

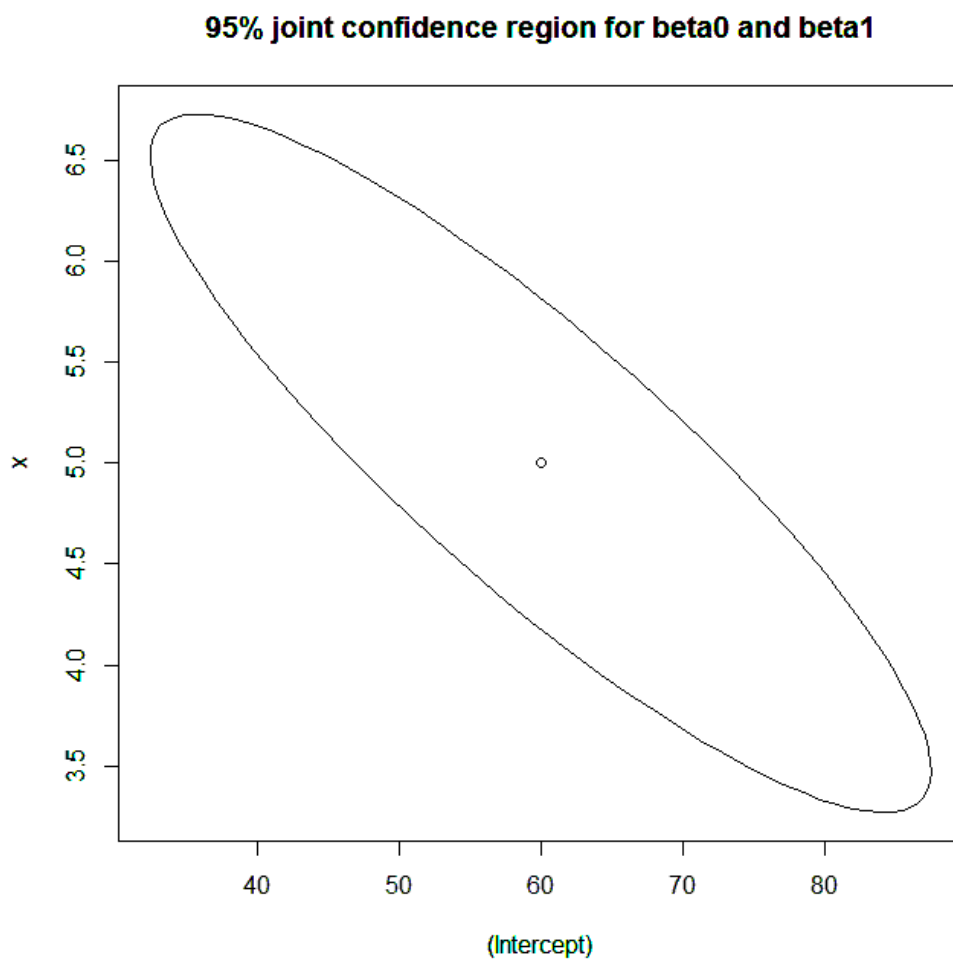
## STAT 420

```
> x = c( 2, 6, 8, 8, 12, 16, 20, 20, 22, 26)
> y = c( 58,105, 88,118,117,137,157,169,149,202)
> N = length(x)
>
> Xmat = cbind(rep(1,N), x)
> Xmat
      x
[1,] 1  2
[2,] 1  6
[3,] 1  8
[4,] 1  8
[5,] 1 12
[6,] 1 16
[7,] 1 20
[8,] 1 20
[9,] 1 22
[10,] 1 26
>
> XX = t(Xmat) %*% Xmat
> XX
      x
    10 140
x 140 2528
>
> XXinv = solve(XX)
> XXinv
      x
 0.44507042 -0.024647887
x -0.02464789  0.001760563
>
> XY = t(Xmat) %*% y
> XY
 [,1]
 1300
x 21040
>
> betahat = XXinv %*% XY
> betahat
 [,1]
  60
x    5
```

