EXERCISE 1

a) · model 1 is a linear model

· model 2 is not a linear model since it is not linear w.r.t. $(\beta_1, \beta_2, \beta_3)$
It can not be transformed to get a linear model.

· model 3 is a linear model

$$\underbrace{\log Y_i}_{Y_i^*} = \beta_2 \underbrace{\frac{x_{i1}}{x_{i2}}}_{x_{i1}^*} + \beta_3 \underbrace{\frac{\log x_{i3}}{x_{i2}}}_{x_{i2}^*} + \varepsilon_i \Rightarrow Y_i^* = \beta_2 x_{i1}^* + \beta_3 x_{i2}^* + \varepsilon_i$$

linear model without intercept

· model 4 is not a linear model
the error term is not additive, not linear w.r.t. (β_1, β_2).
But, with a logarithmic transformation

$$\log Y_i = \log \beta_1 + \log x_{i2}^{\beta_2} + \varepsilon_i$$

$$\underbrace{\log Y_i}_{Y_i^*} = \underbrace{\log \beta_1}_{\beta_1^*} + \underset{\beta_2^*}{\underset{\parallel}{\beta_2}} \underbrace{\log x_{i2}}_{x_{i2}^*} + \underset{\varepsilon_i^*}{\underset{\parallel}{\varepsilon_i}} \Rightarrow Y_i^* = \beta_1^* + \beta_2^* x_{i2}^* + \varepsilon_i^* \qquad \varepsilon_i^* \sim N(0,1) \text{ ind.}$$

b) $\underline{Y}^*$ is an n-dim vector of random variables

$\underline{Y}^* = [Y_1^* \dots Y_n^*]^T = [\log Y_1 \dots \log Y_n]^T$ $\qquad \underline{Y}^* \sim N_n(X^* \underline{\beta}^*, I)$

X^* is an $(n \times 2)$ matrix of known constants

$X^* = [\underline{1}_n \ \underline{x}_2^*]$

$\underline{\beta}^*$ is a 2-dim vector of unknown parameters

$$\underline{\beta}^* = \begin{bmatrix} \beta_1^* \\ \beta_2^* \end{bmatrix}$$

$\underline{\varepsilon}^*$ is an n-dim vector of random variables $\quad \underline{\varepsilon}^* \sim N_n(0, I)$

c) $\hat{\underline{\beta}}^* = (X^{*T} X^*)^{-1} X^{*T} \underline{Y}^*$

with

$\hat{\underline{\beta}}^* \sim N_p(\underline{\beta}, (X^{*T} X^*)^{-1})$

d) $\sum_{i=1}^n e_i = 0$ true, the model includes the intercept

$\sum_{i=1}^n e_i x_{i2} = 0$ false: x_{i2} is not a column of X^*

$\sum_{i=1}^n e_i \log x_{i2} = 0$ true

$\sum_{i=1}^n e_i \log(x_{i2}^2) = \sum_{i=1}^n e_i \cdot 2 \log x_{i2} = 2 \sum_{i=1}^n e_i \log x_{i2} = 0$ true