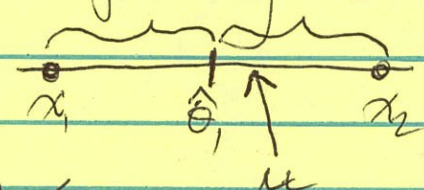
What about $\hat{\theta}_2$?

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

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

These estimators differ by a factor $\frac{n}{n-1}$, so for large n they are nearly the same.

Intuition for why $\hat{\theta}_2$ is not unbiased: $n=2$



$\hat{\theta}_2 = \frac{1}{2}\left((x_1 - \hat{\theta}_1)^2 + (x_2 - \hat{\theta}_1)^2\right)$

$\leq \frac{1}{2}\left((x_1 - \mu)^2 + (x_2 - \mu)^2\right) \qquad \sigma_X^2 = E[(X-\mu)^2]$

So $\hat{\theta}_2$ consistently underestimates σ^2. *

In fact, $E(\hat{\theta}_2) = \frac{n-1}{n}\sigma^2$

Confidence intervals

What is $P(\hat{\theta} = \theta)$? 0, assuming $\hat{\theta}$ is a continuous random variable.

Could we find a small Δ such that θ is in $[\hat{\theta} - \Delta, \hat{\theta} + \Delta]$ with probability 95%, say.

This is called the 95% confidence interval.

---

* More precisely, $E[\hat{\theta}_2] \leq E\left[\frac{1}{2}\left((x_1-\mu)^2 + (x_2-\mu)^2\right)\right]$

$= \frac{1}{2}\left(E[(x_1-\mu)^2] + E[(x_2-\mu)^2]\right)$

$= \frac{1}{2}\left(Var(x_1) + Var(x_2)\right) = \frac{1}{2}\cdot 2\sigma^2 = \sigma^2$

Ex: Consider MLE of the Normal mean μ:

$$\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i, \text{ where } x_1, \dots, x_n \text{ are i.i.d. from } N(\mu, \sigma^2).$$

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

$$E[\hat{\theta}_1] = \mu$$

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

By CLT,

$$\hat{\theta}_1 \to N\left(\mu, \frac{1}{n}\sigma^2\right)$$

$$\frac{\hat{\theta}_1 - \mu}{\sigma/\sqrt{n}} \to N(0,1)$$

$$P\left(-z < \frac{\hat{\theta}_1 - \mu}{\sigma/\sqrt{n}} < +z\right) \approx \Phi(z) - \Phi(-z) = 2\Phi(z) - 1$$

$$P\left(-z < \frac{\mu - \hat{\theta}_1}{\sigma/\sqrt{n}} < +z\right) \approx 2\Phi(z) - 1$$

$$P(\hat{\theta}_1 - z\sigma/\sqrt{n} < \mu < \hat{\theta}_1 + z\sigma/\sqrt{n}) \approx 2\Phi(z) - 1 = 0.95$$

(95% conf. int.)

$$\Phi(z) = 0.975$$

$z \approx 1.96$, from Φ table.

μ is within $1.96\sigma/\sqrt{n}$ of $\hat{\theta}_1$ with probability 95%

The 95% confidence interval for μ is

$$[\hat{\theta}_1 - 1.96\sigma/\sqrt{n}, \hat{\theta}_1 + 1.96\sigma/\sqrt{n}].$$