3.1

(a) $n = 64$.

$H_0: b_2 = 0$

$H_1: b_2 > 0$ #

(b). test statistic:

$$T = \frac{\hat{\beta}_2}{\sqrt{\frac{\hat{\sigma}^2}{\Sigma(X-\bar{X})^2}}} \overset{H_0}{\sim} t(62)$$

#

(c) if alternative hypothesis is true, the distribution will shift to the right relative to the usual t-distribution.

H_0 is true: $b_2 = 0$, H_1 is true: $b_2 > 0$, so the distribution will shift to the right, #

(d) $\alpha = 1\%$

we will reject H_0 if $t \geq 2.388$, do not reject H_0 when $t < 2.388$

#

$\alpha = 1\%$: $t(0.01, 62)$

$= 2.388$

(e) $T_0 = \dfrac{\hat{\beta}_2}{se(\hat{\beta}_2)} = \dfrac{0.01309}{0.00215} \approx 6.0884$.

economic conclusion: $T_0 \in RR$, reject H_0, there is a positive relationship between the number of medals won and GDP.

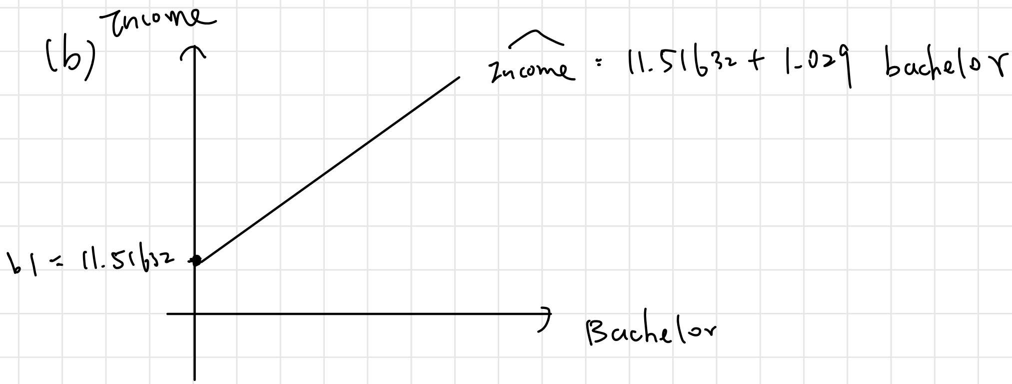
1% level of significance means the type I error, the probability that when H_0 is true, but reject H_0 #

3.7

(a) $t = \dfrac{\hat{b}_1}{se(b_1)}$, $4.31 = \dfrac{b_1}{2.672}$, $b_1 = 11.51632$.

intercept = 11.51632 #

(b)



Income and bachelor have a positive relationship and increasing at an constant rate because the b_2 coefficient is $+1.029$ #

(c) $t = \dfrac{\hat{b_2}}{se(b_2)}$, $10.75 = \dfrac{1.029}{se(b_2)}$

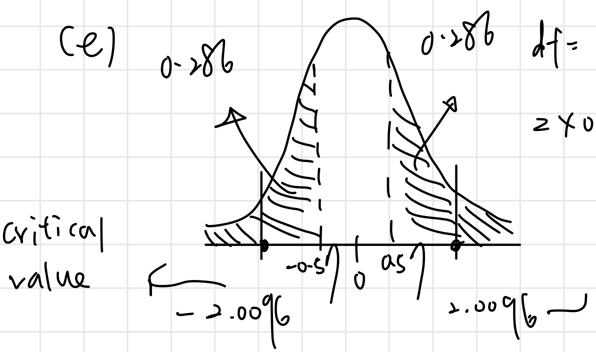
$se(b_2) = 0.0957$ #

(d) $H_0 : b_1 = 10$

t - statistic :

$$\frac{\hat{b_1} - b_1}{se(b_1)} = \frac{11.51632 - 10}{2.672} = 0.567$$ #

(e)



df = 49.

