Solution TMA 4267 August 2012

1a)

X_1, X_2 and X_2, X_3 are independent since they are simultaneous normally distributed and uncorrelated.

$$f_{X_1}(x) = \frac{1}{\sqrt{2\pi}\cdot 2} e^{-\frac{1}{2}\left(\frac{x-1}{2}\right)^2}$$

$$E[Y] = E[X_1] + 3E[X_2] - 2E[X_3] = 1 - 3 - 4 = -6$$

$$Var[Y] = [1, 3, -2]\begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} = [1, 3, -2]\begin{bmatrix} 6 \\ 15 \\ -5 \end{bmatrix} = 61$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sqrt{61}} e^{-\frac{1}{2}\left(\frac{y+6}{\sqrt{61}}\right)^2}$$

b)

$$\begin{bmatrix} X_1 \\ X_3 \end{bmatrix} \sim N_2\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}\right)$$

$$E[X_1 \mid X_3 = x_3] = 1 - \frac{1}{2}(x_3 - 2) = 2 - \frac{x_3}{2}$$

$$Var[X_1 \mid X_3 = x_3] = 4 - \frac{(-1)\cdot(-1)}{2} = \frac{7}{2}$$

Hence $X_1 \mid X_3 = x_3 \sim N\left(2 - \frac{x_3}{2}, \frac{7}{2}\right)$

$$X_1 = 2 - \frac{X_3}{2} + \varepsilon, \qquad \alpha = 2, \ \beta = -\frac{1}{2}, \ \sigma_\varepsilon^2 = \frac{7}{2}$$

2a) 16 runs

4 main effects, 6 two-factor interactions, 4 three-factor interactions, 1 four-factor interaction.

Label the four factors A, B, C and D. An estimator for the variance of an effect is: $\frac{\hat{ABC}^2 + \hat{ABD}^2 + \hat{ACD}^2 + \hat{BCD}^2 + \hat{ABCD}^2}{5} = S^2_{effect}$

Test Compare $|\hat{Effect}|$ with $t_{\frac{\alpha}{2},5} \cdot S_{effect}$

b) $H_0: \beta_1 = \cdots = \beta_{14} = 0$.

$F = \frac{SS_R}{14} \Big/ \frac{SS_E}{13}$. $P(F \geq F_{obs}) = 4.489 \cdot 10^{-11}$ which implies that the regression is significant.

$SS_E = (0,09219)^2 \cdot 13 = 0,1105$

$SS_R = 14 \cdot F \cdot (0,09219)^2 = 14 \cdot 107.3 \cdot (0,09219)^2 = 12,767$

c) $\hat{y} = 5.79 + 0.47\, x_1 + 0.27\, x_2 + 0,46\, x_3 + 0,15\, x_4 - 0.08\, x_1 x_3 - 0.05\, x_1^2$

The design columns for 1. order and crossed terms are all orthogonal to the columns for the quadratic terms. Thereby the estimate are not influenced of what quadratic terms there are in the model if any.

The design columns for quadratic terms are not orthogonal to the column for the intercept and ~~for design~~ the columns for other quadratic terms. Thereby ~~they~~ the estimates will be influenced.

d)

$\underline{J}_c$ and $\underline{H}$ are symmetric. Thereby $(\underline{I} - \underline{J}_c)$ and $(\underline{J}_c - \underline{H})$ are also symmetric.

$$(\underline{I} - \underline{J}_c)^2 = (\underline{I} - \underline{J}_c)(\underline{I} - \underline{J}_c) = \underline{I} - \underline{J}_c - \underline{J}_c + \underbrace{\underline{J}_c \cdot \underline{J}_c}_{\underline{J}_c} = \underline{I} - \underline{J}_c$$

$$(\underline{J}_c - \underline{H})^2 = (\underline{J}_c - \underline{H})(\underline{J}_c - \underline{H}) = \underline{J}_c \cdot \underline{J}_c - \underline{H}\,\underline{J}_c - \underline{J}_c \underline{H} + \underline{H} \cdot \underline{H} = \underline{J}_c - \underline{H}$$

Thereby $(\underline{I} - \underline{J}_c)$ and $(\underline{J}_c - \underline{H})$ are projection matrices

$$\text{Rank}\,(\underline{I} - \underline{J}_c) = tr(\underline{I}) - tr(\underline{J}_c) = 28 - 25 = \underline{3}$$

e)

$$(\underline{I} - \underline{J}_c)\underline{Y} = \begin{bmatrix} Y_1 - Y_1 \\ \vdots \\ Y_{24} - Y_{24} \\ Y_{25} - \bar{Y}_c \\ \vdots \\ Y_{28} - \bar{Y}_c \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ Y_{25} - \bar{Y}_c \\ \vdots \\ Y_{28} - \bar{Y}_c \end{bmatrix}$$

$$(\underline{I} - \underline{J}_c)\begin{bmatrix} E(Y_1) \\ \vdots \\ E(Y_{24}) \\ E(Y_{25}) \\ \vdots \\ E(Y_{28}) \end{bmatrix} = \begin{bmatrix} E(Y_1) - E(Y_1) \\ \vdots \\ E(Y_{24}) - E(Y_{24}) \\ \beta_0 - \beta_0 \\ \vdots \\ \beta_0 - \beta_0 \end{bmatrix} = \underline{0}$$

We then have $\dfrac{SS_{EC}}{\sigma^2} = \dfrac{(\underline{Y} - E(\underline{Y}))^t(\underline{I} - \underline{J}_c)(\underline{Y} - E(\underline{Y}))}{\sigma^2}$ and since

$\underline{I} - \underline{J}_c$ is a projection matrix, we have

$$\frac{SS_{EC}}{\sigma^2} \sim \chi^2(3)$$

f) $SS_R = \underline{Y}^t(\underline{H} - \frac{1}{m}\underline{J})\underline{Y}$ and under the assumption that all effects are 0 we have that $\frac{\underline{Y}^t(\underline{H} - \frac{1}{m}\underline{J})\underline{Y}}{\sigma^2}$ $\sim \chi^2(\text{rank}(H - \frac{1}{m}J))$ i.e. $\chi^2(15-1)$ or $\chi^2(14)$

$$(\underline{H} - \tfrac{1}{m}\underline{J})(\underline{I} - \underline{J}_c) = \underline{H} - \tfrac{1}{m}\underline{J} - \underbrace{\underline{H}\,\underline{J}_c}_{H} + \underbrace{\tfrac{1}{m}\underline{J}\cdot\underline{J}_c}_{\frac{1}{n}\underline{J}} = \underline{0}$$

Hence SS_R and SS_{EC} are independent

Thereby $\frac{SS_R}{\sigma^2 k_1} \Big/ \frac{SS_{EC}}{\sigma^2 k_2}$ under the assumptions of zero effects are F distributed with 14 and 3 degrees of freedom if. $k_1 = 14$ and $k_2 = 3$.

$$F_{obs} = \frac{12.767}{14} \Big/ \frac{0.000477}{3} = 5735$$

i.e. the regression is significant on all possible significance levels.