1. Theoretical Problems

True or False

a. **True**

$R^2 = \frac{SSE}{SST} = \left(1 - \frac{SSR}{SST}\right)$: The portion of variations in y explained by the linear model.

$R^2 = 1$ $implies\ SSR = \sum(y_i - \hat{y}_i)^2 = y_i - \hat{y}_i = \hat{u}_i = residuals\ = 0$ for all i

$R^2 = 1$ means we can explain all variations in y and residuals are all zero

b. **True**

Let $Y_i = \beta_0 + \beta_1 X_i + U_i$ and $X_i = \alpha_0 + \alpha_1 Y_i + U_i$

$\widehat{\beta_1} = \frac{Cov(X,Y)}{Var(X)}$ $\widehat{\alpha_1} = \frac{Cov(X,Y)}{Var(Y)}$

since $\mathrm{Var(X)} = \mathrm{Var(Y)}$ $by\ given,$

$\widehat{\beta_1} = \widehat{\alpha_1}$ $represents\ the\ estimated\ slope\ in\ two\ regression\ model\ are\ equal$

c. **False**

$R^2 = 0$ represents the explanatory variable X in **linear model** cannot explain the variation of dependent variable Y. In other words, there is **no linear relationship** between X and Y

There may have some other non-linear relationships between X and Y.

d. **False**

Sum of residuals are zero is not a crucial assumption of the linear model. We can prove it by properties of OLS.

e. **False**

$$\sum \widehat{U_i^2} = \sum\left(Y_i - \widehat{Y_i}\right)^2 = \sum\left(Y_i - \widehat{\beta_0} - \widehat{\beta_1} X_i\right)^2$$

$\widehat{\beta_0} = \bar{Y} - \widehat{\beta_1}\ \bar{X}$ $\widehat{\beta_1} = \frac{\sum(Y_i - \bar{Y})X_i}{\sum(X_i - \bar{X})X_i}$

$E(Y_i) = \beta_0 - \beta_1 X_i; E\left(\widehat{\beta_0}\right) = \beta_0;$ $E\left(\widehat{\beta_1} X_i\right) = \beta_1 X_i$

$E\left(\widehat{U_i}\right) = E\left(Y_i - \widehat{\beta_0} - \widehat{\beta_1} X_i\right) = E(Y_i) - E\left(\widehat{\beta_0}\right) - E\left(\widehat{\beta_1} X_i\right) = \beta_0 - \beta_1 X_i - \beta_0 + \beta_1 X_i = 0$

Without assuming the expected value of the error term is zero, residual still have zero mean by OLS

f. **False**

$$\widehat{\beta_1} = \frac{\sum(Y_i - \bar{Y})X_i}{\sum(X_i - \bar{X})X_i}$$

$$E(\widehat{\beta_1}|X) = \beta_1 + \frac{\sum(X_i - \bar{X})}{\sum(X_i - \bar{X})^2} E(U_i|X)$$

We only need to use assumption $E(U_i|X) = 0$ to let the least − squares estimator is unbiased .

The assumption that the error term is normally distributed is not necessary.

$Y = \log(W)$ $Y = \beta_0 + \beta_1 X + U$ with $E(U) = 0$

a.

By properties of conditional expectation,

$E(UX) = E(E(UX|X))$

$Cov(X, U) = E(XU) - E(X)E(U) = E(XU) = E(UX) = E(E(UX|X))$

$= E(XE(U|X)) = E(X * 0) = 0$

b.

