1a.

$$\mu = \begin{bmatrix} 60 \\ 55 \\ 80 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} 121 & 80 & 10 \\ 80 & 144 & 15 \\ 10 & 15 & 9 \end{bmatrix}$$

$$S = 0.5F + 0.3M + 0.2H$$

$$\therefore w = [0.5 \quad 0.3 \quad 0.2]$$

$$E[S] = [0.5 \quad 0.3 \quad 0.2] \begin{bmatrix} 60 \\ 55 \\ 80 \end{bmatrix}$$

$$= 0.5(60) + 0.3(55) + 0.2(80)$$

$$= 30 + 16.5 + 16$$

$$= 62.5$$

$$Var(S) = [0.5 \quad 0.3 \quad 0.2] \begin{bmatrix} 121 & 80 & 10 \\ 80 & 144 & 15 \\ 10 & 15 & 9 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.3 \\ 0.2 \end{bmatrix}$$

$$a_{1,1} = 0.5(121) + 0.3(80) + 0.2(10) = 86.5$$

$$a_{1,2} = 0.5(80) + 0.3(144) + 0.2(15) = 86.2$$

$$a_{1,3} = 0.5(10) + 0.3(15) + 0.2(9) = 11.3$$

$$= [86.5 \quad 86.2 \quad 11.3] \begin{bmatrix} 0.5 \\ 0.3 \\ 0.2 \end{bmatrix}$$

$$= 86.5(0.5) + 86.2(0.3) + 11.3(0.2)$$

$$= 43.25 + 25.86 + 2.26$$

$$= 71.37$$

$$\therefore S \sim Normal(62.5, 71.37)$$

1b. Least squares predictor of F based on X is linear because X is a linear combination of M and H, and F, M, H follow a MVN distribution. The least squares predictor of F based on X is guaranteed to be linear in X.

$$X = 0.3M + 0.2H$$

$$E[X] = \begin{bmatrix} 0 & 0.3 & 0.2 \end{bmatrix} \begin{bmatrix} 60 \\ 55 \\ 80 \end{bmatrix}$$

$$= 0 + 0.3(55) + 0.2(80)$$

$$= 32.5$$

$$\begin{aligned} Var(X) &= Var(0.3M + 0.2H) \\ &= 0.09\,Var(M) + 0.04\,Var(H) + 2(0.3)(0.2)\,Cov(M, H) \\ &= 0.09(144) + 0.04(9) + 0.12(15) \\ &= 12.96 + 0.36 + 1.8 \\ &= 15.12 \end{aligned}$$

$$\begin{aligned} Cov(F, X) &= Cov(F, 0.3M + 0.2H) \\ &= Cov(F, 0.3M) + Cov(F, 0.2H) \\ &= 0.3\,Cov(F, M) + 0.2\,Cov(F, H) \\ &= 0.3(80) + 0.2(10) \\ &= 24 + 2 \\ &= 26 \end{aligned}$$

$$\begin{aligned} \hat{F} &= 60 + \frac{26}{15.12}(X - 32.5) \\ &= 60 + 1.7196(X - 32.5) \\ &= 60 + 1.7196X - 55.887 \\ &= 1.7196X + 4.113 \end{aligned}$$

1c. $RMSE(F, \hat{F}) = \sqrt{MSE(F, \hat{F})}$

$$MSE(F, \hat{F}) = E[(F, \hat{F})^2]$$

$$= Var(F \mid x)$$

$$= 121 - \frac{(26)^2}{15.12}$$

$$= 121 - \frac{676}{15.12}$$

$$= 121 - 44.7$$

$$= 76.3$$

$$\therefore RMSE = \sqrt{76.3}$$

$$\approx 8.73$$

2a. $\text{Cov}(D_i, \bar{X}) = \text{cov}(X_i - \bar{X}, \bar{X})$

$= \text{cov}(X_i, \bar{X}) - \text{cov}(\bar{X}, \bar{X})$

$\text{Cov}(X_i, \bar{X}) = \text{Cov}(X_i, \frac{1}{n}\sum_{j=1}^{n} X_j)$

$= \frac{1}{n}\sum_{j=1}^{n} \text{cov}(X_i, X_j)$

$= \begin{cases} \sigma^2 & i = j \\ 0 & i \neq j \end{cases}$

$= \frac{\sigma^2}{n}$

$\text{Cov}(\bar{X}, \bar{X}) = \text{Var}(\bar{X})$

$= \frac{\sigma^2}{n}$

$\therefore \text{Cov}(D_i, \bar{X}) = \frac{\sigma^2}{n} - \frac{\sigma^2}{n}$

$= 0$

2b. $X_i \sim \text{Normal}(\mu, \sigma^2)$

$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$

$D_i = X_i - \bar{X}$

$\bar{X}, D_1, D_2, \ldots, D_{n-1} \sim MVN$

↳ each variable is a linear combination of X_i's and X_i is normal distribution

$D_1 + D_2 + \ldots + D_n = (X_1 - \bar{X}) + (X_2 - \bar{X}) + \ldots + (X_n - \bar{X})$