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix} \quad \mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}, \quad \text{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}.$$

```
> # joint confidence region for the parameters
>
> fit = lm(y ~ x)
>
> library(ellipse)
> plot(ellipse(fit,c(1,2),level=0.95),type="l")
> # plot(ellipse(fit),type="l")
> title("95% joint confidence region for beta0 and beta1")
> points(fit$coeff[1],fit$coeff[2])
```



**1** Consider the following data set:

$x_1$	$x_2$	$y$
0	1	11
11	5	15
11	4	13
7	3	14
4	1	0
10	4	19
5	4	16
8	2	8

Consider the model

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + e_i, \\ i = 1, \dots, 8.$$

where  $e_i$ 's are i.i.d.  $N(0, \sigma_e^2)$ .

Then

$$\mathbf{X}^T \mathbf{X} = \begin{bmatrix} 8 & 56 & 24 \\ 56 & 496 & 200 \\ 24 & 200 & 88 \end{bmatrix}, \quad \mathbf{X}^T \mathbf{Y} = \begin{bmatrix} 96 \\ 740 \\ 336 \end{bmatrix},$$

$$\mathbf{C} = (\mathbf{X}^T \mathbf{X})^{-1} = \begin{bmatrix} 0.7125 & -0.025 & -0.1375 \\ -0.025 & 0.025 & -0.05 \\ -0.1375 & -0.05 & 0.1625 \end{bmatrix},$$

a) Obtain the least-squares estimates  $\hat{\beta}_0$ ,  $\hat{\beta}_1$ , and  $\hat{\beta}_2$ .

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} = \begin{bmatrix} 0.7125 & -0.025 & -0.1375 \\ -0.025 & 0.025 & -0.05 \\ -0.1375 & -0.05 & 0.1625 \end{bmatrix} \begin{bmatrix} 96 \\ 740 \\ 336 \end{bmatrix} = \begin{bmatrix} 3.7 \\ -0.7 \\ 4.4 \end{bmatrix}.$$

$$SYY = \sum (y - \bar{y})^2 = 240, \quad RSS = \sum (y - \hat{y})^2 = 76.4,$$

b) Perform the significance of the regression test at a 5% level of significance.

$$H_0: \beta_1 = \beta_2 = 0$$

$H_a$ : at least one of  $\beta_1$  and  $\beta_2$  is significantly different from 0.

Source	SS	DF	MS	F
Regression	163.6	$p - 1 = 2$	81.8	<b>5.3534</b>
Error (Residual)	76.4	$n - p = 5$	15.28	
Total	240	$n - 1 = 7$		

Critical Value:  $F_{0.05}(2, 5) = \mathbf{5.79}$ . Reject  $H_0$  if  $F > 5.79$ .

Decision: **Do NOT Reject  $H_0$ .**

- c) Test  $H_0 : \beta_1 = 0$  vs.  $H_a : \beta_1 \neq 0$  at  $\alpha = 0.10$ . Find the p-value.

$$\text{Var}(\hat{\beta}_1) = C_{11} \times s^2 = 0.025 \times 15.28 = 0.382.$$

$$\text{Test Statistic: } t = \frac{-0.7 - 0}{\sqrt{0.382}} \approx \mathbf{-1.1326}.$$

$$n - p = \mathbf{5} \text{ d.f.}$$

$$\begin{array}{ccccc} -1.476 & < & -1.1326 & < & -0.727 \\ -t_{0.10}(5) & < & t & < & -t_{0.25}(5) \end{array}$$

$$0.10 < \text{left tail} < 0.25$$

$$\text{p-value} = 2 \text{ tails. } \quad \mathbf{0.20 < \text{p-value} < 0.50.}$$

$$(\text{p-value} \approx 0.308765.)$$

**Do NOT Reject  $H_0$ .**

- d) Test  $H_0 : \beta_2 = 0$  vs.  $H_a : \beta_2 \neq 0$  at  $\alpha = 0.05$ . Find the p-value.

