# ECO220Y1Y, Test #3, Prof. Murdock SOLUTIONS

**(1) (a)** 2.14 (or -2.14); is; 2; 18.6 (or -18.6); is; is [The value of the standardized test statistic is (0.836 – 0.650)/0.087.]

**(b)** -2.6 (or 2.6); is; 3; 1.3 (or -1.3); is not; is not [The value of the standardized test statistic is (0.302 – 0.315)/0.005.]

**(c)** A Type I error would be concluding that there is a difference between the fraction of those without a high school degree versus those with a high school degree that believe a rise in the Federal funds rate will lower inflation when, in fact, there is no difference: opinions are the same for both the less educated and more educated segments of the public.

**(d)** $P(Z > 1.645) = 0.05$

$$z = \frac{\widehat{P_{cv}} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\widehat{P_{cv}} - \frac{2}{3}}{\sqrt{\frac{\frac{2}{3}\left(1-\frac{2}{3}\right)}{250}}} = \frac{\widehat{P_{cv}} - \frac{2}{3}}{0.0298}$$

$$1.645 = \frac{\widehat{P_{cv}} - \frac{2}{3}}{0.0298}$$

$$\widehat{P_{cv}} = 0.7157$$

Hence, fill in the blank with 71.6.

**(2) (a)**

$H_0\colon p_T - p_C = 0$

$H_1\colon p_T - p_C \neq 0$

$\bar{P} = \frac{X_T + X_C}{n_T + n_C} = \frac{0.406810036*558 + 0.354056344*9442}{558+9442} = \frac{227+3343}{10,000} = \frac{3,570}{10,000} = 0.357$ (or you could read it straight from last row – Grand Total – of the PivotTable)

$$z = \frac{\hat{P}_T - \hat{P}_C}{\sqrt{\frac{\bar{P}(1-\bar{P})}{n_T} + \frac{\bar{P}(1-\bar{P})}{n_C}}} = \frac{0.406810036 - 0.354056344}{\sqrt{\frac{0.357(1-0.357)}{558} + \frac{0.357(1-0.357)}{9442}}} = \frac{0.05275369}{0.02087327} = 2.527 \approx 2.53$$

$$P - value = P(Z < -2.53) + P(Z > 2.53) = 2 * (0.5 - 0.4943) = 0.011$$

| *Percent with any visits*[1] | | |
|---|---|---|
| Percent in Control Group | Effect of Medicaid Coverage | P-value |
| 35.4 | 5.3<br>(2.1) | 0.011 |

**(b)**

$H_0: \mu_T - \mu_C = 0$

$H_1: \mu_T - \mu_C \neq 0$

[In this case, it is reasonable to do either the equal or unequal variances case. You do not need to do both. However, both approaches are shown next.]

Unequal variances case:

$$t = \frac{(\bar{X}_T - \bar{X}_C) - \Delta_0}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_C^2}{n_C}}} = \frac{(1.456989247 - 1.201652192) - 0}{\sqrt{\frac{2.784436014^2}{558} + \frac{2.756549513^2}{9442}}} = \frac{0.25533706}{0.12124} = 2.11$$

| *Number of visits*[2] | | |
|---|---|---|
| Mean Value in Control Group | Effect of Medicaid Coverage | |
| 1.202 (2.757) | 0.255 (0.121) | |

Equal variances case:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = 7.569513752$$

$$t = \frac{(\bar{X}_T - \bar{X}_C) - \Delta_0}{\sqrt{\frac{s_p^2}{n_T} + \frac{s_p^2}{n_C}}} = \frac{(1.456989247 - 1.201652192) - 0}{\sqrt{\frac{7.569513752}{558} + \frac{7.569513752}{9442}}} = \frac{0.25533706}{0.11986} = 2.13$$

| *Number of visits*[2] | | |
|---|---|---|
| Mean Value in Control Group | Effect of Medicaid Coverage | |
| 1.202 (2.757) | 0.255 (0.120) | |

**(3)** From the tabulation, 77/3,487 = 0.0221 so AAA = 0.022.

Next, find the standard error using $SE[\hat{P}] = \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} = \sqrt{\frac{\frac{77}{3487}\left(1 - \frac{77}{3487}\right)}{3{,}487}} = 0.002489$ so BBB = 0.002.

