$X_1, \ldots, X_n$ $\quad exp(X)$

$\bar{X} = \lambda$

$X_1, \ldots, X_n$ $\quad$ iid $U(a, b)$

$\frac{a+b}{2} = \bar{X} \Rightarrow a = 2\bar{X} - b$

$\frac{b-a}{12} = \underbrace{\frac{\sum x^2}{n}}_{E\{x^2\}} - \underbrace{(\bar{x})^2}_{E(x)^2}$

$\frac{b - 2\bar{x} + b}{12} = \frac{\sum x^2}{n} - \bar{x}^2$

$\frac{2b - 2\bar{x}}{12} = \frac{\sum x^2}{n} - \bar{x}^2$

$\frac{b - \bar{x}}{6} = \frac{\sum x^2}{n} - \bar{x}^2 \Rightarrow b = \bar{x} + 6\frac{\sum x^3}{n} - 6\bar{x}^2$

$a = 2\bar{x} - \bar{x} - 6\frac{\sum x^2}{n} + 6\bar{x}^2$

$a = \bar{x} - 6\frac{\sum x^2}{n} + 6\bar{x}^2$

$x_1, x_2, \ldots x_n$ $\quad$ iid $N(\mu, \sigma^2)$ $\quad \bar{X} = \mu$

$\frac{\sum x^2}{n} = \sigma^2 + \mu^2 = E(x^2)$

$\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2$

Likelyhood function

$X_1 = x_1, \ldots X_n = x_n$ from $f_\theta(x) \equiv f(x|\theta)$

$$L(\theta) = f_\theta(x_1) f_\theta(x_2) \ldots f_\theta(x_n)$$

$$= \prod_{i=1}^{n} f_\theta(x_i)$$

function of θ, denoted as Θ

Method of Maximum likelyhood

$$\hat{\theta}_{ME} = \underset{\theta \in \Theta}{\arg\max}\, L(\theta)$$

we are finding the parameter that networks the highest probability for the observed data

$\ell(\theta) = \log L(\theta)$ instead sometimes

$\frac{\partial L(\theta)}{\partial \theta}$ is harder than

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{\partial \left(\sum_{i=1}^{n} \log f_\theta(x_i) \right)}{\partial \theta}$$

ex: suppose we observed $X = 3$, from $Bin(6, p)$ $0 < p < 1$

$$\binom{6}{3} p^3 (1-p)^3 = L(p)$$

$$\ell(p) = \log \binom{6}{3} + 3\log(p) + 3\log(1-p)$$

$$\frac{\partial \ell}{\partial p} = \frac{3}{p} - \frac{3}{1-p} = 0 \Rightarrow \frac{3}{p} = \frac{3}{1-p} \Rightarrow \boxed{\hat{p} = \frac{1}{2}}$$

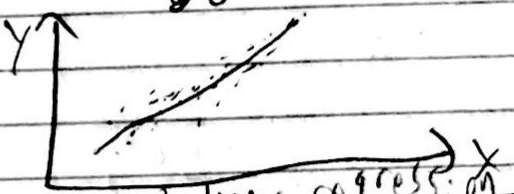
$X_1, \dots, X_n \sim N(\mu, \sigma^2)$ $\mu \in \mathbb{R}$, $\sigma^2 > 0$

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\right\}$$

$$=(2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

$$l = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

take $\frac{\partial l}{\partial \sigma^2} = 0$ and $\frac{\partial l}{\partial \mu} = 0$



