Hayden C. Hancock

3/4/2020

Dr. Hwu

GBA 6230

Homework 4

1. The following model is estimated:

sleep d = 3840.83 − 0.163totwrk + 11.71educ + 8.70age + 0.128age2 + 87.75male,

(235.11) (0.018) (5.86) (11.21) (0.134) (34.33)

n = 706, R2 = 0.123

The variable sleep is total minutes per week spent sleeping at night, totwrk is total weekly minutes spent working, educ and age are measured in years, and male is a gender dummy

a) *All other factors being equal, is there evidence that men sleep more than women? Test the null hypothesis that men sleep the same amount of time as women against the alternative hypothesis that men sleep more than women at 5% significance level.*

Answer: There is evidence that men sleep more than women. Since male is a dummy variable comprised from gender. The Male slope estimator of 87.75 is added to the intercept while holding all else equal.

Using the standard error of the male slope estimator, we get a t-stat of 2.55 which is greater than the critical value of 1.96 at the 5% significance level. We reject the null hypothesis that men sleep the same amount of time as women.

(b) Is there a statistically significant tradeoff between working and sleeping? What is the estimated tradeoff?

Answer: From the totwrk slope estimator we can conclude for each minute increase in total work, the amount of sleep decreases by -.163. This is significant.

2. A model that allows major league baseball player salary to differ by position,

log(salary) = β0 + β1years + β2gamesyr + β3bavg + β4hrunsyr + β5rbisyr +β6runsyr + β7fldperc + β8allstar + β9frstbase +β10scndbase + β11thrdbase + β12shrtstop + β13catcher + u

where outfield is the base group.

(a) *State the null hypothesis that, controlling for other factors, catchers and outfielders earn, on average, the same amount. Test this hypothesis using the data in MLB1 and comment on the size of the estimated salary differential.*

Answer: Ho: bo=bo+b13

Comparing catcher to outfielder we can see that we would fail to reject the null hypothesis at a .05 significance level. That their salaries are different, p-value(.0543). The size of the estimated salary differential is .2535 if the player is a catcher.

```
> mlb1 <- read_excel("C:/Users/Hayden/Downloads/mlb1.xls")
> View(mlb1)
> summary(a <- lm(log(mlb1$salary)~mlb1$years+mlb1$gamesyr+mlb1$bavg+mlb1$hrunsyr+
+                     mlb1$rbisyr+mlb1$runsyr+mlb1$fldperc+mlb1$allstar+mlb1$frstbase+
+                     mlb1$scndbase+mlb1$thrdbase+mlb1$shrtstop+mlb1$catcher))

Call:
lm(formula = log(mlb1$salary) ~ mlb1$years + mlb1$gamesyr + mlb1$bavg +
    mlb1$hrunsyr + mlb1$rbisyr + mlb1$runsyr + mlb1$fldperc +
    mlb1$allstar + mlb1$frstbase + mlb1$scndbase + mlb1$thrdbase +
    mlb1$shrtstop + mlb1$catcher)

Residuals:
     Min       1Q   Median       3Q      Max
-2.42088 -0.42665 -0.03092  0.47925  2.74975

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   11.1295536  2.3044545   4.830 2.07e-06 ***
mlb1$years     0.0584178  0.0122732   4.760 2.87e-06 ***
mlb1$gamesyr   0.0097670  0.0033776   2.892  0.00408 **
mlb1$bavg      0.0004814  0.0011411   0.422  0.67340
mlb1$hrunsyr   0.0191459  0.0159638   1.199  0.23124
mlb1$rbisyr    0.0017875  0.0074755   0.239  0.81116
mlb1$runsyr    0.0118707  0.0045264   2.623  0.00912 **
mlb1$fldperc   0.0002833  0.0023078   0.123  0.90239
mlb1$allstar   0.0063351  0.0028828   2.198  0.02866 *
mlb1$frstbase -0.1328008  0.1309243  -1.014  0.31115
mlb1$scndbase -0.1611010  0.1414296  -1.139  0.25547
mlb1$thrdbase  0.0145271  0.1430352   0.102  0.91916
mlb1$shrtstop -0.0605672  0.1302031  -0.465  0.64210
mlb1$catcher   0.2535592  0.1313128   1.931  0.05432 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7092 on 339 degrees of freedom
Multiple R-squared:  0.6535,    Adjusted R-squared:  0.6403
F-statistic: 49.19 on 13 and 339 DF,  p-value: < 2.2e-16
```

b) *State the test the null hypothesis that there is no difference in average salary across positions, once other factors have been controlled for.*

Answer: Ho: B9=B10=B11=B12=B13=0

The p-value from the hypothesis test is (.1168) If we are assuming the same significance from question A of 5%, we would fail to reject the null hypothesis that there is no difference in average salary across the different positions.

c) *Are the results from part(a) and (b) consistent? If not, explain what is happening*.

Answer: In both cases they are consistent. We would fail to reject the null hypothesis and conclude that there is not a difference between salaries of the different player positions.

3. *Use the data in SLEEP75 for this question. The equation of interest is*

sleep = β0 + β1totwrk + β2educ + β3age + β4age2 + β5yngkid + u

(a) *Estimate this equation separately for men and women and report the results in the usual form. Are there notable differences in the two estimated equations?*

Answer:

For men:

sleep=(3840.85)Bo+(-0.16)totwrk+(-11.71)educ+(-8.697)age+(.12844)age^2+(-0.2280)yngkid+87.75455(male)

For women:
sleep=(3928.60662)Bo+(-0.16342)totwrk+(-11.71)educ+(-8.697)age+(.12844)age^2+(-0.2280)yngkid+(-87.75455)female

(b) *Consider a model including dummy variable male,*
sleep = β0+δ0male+β1totwrk+δ1male∗totwrk+β2educ+δ2male∗educ+β3age +δ3male ∗ age + β4age2 + δ4male ∗ age2 + β5yngkid + δ5male ∗ yngkid + u
Explain what does the null hypothesis δ0 = δ1 = δ2 = δ3 = δ4 = δ5 = 0 imply?

Answer: The given null hypothesis would imply that gender does not affect sleep.

d) C) Test the null hypothesis in part (b) at 5 percent level.

```
Linear hypothesis test

Hypothesis:
sleep75$male = 0
sleep75$male:sleep75$totwrk = 0
sleep75$male:sleep75$educ = 0
sleep75$male:sleep75$age = 0
sleep75$male:sleep75$agesq = 0
sleep75$male:sleep75$yngkid = 0

Model 1: restricted model
Model 2: sleep75$sleep ~ sleep75$male + sleep75$totwrk + sleep75$totwrk *
    sleep75$male + sleep75$educ + sleep75$educ * sleep75$male +
    sleep75$age + sleep75$age * sleep75$male + sleep75$agesq +
    sleep75$agesq * sleep75$male + sleep75$yngkid + sleep75$yngkid *
    sleep75$male

  Res.Df       RSS Df Sum of Sq      F  Pr(>F)
1    700 123267451
2    694 121052555  6   2214896 2.1164 0.04949 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

Answer: Testing our null hypothesis that gender interacting with totwork/educ/age/agesq/yngkid does not affect sleep at the 5% significance level, we would reject the null hypothesis with the resulting p-value of the linear hypothesis testing(.4949).