$Y = \beta_0 + \beta_1 X + U$

$Cov(X, Y) = Cov(\beta_0 + \beta_1 X + U, X) = Cov(\beta_1 X, X) + Cov(U, X) = \beta_1 Var(X) + Cov(U, X)$

$= \beta_1 Var(X) + 0 = \beta_1 Var(X)$

$$\beta_1 = \frac{Cov(X, Y)}{Var(X)}$$

c.

$$\widehat{\beta_1} = \frac{\sum y_i (x_i - \bar{x})}{\sum x_i (x_i - \bar{x})} = \frac{\sum y_i x_i - \sum y_i \bar{x}}{\sum x_i^2 - \sum x_i \bar{x}} = \frac{\sum(\beta_0 + \beta_1 x_i + u) x_i - \bar{x} \sum y_i}{\sum x_i^2 - \bar{x} \sum x_i}$$

$$= \frac{\frac{1}{n} \sum(\beta_0 + \beta_1 x_i + u) x_i - \bar{x} \frac{1}{n} \sum y_i}{\frac{1}{n} \sum x_i^2 - \bar{x} \frac{1}{n} \sum x_i} = \frac{\frac{1}{n} \sum(\beta_0 + \beta_1 x_i + u) x_i - \bar{x}\bar{y}}{\frac{1}{n} \sum x_i^2 - \bar{x}^2} = A$$

By Law of large number, converge in probability

As $\bar{x}\bar{y} \xrightarrow{p} E(x)E(y) = E(x)(\beta_0 + \beta_1 E(x) + E(u)) = \beta_0 E(x) + \beta_1 E(x)^2$

$\beta_1 \frac{1}{n} \sum x_i^2 \xrightarrow{p} \beta_1 E(x^2)$ $\quad$ $\bar{x}^2 \xrightarrow{p} (E[x])^2$ $\quad$ $\beta_0 \bar{x} \xrightarrow{p} \beta_0 E(x)$ $\quad$ $\frac{1}{n} \sum ux \xrightarrow{p} E(xu)$

$$A \xrightarrow{p} \frac{\beta_0 E(x) + \beta_1 E(x^2) + E(xu) - \beta_0 E(x) - \beta_1 E(x)^2}{E(x^2) - E(x)^2}$$

$$\xrightarrow{p} \frac{\beta_1 [E(x^2) - E(x)^2] + E(xu)}{E(x^2) - E(x)^2} = \beta_1 + \frac{Cov(x, u) + E(x)E(u)}{Var(x)}$$

$$= \beta_1 + \frac{Cov(x, u)}{Var(x)} \ (SINCE\ E(U) = 0)$$

d.

Approximate: If x increases by 1 unit, W increase by $\widehat{\beta_1} * 100\ percent$

$\Delta\widehat{W}\% = 100 * \widehat{\beta_1}\ (\Delta X)$

Exact: $\Delta\widehat{W}\% = \frac{W_1}{W_0} - 1$

$\log(\widehat{W}) = \widehat{\beta_0} + \widehat{\beta_1}X + U$

$\widehat{W} = e^{\widehat{\beta_0}+\widehat{\beta_1}X+U}$

$\Delta\widehat{W}\% = \frac{W_1}{W_0} - 1 = \frac{e_1^{\widehat{\beta_0}+\widehat{\beta_1}X_1+U}}{e_0^{\widehat{\beta_0}+\widehat{\beta_1}X_0+U}} - 1 = e^{\widehat{\beta_1}\Delta X} - 1$

e.

$E(\hat{\beta}) = \beta$

note: $e^{x\beta}$ *is* ***not a linear function***, *thus* $E\left(e^{x\hat{\beta}}\right) \neq \left(e^{xE(\hat{\beta})}\right)$

bias $= E\left(e^{x\hat{\beta}} - 1\right) - \left(e^{x\beta} - 1\right) = E\left(e^{x\hat{\beta}}\right) - \left(e^{x\beta}\right) = E\left(e^{x\hat{\beta}}\right) - \left(e^{xE(\hat{\beta})}\right) \neq 0$

Thus, $e^{x\hat{\beta}} - 1$ is a biased estimator for $e^{x\beta} - 1$

Since $\mathrm{Cov}(X, U) = 0$, $\hat{\beta} \xrightarrow{p} \beta$ *by question c*