**(b)** From the first Stata summary, get the mean of 1.014932 so CCC = 1.015.

Next, find the standard error using $SE[\bar{X}] = \frac{s}{\sqrt{n}} = \frac{9.730774}{\sqrt{8{,}351}} = 0.10648$, where the sample s.d. comes from the first Stata summary, so DDD = 0.106.

From the second Stata summary, get the mean of 45.81459 so EEE = 45.815.

Next, find the standard error using $SE[\bar{X}] = \frac{s}{\sqrt{n}} = \frac{47.25831}{\sqrt{185}} = 3.4745$, where the sample s.d. comes from the second Stata summary, so FFF = 3.475.

**(c)** $(\hat{P}_2 - \hat{P}_1) \pm z_{\alpha/2}\sqrt{\frac{\hat{P}_2(1-\hat{P}_2)}{n_2} + \frac{\hat{P}_1(1-\hat{P}_1)}{n_1}}$

$(0.020 - 0.015) \pm 1.645\sqrt{\frac{0.020(1-0.020)}{10{,}029} + \frac{0.015(1-0.015)}{6{,}648}}$

$0.005 \pm 1.645 * 0.002$

$0.005 \pm 0.003$

$LCL = 0.002$

$UCL = 0.008$

We are 90% confident that among people who receive the standard letter with no match, those living in Blue states are between 0.2 percentage points and 0.8 percentage points more likely to donate to the charity compared to those living in Red states using the replication data from Karlan and List (2007). This is a nontrivial difference because only around 2 percent of people donate so in percent terms this is between a about a 10% to 40% higher rate for those living in Blue states. However, the interval is quite wide meaning we are not sure if there is a relatively modest or really quite substantial difference in the propensity to donate comparing Red and Blue states.

[Note: It is fine if the order of Red and Blue are switched so long as the interpretation is consistent with the calculations.]

**(d)** [In this case, it is reasonable to do either the equal or unequal variances case. You do not need to do both. In either case, the degrees of freedom are very large given the very large sample sizes, so no degrees of freedom calculations are needed: use infinity as the approximation.]

Unequal variances case:

$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

$(0.983 - 0.813) \pm 2.576\sqrt{(0.084)^2 + (0.063)^2}$

$0.17 \pm 2.576 * 0.105$

$0.17 \pm 0.27$

$LCL = -0.10$

$UCL = 0.44$

We are 99% confident that compared to people who received the usual letter, with no mention of a match, those who received a letter with a high example match amount on average donated between 10 cents less and 44 cents more using the replication data from Karlan and List (2007). Because the interval spans zero, we are not sure that offering a match and showing it with a high example amount has any effect on the amount of money people donate on average (which includes averaging in the zeros of people who do not donate).

[Note: It is fine if the order of the control group and the high-example-amount are switched so long as the interpretation is consistent with the calculations.]

[The equal variances case is more work but yields the same answer in this case after rounding. The pooled variance estimate is 71.15, the standard error of the difference is 0.103, and the margin of error is 0.27.]

**(4)** There is huge 92 percent chance of an inconclusive result – being unable to prove that less than 45 percent of teenagers identify as male – even if only 43 percent of the teenage population identifies as male, which is two whole percentage points less than 45 percent, with a planned sample size of 500 and a 1 percent significance level. This is an unacceptably high chance of Type II error and unacceptably low power. The researchers will need to plan a larger sample size than 500 teenagers.

**(5)(a)** $n = 378\ (= 109 + 269)$. $a = 6.467890$ and $b = 0.234712\ (= 6.702602 - 6.467890)$

**(b)** On average, in 2022/23 in ECO220Y, students not in a commerce program self-report competitiveness of about 6.54 on a zero to ten scale, with 10 being the most competitive. In contrast, students in a commerce program self-report being about 0.93 units more competitive compared to students who are not in a commerce program. This is substantially more competitive: about 14 percent more competitive.