$$\text{Var}(\hat{\beta}_2) = C_{22} \times s^2 = 0.1625 \times 15.28 = 2.483.$$

$$\text{Test Statistic: } t = \frac{4.4 - 0}{\sqrt{2.483}} \approx \mathbf{2.792}.$$

$$n - p = \mathbf{5} \text{ d.f.}$$

$$\begin{array}{ccccc} 2.517 & < & 2.792 & < & 3.365 \\ t_{0.025}(5) & < & t & < & t_{0.01}(5) \end{array}$$

$$0.01 < \text{right tail} < 0.025$$

p-value = 2 tails.

$$\mathbf{0.02} < \text{p-value} < \mathbf{0.05}.$$

( p-value  $\approx$  0.03834. )

**Reject  $H_0$ .**

a), b), c), d)

```
> x1 = c( 0,11,11, 7, 4,10, 5, 8)
> x2 = c( 1, 5, 4, 3, 1, 4, 4, 2)
> y = c(11,15,13,14, 0,19,16, 8)
>
> fit = lm(y ~ x1 + x2)
>
> summary(fit)
```

Call:

```
lm(formula = y ~ x1 + x2)
```

Residuals:

```
      1      2      3      4      5      6      7      8
2.9 -3.0 -0.6  2.0 -5.3  4.7 -1.8  1.1
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	<b>3.7000</b>	3.2995	1.121	0.3131
x1	<b>-0.7000</b>	0.6181	<b>-1.133</b>	<b>0.3088</b>
x2	<b>4.4000</b>	1.5758	<b>2.792</b>	<b>0.0383</b> *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.909 on 5 degrees of freedom

Multiple R-Squared: 0.6817, Adjusted R-squared: 0.5543

F-statistic: **5.353** on 2 and 5 DF, p-value: **0.05717**

b)  $H_0: \beta_1 = \beta_2 = 0$  vs.  $H_a: \text{not } H_0$ .

$$\text{Full: } Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + e_i,$$

$$\text{Null: } Y_i = \beta_0 + e_i.$$

```

> anova(lm(y~1),fit)
Analysis of Variance Table

Model 1: y ~ 1
Model 2: y ~ x1 + x2
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1       7 240.0
2       5  76.4  2      163.6 5.3534 0.05717 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

c)  $H_0: \beta_1 = 0$  vs.  $H_a: \beta_1 \neq 0$ .

$$\text{Full: } Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + e_i,$$

$$\text{Null: } Y_i = \beta_0 + \beta_2 x_{i2} + e_i.$$

```

> fitc = lm(y ~ x2)
> anova(fitc,fit)
Analysis of Variance Table

Model 1: y ~ x2
Model 2: y ~ x1 + x2
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1       6  96.0
2       5  76.4  1      19.6 1.2827 0.3088

```

d)  $H_0: \beta_2 = 0$  vs.  $H_a: \beta_2 \neq 0$ .