$f(\hat{\beta}) = e^{x\hat{\beta}} - 1$ *is a continuous function,* $\quad f(e^{x\beta}) = e^{x\beta} - 1$

Thus, $f(\hat{\beta}) = e^{x\hat{\beta}} - 1 \xrightarrow{p} f(\beta) = e^{x\beta} - 1$

then $e^{x\hat{\beta}} - 1 \xrightarrow{p} e^{x\beta} - 1$ *by properties of coverge in probability*

2. Computer Based Problems

Determinants of Income

a.

```
. reg loginc female black age agesq educ1 educ2 educ3 educ4
```

Source	SS	df	MS			
				Number of obs	=	3,987
				F(8, 3978)	=	123.70
Model	1072.10954	8	134.013693	Prob > F	=	0.0000
Residual	4309.72891	3,978	1.08339088	R-squared	=	0.1992
				Adj R-squared	=	0.1976
Total	5381.83845	3,986	1.35018526	Root MSE	=	1.0409

loginc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	-.2672268	.0331098	-8.07	0.000	-.3321407	-.202313
black	-.552836	.0565014	-9.78	0.000	-.6636105	-.4420616
age	.0490182	.0053594	9.15	0.000	.0385108	.0595256
agesq	-.0004872	.0000526	-9.27	0.000	-.0005902	-.0003841
educ1	.4489675	.0721566	6.22	0.000	.3075	.5904349
educ2	.705606	.067872	10.40	0.000	.5725387	.8386732
educ3	1.126566	.0713241	15.80	0.000	.9867309	1.266401
educ4	1.503135	.0749474	20.06	0.000	1.356196	1.650074
_cons	9.044574	.1379751	65.55	0.000	8.774066	9.315083

$$\widehat{loginc_i} = 9.044574 - .2672268\ female_i - .552836\ black_i + .0490182\ age_i - .0004872\ age_i^2 - .4489675 educ1_i + .705606\ educ2_i + 1.126566\ educ3_i + 1.503135\ educ4_i + \widehat{\epsilon_i}$$

- Interpretation of educ1:

Holding gender、 race、age constant, the log of income will be 44.89% **higher for the graduated high school people than for the high school dropout** people on average.

- Interpretation of educ4:

Holding gender、 race、age constant, the log of income will be 150.31% **higher for the graduated or professional school people than for the high school dropout** people on average.

b.

```
. reg loginc female black age agesq educ0 educ2 educ3 educ4
```

Source	SS	df	MS			
				Number of obs	=	3,987
				F(8, 3978)	=	123.70
Model	1072.10954	8	134.013693	Prob > F	=	0.0000
Residual	4309.72891	3,978	1.08339088	R-squared	=	0.1992
				Adj R-squared	=	0.1976
Total	5381.83845	3,986	1.35018526	Root MSE	=	1.0409

loginc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	-.2672268	.0331098	-8.07	0.000	-.3321407	-.202313
black	-.552836	.0565014	-9.78	0.000	-.6636105	-.4420616
age	.0490182	.0053594	9.15	0.000	.0385108	.0595256
agesq	-.0004872	.0000526	-9.27	0.000	-.0005902	-.0003841
educ0	-.4489675	.0721566	-6.22	0.000	-.5904349	-.3075
educ2	.2566385	.0466763	5.50	0.000	.1651269	.3481502
educ3	.6775986	.0514378	13.17	0.000	.5767517	.7784455
educ4	1.054167	.0565341	18.65	0.000	.9433288	1.165006
_cons	9.493542	.1289343	73.63	0.000	9.240758	9.746325

$formula$:

$$loginc_i = \beta_0 + \beta_1 female_i + \beta_2 black_i + \beta_3 age_i + \beta_4 age_i^2 + \beta_5 educ0_i + \beta_6 educ2_i + \beta_7 educ3_i + \beta_8 educ4_i + \epsilon_i$$

$$\widehat{loginc_i} = 9.493542 - .2672268 female_i - .552836\, black_i + .0490182\, age_i - .0004872\, age_i^2 - .4489675 educ0_i + .2566385\, educ2_i + .6775986 educ3_i + 1.054167\, educ4_i + \widehat{\epsilon_i}$$

- Interpretation of educ4:

Holding gender、 race、age constant, the log of income will be 105.41% **higher for the graduated or professional school people than for the graduated high school people** on average.

