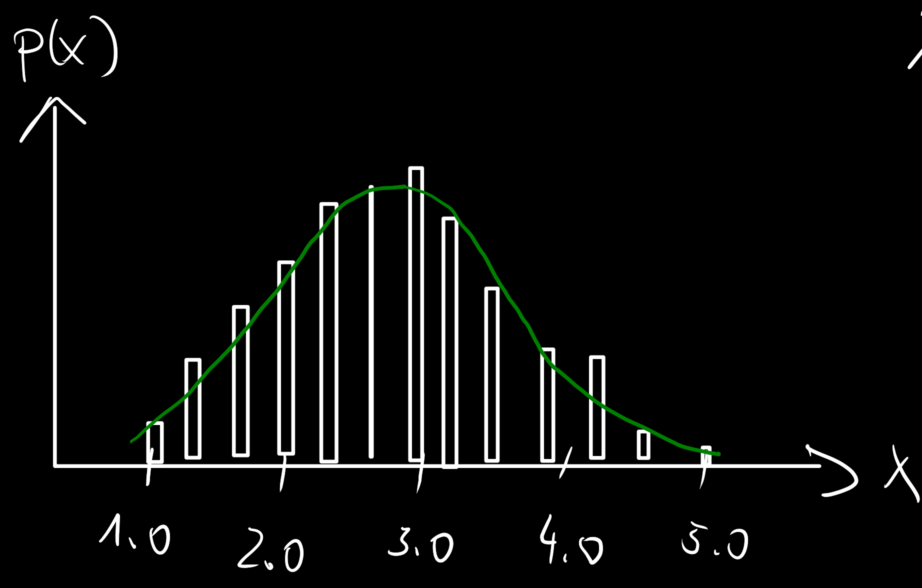
MLE for the Gaussian/Normal Distribution



$X \sim \mathcal{N}(\mu,\sigma)$

μ & σ unknown?

but we have data

$D = \{1.7, 2.5, 2.3, 2.8, 3.3, 1.4, \dots\}$

$$X \sim \mathcal{N}(\mu,\sigma) = p(X;\underline{\theta})$$

$\underline{\theta}$ ... parameter vector

$\underline{\theta} = [\mu,\sigma]^T$

Likelihood

$$\mathcal{L}(D;\underline{\theta}) \overset{i.i.d.}{=} \prod_{i=0}^{N-1} p(X=x^{[i]};\underline{\theta})$$

$$= \prod_{i=0}^{N-1} \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2\sigma^2}\cdot\left(x^{[i]}-\mu\right)^2\right)$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2\right)$$

Log-Likelihood

$$\ell(D;\underline{\theta}) = \log\left(\mathcal{L}(D;\underline{\theta})\right)$$

$$= N\cdot(-1)\cdot\log\left(\sigma\cdot\sqrt{2\pi}\right) - \frac{1}{2\sigma^2}\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2$$

$$\overset{+}{=} -N\cdot\log(\sigma) - \cancel{N\cdot\log\left(\sqrt{2\pi}\right)} - \frac{1}{2\sigma^2}\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2$$

(1) μ

$$\mu^* = \underset{\mu\in\mathbb{R}}{\arg\max}\left(\ell(D;\underline{\theta})\right)$$

(Take derivative and set to 0)

$$\frac{\partial \ell}{\partial \mu} = \cancel{-}\frac{1}{\cancel{2\sigma^2}}\cdot\sum_{i=0}^{N-1}\cancel{2}\cdot\left(x^{[i]}-\mu\right)\cdot\cancel{(-1)} \overset{!}{=} 0$$

$$\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right) = 0$$

$$\sum_{i=0}^{N-1} x^{[i]} - \underbrace{\sum_{i=0}^{N-1}\mu}_{N\cdot\mu} = 0$$

$$\sum_{i=0}^{N-1} x^{[i]} - N\cdot\mu = 0$$

$$\Leftrightarrow \quad \mu = \frac{1}{N}\cdot\sum_{i=0}^{N-1} x^{[i]}$$

(alias the mean)

(2) σ

$$\frac{\partial \ell}{\partial \sigma} = -\frac{N}{\sigma} \overset{+}{\cancel{-}} (\cancel{-2})\cdot\frac{1}{\cancel{2}\sigma^3}\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2 \overset{!}{=} 0 \quad \Big|\cdot\sigma^3$$

$$-N\cdot\sigma^2 + \sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2 \overset{!}{=} 0$$

$$\sigma^2 = \frac{1}{N}\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2$$

$$\sigma^* = \pm\sqrt{\frac{1}{N}\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2}$$

Commonly: $\sigma = +\sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}\left(x^{[i]}-\mu\right)^2}$