HW 0324 Q1

$$Y = \beta X + e, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{1,2} & \cdots & X_{1,k} \\ \vdots & \vdots & \ddots & \\ 1 & X_{n,2} & \cdots & X_{n,k} \end{bmatrix}$$

$$e = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}, \quad e \sim N(0, \sigma^2 I), \quad Y \sim N(X\beta, \sigma^2 I)$$

When $K=2$, $\quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{1,2} \\ \vdots & \vdots \\ 1 & X_{n,2} \end{bmatrix}$

$$\mathbf{b} = (X'X)^{-1} X'Y$$

$$X'X = \begin{bmatrix} 1 & \cdots & 1 \\ X_{1,2} & \cdots & X_{n,2} \end{bmatrix} \begin{bmatrix} 1 & X_{1,2} \\ \vdots & \vdots \\ 1 & X_{n,2} \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} X_{i,2} \\ \sum_{i=1}^{N} X_{i,2} & \sum_{i=1}^{N} X_{i,2}^2 \end{bmatrix}$$

$$X'X = \frac{1}{N\sum_{i=1}^{N} X_{i,2}^2 - \left(\sum_{i=1}^{N} X_{i,2}\right)^2} \begin{bmatrix} \sum_{i=1}^{N} X_{i,2}^2 & -\sum_{i=1}^{N} X_{i,2} \\ -\sum_{i=1}^{N} X_{i,2} & N \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & \cdots & 1 \\ X_{1,2} & \cdots & X_{n,2} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} X_{i,2} y_i \end{bmatrix}$$

$$\mathbf{b} = \frac{1}{N\sum_{i=1}^{N} X_{i,2}^2 - \left(\sum_{i=1}^{N} X_{i,2}\right)^2} \begin{bmatrix} \sum_{i=1}^{N} X_{i,2}^2 & -\sum_{i=1}^{N} X_{i,2} \\ -\sum_{i=1}^{N} X_{i,2} & N \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} X_{i,2} y_i \end{bmatrix}$$

$$= \frac{1}{N\sum_{i=1}^{N} X_{i,2}^2 - \left(\sum_{i=1}^{N} X_{i,2}\right)^2} \begin{bmatrix} \sum_{i=1}^{N} X_{i,2}^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} X_{i,2} \sum_{i=1}^{N} X_{i,2} y_i \\ -\sum_{i=1}^{N} X_{i,2} \sum_{i=1}^{N} y_i + N \sum_{i=1}^{N} X_{i,2} y_i \end{bmatrix}$$

$$\mathbb{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \dfrac{\sum X_i^2 \sum y_i - \sum X_i \sum X_i y_i}{N\sum X_i^2 - (\sum X_i)^2} \\ \dfrac{N\sum X_i y_i - \sum X_i \sum y_i}{N\sum X_i^2 - (\sum X_i)^2} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\sum X_i^2 \sum y_i - \sum X_i \sum X_i y_i}{N\sum X_i^2 - (\sum X_i)^2} \\ \dfrac{\sum X_i y_i - \dfrac{\sum X_i \sum y_i}{N}}{\sum X_i^2 - \dfrac{(\sum X_i)^2}{N}} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\sum y_i}{N} - \dfrac{\sum X_i}{N} \dfrac{\sum (X_i - \bar{X})(y_i - \bar{y})}{\sum (X_i - \bar{X})^2} \\ \dfrac{\sum (X_i - \bar{X})(y_i - \bar{y})}{\sum (X_i - \bar{X})^2} \end{bmatrix}$$

$$\mathrm{Cov}(b_1, b_2) = \frac{\sigma^2}{N\sum_{i=1}^{N} X_{i,2}^2 - \left(\sum_{i=1}^{N} X_{i,2}\right)^2} \begin{bmatrix} \sum_{i=1}^{N} X_{i,2}^2 & -\sum_{i=1}^{N} X_{i,2} \\ -\sum_{i=1}^{N} X_{i,2} & N \end{bmatrix}$$

$$Var(b_1) = \frac{\sigma^2}{\Delta} \sum_{i=1}^{N} X_{i,2}^2 \qquad Var(b_2) = \frac{\sigma^2}{\Delta} N$$

$$\tilde{b} \sim N\left(\hat{\beta}, \sigma^2 \begin{pmatrix} \frac{\sum X_{i,2}^2}{\sim} \\ \frac{\sim}{\Delta} \end{pmatrix}\right)$$