It is possible to obtain the same result using the regression estimated in item (a)

$\widehat{\beta_8} in\ regression\ Q2$(difference between graduated or professional school and graduated high school)

$= \widehat{\beta_8}\ in\ regression\ Q1$(difference between graduated or professional school and high school dropout)

$- \widehat{\beta_5}\ in\ regression\ Q1$(difference between graduated high school and high school dropout)

$= 1.503135 - .4489675 = 1.0541675$

c. $H_0: \beta_3 = \beta_4 = 0$

$H_a: \beta_3 \neq \beta_4 \neq 0$

```
. test (age == 0) (agesq =0)

 ( 1)  age = 0
 ( 2)  agesq = 0

       F(  2,  3978) =   42.98
            Prob > F =    0.0000
```

We use F-test to test whether age has significant impacts on income. **P-value is close to 0**, we need to reject H_0. In other words, there is sufficient evidence that $\beta_3 \neq \beta_4 \neq 0$ and age has significant impacts on income.

$$\widehat{loginc_i} = 9.044574 - .2672268\ female_i - .552836\ black_i + .0490182\ age_i - .0004872\, age_i^2 - .4489675 educ1_i + .705606\, educ2_i + 1.126566\, educ3_i + 1.503135\ educ4_i + \widehat{\epsilon_i}$$

Holding other variables constant, the effect of an increase in age from 34 to 35 on income is

$0.0490182 * 35 - .0004872 * 35^2 - (.0490182 * 34 - .0004872 * 34^2) = 0.0154014$

The age for **maximum in income level**:

$$\frac{dloginc_i}{dage_i} = .0490182 - 2 * .0004872\, age_i = 0$$

$$age_i \approx 50.31$$

Before age 50, the income is increasing as age increasing.

After age 50, the income is decreasing as age increasing.

Economic Convergence

a.

Source	SS	df	MS			
				Number of obs	=	104
				F(1, 102)	=	1.80
Model	.670515709	1	.670515709	Prob > F	=	0.1829
Residual	38.0298019	102	.372841195	R-squared	=	0.0173
				Adj R-squared	=	0.0077
Total	38.7003176	103	.375731239	Root MSE	=	.61061

result1	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
loggdp1975percapita	.0791044	.0589873	1.34	0.183	-.0378965	.1961054
_cons	-.4265494	.4997609	-0.85	0.395	-1.417823	.5647239

regression model: $\log\left(\frac{y_{i,1995}}{y_{i,1975}}\right) = \alpha + \beta \log(y_{i,1975}) + u_{i,1975}$

Estimated:

$$\log\left(\frac{y_{i,1995}}{y_{i,1975}}\right) = -.4265494 + .0791044 \ \log(y_{i,1975}) + u_{i,1975}$$

Interpretation:

Increasing GDP per capita of country i at year 1975 by 1 percent, is associated with 0.079% increase in the growth rate of GDP per capita of country i between year 1975 and 1995.

$\mathrm{H_0}: \beta = 0$ $\quad \mathrm{H_a}: \beta < 0 \ (beta - convergence)$ one side test.

