conclude that the captain exam results are independent of race?**
* **Then does that mean we can conclude that the captain exam is not affected by race?**

**Part XI: Final Analysis**

**Research the “reverse discrimination” lawsuit filed by 18 firefighters from the New Haven Fire Department against the city. The suit was filed because the city of New Haven had decided in 2003 NOT to certify the promotional exam presented by this data. These firefighters (17 white and one Hispanic) were, therefore, not given an opportunity for promotion in that testing cycle. As you do your research, answer the questions posed below.**

**Questions:**

1. **Why did the city decide not to certify the exam?**
2. **Briefly explain the “adverse impact” guidelines which the city of New Haven relied on and the 80% rule used to demonstrate adverse impact.**
3. **This case went all the way to the US Supreme Court. Briefly outline the outcomes of the case as it moved through the justice system.**
4. **The US Supreme Court rules in favor of the firefighters by a 5-4 margin. Briefly explain the ruling of the court.**
5. **In what way could you argue that New Haven was “damned if they did and damned if they didn’t” in making their decision regarding whether to certify the original exam results?**