$$\text{Full: } Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + e_i,$$

$$\text{Null: } Y_i = \beta_0 + \beta_1 x_{i1} + e_i.$$

```

> fitd = lm(y ~ x1)
> anova(fitd,fit)
Analysis of Variance Table

Model 1: y ~ x1
Model 2: y ~ x1 + x2
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1       6 195.54
2       5  76.40  1      119.14 7.797 0.03834 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_{p-1} x_{ip-1} + e_i, \quad i = 1, 2, \dots, n,$$

$$E(e_i) = 0, \quad \text{Var}(e_i) = \sigma^2, \quad \text{Cov}(e_i, e_j) = 0, \quad i \neq j.$$

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{e}, \quad E(\mathbf{e}) = \mathbf{0}, \quad \text{Var}(\mathbf{e}) = (\text{Cov}(e_i, e_j))_{ij} = \sigma^2 \mathbf{I}_n.$$

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p-1} \\ 1 & x_{21} & x_{22} & \dots & x_{2p-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np-1} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_{p-1} \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix}.$$

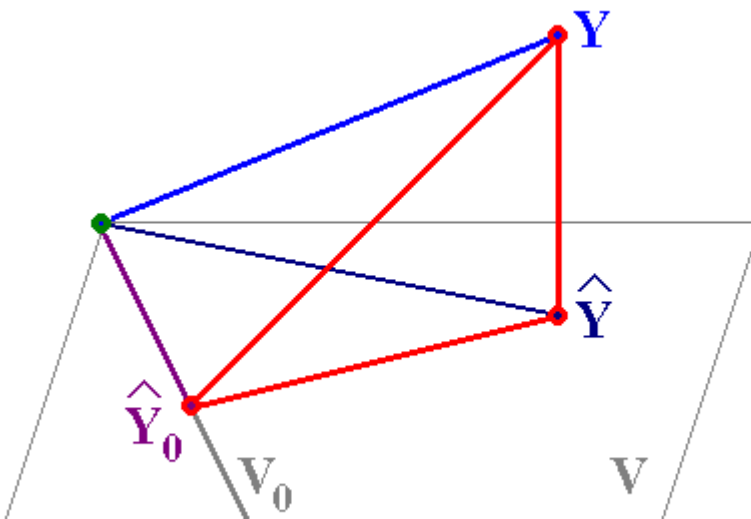
$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}, \quad \text{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}.$$

$$\text{Let } V_0 = \{ a \mathbf{1}, a \in \mathbf{R} \}, \quad \text{where } \mathbf{1} = [1, 1, \dots, 1]^T \in \mathbf{R}^n,$$

$$V = \{ a_0 \mathbf{1} + a_1 \mathbf{x}_1 + \dots + a_{p-1} \mathbf{x}_{p-1}, \quad a_0, a_1, \dots, a_{p-1} \in \mathbf{R} \}.$$

$$\dim(V_0) = 1, \quad \dim(V) = p.$$



$$\sum_{i=1}^n (Y_i - \hat{Y}_{0i})^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \sum_{i=1}^n (\hat{Y}_i - \hat{Y}_{0i})^2 \quad \hat{\mathbf{Y}}_0 = [\bar{Y}, \bar{Y}, \dots, \bar{Y}]^T$$

ANOVA table:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{p-1} = 0$$

Source	SS	df	MS	F
Regr	$\sum_{i=1}^n (\hat{Y}_i - \hat{Y}_{0i})^2$	$\dim(V) - \dim(V_0)$ $p - 1$	$\frac{SS \text{ Regr}}{df \text{ Regr}}$	$\frac{MS \text{ Regr}}{MS \text{ Resid}}$
Resid	$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$	$n - \dim(V)$ $n - p$	$\frac{SS \text{ Resid}}{df \text{ Resid}}$	– estimator for $\sigma^2$
Total	$\sum_{i=1}^n (Y_i - \hat{Y}_{0i})^2$	$n - \dim(V_0)$ $n - 1$		

More general:

$$q < p$$

$$H_0: \beta_q = \beta_{q+1} = \dots = \beta_{p-1} = 0$$

$$\text{Let } V_0 = \{ a_0 \mathbf{1} + a_1 \mathbf{x}_1 + \dots + a_p \mathbf{x}_{q-1}, \quad a_0, a_1, \dots, a_{q-1} \in \mathbf{R} \},$$

$$V = \{ a_0 \mathbf{1} + a_1 \mathbf{x}_1 + \dots + a_p \mathbf{x}_{p-1}, \quad a_0, a_1, \dots, a_{p-1} \in \mathbf{R} \}.$$

	SS	df	MS	F
Diff.	$\sum_{i=1}^n (\hat{Y}_i - \hat{Y}_{0i})^2$	$\dim(V) - \dim(V_0)$ $p - q$	...	...
Full	$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$	$n - \dim(V)$ $n - p$	...	
Null	$\sum_{i=1}^n (Y_i - \hat{Y}_{0i})^2$	$n - \dim(V_0)$ $n - q$		



d 1/4)\* Test  $H_0: \beta_0 = \beta_2$  vs.  $H_a: \beta_0 \neq \beta_2$  at  $\alpha = 0.10$ . Find the p-value.

$$H_0: \beta_0 - \beta_2 = 0 \text{ vs. } H_a: \beta_0 - \beta_2 \neq 0$$

$$X_0^T = [1 \ 0 \ -1] \quad \beta_0 - \beta_2 = X_0^T \boldsymbol{\beta}. \quad X_0^T \hat{\boldsymbol{\beta}} = -0.7.$$

$$X_0^T \mathbf{C} X_0 = [1 \ 0 \ -1] \cdot \begin{bmatrix} 0.7125 & -0.025 & -0.1375 \\ -0.025 & 0.025 & -0.05 \\ -0.1375 & -0.05 & 0.1625 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 1.15.$$

$$\text{Var}(X_0^T \hat{\boldsymbol{\beta}}) = [X_0^T \mathbf{C} X_0] s^2 = 1.15 \times 15.28 = 17.572.$$

$$\text{Test Statistic:} \quad t = \frac{-0.7 - 0}{\sqrt{17.572}} \approx -0.167.$$

$$n - p = 5 \text{ d.f.}$$

$$\begin{array}{ccc} -0.167 & > & -0.267 \\ t & > & -t_{0.40}(5) \end{array}$$

left tail > 0.40

p-value = 2 tails.

p-value > **0.80**.

( p-value  $\approx 0.8739$ .)

**Do NOT Reject  $H_0$ .**

