1 Multivariate normal

a) $X_1 - 2X_2$ is (univariate) normal with mean

$$\mathrm{E}(X_1 - 2X_2) = 1 - 2(-2) = 5$$

and variance

$$\mathrm{Var}(X_1 - 2X_2) = 1 \cdot 1 + (-2)^2 \cdot 2 + 2 \cdot 1 \cdot (-2) \cdot 1 = 5.$$

b) $X_1 | X_2 = x_2$ is normal with mean

$$\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) = 1 + \frac{1}{2}(x_2 + 2) = 2 + \frac{x_2}{2}$$

and variance

$$\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 1 - \frac{1}{2} = \frac{1}{2}.$$

c) Use $\mathrm{Cov}(\mathbf{AX}, \mathbf{BY}) = \mathbf{A}\mathrm{Cov}(\mathbf{X}, \mathbf{Y})\mathbf{B}^T$. Define $\mathbf{A} = [1 \;\; 0]$ and $\mathbf{B} = [1 \;\; c]$, so that $\mathbf{AX} = X_1$ and $\mathbf{BX} = X_1 + cX_2$. Then,

$$\mathrm{Cov}(X_1, X_1 + cX_2) = \mathrm{Cov}(\mathbf{AX}, \mathbf{BX}) = \mathbf{A}\Sigma\mathbf{B}^T = \Sigma_{11} + c\Sigma_{12} = 1 + c$$

Setting $c = -1$ gives $\mathrm{Cov}(X_1, X_1 + cX_2) = 0$ which implies independence.

2 Regression

a)

The missing entries are (1) the Std. Error for $\hat{\beta}_3$ which is $-1.0141/-3.377 = 0.3003$, (2) the R^2 value which we can find from $R^2_{adj} = 1 - \frac{n-1}{n-p}(1 - R^2)$ where $R^2_{adj} = 0.4597$, $n - 1 = 64$, and $n - p = 60$, giving $R^2 = 0.4935$, (3) the degrees of freedom for the F-statistic which is $p - 1 = 4$.

95% **CI for** β_j**:** Let $t_{\alpha,n-p}$ be a critical value such that $P(T_j > t_{\alpha,n-p}) = \alpha$. Then

$$\begin{aligned}
&P(|T_j| > t_{0.025,60}) = 1 - 2 \cdot 0.025 = 0.95 \\
&P(-t_{0.025,60} < \frac{\hat{\beta}_j - \beta_j}{\widehat{\mathrm{SE}}(\hat{\beta}_j)} < t_{0.025,60}) = 0.95 \\
&P(-\hat{\beta}_j - t_{0.025,60} \cdot \widehat{\mathrm{SE}}(\hat{\beta}_j) < -\beta_j < -\hat{\beta}_j + t_{0.025,60} \cdot \widehat{\mathrm{SE}}(\hat{\beta}_j)) = 0.95 \\
&P(\hat{\beta}_j - t_{0.025,60} \cdot \widehat{\mathrm{SE}}(\hat{\beta}_j) < \beta_j < \hat{\beta}_j + t_{0.025,60} \cdot \widehat{\mathrm{SE}}(\hat{\beta}_j)) = 0.95
\end{aligned}$$

Using $t_{0.025,60} = 2$, the 95% CI is

$$\left[\hat{\beta}_j - 2 \cdot \widehat{\mathrm{SE}}(\hat{\beta}_j), \hat{\beta}_j + 2 \cdot \widehat{\mathrm{SE}}(\hat{\beta}_j)\right].$$

For β_1 we have $\hat{\beta}_1 = 2.4094$ and $\widehat{\text{SE}}(\hat{\beta}_1) = 0.4262$. The 95% confidence interval is therefore [1.557, 3.262].

Models A and B can be compared using R^2_{adj}. Since model B has the highest R^2_{adj} we prefer model B over model A.

b)

The distribution of $\hat{\boldsymbol{\beta}}$ is multivariate normal with mean $\boldsymbol{\beta}$ and covariance matrix $\sigma^2(X^TX)^{-1}$. For some new point $\mathbf{x}_0$, $\hat{Y}_0 = \mathbf{x}_0^T\hat{\boldsymbol{\beta}}$ is (univariate) normal with mean $\mathbf{x}_0^T\boldsymbol{\beta}$ and variance $\sigma^2\mathbf{x}_0^T(X^TX)^{-1}\mathbf{x}_0$.

