Point Estimation 2

Day 8 (1/30/20)

4.1.1 Point Estimators Con't.

Note: The MLE works for more than one parameter: $\theta = [\theta_1, \theta_2]$

Example 4.1.3 On page 229, we are given $n-24$ sulfur dioxide measurements from a Bavarian forest. Assuming the population may be modeled by $N(\mu, \sigma^2)$, find the MLE of (μ, σ^2).

Solution: Let $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$. Let $\theta = [\theta_1, \theta_2] = \left[\mu, \sigma^2\right]$. Then

$$f(x_i; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$L(\theta) = \prod f(x_i; \theta) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2}$$

$$l(\theta) = \ln L(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum(x_i-\mu)^2$$

Taking partial derivatives with respect to μ and σ^2, respectively

1. $\frac{\partial}{\partial\mu} l(\theta) = -\frac{1}{2\sigma^2}\sum 2(x_i-\mu)(-1)$
2. $\frac{\partial}{\partial\sigma^2} l(\theta) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(x_i-\mu)^2$

Setting both equations to 0 and solving for $\hat{\mu}$ and $\hat{\sigma}^2$,

$$\sum(x_i - \hat{\mu}) = 0 \quad \Rightarrow \quad \sum x_i = n\hat{\mu} \quad \Rightarrow \quad \hat{\mu} = \overline{x}$$

$$\frac{n}{\hat{\sigma}^2} = \frac{1}{(\hat{\sigma}^2)^2}\sum(x_i-\overline{x})^2 \quad \Rightarrow \quad \hat{\sigma}^2 = \frac{\sum(x_i-\overline{x})^2}{n}$$

so the MLE is

$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{bmatrix} = \begin{bmatrix} \overline{x} \\ \frac{\sum(x_i-\overline{x})^2}{n} \end{bmatrix} = \begin{bmatrix} 53.91667 \\ 97.25139 \end{bmatrix}$$

```
> # CALCULATIONS IN R
> sulfurdioxide<-c(33.4,38.6,41.7,43.9,44.4,45.3,46.1,47.6,50.0,52.4,52.7,
  53.9,54.3,55.1,56.4,56.5,60.7,61.8,62.2,63.4,65.5,66.6,70.0,71.5)
> mean(sulfurdioxide)
[1] 53.91667
> (1/24)*sum((sulfurdioxide-53.91667)^2)
[1] 97.25139
```

Comments:

- MLE of σ^2 is not unbiased.

- If we let $\theta_2 = \sigma$ instead of $\theta_2 = \sigma^2$, we can show that ...

$$\text{MLE of } \hat{\sigma} = \sqrt{\frac{\sum(x_i - \overline{x})^2}{n}}$$

so that the MLE of $\sigma^2 = [\text{MLE of } \sigma]^2$.

This is a general property of the MLE, i.e.

$$\text{MLE of } g(\theta) = g(\text{MLE of } \theta)$$

This is useful when we want to estimate functions of simple parameters.

Example 1 Exponential(θ): Given the data from Example 4.1.1 (operating hours before failure of air-conditioning units), estimate the

1. probability that operating hours exceeds 300

   Solution: $P(X > 300) = \int_{300}^{\infty} \frac{1}{\theta} e^{-x/\theta}\, dx = e^{-300/\theta}$. Since $\hat{\theta}_{MLE} = \overline{x} = 163.5385$, the MLE of $e^{-300/\theta}$ is $e^{-300/163.5385} = 0.1597$.

2. 90th percentile of operating hours, i.e. find q such that $F(q) = .90$.

   Solution: $.90 = \int_0^q \frac{1}{\theta} e^{-x/\theta}\, dx = 1 - e^{-q/\theta}$ implies that $q = -\theta \ln(1 - .90)$. The MLE is $-163.5385 \ln(1 - .90) = 376.56$ hours.

