Meaning

- [[Maximum-Likelihood-Schätzer]] and assumption error is [[Normalverteilung]]
  - $$\epsilon_i|X_i \sim N(0, \sigma^2)$$
- $$Y_i|X_i \ N(\mu_i, \sigma^2), \text{ where } \mu_i = \beta_0 + \beta_1 X_i$$
- likelihood function
  - $$\begin{aligned} L(\beta_0, \beta_1, \sigma) &= \prod_{i=1}^{n} f(X_i, Y_i) = \prod_{i=1}^{n} f_X(X_i) f_{Y|X}(Y_i|X_i) \\ &= \prod_{i=1}^{n} f_X(X_i) \times \prod_{i=1}^{n} f_{Y|X}(Y_i|X_i) \end{aligned}$$
  - first term does not depend on β_0 and β_1
    - can be omitted
  - second term proportional to
    - $$L_2(\beta_0, \beta_1, \sigma) \propto \frac{1}{\sigma^n} e^{-\frac{1}{2\sigma^2} \sum_i (Y_i - \mu_i)^2}$$
  - log likelihood
    - $$\mathcal{L}(\beta_0, \beta_1, \sigma) = -n log \sigma - \frac{1}{2\sigma^2} \sum_i (Y_i - (\beta_0 + \beta_1 X_i))^2$$
    - maximizing equivalent to minimizing SSE
      - sum of squared errors [[Lineare Regression]]

Prediction

- We start with our data as before: $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$
- We estimate $\hat{r}(x) = \hat{\beta}_0 + \hat{\beta}_1 x$ from that data
- We observe a new value $X = x_*$
- Then an estimate or a prediction of Y_* is given by:

$$\hat{Y}_* = \hat{\beta}_0 + \hat{\beta}_1 x_*$$

- example

We estimate $\hat{\beta}_0 = 2.2593$ and $\hat{\beta}_1 = 1.6289$ from the 160 sold video games globally versus North America. With these estimates we predict global sales from the North America sales for 40 new games.