Q5.3

a.

i. $t = \dfrac{b_1}{SE(b_1)} = \dfrac{1.4515}{2.2019} = 0.6592$

ii. $SE(b_2) = \dfrac{2.7648}{5.7103} = 0.4842$

iii. $b_3 = \dfrac{-3.9376}{0.3695} = -1.4558$

iv. $R^2 = 1 - \dfrac{SSE}{SST} = 1 - \dfrac{46221.62}{6.395477^2(1200-1)} = 0.0573$

v. $\hat{\sigma} = \sqrt{MSE} = \sqrt{\dfrac{46,221.62}{1200-4}} = 6.216$

b.

$b_2 = 2.7648 = \dfrac{\partial WALC}{\frac{\partial TOTEXP}{TOTEXP}}$, each 1% increase in home expenditure will raise 2.7648% of household's spending on alcohol.

$b_3 = -1.4558$, each 1 more kid will drop 1.46% of household's spending on alcohol.

$b_4 = -0.1503$, each 1 year old increase in household age will drop 0.1503% of household's spending on alcohol.

c.

$$y = \frac{b_4}{SE(b_4)} = \frac{-0.1503}{0.0235} \overset{A}{\sim} N(0,1)$$

$$95\%\ CI \text{ for } \beta_4 = [\, b_4 \pm Z_{0.025} SE(b_4) \,]$$

$$= -0.1503 \pm 1.96 \times 0.0235$$

$$= [-0.1963, -0.1043]$$

We are 95% confident that the percentage change of spending on alcohol caused by the 1 unit change of household age won't fall out of the range of $[-0.1963, -0.1043]$

d. $\alpha = 0.05$,

p-v_1 > 0.05 insignificant, p-v_2, p-v_3, p-v_4 < 0.05 significant

e. $H_0: \beta_3 = -2$ against $\beta_3 \neq -2$

$$\varphi = \frac{b_3 - (-2)}{SE(b_3)} \overset{A}{\sim} N(0,1)$$

$$RR = \{\varphi \mid |\varphi| > 1.96\}, \quad \varphi^* = 1.4775 \notin RR$$

Under 5% significance, we cannot reject that $\beta_3 = -2$.

Q23.

a. β_2: $-$ (Whole sale) β_3: $+$ (quality premium)

β_4: ? (Inflation, Regulation restriction)

b.

```
Call:
lm(formula = price ~ quant + qual + trend, data = dat
a)

Residuals:
    Min      1Q  Median      3Q     Max 
-43.479 -12.014  -3.743  13.969  43.753 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 90.84669    8.58025  10.588 1.39e-14
quant       -0.05997    0.01018  -5.892 2.85e-07
qual         0.11621    0.20326   0.572   0.5700
trend       -2.35458    1.38612  -1.699   0.0954

(Intercept) ***
quant       ***
qual        
trend       .
```

shift

$\beta_2 = -0.06$ is negative and significant

$\beta_3 = 0.11$ is positive and insignificant

$\beta_4 = -2.3$ is negative and insignificant

c. $R^2 = 0.50965 \rightarrow 51\%$ of total variation can be jointly decided by the variations of quantity, quality and time.

d. $H_0: \beta_2 = 0$ against $H_1: \beta_2 < 0$

p-value $= 1.4775e-07$ is significant under each level of significance.

e. $H_0: \beta_3 = 0$ against $H_1: \beta_3 \neq 0$

p-value $= 0.57 > 0.5$ is insignificant under $\alpha = 5\%$

f.

$$\frac{\partial \text{Price}}{\partial \text{Time}} = \beta_4 = -2.3549$$

Potential reasons are:

1. Smuggeling, S↑, P↓
2. Restriction ↑, risk ↑, D↓, P↓