

EC 320 Problem Set 4

Winter 2022

1. (Textbook Question 6.6) 10 points

In a Monte Carlo experiment, a variable Y is generated as a linear function of two variables X_1 and X_2 :

$$Y = 10.0 + 10.0X_1 + 0.5X_2 + u$$

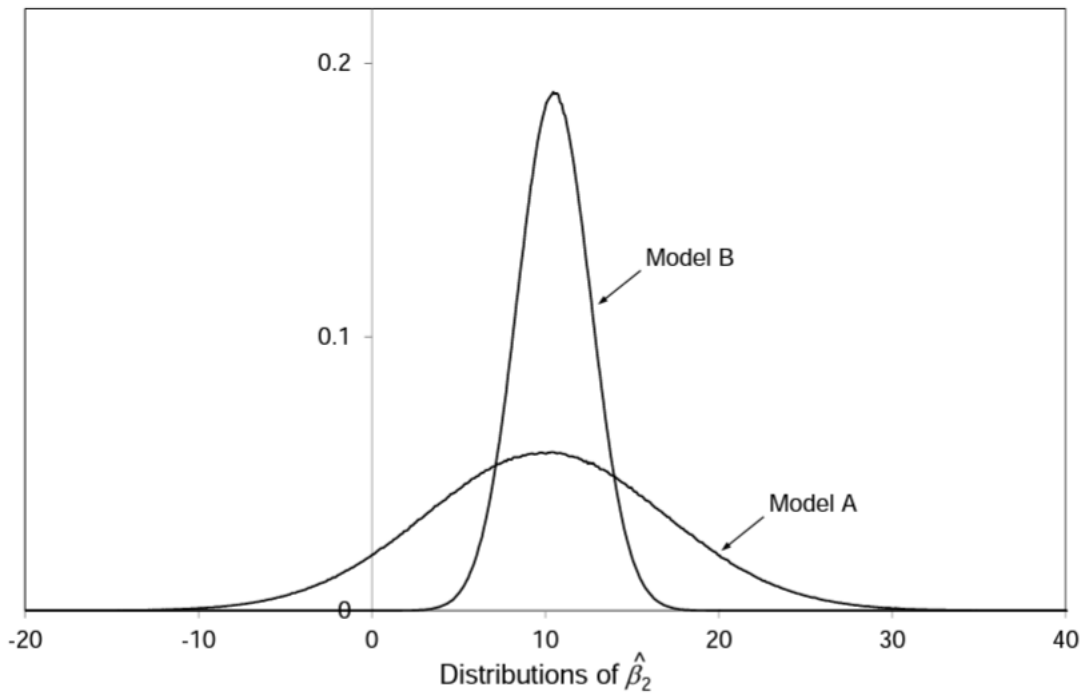
where X_1 is the sequence of integers $1, 2, \dots, 30$, X_2 is generated from X_1 by adding random numbers, and u is a disturbance term with a normal distribution with mean zero and variance 10,000. The correlation between X_1 and X_2 is 0.95. The table shows the result of fitting the following regressions for 10 samples:

$$\text{Model A : } \hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2$$

$$\text{Model B : } \hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1$$

The figure shows the distributions of $\hat{\beta}_1$ for the two models for 10 million samples. In the case of Model A, the distribution of $\hat{\beta}_1$ has mean 10.001 and standard deviation 6.910. For Model B, the mean is 10.500 and the standard deviation is 2.109. Comment on all aspects of the regression results, giving full explanations of what you observe.

Sample	Model A					Model B		
	$\hat{\beta}_2$	s.e. $(\hat{\beta}_2)$	$\hat{\beta}_3$	s.e. $(\hat{\beta}_3)$	R^2	$\hat{\beta}_2$	s.e. $(\hat{\beta}_2)$	R^2
1	10.68	6.05	0.60	5.76	0.5800	11.28	1.82	0.5799
2	7.52	7.11	3.74	6.77	0.5018	11.26	2.14	0.4961
3	7.26	6.58	2.93	6.26	0.4907	10.20	1.98	0.4865
4	11.47	8.60	0.23	8.18	0.4239	11.70	2.58	0.4239
5	13.07	6.07	-3.04	5.78	0.5232	10.03	1.83	0.5183
6	16.74	6.63	-4.01	6.32	0.5966	12.73	2.00	0.5906
7	15.70	7.50	-4.80	7.14	0.4614	10.90	2.27	0.4523
8	8.01	8.10	1.50	7.71	0.3542	9.51	2.43	0.3533
9	1.08	6.78	9.52	6.45	0.5133	10.61	2.11	0.4740
10	13.09	7.58	-0.87	7.21	0.5084	12.22	2.27	0.5081



2. (Textbook Question 6.8) 10 points

Following is the results of regressing *LGEARN* on *S*, *EXP*, *ASVABC*, *MALE*, *ETHBLACK*, *ETHHISP*. Now we repeat the regression adding *AGE*. (*LGEARN* denotes the logged hourly earnings, *S* represents years of schoolings, *EXP* represents the total out-of-school work experience (years), *ASVABC* represents scaled score on a component of the *ASVAB* test, *MALE* is a binary variable denoting male, *ETHBLACK*, *ETHHISP* are binary variables denoting certain ethnicity.)

