

Problem 1

a) $Y_2 \sim N(-1, 2)$

$$\begin{aligned} E(Y_2 | Y_1 = y_1) &= \mu_2 + \Sigma_{21} \Sigma_{11}^{-1} (y_1 - \mu_1) \\ &= -1 + 1 \cdot 1 (y_1 - 1) \\ &= -2 + y_1 \quad \beta_0 = -2 \quad \beta_1 = 1. \end{aligned}$$

b) $\vec{Z} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & -1 \end{bmatrix} \vec{Y}$

$\vec{Z}$ is $MVN_2(\vec{\mu}_Z, \Sigma_Z)$ where

$$\vec{\mu}_Z = E(\vec{Z}) = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \Sigma_Z = \text{cov}(\vec{Z}) &= \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 4 & 4 \\ 4 & 4 & -2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} 10 & 6 \\ 6 & 18 \end{bmatrix}$$

Problem 2

- a) 2 levels each factor (since $Df=1$)
10 experiments at each combination
since total $n = 36 + 3 + 1 = 40$

$$?1 : 1$$

$$?2 : 704.64 / 36 = 19.5733$$

$$?3 : 23.46 / 19.57 = 1.1986$$

Problem 3

$$\begin{aligned} a) \quad H \cdot H &= X \underbrace{(X^T X)^{-1} X^T}_{I} X (X^T X)^{-1} X^T \\ &= X (X^T X)^{-1} X^T = H \quad (\text{idempotent}) \end{aligned}$$

$$\begin{aligned} \text{rank}(H) &= \text{tr}(H) = \text{tr}(X (X^T X)^{-1} X^T) \\ &= \text{tr}(X^T X (X^T X)^{-1}) = \text{tr}(I_p) \\ &= p. \end{aligned}$$

Using rank = trace for idempotent

$$\text{tr} \left(\begin{matrix} A & B \\ n \times p & p \times n \end{matrix} \right) = \text{tr} \left(\begin{matrix} B & A \\ n \times p & p \times n \end{matrix} \right)$$

$$\begin{aligned}
 b) \quad E(\vec{Y}^T H \vec{Y}) &= \text{tr}(H \sigma^2 I) + (X\vec{\beta})^T H (X\vec{\beta}) \\
 &= \sigma^2 \text{tr}(H) + \vec{\beta}^T X^T \underbrace{H X}_{X} \vec{\beta} \\
 &= \sigma^2 p + \vec{\beta}^T X^T X \vec{\beta}
 \end{aligned}$$

$$\frac{1}{\sigma^2} \vec{Y}^T H \vec{Y} \sim \chi^2(r, \lambda) \quad r=p$$

$$\lambda = (X\vec{\beta})^T H X\vec{\beta} = \vec{\beta}^T X^T X \vec{\beta} / \sigma^2$$

$$c) \quad U = \frac{1}{\sigma^2} \vec{Y}^T H \vec{Y} \sim \chi^2(p, \lambda) \quad \text{noncentral}$$

$$\frac{1}{\sigma^2} \vec{Y}^T (I-H) \vec{Y} \sim \chi^2(n-p, \lambda_2)$$

$$\lambda_2 = \frac{1}{2\sigma^2} X^T (I-H) X = \frac{1}{2\sigma^2} (X^T X - X^T X) = 0$$

$$V = \frac{1}{\sigma^2} \vec{Y}^T (I-H) \vec{Y} \sim \chi^2(n-p) \quad \text{central}$$

U and V independent since
 $H(I-H) = H - H = 0$.

$$\text{Then } \frac{U/p}{V/(n-p)} \sim F(p, n-p, \lambda).$$

Problem 4 $\vec{Y} = X_1 \vec{\beta}_1 + X_2 \vec{\beta}_2 + \varepsilon \quad \varepsilon \sim \text{MVN}(0, \sigma^2 V)$

$$\tilde{\beta}_1 = (X_1^T X_1)^{-1} X_1^T \vec{Y}$$

$$\begin{aligned} E(\tilde{\beta}_1) &= (X_1^T X_1)^{-1} X_1^T E(\vec{Y}) \\ &= (X_1^T X_1)^{-1} X_1^T (X_1 \vec{\beta}_1 + X_2 \vec{\beta}_2) \\ &= \vec{\beta}_1 + (X_1^T X_1)^{-1} X_1^T X_2 \vec{\beta}_2 \end{aligned}$$

X_2 orthogonal to X_1 . No requirement for V .

Problem 5

a) $\hat{\beta}_3 \approx 78.02$

95% CI, $t_{21, 0.025} = 2.080$

$$78.02 \pm 2.080 \cdot 56.72$$

$$= [-39.9576, 195.9976] = [-40, 196]$$

The estimated coefficient of premium-location is 78.02, representing an expected turnover for a premium location of approx 78000 nok higher than non-premium, distance and location held constant. However, this is estimated with

large uncertainty, and a test of $H_0: \beta_3 = 0$ against $H_1: \beta_3 \neq 0$ can not be rejected at a 5% sign. level (corresponding to CI including 0).

- b) By qqplot: normal assumption looks ok
By $\hat{y}$ vs $\hat{e}$ plot: perhaps some problems with constant variance assumption, but given our few observations, not possible to conclude.

$$\hat{e} = (I - H)\vec{y} \quad \hat{y} = H\vec{y}$$

$$\begin{aligned} \text{cov}(\hat{e}, \hat{y}) &= (I - H) \sigma^2 I H \\ &= \sigma^2 (I - H)H = 0. \end{aligned}$$

- c) With repeated measurements, assumption of uncorrelated / independent observations (errors) likely to be violated.

$$\begin{aligned} \text{Var}(y_{ij}) &= \tau^2 + \sigma^2 \\ \text{cov}(y_{ij}, y_{ij'}) &= \tau^2 \end{aligned}$$

$$\hat{\text{corr}}(y_{ij}, y_{ij}) = \frac{\hat{\tau}^2}{\hat{\tau}^2 + \hat{\sigma}^2} = 0.72$$

0.72 is the estimated correlation between observations (turnover) from the same store