$2 \times 0.286 = 0.572$ (two-tailed p-value)

rejection region for a 5% test are t-values greater than or equal to 2.0096 or less than or equal to -2.0096 #

(f) interval estimate of the slope:

$$\left[\hat{b_2} - t_{\frac{\alpha}{2}(n-2)} se(b_2) \ , \ \hat{b_2} + t_{\frac{\alpha}{2}(n-2)} \cdot se(b_2)\right]$$

$$= [1.029 - 2.680 \cdot 0.0957 , 1.029 + 2.680 \cdot 0.0957]$$

$$= [0.7725 , 1.2855]$$ #

(g) $H_0: b_2 = 1$

$H_a: b_2 \neq 1$

$\alpha = 0.05$

$$\text{t-statistic} = \frac{\hat{b_2} - b_2}{se(b_2)} \overset{H_0}{\sim} t(n-2)$$

$$\text{rejection region} = \{ |T| \geq t_{\frac{\alpha}{2}(n-2)} \} = |T| \geq 2.010$$

$$T = \frac{1.029 - 1}{0.0957} = 0.303$$

$T_0 \notin RR$, do not reject $H_0: b_2 = 1$

即教育程度提升1%，對應所得提升1000美元 #

3.17

(a) ①

$H_0: b_2 \leq 1.8$

$H_a: b_2 > 1.8$

②

$\alpha = 0.05$

③ test statistic $= \dfrac{\hat{b}_2 - b_2}{se(b_2)} \overset{H_0}{\sim} t(n-2)$

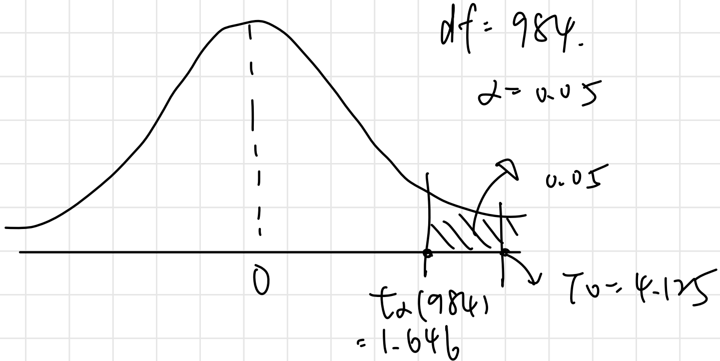
④ rejection region $: T \geq t_\alpha(n-2)$, $T \geq t_{0.05}(984)$

$: T \geq 1.646$

⑤ $T_0 = \dfrac{2.46 - 1.8}{0.16} = 4.125$

$4.125 \geq 1.646$. $T_0 \in RR$.

reject $H_0: b_2 \leq 1.8$ and accept alternative #



(b) EDUC = 16, WAGE = 23.92. (rural)

$t_{\frac{95\%}{2}}(212) \approx 1.971$

interval estimate :

$[23.92 - 1.971 \times 0.833 , 23.92 + 1.971 \times 0.833]$

$= [22.278 , 25.562]$

use -0.761 cov

$SE\ WAGE : \sqrt{[SE(b1)]^2 + (EDUC)^2 [SE(b2)]^2 + 2 \cdot EDUC \cdot cov(b1, b2)}$

$= 1.1035.$

(c) EDUC = 16, WAGE = 28.6 (urban)

COV = -0.345

$SE\ WAGE = \sqrt{(SE(b_1))^2 + (EDUC)^2 \cdot (SE(b_2))^2 + 2(EDUC) \cdot cov(\hat{b}_1, \hat{b}_2)}$

$= 0.8164.$

interval estimate : $t_{\frac{95\%}{2}}(984) \approx 1.962$

$$[28.6 - 0.8164 \cdot 1.962,\ 28.6 + 0.8164 \cdot 1.962]$$

$$= [26.998,\ 30.202]$$

urban's interval estimate is narrower than rural

maybe because it sample is more large. #

(d) rural:

① $H_0: b_1 \geq 4$

$H_a: b_1 < 4$

② $\alpha = 0.01$

③. test statistic: $\dfrac{\hat{b_1} - b_1}{Se(b_1)} \overset{H_0}{\sim} t(212)$

④ rejection region: $T_0 < t_{0.01}(212) = T_0 < 2.344$

⑤ $T_0 = \dfrac{-4.88 - 4}{3.29} = -2.699$

$T_0 \in RR$, reject H_0 and accept H_a. #

3.19.

