MLE

Thursday, September 19, 2024 11:01 AM

In statistics, one of the main goals is to fit probability distributions to Data.
Today, we will be using <u>Maximum Likelihood</u> estimation.
We use this when we have an idea of how the data is distributed.

<u>Basic Ex</u> $P(\text{Rolling two 6s} \mid \text{6-sided die}) = \frac{1}{36}$
$P(\quad " \quad \mid \text{20-sided die}) = \frac{1}{400}$

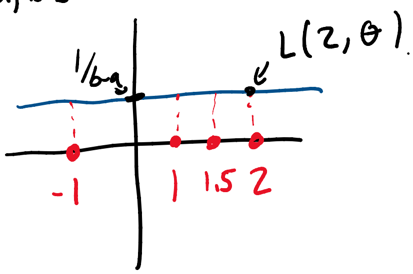
parameters of distribution.

<u>Def</u> Let $X_1 \dots X_n$ be I.I.D. with PDF $f(x, \theta)$. The <u>likelihood</u> function of θ.
<u>function</u> is given by $L_n(\theta) = \prod_{i=1}^{n} f(X_i ; \theta)$.
The <u>log-likelihood function</u> is given by $\ell_n(\theta) = \log(L_n(\theta)) = \sum_{i=1}^{n} \log(f(X_i, \theta))$

<u>Goal</u>: Maximize $L_n(\theta)$. (equivalently $\ell_n(\theta)$).

<u>Recall</u>: $U[a,b] = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{else} \end{cases}$

<u>Ex</u> We have a uniform RNG which returns a random number in $[a,b]$. Given the numbers $-1, 1, 1.5, 2$ what is MLE for a and b.



$$L_n(\theta) = \begin{cases} 0 & \text{if } a > -1 \text{ or } b < 2. \\ 4 \cdot \frac{1}{b-a} & \text{else.} \end{cases}$$

So MLE is $\theta = [-1, 2]$ w/ a likelihood of $4 \cdot \frac{1}{3}$ (note that likelihood is not a probability.

<u>e.g.</u> We Discover a new species of orange. We know fruit weights tend to be normally distributed. But which normal distribution fits?
A normal distribution has two parameters μ and σ. We set

A normal distribution has two parameters μ and σ. We set $\theta = (\mu, \sigma)$.

Fact: MLE for $\theta = (\mu, \theta)$ is given by $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2.$$

Let's sketch out the proof for this!

Recall: • Likelihood function is $L_n(\theta) = \prod_{i=1}^{n} f(X_i ; \theta)$

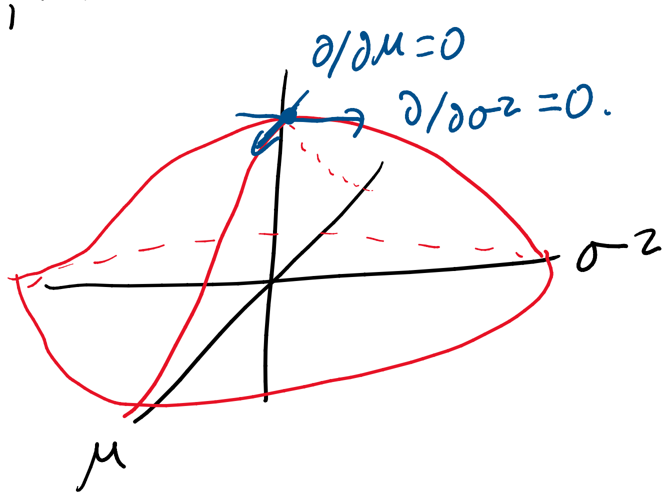
• $N(\mu, \sigma^2)$ has PDF $\frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

Lemma Let $\theta = (\mu, \sigma^2)$ then $L_n(\theta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2}$.

Lemma The log-liklihood is given by:

$$\ell_n(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2.$$

The shape of $\ln(\theta)$ looks like:



• **Mean**: $\frac{\partial}{\partial \mu}\ell_n(\theta) =$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)$$

If $\frac{\partial}{\partial \mu} = 0$ then

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu) = 0 \Rightarrow \sum_{i=1}^{n} X_i - n\mu = 0 \Rightarrow \mu = \frac{\sum_{i=1}^{n} X_i}{n}.$$

$$\frac{\partial}{\partial \sigma^2}\ell_n(\theta) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum (X_i - \mu)^2 \quad \text{Factor } \frac{1}{2\sigma^2}.$$

- <u>variance</u>: $\frac{\partial}{\partial \sigma^2} \ln(\theta) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n} (X_i - \mu)^2$ $\rightsquigarrow$ factor $\overline{2\sigma^2}$.

$$\frac{1}{2\sigma^2} \left(-n + \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 \right).$$

want $\frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 = n. \Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{n}$