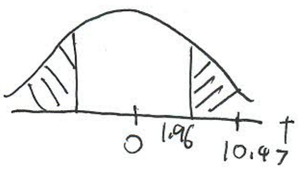
3.1

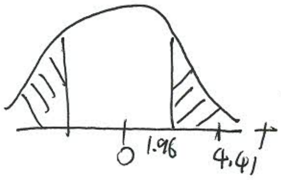
$$\widehat{(lwage_i)} = 5.395 + 0.065\,educ_i + 0.014\,exper_i + 0.0117\,tenure_i + 0.199\,married_i - 0.188\,black_i - 0.09\,South_i + 0.184\,urban_i$$

3.2

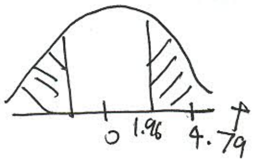
$H_0: \beta_2 = 0$
$H_a: \beta_2 \neq 0$
tcal: 10.47



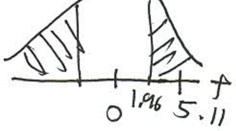
$H_0: \beta_3 = 0$
$H_a: \beta_3 \neq 0$
tcal: 4.41



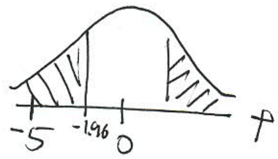
$H_0: \beta_4 = 0$
$H_a: \beta_4 \neq 0$
tcal: 4.79



$H_0: \beta_5 = 0$
$H_a: \beta_5 \neq 0$
tcal: 5.11

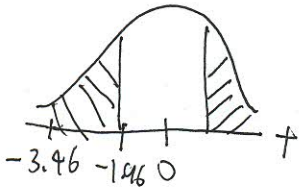


$H_0: \beta_6 = 0$
$H_a: \beta_6 \neq 0$
tcal: -5

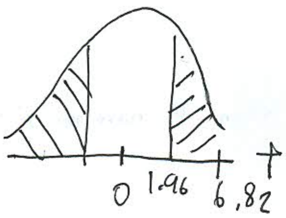


$\boxed{\text{* } |t_{critical}| \text{ is } 1.96}$

$H_0: \beta_7 = 0$
$H_a: \beta_7 \neq 0$
tcal: -3.46

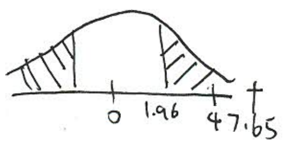


$H_0: \beta_8 = 0$
$H_a: \beta_8 \neq 0$
tcal: 6.82



∴ Since the computed t value (tcal) of all coefficient fall into rejection region at 5% level of significance, we can reject all of H_0 at 5% level of significance. That is, all coefficient is statistically significant difference from 0 at 5% level of significant.

$H_0\ \beta_1 = 0$
$H_a\ \beta_1 \neq 0$
tcal: 47.65



3.3

Holding other factors constant, black people earn $(100)(0.18) = 18\%$ less than non-black, on average.

From 3.2 the difference is statistically difference at 5% level of significance.

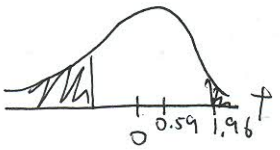
3.4 $(\widehat{lwage})_i = 5.4 + 0.065\, educ_i + 0.14\, exper_i + 0.12\, tenure_i + 0.189\, married_i$
$- 0.24\, black_i - 0.92\, South_i + 0.184\, urban_i + 0.061\, married_i\, black_i$

$H_0\ \gamma_9 = 0$

$H_a\ \gamma_9 \neq 0$

$T_{cal} = 0.59$

assume 5% level of significance
$|t_{cri}| = 1.96$



since T_{cal} doesn't fall into rejection rejion at 5% level of significance,
we cannot reject H_0 at 5% level of significance. That is, the interaction effect
of married and black is not statistically significant difference from 0 at 5% level of significance.

Unlike black dummy variable, the interaction effect seem to have no influence on lwage at 5% level of significance.

