1. (a)

(i) The average of crop yields (in bushels per acre) of Brand 3 with no fertilizer used.

(ii) The difference between the average crop yields of Brand 3 and the average crop yields of Brand 2 which have same amount of fertilizer used.

(iii) The change in the average crop yields for an additional pounds of fertilizer per acre used for all three brands.

(b) Null: $\beta_1 = \beta_2 = \beta_3 = 0$

Alternative: at least one of $\beta_1, \beta_2, \beta_3 \neq 0$

$n = 15 \qquad p = 4$

SS Regression $= \Sigma(y_i - \bar{y})^2 - \Sigma(y_i - \hat{y}_i)^2 = 522 - 113.8667 = 408.1333$

SS error $= 113.8667$

Source	SS	df	MS	F
Regression	408.1333	p-1 = 3	136.0444	13.1425
Error (Residuals)	113.8667	n-p = 11	10.3515	
Total	522	n-2 = 13		

Critical value : $F_{0.05}(3, 11) = 3.587$

Decision : Reject H_0 at $\alpha = 0.05$

(c) Null: $\beta_1 = \beta_2 = 0$

Alternative: At least one of $\beta_1, \beta_2 \neq 0$

$n = 15$

Full model: $Y = \beta_0 + \beta_1 V_1 + \beta_2 V_2 + \beta_3 X + \varepsilon$

$\dim(V) = 4$ $SS_{Resid} = 113.8667$

Null model: $Y = \beta_0 \qquad\qquad + \beta_3 X + \varepsilon$

$\dim(V) = 2$, $SS_{Resid} = ?$

$$\hat{\beta}_3 = 2.0667$$

$$SS_{Regr} = \hat{\beta}_3^2 \, SX_3X_3 = 2.0667^2 \times \left(\Sigma X^2 - \tfrac{1}{15}(\Sigma X)^2\right) = 129.6$$

$$SS_{Resid} = SYY - SS_{Regr} = 522 - 129.6 = 392.4$$

Source	SS	DF	MS	F
H_0 (Diff)	278.5333	4-2=2	139.2667	13.4537
Full	113.8667	15-4=11	10.3515	
Null	392.4	15-2=13		

$F_{0.05}(2, 11) = 3.98$

$\Rightarrow$ Reject H_0 at $\alpha = 0.05$

(d)

null: $\beta_3 = 0$

alternative: $\beta_3 \neq 0$

$\hat{var}(\hat{\beta_3}) = C_{33} \times s^2 = \frac{1}{30} \times 10.3515 = 0.3451$

Test statistic: $T = \dfrac{2.0667 - 0}{\sqrt{0.3451}} = 3.5181$

critical value: $\pm t_{0.025}(11) = \pm 2.201$

$\Rightarrow$ Reject H_0 at $\alpha = 0.05$

(e)

$X_0^T = [\, 1 \quad 0 \quad 1 \quad 2 \,]$

$\hat{Y_0} = 1 \times 15.7333 + 7.9333 + 2 \times 2.0667 = 27.8$

$X_0^T C X_0 = 0.5$

$\hat{Var}(\hat{Y_0}) = [1 + X_0^T C X_0] s^2 = (1 + 0.5) \times 10.3515 = 15.5271$

