ICME Intro to Stats Summer Workshop

Section 4 & 5 Solutions

2023-07-24

Section 4

1.

Answer: (5.52,8.48)
Solution: The sample mean is $x = \frac{(7+7+9+6+8+5)}{6} = 7$ with $n = 6$ and the sample standard deviation is $s = \sqrt(2)$. The sample standard deviation can be calculated as follows:

$$s = \sqrt{\frac{1}{n-1}\sum_n (X_n - \bar{x})^2} = \sqrt{\frac{1}{5}[(7-7)^2 + (7-7)^2 + (9-7)^2 + (6-7)^2 + (8-7)^2 + (5-7)^2]} = \sqrt{2}$$

Therefore the standard error of the mean is $\frac{s}{\sqrt{n}} = \sqrt{\frac{1}{3}}$. For a 95% confidence interval, we use the significance level $\alpha = 0.05$ and $t_{\frac{\alpha}{2},n-1=5} = 2.571$. We divide α by 2 because the confidence interval is two-tailed. Therefore the margin of error is $t_{\frac{\alpha}{2},5} * \frac{s}{\sqrt{n}} = (2.571)\sqrt{\frac{1}{3}} = 1.48$. Therefore the confidence interval is (7-1.48,7+1.48) = (5.52,8.48).

2.

- (a) Answer: A. H0 :iq_med = iq_pop. We are testing whether the medication has any effect, so the opposite of our alternative hypothesis (the one that we are testing) is that the medication has *no* effect, meaning that the average iq is the same for participants who take and do not take mediation.
- (b) Answer: D. H1 :iq_med ≠ iq_pop. The explanation is the same as in part a. We are testing to see whether the mean IQs are any different (or not equal). Note that this will be a two-tailed test.
- (c) Since the variance is known and the number of participants is large enough according to CLT, we can use the Z-distribution. Here is the formula we will follow:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{140 - 100}{\frac{15}{\sqrt{30}}} = 14.606$$

- (d) the p-value can be found by using a Z-table. We should note that since this is a two-tailed test we will have to find the p value as: $p = 2 * P(Z \geq |14.606|) = 1 - 2 * P(Z \leq |14.606|) < 0.0001$. The p-value is very small
- (e) At 5% significance level, we will compare our p-value that is < 0.0001 which is definitely less than 0.05. We can confidently reject the null hypothesis and conclude that the medication does have an effect on IQ.

3.

ANSWER: 0.0006 EXPLANATION: The population standard deviations are unknown and are assumed unequal, since the sample standard deviations are very different. This calls for an unpooled test. Therefore we compute the standard error of the difference of means, $\sigma_{x_1-x_2}$ = 1.1375, which yields the test statistic t* = $\frac{(82-86)}{1.1375}$ = -3.5164, and the p-value = 2P(T > |t*|) = 0.0006, since the null hypothesis calls for a two-tailed test.

The formula for the standard error is as follows:

$$s_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

This t distribution has degrees of freedom equal to:

$$df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1-1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2-1}} = \frac{(\frac{29}{55} + \frac{46}{60})^2}{\frac{(\frac{29}{55})^2}{54} + \frac{(\frac{46}{60})^2}{59}} \approx 112$$

We can use a t-table to find the p-value.

Section 5

1.

ANSWER: Yes

- EXPLANATION: The test statistic is t = $\frac{r\sqrt{(n-2)}}{\sqrt{1-r^2}} = \frac{0.45\sqrt{(25-2)}}{\sqrt{(1-0.45^2)}} = 2.417$. The critical value for $\alpha = 0.05$ for a two-tailed test using the t_{23} distribution is 2.069. t_{23} because DF in this case is $n-2$. The test statistic is greater than this, so the conclusion is that the null hypothesis should be rejected, meaning that the study produced evidence that the variables are significantly correlated. Once again, we can use a t-table to confirm this.

2.

- A. ANSWER: 1.75
  - EXPLANATION: $r\frac{s_y}{s_x} = 0.73(9.6/4.0) = 1.752 \approx 1.75$
- B. ANSWER: 21.76
  - EXPLANATION: a = $\bar{y} - b\bar{x}$ = 141.6 - 1.752(68.4) = 21.7632 $\approx$ 21.76
- C. ANSWER: 140.76
  - EXPLANATION: $\hat{y} = 21.76 + 1.75(68) = 140.76$