$$\frac{\partial}{\partial \theta_2} \ln L(x_1, \cdots, x_n | \theta_1, \theta_2) = \sum_{i=1}^{n} \left(-\frac{\cancel{2\pi}}{\cancel{2} \cdot \cancel{2\pi} \theta_2} + \frac{(x_i - \theta_1)^2}{\cancel{2} \theta_2^2} \right) = 0$$

$$\sum_{i=1}^{n} \left(-\hat{\theta}_2 + (x_i - \hat{\theta}_1)^2 \right) = 0$$

$$-n\hat{\theta}_2 + \sum_{i=1}^{n} (x_i - \hat{\theta}_1)^2 = 0$$

$$\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\theta}_1)^2$$

Is it a maximum? Take $\frac{\partial^2}{\partial \theta_2^2} \ln L$: it isn't negative everywhere, but it is at $\theta_2 = \hat{\theta}_2$, so $\hat{\theta}_2$ is a local maximum.

Recall $\mathrm{Var}(Y) = E[(Y - \mu_Y)^2]$, where $\mu_Y = E[Y]$

$\hat{\theta}_2$ is the sample variance for samples $x_1, x_2, \cdots, x_n$.
$\hat{\theta}_2$ is MLE of the population variance σ^2.

Bias

Defn: An estimator $\hat{\theta}$ of θ is unbiased iff $E[\hat{\theta}] = \theta$.

We are thinking now of the samples $x_1, x_2, \cdots, x_n$ as random variables rather than numbers, and $\hat{\theta}$ as a function of $x_1, x_2, \cdots, x_n$ is also a r.v. with an expectation. If $\hat{\theta}$ is unbiased, that's a desirable property for an estimator of θ.

Ex: Consider $\hat{\theta}_1$, the MLE of θ_1 in the normal distribution application:

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$E[\hat{\theta}_1] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i]$$
$$= \frac{1}{n}\sum_{i=1}^{n}\theta_1 = \frac{1}{n}\cdot n\theta_1 = \theta_1,$$

so $\hat{\theta}_1$ is unbiased.

What about $\hat{\theta}_2$?

$\hat{\theta}_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\theta}_1)^2$ is biased but MLE.

$\hat{\theta}_2' = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\theta}_1)^2$ is unbiased but has slightly lower likelihood.

These differ by a factor of $\frac{n-1}{n}$, so if n is large, they are almost the same.

<u>Confidence Interval</u>

<u>Problem</u>: MLE $\hat{\theta}$ of θ is wrong with probability 1.

$P(\hat{\theta} = \theta) = 0.$

Can we find some Δ such that

$$P(\theta \in [\hat{\theta} - \Delta, \hat{\theta} + \Delta]) = 95\%, \text{ say?}$$

$[\hat{\theta} - \Delta, \hat{\theta} + \Delta]$ is called the <u>95% confidence interval</u>, if so.

<u>Ex</u>: MLE $\hat{\theta}_1$ of normal mean μ.

Assume $x_i \sim N(\mu, \sigma^2)$ are i.i.d.

$\hat{\theta}_1$ is a r.v. with mean and variance

$$Var(\hat{\theta}_1) = Var\left(\frac{1}{n}\sum_{i=1}^{n}(x_i)\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(x_i)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$