```
> x3 = x2 + 1
> x3
[1] 2 6 5 4 2 5 5 3
>
> fite = lm(y ~ x1 + x3 + 0)
> anova(fite, fit)
Analysis of Variance Table

Model 1: y ~ x1 + x3 + 0
Model 2: y ~ x1 + x2
   Res.Df    RSS Df Sum of Sq    F Pr(>F)
1       6 76.826
2       5 76.400  1     0.426 0.0279  0.874
```

d 1/2)\* Test  $H_0: 2\beta_1 + \beta_2 = 0$  vs.  $H_a: 2\beta_1 + \beta_2 \neq 0$  at  $\alpha = 0.05$ . Find the p-value.

$$X_0^T = [0 \ 2 \ 1] \quad 2\beta_1 + \beta_2 = X_0^T \boldsymbol{\beta}. \quad X_0^T \hat{\boldsymbol{\beta}} = 3.0.$$

$$X_0^T \mathbf{C} X_0 = [0 \ 2 \ 1] \cdot \begin{bmatrix} 0.7125 & -0.025 & -0.1375 \\ -0.025 & 0.025 & -0.05 \\ -0.1375 & -0.05 & 0.1625 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = 0.0625.$$

$$\hat{\text{Var}}(X_0^T \hat{\boldsymbol{\beta}}) = [X_0^T \mathbf{C} X_0] s^2 = 0.0625 \times 15.28 = 0.955.$$

Test Statistic:  $t = \frac{3.0 - 0}{\sqrt{0.955}} \approx \mathbf{3.070}.$

$n - p = \mathbf{5}$  d.f.

$$\begin{array}{ccccc} 2.571 & < & 3.070 & < & 3.365 \\ t_{0.025}(5) & < & t & < & t_{0.01}(5) \end{array}$$

$0.01 < \text{right tail} < 0.025$

p-value = 2 tails.

**$0.02 < \text{p-value} < 0.05$ .**

( p-value  $\approx 0.0278$ .)

**Reject  $H_0$ .**

```
> x4 = x1 - 2*x2
> fitf = lm(y ~ x4)
> anova(fitf, fit)
Analysis of Variance Table