3.19 The owners of a motel discovered that a defective product was used during construction. It took 7 months to correct the defects during which approximately 14 rooms in the 100-unit motel were taken out of service for 1 month at a time. The data are in the file *motel*.

a. Plot *MOTEL_PCT* and *COMP_PCT* versus *TIME* on the same graph. What can you say about the occupancy rates over time? Do they tend to move together? Which seems to have the higher occupancy rates? Estimate the regression model $MOTEL\_PCT = \beta_1 + \beta_2 COMP\_PCT + e$. Construct a 95% interval estimate for the parameter β_2. Have we estimated the association between *MOTEL_PCT* and *COMP_PCT* relatively precisely, or not? Explain your reasoning.

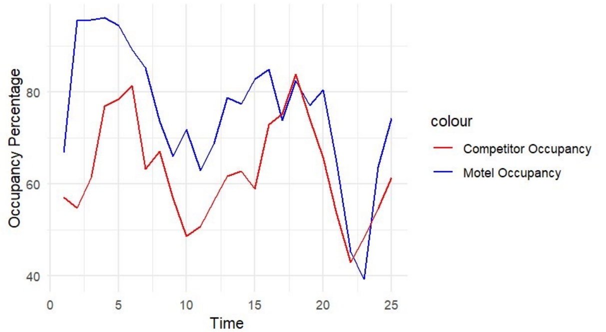
b. Construct a 90% interval estimate of the expected occupancy rate of the motel in question, *MOTEL_PCT*, given that $COMP\_PCT = 70$.

c. In the linear regression model $MOTEL\_PCT = \beta_1 + \beta_2 COMP\_PCT + e$, test the null hypothesis $H_0: \beta_2 \le 0$ against the alternative hypothesis $H_0: \beta_2 > 0$ at the $\alpha = 0.01$ level of significance. Discuss your conclusion. Clearly define the test statistic used and the rejection region.

d. In the linear regression model $MOTEL\_PCT = \beta_1 + \beta_2 COMP\_PCT + e$, test the null hypothesis $H_0: \beta_2 = 1$ against the alternative hypothesis $H_0: \beta_2 \ne 1$ at the $\alpha = 0.01$ level of significance. If the null hypothesis were true, what would that imply about the motel's occupancy rate versus their competitor's occupancy rate? Discuss your conclusion. Clearly define the test statistic used and the rejection region.

e. Calculate the least squares residuals from the regression of *MOTEL_PCT* on *COMP_PCT* and plot them against *TIME*. Are there any unusual features to the plot? What is the predominant sign of the residuals during time periods 17–23 (July, 2004 to January, 2005)?

(a)



their trend sometimes tend to move together, sometimes not (15-20 days), and motel seems to have the higher occupancy rate.

$\beta_2 = 0.8646$，即 competitor 入住率增加1%，motel 入住率在其他變數不變下會增加 0.8646%，相對用圖形看，用 regression model 會更加精確

95%. interval estimate for β_2:

$[0.4453, 1.2840]$ #

(b) MOTEL_PCT = 81.92474.

interval:

$[77.38223, 86.46725]$

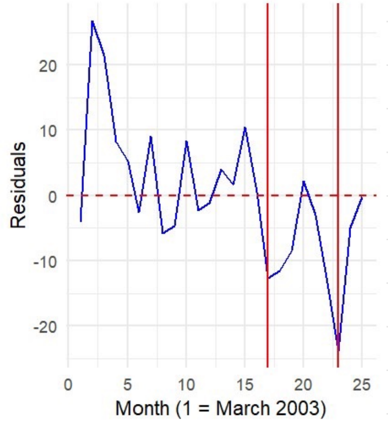
(c) Comp_Pct, t = 4.27, ∈ RR,

reject H_0.

(d) t = -0.67, ∉ RR, do not reject H_0.

because -0.67 is in the non-rejection-region #

(e)



17~23. residuals almost are negative. this reflect maybe some factor not be captured by the regression model