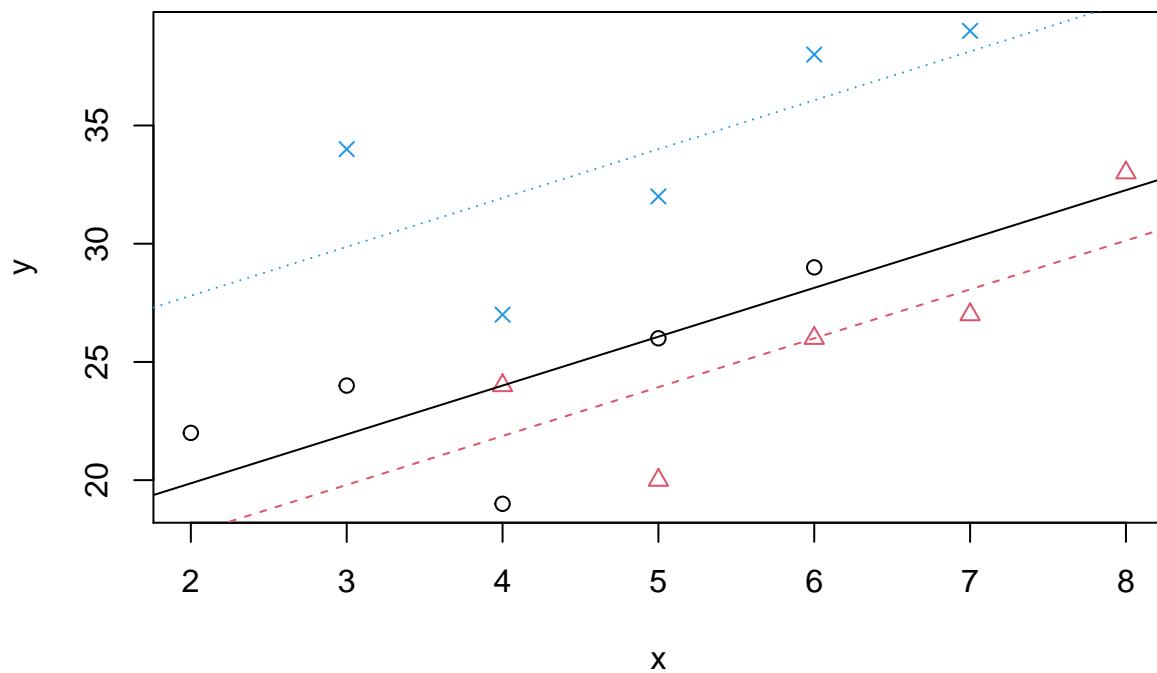$t_{0.025}(11) = 2.201$

$27.8 \pm 2.201 \times \sqrt{15.5271}$

27.8 ± 8.6729

$= (19.1271, 36.4729)$

```
y = c(20, 24, 27, 26, 33, 27, 38, 32, 39, 34, 22, 24, 19, 26, 29)
x = c(5, 4, 7, 6, 8, 4, 6, 5, 7, 3, 2, 3, 4, 5, 6)
v1 = c(1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
v2 = c(0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0)
fit = lm(y~x+v1+v2)
plot(x, y, col = 1+v1+3*v2, pch =1+v1+3*v2 )
abline(fit$coeff[1], fit$coeff[2], col=1, lty=1)
abline(fit$coeff[1]+fit$coeff[3], fit$coeff[2], col=2, lty=2)
abline(fit$coeff[1]+fit$coeff[4], fit$coeff[2], col=4, lty=3)
```


(b)

```
fit1 = lm(y~x+v1+v2)
summary(fit1)
```

```
##
## Call:
## lm(formula = y ~ x + v1 + v2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.0000 -1.5333  0.8667  2.1000  4.1333
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  15.7333     2.7552   5.710 0.000136 ***
## x             2.0667     0.5874   3.518 0.004813 **
```

1

```
## v1             -2.1333     2.3496  -0.908 0.383370
## v2              7.9333     2.1179   3.746 0.003234 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.217 on 11 degrees of freedom
## Multiple R-squared:  0.7819, Adjusted R-squared:  0.7224
## F-statistic: 13.14 on 3 and 11 DF,  p-value: 0.0005877
```

```
anova(lm(y~1), fit1)
```

```
## Analysis of Variance Table
##
## Model 1: y ~ 1
## Model 2: y ~ x + v1 + v2
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     14 522.00
## 2     11 113.87  3    408.13 13.143 0.0005877 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(c)

```
fit2 = lm(y~x)
anova(fit2, fit1)
```

```
## Analysis of Variance Table
##
## Model 1: y ~ x
## Model 2: y ~ x + v1 + v2
##   Res.Df    RSS Df Sum of Sq      F   Pr(>F)
## 1     13 392.40
## 2     11 113.87  2    278.53 13.454 0.001108 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(d)

```
summary(fit1)
```

```
##
## Call:
## lm(formula = y ~ x + v1 + v2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.0000 -1.5333  0.8667  2.1000  4.1333
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  15.7333     2.7552   5.710 0.000136 ***
## x             2.0667     0.5874   3.518 0.004813 **
## v1           -2.1333     2.3496  -0.908 0.383370
## v2            7.9333     2.1179   3.746 0.003234 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 3.217 on 11 degrees of freedom
## Multiple R-squared:  0.7819, Adjusted R-squared:  0.7224
## F-statistic: 13.14 on 3 and 11 DF,  p-value: 0.0005877
```

(e)

```
new = data.frame(v1=0, v2=1, x=2)
predict(fit,new,interval="prediction", level=0.95)
```

```
##    fit      lwr      upr
## 1 27.8 19.12709 36.47291
```