Model 1: y ~ x4
Model 2: y ~ x1 + x2
  Res.Df  RSS Df Sum of Sq    F   Pr(>F)
1      6 220.4
2      5  76.4  1    144.0 9.4241 0.02779 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- e) Construct a 90% prediction interval for the value of Y at  $x_{01} = 2$  and  $x_{02} = 3$ .

$$\mathbf{X}_0^T = [1 \ 2 \ 3] \quad \hat{Y}_0 = 1 \times 3.7 + 2 \times (-0.7) + 3 \times 4.4 = 15.5.$$

$$\mathbf{X}_0^T \mathbf{C} \mathbf{X}_0 = [1 \ 2 \ 3] \cdot \begin{bmatrix} 0.7125 & -0.025 & -0.1375 \\ -0.025 & 0.025 & -0.05 \\ -0.1375 & -0.05 & 0.1625 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 0.75.$$

$$[1 + \mathbf{X}_0^T \mathbf{C} \mathbf{X}_0] s^2 = (1 + 0.75) \times 15.28 = 26.74.$$

$$t_{0.05}(5) = 2.015. \quad 15.5 \pm 2.015 \times \sqrt{26.74} \quad \mathbf{15.5 \pm 10.42}$$

$$\mathbf{(5.08, 25.92)}$$

```
> predict.lm(fit,data.frame(x1=2,x2=3),interval=c("prediction"),level=0.90)
      fit      lwr      upr
[1,] 15.5  5.080037 25.91996
```

- f) Construct a 90% prediction interval for the value of Y at  $x_{01} = 8$  and  $x_{02} = 5$ .

$$\mathbf{X}_0^T = [1 \ 8 \ 5] \quad \hat{Y}_0 = 1 \times 3.7 + 8 \times (-0.7) + 5 \times 4.4 = 20.1.$$

$$\mathbf{X}_0^T \mathbf{C} \mathbf{X}_0 = [1 \ 8 \ 5] \cdot \begin{bmatrix} 0.7125 & -0.025 & -0.1375 \\ -0.025 & 0.025 & -0.05 \\ -0.1375 & -0.05 & 0.1625 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 8 \\ 5 \end{bmatrix} = 0.60.$$

$$[1 + \mathbf{X}_0^T \mathbf{C} \mathbf{X}_0] s^2 = (1 + 0.60) \times 15.28 = 24.448.$$

$$t_{0.05}(5) = 2.015. \quad 20.1 \pm 2.015 \times \sqrt{24.448} \quad \mathbf{20.1 \pm 9.96}$$

