Woradech
6104640641

HW 6

Chapter 6

3i.) $\frac{\partial \widehat{rdintens}}{\partial sales} = 0.0003 - 0.00000014\, sales = 0$

$0.00000014\, sales = 0.0003$

$sales = 21{,}428.57$

$\therefore$ The point that the marginal effect of sales on rdintens becomes negative is 21,428.57.

3ii.) Since the t-stat on $\hat{\beta}_{sales2} = \frac{-0.000000007 - 0}{0.0000000037} = -1.89$

$t_{0.05,29} = -1.7$ (p-value = 0.036), it is significant against the one-sided alternative $H_0: \beta_1 < 0$ at 5% sig level. In conclusion, we probably keep the model in quadratic term.

3iii.) In order to transform from sales to salesbil, we have to divide sales by 1000. Consequently, we have to multiply back to coefficient by 1000 $(0.00030)(1000) = 0.3$ and se will also have to multiply by $1000 = (0.00014)(1000) = 0.14$. For salesbil^2, we have to divide by 1,000,000 to transform from sales^2 to salesbil^2. Therefore, we have to multiply the coefficient back by $1{,}000{,}000 = (0.0000000070)(1{,}000{,}000) = 0.0070$ and the se will be multiply by the same number so the new s.e. will be 0.0039. However, the R-squared will remains constant.

$$\widehat{rdintens} = \underset{(0.429)}{2.613} + \underset{(0.14)}{0.30}\, salesbil - \underset{(0.0039)}{0.0070}\, salesbil^2$$

$n = 32$, $R^2 = 0.1484$

3iv). I prefer the second equation (The equation that has been transform to salesbil and salesbil^2). The reason is it contains less decimal places why the intepretation is still the same since it only changes the scales.

Chapter 7

1i). The coefficient for men is 87.75 which means a man is estimated to sleep almost one and one-half hours more per week than woman. Moreover, the t-stat for men is around 2.56 which is close to $t_{0.005,\infty} = 2.58$.

$\therefore$ The evidence to proof is quite strong.

1ii). t-stat for totwrk $= \dfrac{-0.163-0}{-0.018} = 9.056$ which is very significant. The coefficient shows that if one more hour has been spent on working, the time to sleep will be lessened by $(0.163)(60) = 9.78$ minutes.

1iii). In order to get R^2_r, we need to estimate the model by removing age and age^2 since both of them doesn't make impact (parameter = 0).

8i). $\log(wage) = \beta_0 + \beta_1 usage + \beta_2 educ + \beta_3 exper + \beta_4 exper^2 + \beta_5 female + u$

$\therefore \beta_1(100)$ = The percentage change in way when the usage of marijuana is increased by 1 time per month.

8ii). $\log(wage) = \beta_0 + \beta_1 usage + \beta_2 educ + \beta_3 exper + \beta_4 exper^2 + \beta_5 female + \beta_6 female \cdot usage + u$

We have to choose one from whether men·usage or female·usage but the equation already contains female so we choose the latter one. $H_0: \beta_6 = 0$ (The effect of marijuana won't change by gender).

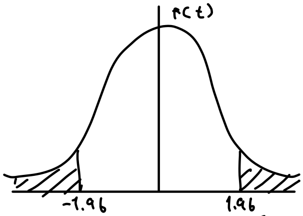
8iii). The base group that is chosen is nonuser and we have to includes three dummy variables (lghtuser, moduser, and hvyuser).

$$\log(wage) + \beta_0 + \delta_1 lghtuser + \delta_2 moduser + \delta_3 hvyuser + \beta_2 educ + \beta_3 exper + \beta_4 exper^2 + \beta_5 female + u$$

8iv.) $H_0: \delta_1 = 0 \quad \delta_2 = 0 \quad \delta_3 = 0$; so the $q = 3$ (Restricted). The sample size $= n-7-1 = n-8$. In this case, we have to use F-test, $F_{q,n-8}$ distribution.

8v.) The error term could contain factors, such as family background that could directly affect wages and also be correlated with marijuana usage. We are interested in the effects of a person's drug usage on his/her wage, so we would like to hold other factors constant.

11 i). The coefficient on male is 3.83 which means as one more male student increases, the average class score will increases by 3.83%, by holding other factors constant.



