ECOM20001: Econometrics 1

Tutorial 4: Suggested Solutions

Part 1: Hypothesis Testing and Confidence Intervals for Sample Means in R

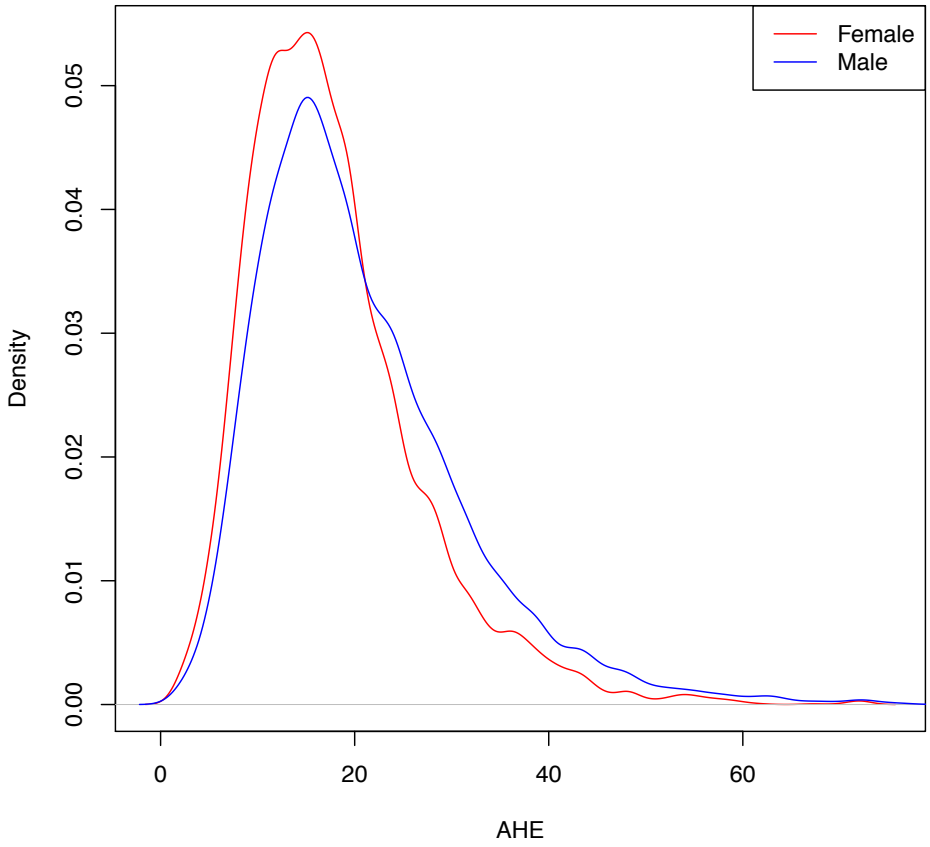
1. Question 1

   - Mean AHE for males is $20.58, Std. Dev. Is $10.55.
   - Mean AHE for females is $17.81, Std. Dev. is $8.87.
   - The difference is $20.58-$17.81=$2.77 average earnings per hour
   - This gender earnings gap is clear visually in ahe_female.pdf by the additional probability mass in the right part of the earnings distribution of males compared to females (next page).
   - What are potential economic explanations for this gender earnings gap?
     - The impact of differential child-rearing on long-run career earnings of females vs males. This makes it more difficult/costly in the long-run for females to have the same average hourly earnings as males.
     - The “glass ceiling” at work. Among other explanations, it is often speculated that females on average tend to be less aggressive in negotiating pay with employers than males for equal jobs (all else equal). That is, it can be thought that is potentially more costly for females to negotiate than males in the workplace