P-value = 0.183/2= 0.0915 >0.05 (significance level)

Fail to reject H0, there is insufficient evidence of beta-convergence.

b.

```
. reg result loggdp1975percapita hci1975
```

Source	SS	df	MS			
				Number of obs	=	104
				F(2, 101)	=	11.56
Model	7.20626904	2	3.60313452	Prob > F	=	0.0000
Residual	31.4940485	101	.311822263	R-squared	=	0.1862
				Adj R-squared	=	0.1701
Total	38.7003176	103	.375731239	Root MSE	=	.55841

result1	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
loggdp1975percapita	-.2737754	.0940804	-2.91	0.004	-.4604056	-.0871452
hci1975	.7028936	.1535307	4.58	0.000	.3983299	1.007457
_cons	1.299738	.5925073	2.19	0.031	.1243628	2.475113

regression model: $\log\left(\frac{y_{i,1995}}{y_{i,1975}}\right) = \alpha + \beta_1 \log(y_{i,1975}) + \beta_2 HC_{i,1975} + u_{i,1975}$

Estimated:

$$\log\left(\frac{y_{i,1995}}{y_{i,1975}}\right) = 1.299738 - .2737754 \log(y_{i,1975}) + .7028936 HC_{i,1975} + u_{i,1975}$$

Interpretation:

β_1 : Increasing GDP per capita of country i at year 1975 by 1 percent, is associated with 0.273% decrease in the growth rate of GDP per capita of country i between year 1975 and 1995.

β_2 : Increasing human capital index of country i at year 1975 by 1 percent, is associated with 0.702% decrease in the growth rate of GDP per capita of country i between year 1975 and 1995

$H_0: \beta = 0 \quad H_a: \beta < 0$ (*conditionally beta − convergence*) one side test.

P-value = 0.004/2= 0.002 < 0.05

Controlling the human capital index. Reject H0, there is sufficient evidence of beta-convergence.

P-value is **much smaller** compare to question a, we can prove the statistically significance of conditionally beta-convergence in question b.

c.

```
. reg result loggdp1975percapita  gcf1975 hci1975
```

Source	SS	df	MS
Model	8.27402029	3	2.75800676
Residual	30.4262973	100	.304262973
Total	38.7003176	103	.375731239

Number of obs	=	104
F(3, 100)	=	9.06
Prob > F	=	0.0000
R-squared	=	0.2138
Adj R-squared	=	0.1902
Root MSE	=	.5516

result1	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
loggdp1975percapita	-.3519952	.1018823	-3.45	0.001	-.5541268	-.1498635
gcf1975	1.015701	.542195	1.87	0.064	-.0599982	2.091401
hci1975	.7384911	.1528442	4.83	0.000	.4352526	1.04173
_cons	1.656595	.615502	2.69	0.008	.4354566	2.877733

regression model: $\log\left(\frac{y_{i,1995}}{y_{i,1975}}\right) = \alpha + \beta_1 \log(y_{i,1975}) + \beta_2 GCF_{i,1975} + \beta_3 HC_{i,1975} + u_{i,1975}$

Estimated:

$$\log\left(\frac{y_{i,1995}}{y_{i,1975}}\right) = 1.656595 - .3519952 \log(y_{i,1975}) + 1.015701\, GCF_{i,1975} + .7384911\, HC_{i,1975} + u_{i,1975}$$

Interpretation:

β_1 : Increasing GDP per capita of country i at year 1975 by 1 percent, is associated with 0.352% decrease in the growth rate of GDP per capita of country i between year 1975 and 1995.

β_2 : Increasing Gross capital formation shares of country i at year 1975 by 1 percent, is associated with 1.01% decrease in the growth rate of GDP per capita of country i between year 1975 and 1995

β_3 : Increasing Human capital index of country i at year 1975 by 1 percent, is associated with 0.738% decrease in the growth rate of GDP per capita of country i between year 1975 and 1995

$\mathrm{H_0}: \beta = 0 \quad \mathrm{H_a}: \beta < 0\ (conditionally\ beta - convergence)$ one side test.

P-value = 0.001/2= 0.0005 < 0.05(significance level)

Controlling the human capital index and Gross capital formation shares. Reject H0, there is sufficient evidence of beta-convergence.