$$\mathbf{(10.14, 30.06)}$$

```
> predict.lm(fit,data.frame(x1=8,x2=5),interval=c("prediction"),level=0.90)
      fit      lwr      upr
[1,] 20.1 10.13661 30.06339
```

```

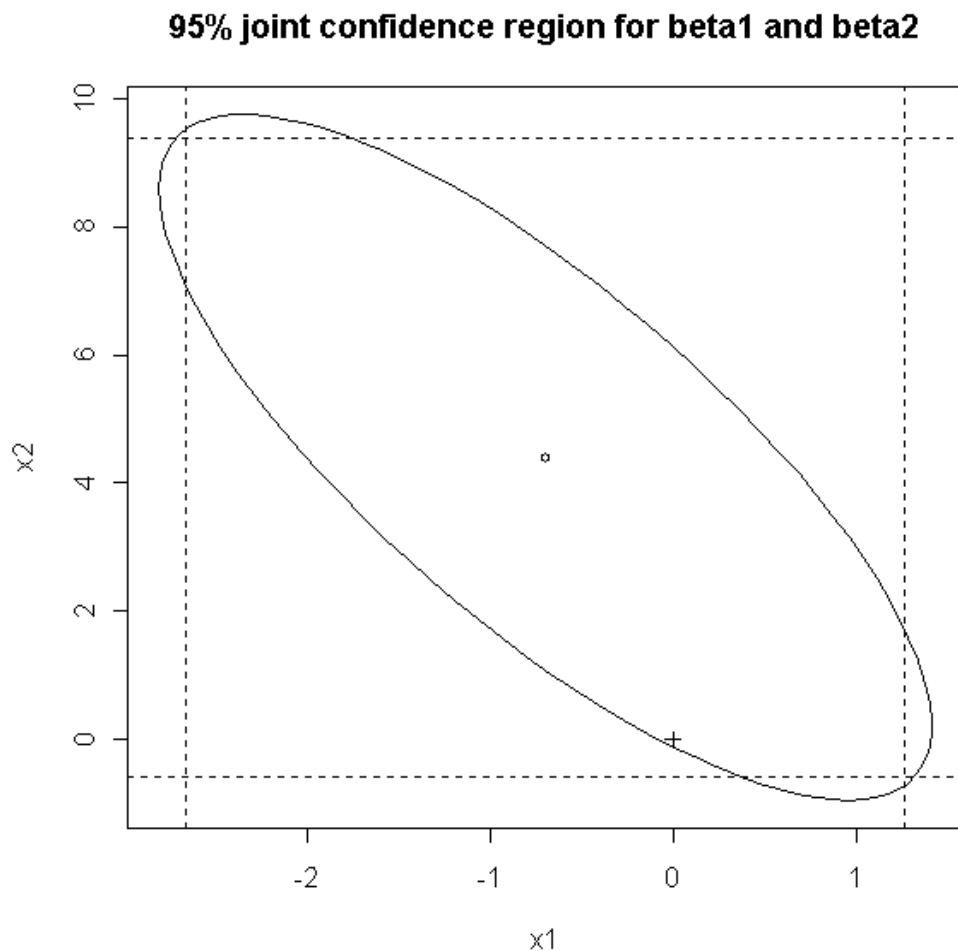
> x1 = c( 0,11,11, 7, 4,10, 5, 8)
> x2 = c( 1, 5, 4, 3, 1, 4, 4, 2)
> y = c(11,15,13,14, 0,19,16, 8)
>
> X = cbind( rep(1,8), x1, x2 )
> X
      x1 x2
[1,] 1  0  1
[2,] 1 11  5
[3,] 1 11  4
[4,] 1  7  3
[5,] 1  4  1
[6,] 1 10  4
[7,] 1  5  4
[8,] 1  8  2
>
> t(X) %*% X
      x1  x2
      8 56 24
x1 56 496 200
x2 24 200  88
>
> solve( t(X) %*% X )
      x1      x2
      0.7125 -0.025 -0.1375
x1 -0.0250  0.025 -0.0500
x2 -0.1375 -0.050  0.1625
>
> t(X) %*% y
[,1]
      96
x1  740
x2  336
>
> solve( t(X) %*% X ) %*% t(X) %*% y
[,1]
      3.7
x1 -0.7
x2  4.4

```

```

> plot(ellipse(fit,c(2,3),level=0.95),type="l")
> title("95% joint confidence region for beta1 and beta2")
> points(fit$coeff[2],fit$coeff[3])
> points(0,0,pch=3)
>
> confint(fit,"x1",level=0.975)
      1.25 %   98.75 %
x1 -2.655164  1.255164
> confint(fit,"x2",level=0.975)
      1.25 %   98.75 %
x2 -0.5847103  9.38471
> abline(v=confint(fit,"x1",level=0.975),lty=2)
> abline(h=confint(fit,"x2",level=0.975),lty=2)

```



Consider  $\mathbf{H}_0: \beta_1 = \beta_2 = 0$  vs.  $H_1$ : at least one of  $\beta_1$  and  $\beta_2$  is not zero.

Recall part (b): **Do NOT Reject  $\mathbf{H}_0$**  at  $\alpha = 0.05$  ( p-value = 0.05717 ).

Note that the origin ( 0, 0 ) lies **inside** the 95% joint confidence region for  $\beta_1$  and  $\beta_2$ .

