EXAM 1

Convocatoria Extraordinaria

Miguel A. Delgado

02/07/20

Question 1 (Solution) We can get the average directly, but we prefer to use the regression model in order to answer both questions at once,

$$children = \gamma_0 + \gamma_1 electric + error.$$

The population mean for those with electricity is $\mu_1 = \gamma_0 + \gamma_1$ and the mean for those without electricity is $\mu_0 = \gamma_0$. Therefore, $\gamma_1 = \mu_1 - \mu_0$ is the difference between the means. The estimated model (heteroskedasticity robust SE in parenthesis) is

$$\widehat{children}_i = \underset{(0.0372089)}{2.32773} - \underset{(0.0818574)}{0.429202} \cdot electric_i.$$

Thereforte, the average children for those with electricity is $\hat{\mu}_1 = 2.32773 - 0.42920 = 1.8985$, and the average for those without electricity is $\hat{\mu}_0 = 2.32773$. In order to test that the population means are the same, it suffices to test

$$H_0 : \gamma_1 = 0 \text{ vs } H_1 : \gamma_1 \neq 0,$$

The t ration is $t = -0.429202/\, 0.0818574 = -5.2433$. Therefore, we reject H_0 at any significance level. But we cannot infer that electricity "causes" that women have more children. The absence of electricity reflects a degree of deprivation, highly correlated with a lack of education, as well as cultural and behavioural issues. Therefore, there is an omitted variable problem, and the significance of *electric* may be *spurious*. Therefore, we cannot conclude that there is a causal relation between *children* and electricity.

Question 2 (Solution) In the model

$$children = \beta_0 + \beta_1 age + \beta_2 educ + \beta_3 electric + U,$$

the $\beta_3's$ OLS estimate of is -0.361758, smaller in absolute value than in 1, with a robust SE equal to 0.0637644. Therefore, the $electric's$ effect is significant at any significance level. The augmented model is

$$\begin{aligned} children &= \beta_0 + \beta_1 age + \beta_2 educ + \beta_3 electric + \beta_4 age^2 + \beta_5 urban \\ &\quad + \beta_6 catholic + \beta_7 protest + \beta_8 spirit + error, \end{aligned}$$

The $\beta_3's$ OLS estimate is now -0.305719 with SE 0.0640662. The partial effect is still significant. The one side hypothesis for the equality of $catholic$ and $protest$ coefficients is expressed as

$$H_0 : \beta_6 = \beta_7 \text{ vs } H_1 : \beta_6 > \beta_7.$$

Define the artificial parameter $\theta = \beta_6 - \beta_7$, once we sustitute $\beta_6 = \theta + \beta_7$ we get

$$\begin{aligned} children &= \beta_0 + \beta_1 age + \beta_2 educ + \beta_3 electric + \beta_4 age^2 + \beta_5 urban \\ &\quad + (\theta + \beta_7) \cdot catholic + \beta_7 protest + \beta_8 spirit + error \\ &= \beta_0 + \beta_1 age + \beta_2 educ + \beta_3 electric + \beta_4 age^2 + \beta_5 urban \\ &\quad + \theta \cdot catholic + \beta_7 \cdot (catholic + protest) + \beta_8 spirit + error \end{aligned}$$

The $\theta's$ OLS estimate is 0.0419578 with robust SE 0.0787032, the t ratio is 0.5331 and the $p - value$ for the one sided test is 0.297. Therefore, we don't reject H_0 at any reasonable significance level.

Question 3 (Solution) Now the model is

$$\begin{aligned} children &= \beta_0 + \beta_1 age + \beta_2 educ + \beta_3 electric + \beta_4 age^2 + \beta_5 urban \\ &\quad + \beta_6 catholic + \beta_7 protest + \beta_8 spirit + \beta_9 educ \cdot electric + error, \end{aligned}$$

The significance of *electric* partial effect hypothesis is

$$H_0 : \beta_3 = \beta_9 = 0 \text{ vs } H_1 : \beta_3 \neq 0 \text{ or/and } \beta_9 \neq 0,$$

the robust F statistic is $F = 16.4238$ with $p-value = 7.83508 \cdot 10^{-8}$. Therefore, we reject H_0 at any significance level. The difference between the two women is

$$\delta = \beta_2(4-7) + \beta_1(30-25) + \beta_4(30^2 - 25^2) = 5\beta_1 - 3\beta_2 + 275\beta_4.$$

Now, using OLS $\hat{\delta} = 5\,(0.341955) - 3\,(-0.0761469) + 275\,(-0.00276111) = 1.1789$, and

$$\begin{aligned}
\widehat{Var}\left(\hat{\delta}\right) &= 5^2 \cdot \widehat{Var}\left(\hat{\beta}_1\right) + 3^2 \cdot \widehat{Var}\left(\hat{\beta}_2\right) + 275^2 \cdot \widehat{Var}\left(\hat{\beta}_4\right) \\
&\quad -2 \cdot 5 \cdot 3 \cdot \widehat{Cov}\left(\hat{\beta}_1, \hat{\beta}_2\right) + 2 \cdot 5 \cdot 275 \cdot \widehat{Cov}\left(\hat{\beta}_1, \hat{\beta}_4\right) \\
&\quad -2 \cdot 3 \cdot 275 \cdot \widehat{Cov}\left(\hat{\beta}_2, \hat{\beta}_4\right) \\
&= 5^2 \cdot \left(3.68601 \cdot 10^{-4}\right) + 3^2 \left(4.12954 \cdot 10^{-5}\right) + 275^2 \left(1.23239 \cdot 10^{-7}\right) \\
&\quad -2 \cdot 5 \cdot 3 \cdot \left(-8.26600 \cdot 10^{-7}\right) + 2 \cdot 5 \cdot 275 \cdot \left(-6.67575 \cdot 10^{-6}\right) \\
&\quad -2 \cdot 3 \cdot 275 \cdot \left(5.47479 \cdot 10^{-8}\right) \\
&= 4.8278 \times 10^{-4}
\end{aligned}$$

Therefore, the confidence interval is

$$1.1789 \pm 1.96 \cdot \sqrt{4.8278 \times 10^{-4}} = (1.1358, 1.2220)\,.$$

Full credits if there are errors in the calculations, but the expression for the confidence interval is correct.

Variables in FERTIL2

1. *mnthborn* : month woman born
2. *yearborn* : year woman born
3. *age* : age in years
4. *electric* =1 if has electricity
5. *radio* =1 if has radio
6. *tv* =1 if has tv
7. *bicycle* =1 if has bicycle
8. *educ*: years of education
9. *ceb*: children ever born
10. *agefbrth*: age at first birth
11. *children*: number of living children
12. *knowmeth* =1 if know about birth control
13. *usemeth* =1 if ever use birth control
14. *monthfm*: month of first marriage
15. *yearfm*: year of first marriage
16. *agefm*: age at first marriage
17. *idlnchld*: 'ideal' number of children
18. *heduc*: husband's years of education
19. *agesq*: age^2
20. *urban*=1 if live in urban area
21. *urbeduc*: urban*educ
22. *spirit*=1 if religion == spirit
23. *protest*=1 if religion == protestant
24. *catholic* =1 if religion == catholic
25. *frsthalf*=1 if mnthborn <= 6
26. *educ*0 =1 if educ == 0

27. $evermarr$ =1 if ever married