$E(lwage \mid married = 1,\ black = 1)$
$= (5.4 + 0.189 - 0.24 + 0.061) + 0.065\, educ_i + 0.14\, exper_i + 0.12\, tenure_i - 0.92\, south_i$
$+ 0.184\, urban_i$

$= 5.41 + 0.065\, educ_i + 0.14\, exper_i + 0.12\, tenure_i - 0.92\, south_i + 0.184\, urban_i$

3.5
$E(lwage \mid married = 1,\ black = 1)$ => married black
$= 5.41 + 0.065\, educ_i + 0.14\, exper_i + 0.12\, tenure_i - 0.92\, south_i + 0.184\, urban_i$ ①
$E(lwage \mid married = 1,\ black = 0)$ => married non-black
$= \underbrace{(5.4 + 0.189)}_{5.589} + 0.065\, educ_i + 0.14\, exper_i + 0.12\, tenure_i - 0.92\, south_i + 0.184\, urban_i$ ②

① - ② ; -0.179 => -17.9%

∴ The difference of estimated wage of married black and married non-black is 17.9% on average.

4.1

$\widehat{(hrs)} = 1904.577 - 93.75\, rate_i + 0.0002\, erspi - 0.215\, erno_i + 0.157\, nein_i + 0.1557\, asset_i - 0.3486\, age_i + 20.7\, dep_i + 37.35\, school_i$

4.2

Yes there is a multicollinearity problem because

① There exist a high correlation between independent variables according to table 4.2.
(ex. asset and rate, asset and nein, age and erno)

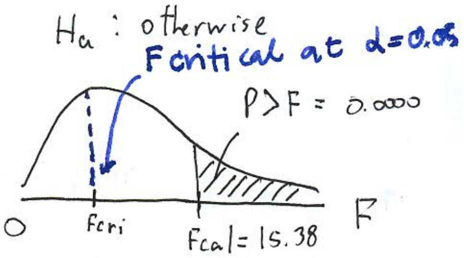
② Although the R^2 is high and F statistic test (overall significant test) is rejecting H_0 that $\beta_2 = \beta_3 = \beta_4 = \dots = \beta_9 = 0$ at 0.05 level of significance, some of the individual test is not statistically significance at 0.05 level of significance.

F-test

$H_0 : \beta_2 = \beta_3 = \beta_4 = \dots = \beta_9$

H_a : otherwise

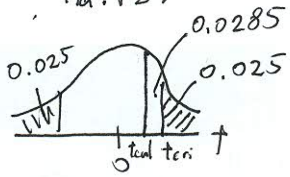
$\alpha = 0.05$



Individual test $\alpha = 0.05$

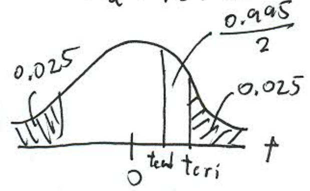
rate $H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$



ersp $H_0 : \beta_3 = 0$

$H_a : \beta_3 \neq 0$



③ From table 4.3, some of VIF excess 10.

④ Normally wealth should be correlated with income. In this case, it is an asset and nein.

Actions when facing multicollinearity

① Drop variable => could lead to hetero or autocorrelation or omitted variable bias

② Transform regression function => " ——————— "

③ Correct more data => high cost for collecting data

④ Do nothing

Therefore, it depends on situation to choose which action to use because every action comes with a cost.

4.3

Both VIF and tolerance measure a level/degree of correlation of independent varables.

If VIF is more than 10 like in this case, there exist a multicollinearity problem.

If tolerance is more than 0.1 " ——— "

~~of X_i between independent variables~~

$$VIF = \frac{1}{1-r^2}$$; r^2 is a correlation coefficient squared of X_i and other independent variables

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2} VIF$$

∴ If r^2 is high, VIF will be large leading to high $Var(\hat{\beta}_2)$.

Same go for $\hat{\beta}_3, \hat{\beta}_4, \dots, \hat{\beta}_q$.

4.4 False

ⓐ If there is a perfect multicollinearity, we cannot even compute the OLS estimators. Also VIF will be infitity as well as $Var(\hat{\beta}_i)$.

ⓑ False

In the pressence of high multicollinearity, we still can estimate OLS estimators as well as testing the individual significance, but the result will not be reliable because

① abnormally high variance leading to ~~acceptance~~ Cannot reject of zero null hypothesis.

② small change in data ~~will~~ can result in large change in OLS estimator and standard error.