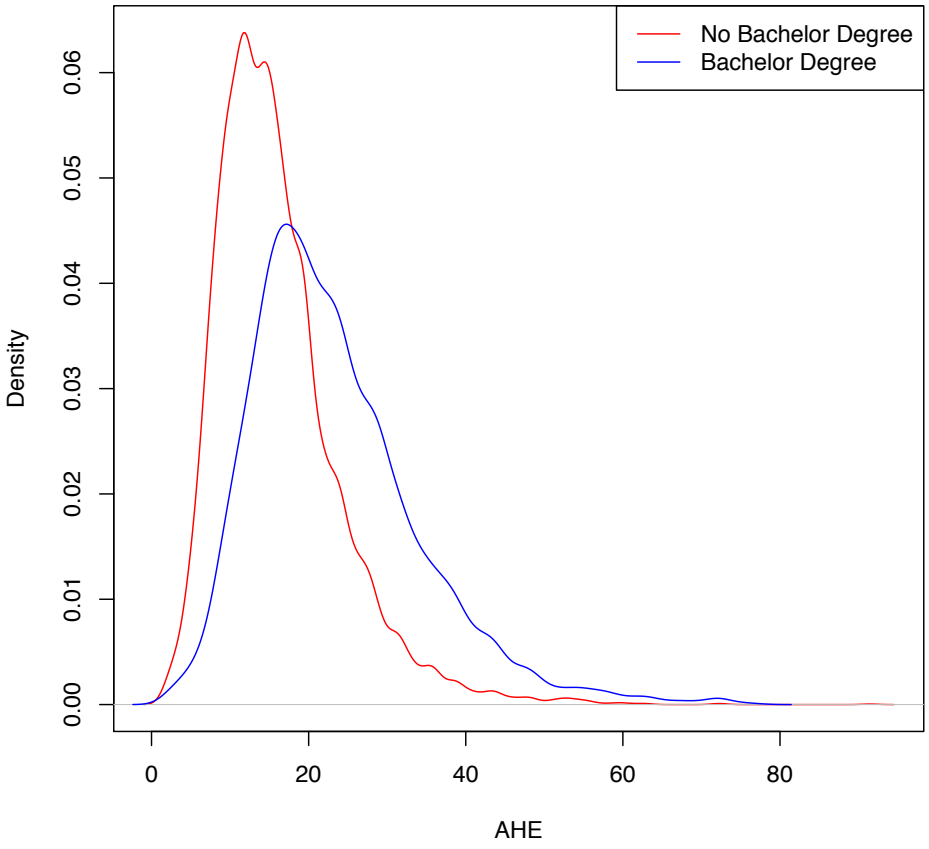
2. Question 2

   - Mean AHE for bachelor degree is $23.35, Std. Dev. Is $10.72.
   - Mean AHE for no bachelor degree is $16.05, Std. Dev. is $7.86.
   - The difference is $23.35-$16.05=$7.30 average earnings per hour
   - The gap in the earnings distributions for people without and with bachelor degrees in ahe_bachelor.pdf makes the gap very clear visually (next page).

Gender and Earnings



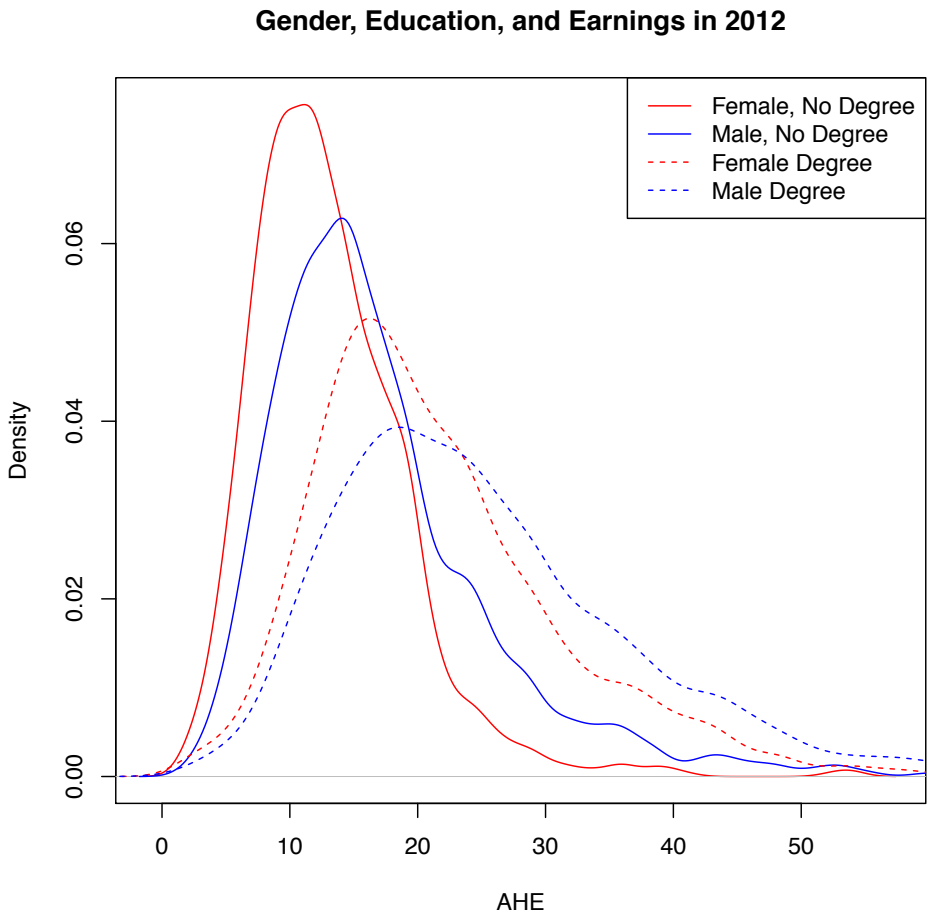
Education and Earnings



- What are potential economic explanations for this education earnings gap?
  - Individuals with bachelor degrees are able to signal higher effort levels and worker quality to employers and able to compete for better paying jobs in the labour market. That is, there is more expected benefit from hiring students with bachelor degrees in the labor market for better paying jobs
  - Individuals with bachelor degrees have more intrinsic ability on average, and are able to progress more rapidly and higher up the career ladder. That is, it is intrinsically less costly for people with bachelor degrees to progress into higher paying roles in their career.

