

# Simple Linear Regression

Suppose we have datapoints  $(x_i, y_i)$

Assumption:  $Y_i \sim G(\mu(x_i), \sigma^2)$

where  $\mu(x_i) = \alpha + \beta x_i$

How can we get estimates for  $\alpha$  and  $\beta$ ?

Since our distribution is Gaussian:

$$L(\alpha, \beta, \sigma) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2} (y_i - \alpha - \beta x_i)^2\right]$$

Getting log likelihood and solving

$$\frac{d\ell}{d\alpha} = 0 \quad \frac{d\ell}{d\beta} = 0 \quad \frac{d\ell}{d\sigma} = 0$$

we have the MLEs

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sum_{i=1}^n x_i (x_i - \bar{x})} = \frac{S_{xy}}{S_{xx}}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}x_i)^2 = \frac{1}{n} (S_{yy} - \hat{\beta}S_{xy})$$

$s_e^2$  is an unbiased estimator for  $\sigma^2$  since  $E(s_e^2) = \sigma^2$