Handle Categorical Variables

Consider a categorical predictor, `Size`, taking values from $\{S, M, L\}$, which needs to be coded as two numerical predictors.

$$\begin{pmatrix} S \\ S \\ M \\ M \\ L \\ L \end{pmatrix} \Longrightarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}_{6\times 2}$$

- 1st column: indicator for value ”M”.
- 2nd column: indicator for value ”L”.
- No need to code ”S”, which is chosen as the **reference level** and its effect is absorbed into the intercept. (You can choose any value as the reference group.)
- In general, code a categorical predictor with K values as $(K-1)$ binary vectors.