simple linear regression → intercept

$\mu_{Y|X}(x) = \beta_0 + \beta_1 x$ → slope

$Var(Y|X=x) = \sigma^2 > 0$ ↳ variance

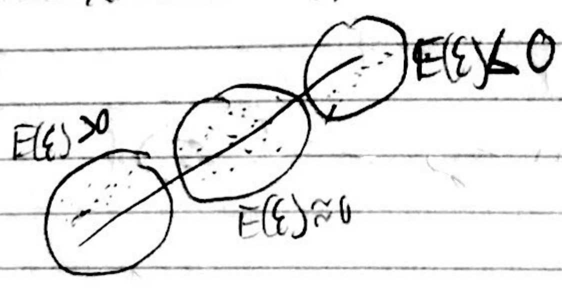
Y response or ~~dependent~~ variable

X predictor or indep. variable

$Y = \beta_0 + \beta_1 X + \varepsilon$

$E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$

ε and X are independent



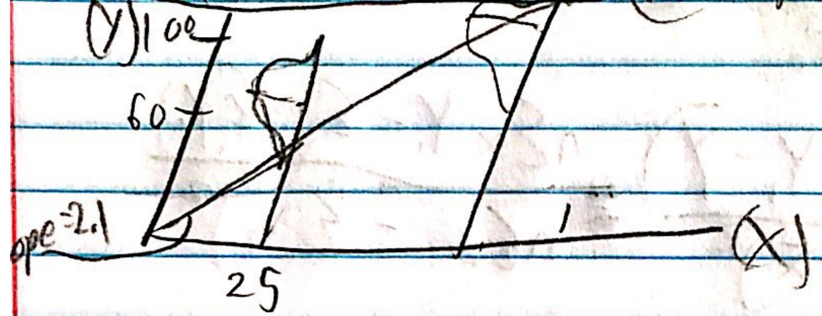
Sometimes we may even impose $\varepsilon \sim N(0, \sigma^2)$

$\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$

number of bids requested (X)

the time required to complete them (Y)

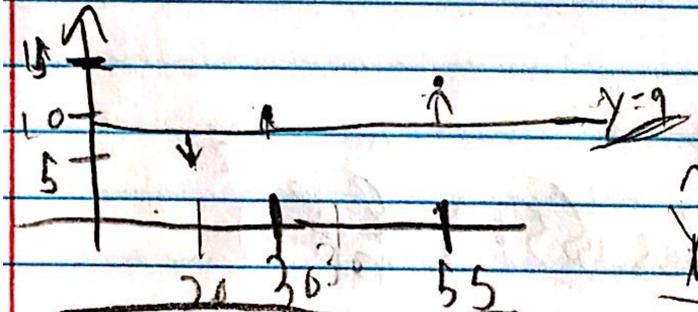
$\boxed{Y = 9.5 + 2.1X + \varepsilon_i}$ (in days)



$\mu = 62$

ex.

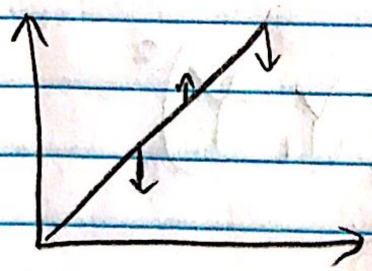
age X_i	20	55	30
# of attempts (Y_i) to complete task	15	12	10



$\hat{Y} = 9 + 0X + \varepsilon$

$\boxed{Q = 16 + 1 + 9 = 26}$

$\hat{Y} = 2.81 + .177X$



$Q = 5.7$

Method of least squares

Method of least squares

$$Q(\beta_0, \beta_1) = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2$$

$$(\hat{\beta}_0, \hat{\beta}_1) = \arg\min Q(\beta_0, \beta_1)$$

LSE (least square estimate)

$$\frac{\partial Q}{\partial \beta_1} = 0 \qquad \frac{\partial Q}{\partial \beta_0} = 0$$

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$E(Y) = E(\beta_0 + \beta_1 X_i + \varepsilon)$$

$$E(Y) = \beta_0 + \beta_1 E(X) + 0$$

Fitted Value

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

residuals $e_i = Y_i - \hat{Y}_i$

Error Sum of Squares (SSE) $\sum_{i=1}^{n} (e_i)^2$

Regression Sum of Squares

(SSR)

$$\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$$

Total Sum of Squares (SSTO) $\sum_{i=1}^{n} (Y_i - \bar{Y})^2$

Analysis of Variance Decomposition

SSTO = SSR + SSE

Point estimators of σ^2

$$\hat{\sigma}^2 = \frac{SSE}{n-2}$$

$$= \frac{1}{n-2} \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$$