Questions 12 - 17

Labor economists have estimated Mincer equations that include not only total years of schooling, but also total experience as explanatory variables of the wage. Assume now that you want to estimate the following model:

$$log(wage_i) = \beta_0 + \beta_1 yrs\_school_i + \beta_2 total\ experience + \varepsilon_i$$

Question 12

1.0/1.0 point (graded)

If you run this model in R, what would be the value of the R^2?

Please round your answer to the third decimal place, i.e. if your answer is 0.7283, please round to 0.728 and if it is 0.7289, round to 0.729.

0.267 ✔ **Answer:** 0.267

0.267

Explanation

If we run the following code:

```r
#multivariable regression
multi1 <- lm(lwage ~ yrs_school + ttl_exp, data = nlsw88)
summary(multi1) # show results
```

This is the output that we get:

```
Call:
lm(formula = lwage ~ yrs_school + ttl_exp, data = nlsw88)

Residuals:
     Min       1Q   Median       3Q      Max
-2.09807 -0.29945 -0.00571  0.25158  2.49949

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.336944   0.057308    5.88 4.73e-09 ***
yrs_school  0.079148   0.004150   19.07  < 2e-16 ***
ttl_exp     0.039559   0.002296   17.23  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4921 on 2243 degrees of freedom
Multiple R-squared:  0.2671,    Adjusted R-squared:  0.2664
F-statistic: 408.7 on 2 and 2243 DF,  p-value: < 2.2e-16
```

From there, we know that the R^2 is 0.2671. This implies that 26.71% of the total variance in the logarithm of the wage is explained by the years of schooling and the total experience.

The Multivariate Linear Model
due Nov 28, 2016 05:00 IST

Module 9: Homework
due Nov 21, 2016 05:00 IST

▸ Exit Survey

Submit You have used 1 of 2 attempts

Some young folks are claiming that they prefer to drop out from school since each additional year of schooling changes the log of the wage in the same amount as one half year of experience. A group of parents are really worried. They ask you to conduct a formal test over this sample.

Question 13

1.0/1.0 point (graded)

What would be the null hypothesis of this test?

- a. The null hypothesis of this test is: $\beta_1 = 2\beta_2$
- b. The null hypothesis of this test is: $\beta_1 = \beta_2 + \beta_1$
- c. The null hypothesis of this test is: $\beta_1 + \beta_2 = \beta_2$
- d. The null hypothesis of this test is: $2\beta_1 = \beta_2$ ✔

Explanation

If the effect of one year of experience is equivalent to two years of education over the log of the wage. Then, the null hypothesis of this test is that $2\beta_1 = \beta_2$.

Submit You have used 1 of 2 attempts

Question 14

1.0/1.0 point (graded)

Which of the following would correspond to the restricted model under this null hypothesis? (Select all that apply)

- [ ] a. The model $log(wage_i) = \beta_0 + \beta_2(yrs\_school_i + 2total\ experience_i) + \varepsilon_i$
- [ ] b. The model $log(wage_i) = \beta_0 + \beta_1(\frac{1}{2}yrs\_school_i + total\ experience_i) + \varepsilon_i$
- [x] c. The model $log(wage_i) = \beta_0 + \beta_1(yrs\_school_i + 2total\ experience_i) + \varepsilon_i$
- [ ] d. The model $log(wage_i) = \beta_0 + (\beta_1 + 2\beta_2)yrs\_school_i + \varepsilon_i$
- [ ] e. The model $log(wage_i) = \beta_0 + (2\beta_1 + \beta_2)yrs\_school_i + \varepsilon_i$
- [x] f. The model $log(wage_i) = \beta_0 + \beta_2\left(\frac{1}{2}yrs\_school_i + total\ experience_i\right) + \varepsilon_i$

✔

Explanation

If we substitute the null hypothesis $(2\beta_1 = \beta_2)$ in the equation $log(wage_i) = \beta_0 + \beta_1 yrs\_school_i + \beta_2 total\ experience + \varepsilon_i$, then we have that:

$$log(wage_i) = \beta_0 + \beta_1 yrs\_school_i + \beta_2 total\ experience + \varepsilon_i$$

$$log(wage_i) = \beta_0 + \beta_1 yrs\_school_i + 2\beta_1 total\ experience + \varepsilon_i$$

$$log(wage_i) = \beta_0 + \beta_1 (yrs\_school_i + 2total\ experience_i) + \varepsilon_i$$