P-value is **much smaller** compare to question a and b, we can prove the statistically significance of conditionally beta-convergence in question b & c.

```
. test (gcf1975 == 0) (hci1975 == 0)

 ( 1)  gcf1975 = 0
 ( 2)  hci1975 = 0

       F(  2,   100) =   12.49
            Prob > F =    0.0000
```

$\mathrm{H_0}: \beta_2 = \beta_3 = 0 \quad \mathrm{H_a}: \beta_2 \neq \beta_3 \neq 0$

We use F-test, with degree of freedom for the numerator is 2, for the denominator is 104-3-1=100

P-value is close to 0, P-value< 0.05(significance level). Reject H0, there is sufficient evidence that both types of capitals jointly important to explain future growth.

Monte Carlo Simulation

a.

```
. sum p_value b1 b0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
     p_value |      1,000    .5072577    .2864259   .0001011   .9990631
          b1 |      1,000    5.114218    1.712932   .1414761   11.11117
          b0 |      1,000   -10.07301    .9718946  -12.79138  -7.294732

.
. count if p_value <0.05
  49
```

H_0: $\beta_1 = 5$

H_a: $\beta_1 \neq 5$

We reject $\mathrm{H_0}$ whenever $\mathrm{P - value} \leq \alpha = 0.05(\text{siginificance level})$

In 1000 simulations, we reject the null hypotheses 49 times.

The fraction of the simulations we can rejut the null hypotheses:

$\frac{49}{1000} * 100\% = 4.9\%$ 4.9% is close to 5%

For those 49 times, you have a Type 1 Error.

$\boldsymbol{Probabilty\ of\ Type\ 1\ error}$

: $\alpha(significance\ level)\ represent\ the\ probability\ that\ reject\ H_0\ given\ H_0\ is\ ture.$

$P(reject\ H_0|H_0:\ \beta_1 = 5\ \ is\ true) = \alpha = 0.05 = 5\%$

b.

For $H_0: \beta_1 = 4.5$ $H_a: \beta_1 \neq 4.5$

```
. sum p_value b1 b0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
     p_value |      1,000    .4914149    .2950307   .0000292   .9990672
          b1 |      1,000    5.114218    1.712932   .1414761   11.11117
          b0 |      1,000   -10.07301    .9718946  -12.79138  -7.294732

.
. count if p_value <0.05
  65
```

In 1000 simulations, we reject the null hypotheses 65 times.

The fraction of the simulations we can rejut the null hypotheses:

$\frac{65}{1000} * 100\% = 6.5\%$ **6.5% is close to 5%**

For $H_0: \beta_1 = 0 \quad H_a: \beta_1 \neq 0$

```
. sum p_value b1 b0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
p_value	1,000	.0367694	.0922712	5.45e-11	.9399751
b1	1,000	5.114218	1.712932	.1414761	11.11117
b0	1,000	-10.07301	.9718946	-12.79138	-7.294732

```
.
. count if p_value <0.05
  829
```

In 1000 simulations, we reject the null hypotheses 829 times.

The fraction of the simulations we can rejut the null hypotheses:

$\frac{829}{1000} * 100\% = 82.9\%$ **82.9% is NOT close to 5%**

When $H_0: \beta_1 = 0 \quad H_a: \beta_1 \neq 0$, we get more significant fractions of the simulations we can reject each null hypothesis.

In those two cases, we get the **Power** of the test.

Power represent the probability that reject H_0 given H_a is ture.

Power: $P(\text{reject } H_0 | H_a \text{ is true}) = 6.5\%$ for $H_0: \beta_1 = 4.5 \quad H_a: \beta_1 \neq 4.5$

$= 82.9\%$ for $H_0: \beta_1 = 0 \quad H_a: \beta_1 \neq 0$

Since 4.5 is close to 5, it has a smaller power. 0 is much smaller than 5, it has a greater power.