(c) Predict the number of chirps per second at temperature 60 degrees Fahrenheit. 2 Marks

(d) Under the usual Normality assumptions, $(1-\alpha) \times 100\%$ *confidence* and *prediction* intervals for the predicted response at $x = x^{\text{new}}$ take the form

$$\widehat{y}^{\text{new}} \pm t_{\alpha/2,n-2} \times \widehat{\sigma}\sqrt{\left(\frac{1}{n} + \frac{(x^{\text{new}} - \overline{x})^2}{S_{xx}}\right)} \qquad \text{and} \qquad \widehat{y}^{\text{new}} \pm t_{\alpha/2,n-2} \times \widehat{\sigma}\sqrt{\left(1 + \frac{1}{n} + \frac{(x^{\text{new}} - \overline{x})^2}{S_{xx}}\right)}$$

respectively. Values for $\overline{x}$ and S_{xx} for the study data are given on lines 19 and 21 respectively. The 0.975 quantile of the Student-t distribution with 13 degrees of freedom is 2.1604.

Compute the confidence and prediction intervals for the prediction in (c).

4 Marks

(e) Would the confidence interval for the prediction at 70 degrees be wider or narrower than that computed in (d) ? Justify your answer.

2 Marks

(f) Suppose the units of temperature were changed from degrees Fahrenheit (x) to Celsius (x_2) using the transform

$$x_2 = \frac{5}{9}(x - 32)$$

Compute the values of the estimates of the coefficients in the fit of a simple linear regression of y on x_2.

4 Marks

Answer to Q2: Here is the complete output:

```
> fit.Chirp<-lm(y~x)
> abline(fit.Chirp)
> summary(fit.Chirp)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.45931    2.98920   0.154 0.880239
x            0.20300    0.03754   5.408 0.000119 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05    .    0.1        1

Residual standard error: 0.986 on 13 degrees of freedom
Multiple R-squared:  0.6923,    Adjusted R-squared:  0.6686
F-statistic: 29.25 on 1 and 13 DF,  p-value: 0.0001195

> SST<-sum((y-mean(y))^2)
> SST
[1] 41.07333
> mean(x)
[1] 79.34667
> sum((x-mean(x))^2)
[1] 690.0173
```

(a) Based on the computed test statistic $0.20300/0.03754 = 5.408$ which is clearly strongly significantly positive, so therefore temperature is a strongly significant predictor. There is a positive dependence, and chirp rate increases with temperature.

4 Marks

(b) The R^2 value can be computed as

$$R^2 = 1 - \frac{\text{SS}_{\text{Res}}}{\text{SS}_{\text{T}}} = 1 - \frac{13 \times 0.986^2}{41.07333} = 0.6922931.$$

which confirms the result.

4 Marks

(c) Prediction is

$$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \times 60 = 0.45931 + 0.20300 \times 60 = 12.63931.$$

2 Marks

(d) Confidence interval:

$$12.63931 \pm 2.1604 \times 0.986\sqrt{\left(\frac{1}{15} + \frac{(60 - 79.34667)^2}{690.0173}\right)} = (10.97684, 14.30176)$$

Prediction interval:

$$12.63931 \pm 2.1604 \times 0.986\sqrt{\left(1 + \frac{1}{15} + \frac{(60 - 79.34667)^2}{690.0173}\right)} = (9.937231, 15.34137)$$

Check in R:

```
> predict(fit.Chirp,newdata=data.frame(x=new.temp),interval='conf')
       fit      lwr      upr
1 12.6393 10.97684 14.30176
> predict(fit.Chirp,newdata=data.frame(x=new.temp),interval='pred')
       fit      lwr      upr
1 12.6393 9.937231 15.34137
```

4 Marks

(e) It would be **narrower**, due to the dependence on $(x^* - \overline{x})^2$ in the formula; clearly 70 is closer to the mean 79.34667 than 60 is.

2 Marks

(f) Writing $x = 9x_{new}/5 + 32$, we have that

$$\beta_0 + \beta_1 x = \beta_0 + \beta_1(9x_{new}/5 + 32) = (\beta_0 + 32\beta_1) + (9\beta_1/5)x_{new} = \beta_0^{new} + \beta_1^{new} x_{new}$$

so that $\beta^{\text{new}} = \mathbf{A}\beta$ with $\mathbf{A}$ given by

$$\mathbf{A} = \begin{bmatrix} 1 & 32 \\ 0 & 9/5 \end{bmatrix}$$

So, for the regression on x_{new}, the estimate would be $\mathbf{A}\widehat{\beta}$, and sampling distribution variance would be

$$\sigma^2 \mathbf{A}(\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{A}^\top$$

so that the standard errors would be the square root of the diagonal elements of the this matrix.

4 Marks

Numerically, we have that

$$\mathbf{A}\widehat{\beta} = \begin{bmatrix} 1 & 32 \\ 0 & 9/5 \end{bmatrix} \begin{bmatrix} 0.45931 \\ 0.20300 \end{bmatrix} = \begin{bmatrix} 6.9553061 \\ 0.3653995 \end{bmatrix}$$

For the standard errors: we know that

$$(\mathbf{X}^\top \mathbf{X})^{-1} = \begin{bmatrix} n & \sum_i x_{i1} \\ \sum_i x_{i1} & \sum_i x_{i1}^2 \end{bmatrix}^{-1}$$

and from the earlier output

$$\sum_i x_i = 15 \times 79.34467 = 1190.2 \qquad \sum_i x_i^2 = S_{xx} + n\overline{x}^2 = 690.0173 + 15 \times 79.34467^2 = 95123.67$$

Therefore

$$(\mathbf{X}^\top \mathbf{X})^{-1} = \begin{bmatrix} 15 & 1190.2 \\ 1190.2 & 95123.67 \end{bmatrix}^{-1} = \begin{bmatrix} 9.1909208 & -0.114992280 \\ -0.1149923 & 0.001449239 \end{bmatrix}$$

and as $\widehat{\sigma} = 0.986$, we have

$$\widehat{\sigma}^2 \mathbf{A}(\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{A}^\top = \begin{bmatrix} 3.2232369 & -0.120075265 \\ -0.1200753 & 0.004564957 \end{bmatrix}$$

Thus the standard errors are the square root diagonal elements, that is 1.7953 and 0.0676.

Check in R

```
> xnew<-5*(x-32)/9

> summary(lm(y~xnew))
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.95531    1.79534   3.874 0.001918 **
xnew         0.36540    0.06756   5.408 0.000119 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05    .    0.1       1

Residual standard error: 0.986 on 13 degrees of freedom
Multiple R-squared:  0.6923,    Adjusted R-squared:  0.6686
F-statistic: 29.25 on 1 and 13 DF,  p-value: 0.0001195
```