• Chonwanee katemukda 6004641061

Charpter 6 #3

i) we can take derivative of rdintenes with respect to sale and set it equal to 0

$0.000000014\ sales^{*} = 0.0003$

$sales = 21{,}428.57$

The we know that at $sales = 21{,}428.57$ the marginal effect of sales on rdindtens become (-).

ii) The t-static on $\hat{\beta}_{sales^2}$ is $-.00000007 / .0000000037 = -1.89$. which significant against the one-sided alternative $H_1: \beta_{sales^2} < 0$.

iii) $rdintens = 2.613 + .0003\,sales - .000000007\,sales^2$

$= 2.613 + 0.0003(1000 * salesbil) - .000000007(1000 * salesbil)^2$

$= 2.613 + 0.3\,salesbil - .007\,salesbil^2$

(S.d.) (.429) (.14) (.0037)

recall that $s.e.(\hat{\beta}_j) = \dfrac{\hat{\sigma}^2}{[SST_j(1-R^2)]^{\frac{1}{2}}}$ (3.58). Rescaling sale will have no effect on $\hat{\sigma}$ or R^2

since it does not change the fit of the regression. It will, however, affect SST_{sales} and SST_{sale^2}. Specifically, $SST_{salebil} = \sum_{i=1}^{n}(salesbil - \overline{salebil})^2$

$= \dfrac{\sum_{i=1}^{n}(sales - \overline{sales})^2}{1000^2}$

$= \dfrac{SST_{sales}}{1000^2}$ similarly $SST_{salesbil^2} = \dfrac{SST_{sales^2}}{1000^4}$

∴ we need to scale the standard error of $\hat{\beta}_{salebil}$ and $\hat{\beta}_{salebil^2}$ up by 1000 and 1000^2 respectivly.

iv) equation in part iii because it's easier to read due to it contain fewer zeros to the right of decimal. Of course the interpretation of the two equations is idetical once of different scales are accounted for.

Chapter 7 #1

i) in the regression equation, the dependant variable is sleep and the independent variables are total work, education, age and dummy male. The coefficient on male is 87.75, so a man is estimated to sleep almost one and one-half more per week than a comparable woman. Further, $t_{male} = \dfrac{87.75}{34.33} \approx 2.56$, which is close to the 1% critical value against a two-sided alternative (about 2.58). Thus, the evidence for a gender differential is fairly strong.

ii) The t static on totwrk is $\dfrac{-.163}{.018} \approx -9.06$, which is very statistically significant. The coefficient implies that one more hour of work (60 minutes) is associated with $0.163(60) = 9.8$ mins. less sleep.

iii) The null we are interested in testing here is that the coefficients on age and age^2 are jointly zero. So we would need to run a restricted version of the regression above where these two variables are omitted and calcul atate.

Chonwanee Katemukda 6004641061

Chapter 7 #8

i) $\log(wage) = \beta_0 + \beta_1 useage + \beta_2 exper + \beta_4 female + u$

ii) $\log(wage) = \beta_0 + \beta_1 useage + \beta_2 educ + \beta_3 exper + \beta_4 female + \beta_5 useage \cdot female + u$

Testing that there are no difference in the effect of drug useage for men and women would involve testing $H_0: \beta_5 = 0$ against

$H_a: \beta_5 \neq 0$

iii) Assuming no interaction effect between useage and sex, the model would look like.

$\log(wage) = \beta_0 + \beta_1 light + \beta_2 moderate + \beta_3 heavy + \beta_4 educ + \beta_5 exper + \beta_6 female + u$

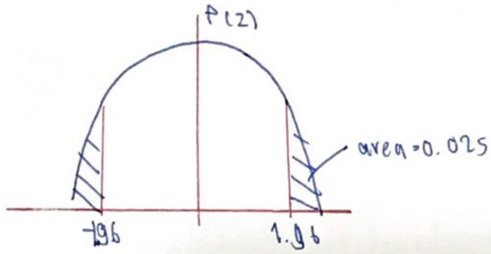
in this model, non-user is the omitted category.

iv) The null hypothesis here is $H_0: \beta_1 = \beta_2 = \beta_3 = 0$. Naturally this is going to be an F-test on $q = 3$ restrictions. we are also going to have degrees of freedom $df = n-6-1$ for a sample of size n, since we have 6 independent variables in the unrestricted model. So we would be obtaining a critical value from the $F_{q,n-7}$ distribution.

v) - There may be omitted variables which determine both marijuana useage and wages. For example, people living in urban areas may have easier access to marijuana and may earn higher wages on average. In this example our estimate would be downward biased.

Chapter 7 #11

i) A male student have 3.83 more score than a female student.



d.f. = 856 - 2 - 1 = 853

95% confidence interval = [3.83 - 1.96 (0.74), 3.83 + 1.96 (0.74)]
= [2.3796, 5.2804]

95% confidence interval for β_{male} is excluding zero.

ii) from 3rd model, $H_0: \beta_1 = \beta_3 = 0$
H_a: otherwise

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/n-k-1} = \frac{R^2_{ur} - R^2_r / q}{(1-R^2_{ur})/n-k-1} = \frac{(0.349 - 0.329)/2}{(1-0.349)/856-3-1} = 13.0876$$

$F = 13.0876 > 2.6$, we reject H_0 at 5% level so gender has an impact on score.

iii) the interaction between male and deviation of colgpa from mean may be better explain

Chonwanee Katemuked

Chapter 7

i) the two signs that are pretty clear are $\beta_3 < 0$ (because hsperc is defined so that the smaller the number the better the student) and $\beta_4 > 0$. The effect of size of graduating class (β_1 and β_2) is not clear. It is also unclear whether males and females (β_5) have systematically different GPAs. we may think that $\beta_6 < 0$, that is, athletes do worse than other students with comparable characteristics. But remember, we are controlling for ability to some degree with hsperc and sat.

ii) $colgpa = 1.241 - 0.569\ hsize + .00468\ hsize^2 - .0132\ hsperc + 0.00165\ sat + 0.155\ female + 0.169\ athlete$

$n = 4{,}137,\ R^2 = 0.293.$

Holding other factors fixed, an athlete is predicted to have a GPA about .169 points higher than a nonathlete. The t statistic $\frac{0.169}{0.042} = 4.02$ which is very significant

iii) with sat dropped from the model, the coefficient on athlet becomes about .0054 (se ≈ 0.0448). which is not significant. This happens because we don't control for SAT scores, and athletes score lower on average than nonathletes. part ii shows that, once we account for sat differences, athletes do better than nonathletes. Even if we don't control for SAT score, There is no difference

iv) To facilitate testing the hypothesis that there is no difference between woman athletes and women nonathletes, we should choose one of these as the base group. we choose female nonathletes.

The estimate equation is

$$colgpa = 1.396 - 0.0568\ hsize + .00467\ hsize^2 - 0.0132\ hsperc + 0.00165\ sat + 0.175\ femath + 0.013\ maleath - 0.155\ malenonath$$

$n = 4{,}137 \quad R^2 = .293$

The coefficient on female · athlete shows that colgpa is predicted to be about .175 points higher for a female athlete than female nonathlete. other variable in the equation fixed. The hypothesis that there is no difference between female athlete and female non-athletes is testing by using the t statistic on femath. In this case, t = 2.08, which is statically significant at the 5% level against a 2 sides altenative.

v) whether we add the interaction female · sat to the equation in part (ii) or part (iv), the outcome is practically the same. For example, when female · sat is added to the equation in part (ii), its coefficient is about .000051 and its t statistic is about .40. There is very little evidence that the effect of sat differs by gender.