- $Y_i = B_0 + B_1x_i + e_i$
    - $B_0 + B_1x_i$ = systematic
    - e_i = random
        - Aggregation of all the nudges
        - The balance of the nudges
- e_i~ N(0, σ^2)
    - Mean = 0
    - Variance = σ^2
- Chain of reasoning (depends upon the notion that each nudge is the same size and independent of each other)
    - Residuals are aggregations of nudges
    - Each nudge is a coin flip up or down
        - Therefore, the sum of nudges is described by a binomial distribution
            - 10 flips (N)
                - X = number of heards
                - $\binom{10}{K} P^k(1-p)^{(10-k)}$
                - P(x=k) = $\binom{10}{K} P^k(1-p)^{(10-k)}$ if probability (heads) = P
                    - $\binom{10}{K}$= $\binom{n}{K}$
    - Binomial distribution is approximately normal for large N (number of nudges)
- Normal distribution is designed for minnows and not sharks
    - Not for dominant features
    - IBM stock distributions don't follow a normal distribution
        - More like the shark
    - DOW Jones
        - Averaging effect allows us to interpret why normal distributions work better for DOW jones than IBC
        - But normal is still not great for DOW jones
            - Big economic effects that affect the stock in general
- Normal linear regression model
    - Linearity: $\hat{y}_i = \beta_0 + \beta_1 x_i \ (+\beta_2 x_{i2} + \cdots)$
    - Assumptions
        - Normality
                - $Y_i = B_0 + B_1x_i + e_i$
                - e_i~ N(0, σ^2)
        - Independence: no residual provides information about any other
        - Constant variance
            - Not spread out for certain parts of the line
            - Homoscedasticity
            - Heteroscedasticity
- NLRM → explicitly formulas for standard errors and prediction intervals
    - → = fancy math

- Page 137-138 for more details