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      2.13   11.53   15.49   18.44   21.79   132.89
```

```
## # A tibble: 7 x 5
##   term          estimate std.error statistic  p.value
##   <chr>          <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept)    0.977    0.194     5.04 6.62e- 7
## 2 S              0.0954   0.0106     8.99 5.35e-18
## 3 EXP            0.0431   0.00893    4.83 1.81e- 6
## 4 ASVABC         0.0478   0.0283     1.69 9.18e- 2
## 5 MALE           0.195    0.0443     4.41 1.28e- 5
## 6 ETHBLACK      -0.0448   0.0747    -0.600 5.49e- 1
## 7 ETHHISP        0.123    0.0693     1.77 7.72e- 2
```

```
## # A tibble: 8 x 5
##   term          estimate std.error statistic  p.value
##   <chr>          <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept)    1.29    0.475     2.71 6.94e- 3
## 2 S              0.0985   0.0115     8.57 1.30e-16
## 3 EXP            0.0473   0.0107     4.44 1.12e- 5
## 4 ASVABC         0.0450   0.0286     1.57 1.16e- 1
## 5 MALE           0.194    0.0444     4.36 1.57e- 5
## 6 ETHBLACK      -0.0398   0.0751    -0.530 5.96e- 1
## 7 ETHHISP        0.122    0.0693     1.76 7.91e- 2
## 8 AGE           -0.0132   0.0185    -0.715 4.75e- 1
```

Correlations between *AGE* and other explanatory variables are given as follows:

```
##           S      EXP  ASVABC  MALE ETHBLACK ETHHISP  AGE
## S          1.0000 -0.5003  0.5338 -0.1852 -0.0891 -0.1215  0.0748
## EXP        -0.5003  1.0000 -0.2119  0.0990 -0.0804  0.0607  0.4165
## ASVABC      0.5338 -0.2119  1.0000 -0.0902 -0.3162 -0.1328 -0.0511
## MALE        -0.1852  0.0990 -0.0902  1.0000 -0.0381 -0.0558 -0.0581
## ETHBLACK    -0.0891 -0.0804 -0.3162 -0.0381  1.0000 -0.1299  0.0417
## ETHHISP     -0.1215  0.0607 -0.1328 -0.0558 -0.1299  1.0000 -0.0196
## AGE         0.0748  0.4165 -0.0511 -0.0581  0.0417 -0.0196  1.0000
```

Compare the results of the two regressions.

3. (Textbook Question 5.10) 10 points

Does belonging to a union have an impact on earnings? In the output below, *COLLBARG* is a dummy variable defined to be 1 for workers whose wages are determined by collective bargaining and 0 for the others. Provide an interpretation of the regression coefficients and perform appropriate statistical tests.

```
## # A tibble: 5 x 5
##   term          estimate std.error statistic  p.value
##   <chr>          <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept)    1.04      0.197      5.29 1.87e- 7
## 2 S              0.0932    0.0101      9.19 1.11e-18
## 3 EXP            0.0423    0.00940     4.50 8.61e- 6
## 4 MALE           0.172     0.0452      3.79 1.67e- 4
## 5 COLLBARG       0.258     0.0631      4.08 5.19e- 5
```

4. (Textbook Question 5.19) 10 points

Is the effect of education on earnings different for members of a union? In the output below, *COLLBARG* is a dummy variable defined to be 1 for workers whose wages are determined by collective bargaining and 0 for the others. *SBARG* is a slope dummy variable defined as the product of *S* and *COLLBARG*. Provide an interpretation of the regression coefficients, comparing them with those in question 3, and perform appropriate statistical tests.

```
## # A tibble: 6 x 5
##   term          estimate std.error statistic  p.value
##   <chr>          <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept)    1.03      0.205     5.05 6.24e- 7
## 2 S              0.0937    0.0108     8.66 6.65e-17
## 3 EXP            0.0423    0.00941    4.49 8.75e- 6
## 4 MALE           0.171     0.0454     3.78 1.78e- 4
## 5 COLLBARG       0.298     0.357     0.835 4.04e- 1
## 6 SBARG        -0.00261    0.0227    -0.115 9.08e- 1
```