Homework 14: ECO220Y – SOLUTIONS

Required Problems:

(1) (a) Remember the P-value measures the strength of the evidence in favor of your research hypothesis: that evidence may be weak, strong, or non-existent depending on how your data turned out and what H_1 had said. The P-value is the probability that sampling error could have caused the sample to appear to be inconsistent with the null hypothesis (in the direction of your research hypothesis). For example recall Slide 25 (Lecture 13) regarding the sex of first babies (parity = 0) born in ON to Chinese-born moms where H0: p = 0.512; H1: p > 0.512. In words, the null hypothesis says that the proportion of males for first-time Chinese-born moms in ON is no different from what human biology would dictate. The research hypothesis says that the proportion of males is higher than the natural rate. Our sample of n = 12,339 babies born where 6,429 are male – the sample proportion of males (P-hat) is 0.521 – provides some support for our research hypothesis. Why? Because the proportion of males in our sample is in fact higher than what nature would predict. But is that enough evidence? Maybe it is just sampling error (by chance, slightly more of the 12,339 babies turned out to be male)? The P-value allows us to assess the strength of this evidence. Hence we ask, what is the probability of having such a high proportion of males if, in fact, the only thing affecting the sex is chance (sampling error)?

P-value = P(P-hat >= 0.521 | p = 0.512, n = 12,339) = ?

We use the sampling distribution of P-hat UNDER THE PRESUMPTION THAT H_0 IS TRUE. If H_0 were true then P-hat ~ Normal with mean 0.512 and s.d. 0.0045 (=(0.512)(1-0.512)/12,339)^0.5). Hence we standardize and find P-value = P(Z > 2.007) = 0.022. That means that we have fairly strong evidence in favor of our research hypothesis: it is pretty unlikely (2.2% chance) that we could end up with a random sample of 12,339 babies where so many were male just by chance given the natural birth rate of males. That is pretty unlikely, but not overwhelming evidence. If we had set the research hypothesis to be two-directional (H1: p ≠ 0.512) then our evidence would be weaker. This is because in this case the P-value must reflect the chance the sample proportion differs from the null by as much as it does (i.e. in either direction). The difference between 0.521 and 0.512 is 0.009 (i.e. just shy of one percentage point). Hence the P-value if the research hypothesis were two-tailed would be P-value = P(P-hat >=0.521) + P(P-hat <= 0.503). The chance the data would come out to be so different from the natural birth rate (so different naturally means in either direction) is 0.044 (=0.022*2).

(b) Recall our null and research hypotheses: H_0: p = 0.512 and H_1: p > 0.512. In the case of first babies of Indian-born moms the sample proportion (P-hat) of males is 0.510 for those 14,789 babies. This is an example where we have NO evidence in favor of our research hypothesis: the observed proportion is actually LESS THAN 0.512. Hence our P-value will be very big. How do we calculate it? P-value = P(P-hat > 0.5102441 | p = 0.512, n = 14,789) = P(Z > (0.510 – 0.5102441)/(0.512*(1-0.512)/14,789)^0.5) = P(Z > -0.43) = 0.665. In fact, in this example there is no point in even calculating the P-value because we obviously will be unable to conclude that there is gender selection in favor of boys if the proportion of boys is LESS than the biological norm. There is no way we could reject the null in favor of our research hypothesis: we do not need any fancy calculations to see that.

(c) Part (b) is a perfect example. (See also the HW 13 question on this example.) To give another example, suppose we wished to prove that a political candidate has more than a majority share of the votes: $H_0: p = 0.5$ vs. $H_1: p > 0.5$. If a poll (with sample size n) reveals s/he actually has only 30% support ($\hat{P} = 0.30$), then the P-value for $H_1: p > 0.5$ will be greater than 0.5 (we clearly have no evidence to support our research hypothesis). However, the P-value for the two-tailed test ($H_0: p = 0.5$ vs. $H_1: p \neq 0.5$) is not twice the P-value from the original one-tailed test (which would be a probability greater than 1!). Instead, we would compute the P-value for the two-tailed test as P-value = $P\left(\hat{P} < 0.30 \middle| p = 0.5, n\right) + P(\hat{P} > 0.70 | p = 0.5, n)$.

(2) Only the rejection region and not the test statistic.

(3)(a) $H_0: (p_{Match} - p_{No\ Match}) = 0$ versus $H_1: (p_{Match} - p_{No\ Match}) > 0$

(b) $H_0: (p_{Match} - p_{No\ Match}) = 0$ versus $H_1: (p_{Match} - p_{No\ Match}) < 0$

(c) $H_0: (p_{Match} - p_{No\ Match}) = 0$ versus $H_1: (p_{Match} - p_{No\ Match}) \neq 0$

(d) That means that in the data the response rate for those with a match was actually lower than the response rate of those with no match (in other words, a direct contradiction of our research hypothesis).

(e) The rejection region would be getting a z test statistic greater than 1.645.

(f) The rejection region would be getting a z test statistic less than -1.28.

(g) The rejection region would be getting a z test statistic less than -2.576 OR a z test statistic greater than 2.576.

(4) (a) This result is highly statistically significant, but is unlikely to be economically significant. Who cares if the redemption rate for coupons is 0.1 percentage points higher than the claim (15.1 versus 15 percent redeemed)? That is a tiny difference and unlikely to affect decisions. You may wonder how such a tiny difference can be statistically different from zero. The trick is the huge samples size (3 million!). With huge samples, even tiny differences will be statistically significant, but that does *not* mean that they are significant. Remember that significant means both statistically significant and economically significant.

(b) This result is highly economically significant but is not statistically significant. A male birth rate that is more than 10 percentage points higher than the natural rate would point to *extreme* sex selection in favor of male infants. This would be a demographic catastrophe. However, the result is not significant. You may wonder how such a huge difference is not statistically different from zero. The trick is the sample size (which the question did not disclose) must have been tiny. For example, suppose a family had 5 kids and 3 were male, which is 60 percent male. Obviously, we have no basis to conclude that that family engaged in sex selection. With a small sample size ($n = 5$), we could easily observe 60 percent male (or even 80 percent or 100 percent). However, we would not be able to reject the null and conclude there is sex selection.

(c) This result is significant: both statistically significant and economically significant. How do we know it is economically significant? As always, this assessment depends on the context. In the case of the male birth rate, a more than 2 percentage point higher male birth rate would definitely concern policy makers: that will be a lot of missing girls in a country's population. (In other contexts, 2 percentage points may not be economically significant.)

(d) This result is significant: both economically and statistically significant. (For example, Group 1 is Canadian born adults of working age and Group 2 is non-Canadian born adults of working age.)

(e) This result is not significant. It is not statistically significant and it is not economically significant. Why say it is not economically significant? Who cares about 22 calories of difference? A typical adult in Canada consumes about 2,000 calories per day: 22 is a mere 1 percent. That seems pretty small given the considerable costs of redoing menus to provide calorie information.