Analogously we have that:

$$log(wage_i) = \beta_0 + \beta_1 yrs\_school_i + \beta_2 total\ experience + \varepsilon_i$$

$$log(wage_i) = \beta_0 + \frac{\beta_2}{2} yrs\_school_i + \beta_2 total\ experience + \varepsilon_i$$

$$log(wage_i) = \beta_0 + \beta_2 \left(\frac{1}{2} yrs\_school_i + total\ experience\right) + \varepsilon_i$$

Submit You have used 1 of 2 attempts

Question 15

1.0/1.0 point (graded)

Estimate the restricted model in R. What is the value that you obtain for $\hat{\beta}_1$ in the restricted model? For the restricted model, use the first correct model in the list from question 15, i.e. if both (a) and (d) are correct, use (a) or if both (b) and (c) are correct, use (b).

Please round your answer to the fourth decimal place, i.e. if your answer is 0.78244, please round to 0.7824, and if it is 0.78247, please round to 0.7825.

0.0263 ✔ **Answer:** 0.0263

0.0263

Explanation

If we run the following code:

```
#Restricted model
nlsw88$newvar <- nlsw88$yrs_school + 2*nlsw88$ttl_exp
restricted <- lm(lwage ~ newvar, data = nlsw88)
summary(restricted) # show results
```

This is the output that we get:

```
Call:
lm(formula = lwage ~ newvar, data = nlsw88)

Residuals:
     Min       1Q   Median       3Q      Max
-1.79637 -0.32172 -0.02268  0.27505  2.39896

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.865430   0.042395   20.41   <2e-16 ***
newvar      0.026292   0.001075   24.47   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5106 on 2244 degrees of freedom
Multiple R-squared:  0.2106,    Adjusted R-squared:  0.2102
F-statistic: 598.6 on 1 and 2244 DF,  p-value: < 2.2e-16
```

Submit You have used 1 of 2 attempts

Question 16

0.0/1.0 point (graded)

Use the **anova** command in R to calculate the test $\frac{SSR_r - SSR_u}{r} / \frac{SSR_u}{N-K-1}$, what is the value of the test?

Please round your answer to the second decimal places, i.e. if your answer is 89.28397, please round to 89.28 and if it is 89.28997, round to 89.29

172.96 ✔ **Answer:** 172.96

172.96

Explanation
If we run the following code:

```
#multivariable regression
multi <- lm(lwage ~ yrs_school + ttl_exp, data = nlsw88)
summary(multi) # show results
anova_unrest <- anova(multi)

#Restricted model
nlsw88$newvar <- nlsw88$yrs_school + 2*nlsw88$ttl_exp
restricted <- lm(lwage ~ newvar, data = nlsw88)
summary(restricted) # show results
anova_rest <- anova(restricted)

#Test
statistic_test <- (((anova_rest$`Sum Sq`[2]-anova_unrest$`Sum Sq`[3])/1)
                   /((anova_unrest$`Sum Sq`[3])/anova_unrest$Df[3]))
statistic_test
pvalue <- df(statistic_test, 1, anova_unrest$Df[3])
pvalue
```

This is the output that we get:

```
> statistic_test
[1] 172.9599
> pvalue <- df(statistic_test, 1, anova_unrest$Df[3])
> pvalue
[1] 1.930469e-38
```

Submit You have used 2 of 2 attempts

✔ Correct (1/1 point)

Question 17

1.0 point possible (graded)

Do you reject or not reject this null hypothesis at a confidence level of 95%?

- (•) a. Reject ✔
- ( ) b. Do not reject

Explanation

The p-value associated with this test is less than 0.05. Then, we can reject the null hypothesis at this confidence level. You can also use the **car** package in R and the following code to perform the test directly:

```
matrixR <- c(0, -2, 1)
linearHypothesis(multi, matrixR)
```

```
Hypothesis:
- 2 yrs_school  + ttl_exp = 0

Model 1: restricted model
Model 2: lwage ~ yrs_school + ttl_exp

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1   2244 585.09
2   2243 543.20  1    41.887 172.96 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)