3. Question 3
   - H0: mean(AHE) = 19.5 vs H1: mean(AHE) != 19.5 (!= means “not equal”)
     - T-statistic is -1.22, and |-1.22| < 1.96 (95% critical value for a 2-sided test), so fail to reject the null
   - H0: mean(AHE) = 19.5 vs H1: mean(AHE) > 19.5
     - T-statistic is -1.22, and -1.22 < 1.65 (95% critical value for a one-sided test for a greater than test), so fail to reject the null
   - H0: mean(AHE) = 19.5 vs H1: mean(AHE) < 19.5
     - T-statistic is -1.22, and -1.22 > -1.65 (95% critical value for a one-sided test for a less than test), so fail to reject the null
4. Question 4
   - H0: mean(AHE_Female_2012_NoBach) = mean(AHE_Male_2012_NoBach) vs H1: mean(AHE_Female_2012_NoBach) != mean(AHE_Male_2012_NoBach)
     - Difference in sample means is $13.11-$17.04=$3.93
     - T-statistic for the test of difference in means is -15.361, and |-15.361|>1.96 so reject the null at the 5% level that the difference of means equals 0
     - 95% confidence interval (CI) is [-$4.42, -$3.42], which means with 95% confidence we cannot reject a gender earnings gap among people without a bachelor’s degree as low as -$4.42.
5. Now discuss the difference in mean AHE for 2012 from a two-sample t-test for the gender earnings gap among males and females <u>with</u> bachelor degrees:

- H0: mean(AHE_Female_2012_Bach) = mean(AHE_Male_2012_Bach) vs H1: mean(AHE_Female_2012_Bach) != mean(AHE_Male_2012_Bach)
  - Difference in sample means is $21.50-$25.30=$3.80
  - T-statistic for the test of difference in means is -10.79, and |-10.79|>1.96 so reject the null at the 5% level that the difference of means equals 0
  - 95% CI is [-$4.49, -$3.11]
  - So we see our estimated gender earnings gap falls from $3.93 to $3.80 when we go from people without bachelor degrees to people with bachelor degrees. In short education appears to make the gender earnings gap smaller. The graph in ahe_female_bachelor_2012.pdf has a smaller gap in the means among males and females with bachelor degrees.
  - Both the shifts in the distributions of earnings without and with bachelor degrees for males and females again highlight the importance of education for earnings (below).
    - One possible economic explanation for this is among women with bachelor's degree, there is a smaller propensity to have as many children, and hence less disruption in their careers due to children, which would mean a smaller gender earnings gap among people with bachelors degrees.



Part 2: Hypothesis Testing and Confidence Intervals Practice Problems

1. Answers

a. |t_act|=1.66<1.96, so fail to reject null at 5% level. P-value is computed as 2 x Phi(-|1.66|)=0.097 > 0.05, which equivalently means fail to reject null at the 5% level (where Phi() is the cumulative density function of the N(0,1) distribution)

b. 95% CI: [27.83, 28.67]

c. - n=2,500: t=1.17, p-value=0.242, 95% CI=[27.83, 28.67], CI width=0.84
- n=5,000: t=1.66, p-value=0.097, 95% CI=[27.95, 28.55], CI width=0.59
- n=10,000: t=2.35, p-value=0.019, 95% CI=[28.04, 28.46], CI width=0.42
- n=20,000: t=3.32, p-value=0.001, 95% CI=[28.10, 28.40], CI width=0.30

So we see p-values shrink as sample size grows. Intuitively, as our estimate of the sample mean becomes more precise with larger samples (which reduces the standard error of the mean), we are increasingly able to reject even small differences between the sample mean and hypothesised values of the sample mean.

The CI width shrinks as the sample size goes up for the same reason: we maintain the same degree of confidence in the location of the population mean (95%) for a smaller interval because we have a more precise estimate of the sample mean with larger sample sizes. That is, larger samples yield tighter CIs in which the population mean is likely to be.

2. Answers

a. 95% CI: [30.48,31.28], CI width=0.80. We have a wider CI for NSW than VIC because the sample size of VIC is larger and the sample standard deviation of the data is smaller. This implies a more precise estimate of the sample mean for VIC and a tighter CI around it.

b. Two-sample t-test yields a t-statistic of -10.34, with |-10.34| > 1.96 implying we reject the null that mean AHE in VIC and NSW are equal. The p-value for the test is less than 0.000001.

c. 95% CI for the difference in means is [-3.12, -2.13]

3. Answers

a. The 90% CI does not contain $\boldsymbol{m}$=10. With a p-value of 0.07, you would reject the null hypothesis at the 10% level. That is, m=10 does not sit within the region of null values that would not be rejected at the alpha=10% level, which is what defines a (1-alpha) = 90% confidence interval.

b. You cannot determine if $\boldsymbol{m}$=8 lies within the 90% CI with the data provided. By knowing the p-value of 0.07, you can compute the magnitude of the t-statistic for the test from the CDF of the N(0,1) distribution. However, even if you know the t-statistic, you will not know what is in the numerator and the denominator, which would be needed to figure out the sample mean and standard error of the mean, which in turn is required to compute a 90% CI.