Example 2 Normal(μ, σ^2): Given the sulfur dioxide data from Example 4.1.3, estimate the

1. percentage of times that sulfur dioxide concentration is less than 50.0

   Solution: $F(50.0) = \Phi\left(\frac{50-\mu}{\sqrt{\sigma^2}}\right)$. The MLE is $\Phi\left(\frac{50-\hat{\mu}}{\sqrt{\widehat{\sigma^2}}}\right) = \Phi\left(\frac{50-53.91667}{\sqrt{97.25139}}\right) = 0.3456$

2. 90th percentile of sulfur dioxide concentrations

   Solution: $q_{.90} = F^{-1}(.90) = \mu + 1.2815\sqrt{\sigma^2}$, so the MLE is $53.916 + 1.28\sqrt{97.25} = 66.55$.

5 Estimating a density (nonparametric estimates vs MLE)

Case 1: Discrete $p(x)$

Example 4.1.6 Poisson(λ): For $j = 1, 2, 3, 4, 5, 6$

$$\hat{p}(j) = \text{proportion of sample equal to } j = \frac{\#\{x_i = j\}}{n} = \frac{\sum \mathbf{I}(x_i = j)}{n}$$

where $\mathbf{I}(E) = \begin{cases} 1, & \text{if } E \text{ occurs} \\ 0, & \text{otherwise} \end{cases}$

For example, $p(3) = 5/30 = .167$. Note that the MLE is $e^{-\hat{\lambda}}\hat{\lambda}^3/3!$ where $\hat{\lambda} = \overline{x}$.

Case 2: Continuous $f(x)$.

$$\begin{aligned} P[x-h<X<x+h] &= \int_{x-h}^{x+h} f(t)dt \\ &= 2hf(\epsilon) \ \text{ for some } \epsilon \text{ in the interval } [x-h, x+h] \\ &\doteq 2hf(x) \end{aligned}$$

Therefore

$$\begin{aligned} \hat{f}(x) &= \frac{\hat{P}(x-h<X<x+h)}{2h} = \frac{\#\{x-h<x_i<x+h\}}{n2h} \\ &= \frac{1}{n2h}\sum_{i=1}^{n} \mathbf{I}(x-h \le x_i \le x+h) \end{aligned}$$

Comments:

1. This is a kernel density estimator (KDE) using a rectangular kernel
2. May be generalized to nonrectangular 'smoother' kernels
3. $2h$ is called the *bandwidth*. There is a lot of research on optimal choice of bandwidth.

Example 4.1.7

```
> density(sulfurdioxide)

Call:
 density.default(x = sulfurdioxide)

Data: sulfurdioxide (24 obs.); Bandwidth 'bw' = 4.802

       x               y            
 Min.   :19.00   Min.   :3.977e-05  
 1st Qu.:35.72   1st Qu.:2.325e-03  
 Median :52.45   Median :1.325e-02  
 Mean   :52.45   Mean   :1.493e-02  
 3rd Qu.:69.18   3rd Qu.:2.742e-02  
 Max.   :85.90   Max.   :3.331e-02  

> a<-density(sulfurdioxide)
> head(cbind(a$x,a$y))
         [,1]         [,2]
[1,] 18.99507 3.976838e-05
[2,] 19.12601 4.316185e-05
[3,] 19.25695 4.684999e-05
[4,] 19.38789 5.071897e-05
[5,] 19.51883 5.497508e-05
```

```
[6,] 19.64977 5.951386e-05
> tail(cbind(a$x,a$y))
          [,1]         [,2]
[507,] 85.25023 8.337117e-05
[508,] 85.38117 7.681850e-05
[509,] 85.51211 7.069109e-05
[510,] 85.64305 6.514269e-05
[511,] 85.77399 5.986374e-05
[512,] 85.90493 5.502299e-05

> hist(sulfurdioxide,xlim=c(20,80))
> plot(density(sulfurdioxide),xlim=c(20,80))
```