The prediction error $\hat{\varepsilon}_0$ is univariate normal with mean 0 and variance $\sigma^2 + \sigma^2\mathbf{x}_0^T(X^TX)^{-1}\mathbf{x}_0$. The latter follows from Y_0 and $\hat{Y}_0$ independent.

The prediction is $3.5642 + 3 \cdot 1.2523 = 7.3211$. The 95% prediction interval is given by

$$\left[\hat{y}_0 - t_{61,0.025}\sqrt{\hat{\sigma}^2 + \hat{\sigma}^2\mathbf{x}_0^T(X^TX)^{-1}\mathbf{x}_0},\ \hat{y}_0 + t_{61,0.025}\sqrt{\hat{\sigma}^2 + \hat{\sigma}^2\mathbf{x}_0^T(X^TX)^{-1}\mathbf{x}_0}\right].$$

We use $t_{61,0.025} \approx 2$, and note that $\hat{\sigma}^2(X^TX)^{-1} = \widehat{\text{Cov}}(\hat{\boldsymbol{\beta}})$ which is given in the exercise. Further $\hat{\sigma} = 3.857$ can be found in the R output. The solution is

$$[7.3211 - 2\sqrt{14.87645 + 1.2},\ 7.3211 + 2\sqrt{14.87645 + 1.2}] \approx [-0.698,\ 15.340].$$

3 Partial F test

a)

We know that for some random vector $\mathbf{Y}$ which is multivariate normal with mean $\boldsymbol{\mu}$ and covariance $\sigma^2\mathbf{I}$, and for some symmetric and idempotent matrix $\mathbf{A}$ with rank q, then

$$\frac{1}{\sigma^2}(\mathbf{Y} - \boldsymbol{\mu})^T\mathbf{A}(\mathbf{Y} - \boldsymbol{\mu}) \sim \chi^2_q.$$

For some symmetric and idempotent matrix $\mathbf{B}$ with rank r such that $\mathbf{AB} = \mathbf{0}$, we also know that

$$\frac{(\mathbf{Y} - \boldsymbol{\mu})^T\mathbf{A}(\mathbf{Y} - \boldsymbol{\mu})/q}{(\mathbf{Y} - \boldsymbol{\mu})^T\mathbf{B}(\mathbf{Y} - \boldsymbol{\mu})/r} \sim F_{q,r}.$$

These two results will be used to solve the exercise.

The difference in error sums of squares is

$$SSE_0 - SSE = \mathbf{Y}^T(\mathbf{I} - \mathbf{H}_0)\mathbf{Y} - \mathbf{Y}^T(\mathbf{I} - \mathbf{H})\mathbf{Y} = \mathbf{Y}^T(\mathbf{H} - \mathbf{H}_0)\mathbf{Y}.$$

We know that for any column $\mathbf{x}_j$ of $\mathbf{X}$, $\mathbf{H}\mathbf{x}_j = \mathbf{x}_j$, and so $\mathbf{H}\mathbf{X}_0 = \mathbf{X}_0$. Then,

$$\mathbf{H}\mathbf{H}_0 = \mathbf{H}\mathbf{X}_0(\mathbf{X}_0^T\mathbf{X}_0)^{-1}\mathbf{X}_0 = \mathbf{X}_0(\mathbf{X}_0^T\mathbf{X}_0)^{-1}\mathbf{X}_0 = \mathbf{H}_0.$$

Furthermore, since both $\mathbf{H}$ and $\mathbf{H}_0$ are symmetric

$$\mathbf{H}_0^T = (\mathbf{H}\mathbf{H}_0)^T = \mathbf{H}_0\mathbf{H} = \mathbf{H}_0.$$

It follows that

$$(\mathbf{H} - \mathbf{H}_0)(\mathbf{H} - \mathbf{H}_0) = \mathbf{H}\mathbf{H} - \mathbf{H}\mathbf{H}_0 - \mathbf{H}_0\mathbf{H} + \mathbf{H}_0\mathbf{H}_0 = \mathbf{H} - \mathbf{H_0}.$$