95% CI

The 95% CI for $\hat{\beta}_{male} = [3.83 \pm 1.96(0.74)]$

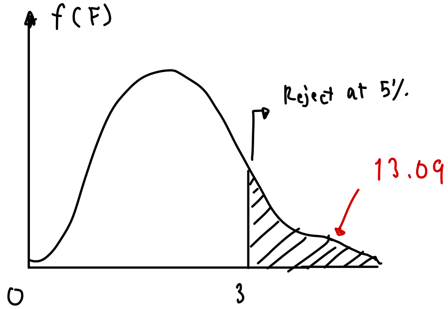
$= 5.2804, 2.3796$

$\therefore$ The CI exclude zero since it does not lies between 2.3796 and 5.2804.

11 ii) It is imprecise since there're too low number of variables so it might includes the error term. We can conclude that no gender has differences in score after controlling for colgpa by let $H_0: \hat{\beta}_{male} = \hat{\beta}_{male \cdot colgpa} = 0$

$F_{2,\infty}$ at $\alpha = 0.05 = 3$

$$F = \frac{R^2_{ur} - R^2_r / q}{1 - R^2_{ur} / (n-k-1)} = \frac{(0.349 - 0.329)/2}{(1-0.349)/(856-3-1)} \equiv \frac{0.01}{0.000761} \doteq 13.09$$



$\therefore$ We reject H_0 at 5% significant level which means there's gender difference on score at 5% significant level

11 iii). As it is given in the question that average colgpa = 2.81 so if we substitute the number into the last equation, 0.479 male (colgpa - 2.81) will be cancelled out. Therefore, the equation will be similar to second equation so that's why the coefficient on male for both equation are much closer compared to third equation. In the third equation, there're more variables which means the coefficient will be more fluctuate. Moreover, the number will be just as precisely estimated since there're some changes in the coefficients by include more significant variable to make the number more precise.

Computer

C4 i) β_1, β_3, and β_6 are all expected to have negative effect since when all of them increase, colgpa would decrease while β_4 is expected to have positive effect. β_2 and β_5 are unclear since they are quadratic form (square) and dummy variable (qualitative) respectively.

C4 ii).

$$\widehat{colgpa} = \underset{(0.079)}{1.241} - \underset{(0.0164)}{0.059}\,hsize + \underset{(0.00225)}{0.00468}\,hsize^2 - \underset{(0.0006)}{0.0132}\,hsperc + \underset{(0.00007)}{0.00165}\,sat + \underset{(0.018)}{0.155}\,female + \underset{(0.042)}{0.169}\,athlete$$

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

Athelete is predicted to have higher GPA than a nonathlete by about 0.169 points, at 5% significant level, by holding other factors constant.

C4 iii). After sat is dropped from the model, the coefficient on athlete now becomes about 0.0054 (the $P>|t|$ is 0.903 which is higher than α so it is not significant). This occurs because we don't control SAT scores. However, there'll no differences whether we control SAT scores or not.

C4 iv). Since the allowance to show effect of the difference being athlete by gender, we will divide into 4 hypotheses (maleath, malenonath, femath, and femnonath) but we will choose femnoath as base group.

$$\widehat{colgpa} = \underset{(0.076)}{1.396} - \underset{(0.0164)}{0.568}\,hsize + \underset{(0.00225)}{0.00467}\,hsize^2 - \underset{(0.0006)}{0.0132}\,hsperc + \underset{(0.00007)}{0.00165}\,sat + \underset{(0.084)}{0.175}\,femath + \underset{(0.049)}{0.013}\,maleath - \underset{(0.018)}{0.155}\,malenonath$$

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

The coefficient on female = femath shows that colgpa is predicted to be about 0.175 points higher for a female athlete than a femnonath, by holding other factors constant. The t-stat for femath is 2.08 so it is significant at the 5% significant level against a two-sided alternative which shows that there's no different between femath and femnonath.

C4 v). In order to justify the answer, we have to add female · sat to the equation. Therefore, there's effect of sat on colgpa differ by gender since 14.92 > 1.96 so we reject $H_0: \beta_{femalesat} = 0$ at 5% significant level.