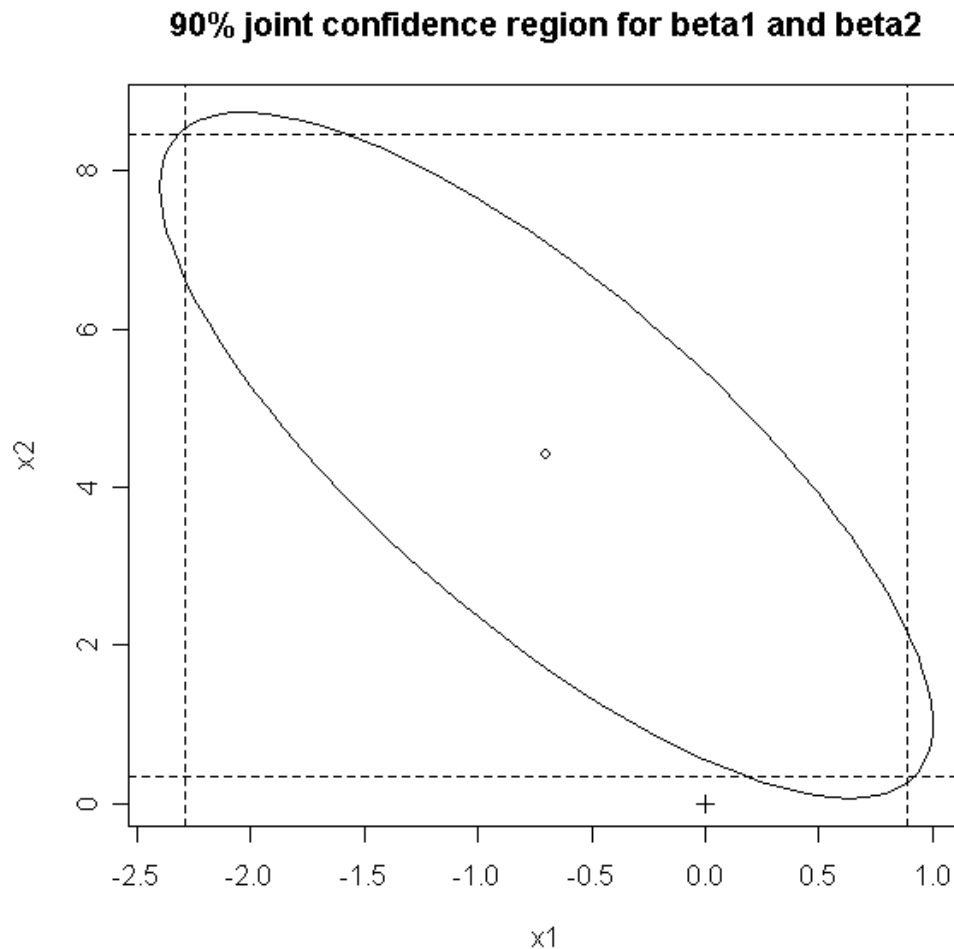
```

> plot(ellipse(fit,c(2,3),level=0.9),type="l")
> title("90% joint confidence region for beta1 and beta2")
> points(fit$coeff[2],fit$coeff[3])
> points(0,0,pch=3)
>
> confint(fit,"x1",level=0.95)
      2.5 %    97.5 %
x1 -2.288778 0.8887776
      ⇒ Do NOT Reject  $H_0: \beta_1 = 0$  at  $\alpha = 0.05$ 

> confint(fit,"x2",level=0.95)
      2.5 %    97.5 %
x2 0.3493959 8.450604
      ⇒ Reject  $H_0: \beta_2 = 0$  at  $\alpha = 0.05$ 

> abline(v=confint(fit,"x1",level=0.95),lty=2)
> abline(h=confint(fit,"x2",level=0.95),lty=2)

```



**Reject  $H_0: \beta_1 = \beta_2 = 0$  at  $\alpha = 0.10$ .**

Note that the origin  $(0, 0)$  lies **outside** the 90% joint confidence region for  $\beta_1$  and  $\beta_2$ .