Thus, $\mathbf{H} - \mathbf{H_0}$ is an $n \times n$ symmetric and idempotent matrix. The rank of $\mathbf{H} - \mathbf{H_0}$ is $tr(\mathbf{H}) - tr(\mathbf{H}_0) = p - r$. Finally, when H_0 is true and $\boldsymbol{\mu} = \mathbf{X}_0\boldsymbol{\beta}_0$, then

$$(\mathbf{H} - \mathbf{H}_0)(\mathbf{Y} - \mathbf{X}_0\boldsymbol{\beta}_0) = (\mathbf{H} - \mathbf{H}_0)\mathbf{Y} - (\mathbf{H} - \mathbf{H}_0)\mathbf{X}_0\boldsymbol{\beta}_0) = (\mathbf{H} - \mathbf{H}_0)\mathbf{Y}.$$

Then, if H_0 is true,

$$\frac{1}{\sigma^2}(SSE_0 - SSE) = \frac{1}{\sigma^2}\mathbf{Y}^T(\mathbf{H} - \mathbf{H_0})\mathbf{Y} = \frac{1}{\sigma^2}(\mathbf{Y} - \boldsymbol{\mu})^T(\mathbf{H} - \mathbf{H_0})(\mathbf{Y} - \boldsymbol{\mu}) \sim \chi^2_{p-r}.$$

Next, $SSE = \mathbf{Y}^T(\mathbf{I} - \mathbf{H})\mathbf{Y}$, with $(\mathbf{I} - \mathbf{H})$ symmetric, idempotent and with rank $n - p$. Further $(\mathbf{I} - \mathbf{H})\mathbf{Y} = (\mathbf{I} - \mathbf{H})(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ since $\mathbf{I}\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta}$ and $\mathbf{H}\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta}$. Then,

$$\frac{1}{\sigma^2}SSE = \frac{1}{\sigma^2}\mathbf{Y}^T(\mathbf{I} - \mathbf{H})\mathbf{Y} = \frac{1}{\sigma^2}(\mathbf{Y} - \boldsymbol{\mu})^T(\mathbf{I} - \mathbf{H})(\mathbf{Y} - \boldsymbol{\mu}) \sim \chi^2_{n-p}.$$

Finally, $(\mathbf{H} - \mathbf{H_0})(\mathbf{I} - \mathbf{H}) = \mathbf{H} - \mathbf{H_0} - \mathbf{H}\mathbf{H} + \mathbf{H_0}\mathbf{H} = \mathbf{H} - \mathbf{H_0} - \mathbf{H} + \mathbf{H_0} = 0$. It follows that $F_1 \sim F_{p-r,n-p}$ when H_0 is true.

b)

Our null hypothesis $H_0 : \boldsymbol{\beta}_1 = \mathbf{0}$ can be expressed as a general linear hypothesis $\mathbf{C}\boldsymbol{\beta} = \mathbf{d}$ with $\mathbf{d} = \mathbf{0}$ and

$$\mathbf{C} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ & & & & & \ddots & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Then $\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{d} = \hat{\boldsymbol{\beta}}_1$. By left-multiplying both sides of the equation of the hint by $\mathbf{Y}$, we obtain

$$\mathbf{Y}^T(\mathbf{I} - \mathbf{H}_0)\mathbf{Y} = \mathbf{Y}^T(\mathbf{I} - \mathbf{H})\mathbf{Y} + \mathbf{Y}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T(\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T)^{-1}\mathbf{C}\hat{\boldsymbol{\beta}}$$
$$\downarrow$$
$$SSE_0 = SSE + (\mathbf{C}\hat{\boldsymbol{\beta}})^T(\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T)^{-1}\mathbf{C}\hat{\boldsymbol{\beta}}$$
$$\downarrow$$
$$SSE_0 - SSE = \hat{\boldsymbol{\beta}}_1(\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T)^{-1}\hat{\boldsymbol{\beta}}_1$$

