ECO220Y1Y, Term Test #4, Prof. Murdock: SOLUTIONS

March 2, 2018, 9:10 – 11:00 am

(1) (a) At the University of Waterloo and among employees making at least $100,000, male professors have 2016 salaries that are on average $12,748 higher than female professors.

(b) salary-hat = 162.485 – 12.748*Female

(c) On average, the salaries of male professors are approximately 8.4% higher than female professors.

(2) (a) $H_0: (\mu_M - \mu_F) = 0$ versus $H_1: (\mu_M - \mu_F) \neq 0$. (Note: given the negative test statistic, it must be male minus female.) Given $t = -1.38$ and $\nu = 826$, we consult the Student t table, which for either 750 or 1,000 degrees of freedom indicates that that test statistic is between $-t_{0.10}$ and $-t_{0.05}$. Given that the reported P-value is 0.17, the authors must be using a two-tailed test, which implies a P-value between 0.10 and 0.20 from the table.

(b) Given the reported P-value of 0.02, the difference in marriage rates is statistically significant a 5% significance level, but not a 1% sig. level. The difference of 6 percentage points – 81% of male lawyers are married whereas only 75% of female lawyers are – is big enough to be economically significant. Hence, the difference is significant overall.

(c) Given the reported P-value of 0.59, the difference in tenure is not statistically significant at even a 10% significance level. The difference of 0.08 years – 5.18 years for male lawyers versus 5.26 for female lawyers – is too small to be economically significant.

(d) $(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ w/ $\nu = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{1.24^2}{684}+\frac{0.91^2}{441}\right)^2}{\frac{1}{684-1}\left(\frac{1.24^2}{684}\right)^2+\frac{1}{441-1}\left(\frac{0.91^2}{441}\right)^2} = \frac{0.000017022}{0.000000015} = 1104.4$

$(1.22 - 0.82) \pm 1.962\sqrt{\frac{1.24^2}{684} + \frac{0.91^2}{441}}$

$0.40 \pm 1.962\sqrt{0.004125731}$

$0.40 \pm 1.962 * 0.064$

0.40 ± 0.126

LCL = 0.274 and UCL = 0.526

We are 95% confident that male lawyers on average have from 0.3 to 0.5 more children than female lawyers.

(e) $H_0: p = 0.2$ versus $H_1: p > 0.2$. Must use a rejection region approach. We can choose any reasonable significance level. For $\alpha = 0.05$, we will need $z > 1.645$, where $z = \frac{\hat{P}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$. Plugging in and solving for the critical value yields: $1.645 = \frac{\hat{P}_{c.v.}-0.2}{\sqrt{\frac{0.2(1-0.2)}{441}}}$ where $\hat{P}_{c.v.} = 0.231$. Hence, we need a sample proportion of at least 23.1% to prove that more than 20% have the highest aspirations at a 5% significance level.

(3) $(\hat{P}_2 - \hat{P}_1) \pm z_{\alpha/2}\sqrt{\frac{\hat{P}_2(1-\hat{P}_2)}{n_2} + \frac{\hat{P}_1(1-\hat{P}_1)}{n_1}} = (0.74 - 0.66) \pm 1.645\sqrt{\frac{0.74(1-0.74)}{90} + \frac{0.66(1-0.66)}{90}}$

$0.08 \pm 1.645 * 0.06805$

0.08 ± 0.112

LCL = -0.03 and UCL = 0.19

We are 90% confident that the percent of people agreeing that autonomous vehicles (AVs) should sacrifice the life of the driver to save 20 lives is somewhere between 3 percentage points lower to 19 percentage points higher compared to the percent agreeing that AVs should sacrifice the driver to save only 5 lives. This is an exceptionally broad interval: we are not even confident that more people would agree to the sacrifice to save 20 lives than 5 lives. Despite the positive point estimate of the difference – the proportion agreeing is 8 percentage points higher – the margin of error is huge at 11 percentage points. [**Note:** It would also be correct to give a causal interpretation here. The number of lives saved was randomly assigned and, hence, any difference in the proportion agreeing would be caused by that. However, this is a bit of grey area because the CI spans both a positive and negative range. Hence, we are not sure if there is any effect at all or it is just sampling error. This is why the suggested interpretation is not stronger on causality.]

(4) (a) $H_0: \mu = 50$ versus $H_1: \mu > 50$. The relevant test statistic is $t = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{51.22-50}{\frac{24.21}{\sqrt{394,111}}} = \frac{1.22}{0.038564305} = 31.6$ with $\nu = n - 1 = 394{,}111 - 1 = 394{,}110$. The P-value is 0. Hence, we definitely have sufficient evidence to reject the notion that people pay *no* attention to interest rates when deciding how to allocate money across repayment of credit cards with varying APRs. However, many pay little attention to interest rates. The only reason our result is statistically significant is because the sample size is so huge, which makes being even a tiny bit above 50% statistically higher than 50%: the standard error of the sample mean is tiny. This is an example where a result is not economically significant (the difference between 51% and 50% is negligible) but it is statistically significant because of the huge sample size.

(b) The formula and approach given in the question are wrong because they apply to *independent samples* and this case is clearly about *paired data*. Each person has their own credit card situation with respect to APRs, minimum payments due, balances owed, etc. Each person's actual payment will not be independent of the optimal payment for her/him to allocate to the high APR card that month. At a minimum, people in substantial debt will have both higher actual and optimal payments. In fact, the standard deviation of the difference is much smaller than the standard deviation of either the actual or optimal payments, which means there is a *strong* positive correlation. Instead, we should use a paired data approach: $\bar{X}_d \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}}$ with $\nu = n - 1$, which plugging in values will be $117.54 \pm 1.960\frac{422.14}{\sqrt{394,111}}$ for a 95% confidence interval estimate of the difference (noting that the very large number of degrees of freedom allows us to use 1.960).

(5) $H_0: (p_T - p_C) = 0$ versus $H_1: (p_T - p_C) > 0$, where T stands for treatment group (got the policy intervention of a notification letter) and C stands for control group (did not get the policy intervention). A Type I error would be concluding that the notification letters *are effective* in increasing the fraction of people taking vitamin D supplements when, in fact, the letters are useless (a waste of taxpayer money and paper). A Type II error would be not having sufficient evidence to prove that the letters are effective even though they really are boosting the rate at which people take vitamin D supplements.

(6) For each of the 160 countries we need to compute the growth rates for two decades (from 1996 – 2006 and from 2006 – 16), which means 320 (=160*2) regressions must be run on the raw data. The y variable is the natural logarithm of annual real GDP per capita, the x variable is the year, and the sample size is 11 (as it includes both endpoints for each decade: e.g. 1996 and 2006), which gives ln(GDP per capita) = a + b*year for n = 11. We need the OLS slope coefficient, b, which is a measure of the average annual growth rate of GDP per capita for that country in that decade, which will populate the variables in the data on which the regressions reported in Table 1 are run. (It makes no difference whether we take b or b*100, where the latter corresponds to the percentage growth rate.)