ECO220Y: Homework 13

Required Exercises: Chapter 12: 1, 3, 21, 23, 33, 35, 39, 81

Required Problems:

(1) A researcher wishes to prove that *fewer than* 20% of Canadians are in favor of legalizing heroin. Suppose that in a random sample of 400 Canadians, 40 are in favor. What should we conclude? Make sure to include the hypotheses in formal notation, a quantitative analysis of the strength of the evidence, and an interpretation of the results.

(2) A researcher wishes to prove that *more than* 15% of e-mails are junk. Suppose that in a random sample of 100 e-mails, 16 are junk. What should we conclude? Make sure to include the hypotheses in formal notation, a quantitative analysis of the strength of the evidence, and an interpretation of the results.

(3) A researcher wishes to prove that the fraction of Ontario Grade 3 students scoring above the 98^{th} percentile on a standardized intelligence test *differs* from 2%. Suppose that in a random sample of 1,000 Grade 3 Ontario students, 30 score above the 98^{th} percentile. What should we conclude? Make sure to include the hypotheses in formal notation, a quantitative analysis of the strength of the evidence, and an interpretation of the results.

(4) An economist wants to show that the supply of electricity is inelastic. What are the null and research hypotheses?

(5) An economist wants to show that the demand of electricity is inelastic. What are the null and research hypotheses?

(6) Recall the sex ratio at birth example from lecture: we used complete data for Ontario births from 2002 – 2007 from the article "Sex ratios among Canadian liveborn infants of mothers from different countries" published in the *Canadian Medical Association Journal* in 2012 by Ray et al http://www.cmaj.ca/content/184/9/E492. The natural human proportion of boys born is 0.512 (105 boys for every 100 girls). One subgroup we looked at is mothers (now in Ontario) who were born in India. We broke this out by parity (how many children the woman already had delivered at the time of the birth in question). Here are some of the results that we looked at:

> Parity = 0, Indian, n = 14,789 births, 7,546 males; $\hat{P}$ = 0.510, P-value = 0.665
> Parity = 1, Indian, n = 13,076 births, 6,873 males; $\hat{P}$ = 0.526, P-value = 0.001
> Parity = 2, Indian, n = 3,268 births, 1,883 males; $\hat{P}$ = 0.576, P-value = 1.1×10^{-13}

> **(a)** For Parity = 0, write out the relevant hypotheses using formal notation. Compute the P-value yourself. (Note: I got these P-values using STATA but you should get them from your Standard Normal table, so your answer can differ a little. However, you can get the exact value in Excel. Remember that you *cannot* round P-hat off to the third decimal place and still get a P-value accurate to the third decimal place.)

> **(b)** In which case do we have the strongest evidence in favor of the research hypothesis?

> **(c)** Answer this multiple-choice question *and* explain your answer with TWO graphs. One graphs must show the P-value for the original right-tailed test and another must show the P-value for the proposed two-tailed test.

> For the first babies (parity = 0) of Indian-born mothers, for H_0: p = 0.512 and H_1: p > 0.512 the P-value is 0.6654. Given this, what would the P-value be if H_0: p = 0.512 and H_1: $p \neq 0.512$?

> **(A)** 0.3327
> **(B)** 0.3346
> **(C)** 0.6654
> **(D)** 0.6692
> **(E)** 1.3308

(7) A researcher is testing the hypothesis that more than 10% of parts are defective using a random sample of 100 parts. She obtains a z test statistic of -1.33. Write down the formal hypotheses in standard notation. What must have been the sample proportion? What is the P-value? Conclusion?

(8) Consider this quote from R. A. Fisher:

> "In relation to any experiment we may speak of this hypothesis as the 'null hypothesis,' and it should be noted that the null hypothesis is never proved or established, but is possibly disproved, in the course of experimentation. Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis."

Suppose in testing the fairness of a coin, you tossed it 100 times and got 50 heads. What are the hypotheses in formal notation? What do you conclude?

(9) Consider this quote from paragraph 32 of TK71 (a copy of that article is available on the Readings page in Quercus). The quote compares and contrasts the two methods of statistical inference: confidence interval estimation (introduced in Lecture 12) and hypothesis testing (introduced in Lecture 13):

> "In the early psychological literature, the convention prevailed of reporting, for example, a sample mean as $\bar{X} \pm ME$, where ME is the margin of error (i.e., the 50% confidence interval around the mean). This convention was later abandoned in favor of the hypothesis testing formulation. A confidence interval, however, provides a useful index of sampling variability, and it is precisely this variability that we tend to underestimate." TK71, ¶32

Which method of statistical inference – confidence interval estimation or hypothesis testing – does TK71 favor?