We use that $\mathbf{Cov}(\mathbf{A}\mathbf{Y}) = \mathbf{A}\mathrm{Cov}(\mathbf{Y})\mathbf{A}^T$ for some random vector $\mathbf{Y}$. Here, $\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$. Then

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}_1) = \mathrm{Cov}(\mathbf{C}\hat{\boldsymbol{\beta}}) = \sigma^2\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T.$$

Futher, $\widehat{\text{Cov}}(\hat{\boldsymbol{\beta}}_1) = \hat{\sigma}^2 \mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T$, where $\hat{\sigma}^2 = SSE/(n-p)$. Therefore,

$$F_1 = \frac{(SSE_0 - SSE)/(p-r)}{SSE/(n-p)} = \frac{\hat{\sigma}^2 \hat{\boldsymbol{\beta}}_1 \widehat{\text{Cov}}(\hat{\boldsymbol{\beta}}_1)^{-1} \hat{\boldsymbol{\beta}}_1/(p-r)}{\hat{\sigma}^2} = \frac{1}{p-r} \hat{\boldsymbol{\beta}}_1 \widehat{\text{Cov}}(\hat{\boldsymbol{\beta}}_1)^{-1} \hat{\boldsymbol{\beta}}_1 = F_2.$$

4 2-level fractional factorial designs

a)

Since ABCD = E, then BCD = AE etc so that 2-factor interactions are aliased with 3-factor interactions. The resolution is five, $R = \text{V}$ since no p-factor effect is aliased with an effect with less than $R - p$ factors, e.g. 1-factor effects aliased with $5 - 1 = 4$ and 2-factor effects aliased with $5 - 2 = 3$.

b)

Model:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_4 + \beta_{14} x_1 x_4 + \beta_{15} x_1 x_5 + \varepsilon,$$

where $x_i \in \{-1, 1\}$ for factor A at low and high level, etc. The effect of factor A is $2\beta_1$, while $2\beta_{12}$ is the interaction between A and B etc. The error term ε is normal with mean 0 and variance σ^2. The intercept β_0 represents a global mean. The response Y represents a measurement taken at set levels of the factors.

c)

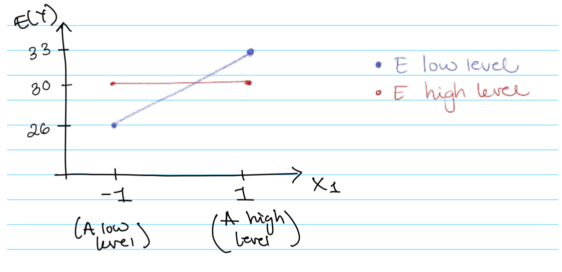
We do 9 tests, so $0.05/9 = 0.00556$ is our local significance level. Then, effects A, D and AE are significant.

To draw the sketch, we can calculate estimated expected outcomes at various settings;

$$\hat{\text{E}}(Y|\text{A and E low level}) = \hat{\beta}_0 - \hat{\beta}_1 - \hat{\beta}_5 + \hat{\beta}_{15} = 26.2$$
$$\hat{\text{E}}(Y|\text{A high, E low level}) = \hat{\beta}_0 + \hat{\beta}_1 - \hat{\beta}_5 - \hat{\beta}_{15} = 33.5$$
$$\hat{\text{E}}(Y|\text{A low, E high level}) = \hat{\beta}_0 - \hat{\beta}_1 + \hat{\beta}_5 - \hat{\beta}_{15} = 29.3$$
$$\hat{\text{E}}(Y|\text{A and E high level}) = \hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_5 + \hat{\beta}_{15} = 30.2$$

The sketch is given below.

The interaction coefficient $\hat{\beta}_{15}$ is the difference in slopes between the two lines in the figure. We observe that when E is at high level, the level of A is 'irrelevant', while when E is at low level, a high level of A results in a greater expected outcome than if A is at low level.



We have that $SSR = n\hat{\beta}_1^2 + n\hat{\beta}_2^2 + \cdots$. And so the propotion of the total sums of squares that is accoutned for by the main effect A in the model is

$$n\hat{\beta}_1^2 / SST = 16 \cdot 2.03032^2 / 205.0 \approx 0.32.$$