DANA 4800 HW4 Answer Keys

1. According to previous study in 2016, 42% of Canadian employees had extended health care coverage from their employers. With organization putting more focus on the employees' physical and mental health, an HR personnel would like to see if the number has gone up since. A recent study found that 180 out of 400 randomly sampled workers have the extended health care coverage from their employer's health care plan.

   Subjects: Canadian employees
   # of variable: 1
   Variable: whether or not Canadian employees have the extended health care coverage from their employers (classes: Yes, No)
   Type: categorical variable

   a) Identify the parameter of interest. **[2 bonus marks]**
      Define p as the proportion of all Canadian employees who have the extended health care coverage from their employers [2]

   b) Set up the null hypothesis and the alternative hypothesis. **[2 bonus marks]**
      $H_0: p \leq 0.42$ **vs.** $H_a: p > 0.42$

   c) Calculate the test statistic and find the p-value. **[4 bonus marks]**
      Note: Make sure you understand checking conditions is part of the procedure.
      $$\bar{p} = \frac{180}{400} = 0.45, T.S. = \frac{\bar{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.45 - 0.42}{\sqrt{(0.42)(1-0.42)/400}} = 1.216 \text{ [1]}$$
      p-value = 0.112. [1]

      Note that the p-value is only an approximation and the approximation is good because all three conditions are satisfied:
      1) $n = 400$, which is bigger than the requirement of 30,
      2) $n \times p_0 = 400 \times 0.42 = 168$, which is bigger than the requirement of 5,
      3) $n \times (1 - p_0) = 400 \times 0.58 = 232$, which is bigger than the requirement of 5. [2]

   d) Make an appropriate conclusion in plain English. Use $\alpha = 0.01$. **[2 bonus marks]**
      Since the p-value is bigger than the 1% level of significance, we do not have enough statistical evidence to reject the null hypothesis and conclude that proportion of all Canadian employees who have the extended health care coverage from their employers is not significantly higher than 42%. [2]

2. At ABC University, the average scholarship examination scores for freshman applications have been 900. Every year, the Associate Dean uses a sample of applications to determine whether the average examination score for the new freshman applications is different from 900. A random sample of 200 applications was drawn this year and the average score was found to be 935 and a standard deviation of 180. Historically, it is known that the exam score has a left-skewed distribution.
   Subjects: freshman applications at ABC University
   # of variable: 1
   Variable: application score (no units)
   Type: numerical variable
   a) Identify the parameter of interest. **[2 marks]**
      Define μ as the average scholarship examination score of all freshman applications at ABC University [2]

   b) Set up the null hypothesis and the alternative hypothesis. **[2 marks]**
      $$H_0: \mu = 900 \text{ vs. } H_a: \mu \neq 900 \text{ [2]}$$

   c) Calculate the test statistic and find the p-value. **[2 marks]**
      Given: $\bar{X} = 935$, $s = 180$ and $n = 200$.
      Test statistic: $TS = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} = \dfrac{935 - 900}{180/\sqrt{200}} = +2.75$ [1]
      p-value = $2 \times (0.0030) = 0.0060$ [1]

   d) Note that the p-value is only an approximation. Briefly explain why. Also briefly justify why we can still make a valid conclusion in the following part? **[1+2 marks]**
      The p-value is only an approximation because the variable of interest (exam score) is given not to have a Normal distribution. [1]
      Here, because the sample size is bigger than 30, the average exam score still follows approximately a Normal distribution, according to the Central Limit Theorem. [2]

   e) Draw a conclusion so that the Dean could understand. Use 5% significance level. **[2 marks]**
      Since p-value is less than the 5% level of significance, we have enough statistical evidence to reject the null hypothesis and conclude that the average scholarship examination score of all freshman applications at ABC University is significantly different from 900. [2]

3. A local pizza joint has recently hired additional drivers and claims that its average delivery time for orders is under 45 minutes. A random sample of 36 customer deliveries was examined and the average delivery time was found to be 42.3 minutes with a standard deviation of 11.6 minutes. The delivery time generally follows a Normal distribution.
   Subjects: customer deliveries of this local pizza joint
   # of variable: 1
   Variable: delivery time (in minutes)
   Type: numerical variable
   a) Identify the parameter of interest. **[2 marks]**
      Define μ as the average delivery time (in minutes) of all pizza deliveries

   b) Set up the null hypothesis and the alternative hypothesis. **[2 marks]**
      $H_0: \mu \geq 45$ **vs.** $H_a: \mu < 45$

   c) Calculate the test statistic and find the p-value **[2 marks]**
      $TS = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} = \dfrac{42.3 - 45}{11.6/\sqrt{36}} = -1.40$ **[1]**
      Because it's a lower-tailed test, p-value is the left-hand area
      p-value = 0.0808 [1]
      Note: Use Excel function =NORM.DIST(-1.4,0,1,TRUE) → p-value = 0.0808.

   d) Is the p-value above exact or approximate? Briefly justify your answer. **[0+1 mark]**
      The p-value above was exact because the variable of interest (delivery time) is given to follow a Normal distribution.

   e) Draw an appropriate conclusion using 5% significance level. **[2 marks]**
      Since the p-value is bigger than the 5% level of significance, we do not have enough statistical evidence to reject the null hypothesis and conclude that the average delivery time among all pizza deliveries is not significantly shorter than 45 minutes, as claimed.