$x_i \sim N(\theta_1, \theta_2)$, both $\theta_1 = \mu$ and $\theta_2 = \sigma^2$ are unknown.

Find $\hat{\theta}_1, \hat{\theta}_2$ that maximize $L(x_1, x_2, \cdots, x_n \mid \theta_1, \theta_2)$.

$$L(x_1, x_2, \cdots, x_n \mid \theta_1, \theta_2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} e^{-(x_i-\theta_1)^2/(2\theta_2)}$$

$$\ln L(x_1, x_2, \cdots x_n \mid \theta_1, \theta_2) = \sum_{i=1}^{n} \left(-\frac{1}{2}\ln(2\pi\theta_2) - \frac{(x_i-\theta_1)^2}{2\theta_2} \right)$$

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

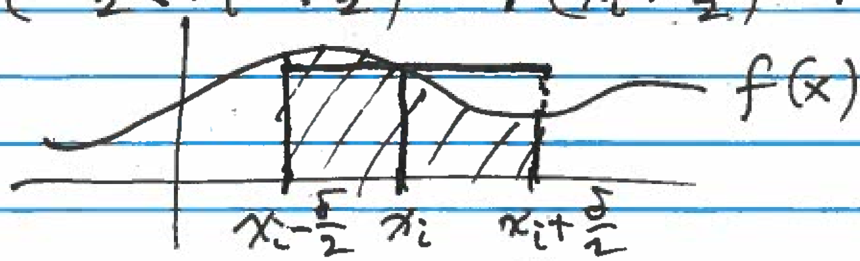
$$\left(\sum_{i=1}^{n} x_i\right) - n\hat{\theta}_1 = 0$$

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

Sample mean is MLE of population mean $\mu = \theta_1$.

Why is L defined as the product of densities in the case of a continuous distribution?

$$P\left(-\frac{\delta}{2} < x_i < \frac{\delta}{2}\right) = F\left(x_i + \frac{\delta}{2}\right) - F\left(x_i - \frac{\delta}{2}\right) \approx \delta f(x_i)$$



$$L(x_1, \cdots x_n \mid \theta_1, \theta_2) \approx \prod_{i=1}^{n} \delta f(x_i)$$

when we take $\ln$ and $\frac{\partial}{\partial\theta_1}$, δ term drops out.

Is $\hat{\theta}_1$ a max?

$$\frac{\partial^2}{\partial \theta_1^2} \ln L(x_1,\dots,x_n|\theta_1,\theta_2) = \sum_{i=1}^{n} -\frac{1}{\theta_2} < 0$$

so $\ln L$ is concave downward everywhere.

Now $\hat{\theta}_2$:

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

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

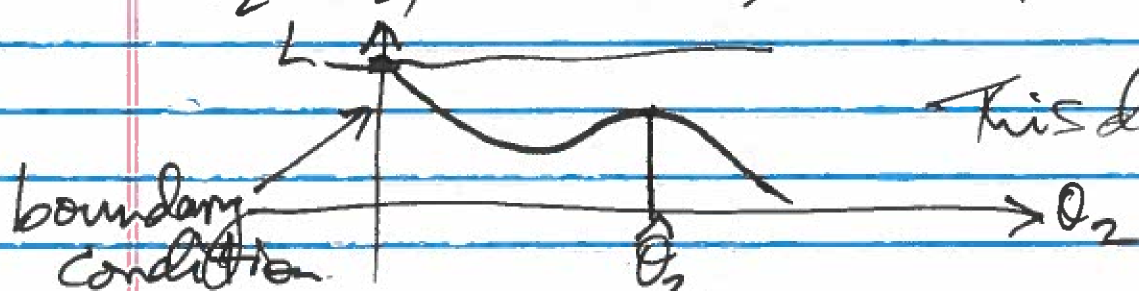
$$n\hat{\theta}_2 = \sum_{i=1}^{n}(x_i-\hat{\theta}_1)^2$$

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

This is (one version of) the sample variance, the MLE of the population variance $\hat{\sigma}^2 = \hat{\theta}_2$.

$$Var(Y) = E[(Y-\mu)^2] \text{ where } \mu = E[Y]$$

Is $\hat{\theta}_2$ a max? If you take $\frac{\partial^2}{\partial \theta_2^2}$, it is not negative everywhere, but it is negative at $\theta_2 = \hat{\theta}_2$, so $\hat{\theta}_2$ is a local maximum.



This doesn't happen.

Bias.

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

This is a desirable property of an estimator.