$= (X_1 + X_2 + \ldots + X_n) - n\bar{X}$

$= n\bar{X} - n\bar{X}$

$= 0$

$\therefore$ D_n is NOT on the list because it is already determined by $D_1, D_2, \ldots, D_{n-1}$ so once you know those values, it doesn't give you new information

2c. $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

$= \frac{1}{n-1} \sum_{i=1}^{n} D_i^2$

known:

- $\bar{X}, D_1, D_2, \dots, D_{n-1} \sim MVN$
- $Cov(D_i, \bar{X}) = 0$

$\therefore \bar{X}$ independent of ALL D_i

$\sum_{i=1}^{n} D_i^2 = \sum_{i=1}^{n-1} D_i^2 + D_n^2$

$\therefore D_n$ is linear combination of $D_1, D_2, \dots, D_n$

$\hookrightarrow S^2$ is function of $D_1, \dots, D_{n-1}$

$\therefore \bar{X}$ is independent of S^2

$\therefore$ True, sample mean and sample variance of an iid normal sample are independent of each other

3a. $R = Z_1^2 + Z_2^2 + \dots + Z_n^2$

$R \sim \chi_n^2$

$Z_i \sim \text{Normal}(0,1)$

$$M_{Z_i^2}(t) = E\left[e^{t \cdot Z_i^2}\right]$$

$$= \frac{1}{\sqrt{1-2t}}$$

$$M_R(t) = \underbrace{\frac{1}{\sqrt{1-2t}}}_{1} \times \underbrace{\frac{1}{\sqrt{1-2t}}}_{2} \times \dots \times \underbrace{\frac{1}{\sqrt{1-2t}}}_{n}$$

$$= \left(\frac{1}{\sqrt{1-2t}}\right)^n \qquad t < \frac{1}{2}$$

3b. $R = V + W$

$R \sim \chi_n^2$

$V \sim \chi_m^2$

V, W are independent

$$M_R(t) = M_V(t)\, M_W(t)$$

$$\left(\frac{1}{\sqrt{1-2t}}\right)^n = \left(\frac{1}{\sqrt{1-2t}}\right)^m \cdot M_W(t)$$

$$M_W(t) = \left(\frac{1}{\sqrt{1-2t}}\right)^{n-m}$$

$$= \chi_{n-m}^2$$

3c. Prove $\sum_{i=1}^{n} (X_i - \alpha)^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 + n(\bar{X} - \alpha)^2$ ✓

$$\sum_{i=1}^{n} (X_i - \alpha)^2 = \sum_{i=1}^{n} ((X_i - \bar{X}) + (\bar{X} - \alpha))^2$$

$$= \sum_{i=1}^{n} (X_i - \bar{X})^2 + 2(X_i - \bar{X})(\bar{X} - \alpha) + (\bar{X} - \alpha)^2$$

$$= \sum_{i=1}^{n} (X_i - \bar{X})^2 + 2(X_i - \alpha) \underbrace{\sum_{i=1}^{n} (X_i - \bar{X})}_{=0} + \underbrace{\sum_{i=1}^{n} (\bar{X} - \alpha)^2}_{= n(\bar{X} - \alpha)^2}$$

idea: sum of deviations from mean is 0

idea: quantity inside sum doesn't depend on i

$$= \sum_{i=1}^{n} (X_i - \bar{X})^2 + n(\bar{X} - \alpha)^2$$ ✓

3d. $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

$$\sum_{i=1}^{n} (X_i - \alpha)^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 + n(\bar{X} - \alpha)^2$$

$$\sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2$$

$$\underbrace{\frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2}_{= \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 = \chi^2_n} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 + \underbrace{\frac{n}{\sigma^2}(\bar{X} - \mu)^2}_{= \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 = \chi^2_1}$$

$$\chi^2_n = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 + \chi^2_1$$

$$\chi^2_n = \underbrace{\frac{1}{\sigma^2}(n-1) S^2}_{= \chi^2_{n-1}} + \chi^2_1$$

$$c S^2 = \chi^2_{n-1}$$

$$c S^2 = \frac{1}{\sigma^2}(n-1) S^2$$

$$c = \frac{n-1}{\sigma^2}$$

4a. $Y = X\beta + \varepsilon$

$$\begin{aligned}
\hat{\beta} &= (X^T X)^{-1} X^T Y \\
&= (X^T X)^{-1} X^T (X\beta + \varepsilon) \\
&= (X^T X)^{-1} X^T X\beta + (X^T X)^{-1} X^T \varepsilon \\
&= \beta + (X^T X)^{-1} X^T \varepsilon
\end{aligned}$$

$$\begin{aligned}
E[\hat{\beta}] &= E[\beta + (X^T X)^{-1} X^T \varepsilon] \\
&= \beta + (X^T X)^{-1} X^T \underbrace{E[\varepsilon]}_{=0} \\
&= \beta
\end{aligned}$$

$$\begin{aligned}
\text{Cov}(Y) &= \text{Cov}(X\beta + \varepsilon) \\
&= \text{Cov}(\varepsilon) \\
&= \sigma^2 I
\end{aligned}$$

$$\begin{aligned}
\text{Cov}(\hat{\beta}) &= (X^T X)^{-1} X^T (\sigma^2 I)((X^T X)^{-1} X^T)^T \\
&= \cancel{(X^T X)^{-1} X^T} (\sigma^2 I)(X^T X)^{-1} \cancel{X} \\
&= \sigma^2 I (X^T X)^{-1} \\
&= \sigma^2 (X^T X)^{-1}
\end{aligned}$$

$\therefore \hat{\beta} \sim \text{Normal}(\beta, \sigma^2 (X^T X)^{-1})$

4b. $\hat{Y} = X(X^T X)^{-1} X^T Y$

$\hat{Y} = HY$

$H = X(X^T X)^{-1} X^T$

$$\begin{aligned}
H^T &= H\,? \\
(X(X^T X)^{-1} X^T)^T &= X(X^T X)^{-1} X^T \\
X(X^T X)^{-1} X^T &= X(X^T X)^{-1} X^T \\
H^T &= H \checkmark
\end{aligned}$$

$\therefore$ H is symmetric

4c. $H^2 = H$ if idempotent

$$
\begin{aligned}
H^2 &= H \cdot H \\
&= X(X^TX)^{-1}\underbrace{X^T \cdot X(X^TX)^{-1}}_{I}X^T \\
&= X(X^TX)^{-1} I X^T \\
&= X(X^TX)^{-1}X^T \\
&= H
\end{aligned}
$$

$\therefore$ H is idempotent

4d.
$$
\begin{aligned}
e &= Y - \hat{Y} \\
&= Y - HY \\
&= (I - H)Y \\
&= (I-H)(X\beta + \varepsilon) \\
&= (I-H)X\beta + (I-H)\varepsilon
\end{aligned}
$$

$$
\begin{aligned}
(I-H)X\beta &= (I-H)X \\
&= X - HX \\
&= X - X\underbrace{(X^TX)^{-1}X^TX}_{=I} \\
&= X - XI \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
e &= 0 + (I-H)\varepsilon \\
&= (I-H)\varepsilon
\end{aligned}
$$

$\varepsilon \sim \text{Normal}(0, \sigma^2 I)$

$$
\begin{aligned}
e &\sim \text{Normal}((I-H)0,\ (I-H)\sigma^2 I (I-H)^T) \\
&\sim \text{Normal}(0,\ \sigma^2 \underbrace{(I-H)(I-H)^T}_{\therefore \text{ symmetric, idempotent}}) \\
&\sim \text{Normal}(0, \sigma^2(I-H))
\end{aligned}
$$

e. known: $e = (I - H)\varepsilon$

$$
\begin{aligned}
cov(e) &= Cov((I-H)\varepsilon) \\
&= (I-H)\,\sigma^2 \underline{I}\,(I-H)^T \\
&= \sigma^2 (I-H)(\underline{I}-H)^T \\
&= \sigma^2 (\underline{I}-H)(\underline{I}-H) \\
&= \sigma^2 (\underline{I}-H)^2 \\
&= \sigma^2 (\underline{I}-H)
\end{aligned}
$$