M4S3: Estimating Parameters

Method 1: Empirical Distribution Estimators

Notation: \( \hat{X} \) denotes the empirical distribution for the given sample of data.

We approximate the parameter \( \Theta \) by its analog \( \hat{\Theta} \) in the empirical distribution.

Example 1: \( X \sim \text{Exp}(\Theta) \)

\( X: 15, 20, 15, 30 \)

Q: Approximate \( \Pr(X > 25) \) = \( \pi \)

A: \[
\begin{array}{c|c|c}
X & Pr & \Pr(X > 25) = \pi \\
15 & .5 & \pi \\
20 & .25 & \pi \\
30 & .25 & \pi \\
\end{array}
\]

By \( \hat{X} \) is based on 4 observations.

Example 2: (Moment Matching)

\( X \sim \text{Exp}(\Theta) \)

\( X: 15, 20, 15, 30 \)

Q: Approximate \( \Theta = E[X] \)

\[ \Theta = E[X] = E[\hat{X}] = 20 = \hat{\Theta}_4 \]
Method 2: MLE (see last class notes)

Example 1: \( N \sim \text{Bin}(m=10, \theta) \) \( 0 \leq \theta \leq 1 \)

\( N: 4, 5, 7 \)

Q: MLE of \( \theta \)?

A: \( L(\theta) = \Pr(N_1=4 \text{ and } N_2=2 \text{ and } N_3=7) \)

\[ = \Pr(N=4) \cdot \Pr(N=2) \cdot \Pr(N=7) \]

Likelihood function:
\[ L(\theta) \propto \theta^4 \cdot (1-\theta)^9 \]
\[ = \widetilde{L}(\theta) \] (ignore constant)

\( \ell(\theta) = \ln(L(\theta)) = \ln(\text{likelihood function}) \)

\[ \ell(\theta) = \ln(\widetilde{L}(\theta)) = 13 \ln(\theta) + 17 \ln(1-\theta) \]

Note: The maximum value of \( L(\theta), \widetilde{L}(\theta), \ell(\theta), \widetilde{\ell}(\theta) \)
are all located at the same \( \theta \)-value.

Focus on \( \widetilde{\ell}(\theta) = 13 \ln(\theta) + 17 \ln(1-\theta) \) (maximize)

\[ \widetilde{\ell}'(\theta) = \frac{13}{\theta} - \frac{17}{1-\theta} = 0 \Rightarrow \frac{13}{\theta} = 17 \]

\[ \Rightarrow \theta = \frac{13}{17} = \hat{\theta} \quad \text{(MLE)} \]

Example 2: \( X \sim \text{Exp}(\theta) \)

\( X: 15, 20, 15, 30 \)

Q: MLE of \( \theta \)?

Since \( X \) is continuous, use pdfs

A: \( L(\theta) = \frac{f(15) \cdot f(20) \cdot f(15) \cdot f(30)}{\Pr(X=15) \cdot \Pr(X=20) \cdot \Pr(X=15) \cdot \Pr(X=30)} \)

\[ = \frac{1}{\theta^4} \cdot e^{-\frac{15}{\theta}} = \theta^{-4} \cdot e^{-\frac{15}{\theta}} \]
\[ l(\theta) = -\frac{y}{\theta} + \frac{80}{\theta^2} \]

\[ l'(\theta) = -\frac{y}{\theta} + \frac{80}{\theta^2} = \frac{-y + 80}{\theta^2} = 0 \]

\[ \Rightarrow \theta = 20 = \hat{\theta} \text{ (MLE)} \]

**Example 3:** \( X \sim \text{Exp}(\theta) \)

\[ X: 15, 20, 15, 30, 40^+, 40^+ \]

**Q:** What is the MLE of \( \theta \)?

**A:**

\[ E[L(\theta)] = \left[ \frac{f(15)}{Pr(X=15)} \cdot \frac{f(20)}{Pr(X=20)} \cdot \frac{f(30)}{Pr(X=30)} \cdot \left[ Pr(X > 40) \right]^2 \right] \]

\[ = \frac{1}{15} \cdot e^{\frac{160}{15}} \cdot \left( e^{\frac{-480}{15}} \right)^2 = \theta^4 \cdot e^{\frac{-160}{15}} \]

\[ l(\theta) = -\frac{y}{\theta} + \frac{160}{\theta^2} \]

\[ l'(\theta) = -\frac{y}{\theta} + \frac{160}{\theta^2} = \frac{-y + 160}{\theta^2} = 0 \]

\[ \Rightarrow \theta = \frac{160}{y} = 40 = \hat{\theta} \]

**Remarks:**

1) From Example 1, \( N \sim \text{Bin}(n=10, \theta) \)

\[ N = 4, 2, 7 \]

\[ E[N] = 10 \theta \]

\[ \frac{E[N]}{n} = \frac{13}{3} \]

Setting \( E[N] = \bar{N} \), gives

\[ 10\theta = \frac{13}{3} \Rightarrow \theta = \frac{13}{30} = \hat{\theta} \]

2) From Example 2, \( X \sim \text{Exp}(\theta) \)

\[ X: 15, 20, 15, 30 \]

\[ E[X] = \theta \]

Setting \( E[X] = \bar{x} \), gives

\[ \frac{E[X]}{n} = \frac{20}{5} = \frac{13}{3} \Rightarrow \theta = 20 = \hat{\theta} \]
3) From Example 3, \( X \sim \text{Exp}(\theta) \) \( \hat{\theta} = 40 \) (MLE)

\[
\begin{align*}
X: & 15, 20, 15, 30, 40^+, 40^+ \\
E[X] &= \theta \\
\bar{X} &= 7 \text{ (how to do with the 2 40\(^+\) observations?)}
\end{align*}
\]
(Come back to this shortly)

\( \hat{\theta} \) Fact: The following are cases for which MLE makes \( E[X] = \text{sample mean} \):

1) Bin\((m, p)\) with \( p \) unknown
2) NB\((r, p)\) with \( p \) unknown
3) P\((\lambda)\)
4) 0-modified (0-truncated) versions of \#1, \#2, \#3
5) Exp\((\theta)\)

\[ \text{Note: Given } X: x_1, x_2, \ldots, x_n, u_1^+, u_2^+, \ldots, u_k^+ \]
\[ \text{then } \hat{\theta} = \frac{x_1 + x_2 + u_1 + u_2 + \cdots + u_k}{n} \]

6) \( \Gamma(\alpha, \theta) \) with \( \theta \) unknown

7) \( N(\mu, \sigma^2) \) with \( \mu \) unknown

(a) If both \( \mu \) and \( \sigma^2 \) are unknown, then \( \hat{\mu} = \bar{X} \) and \( \hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \bar{X})^2 \)

(b) If \( \mu \) is known and \( \sigma^2 \) is unknown, then
\[
\hat{\sigma}^2 = \frac{1}{n-1} \sum (x_i - \mu)^2
\]

Since \( \bar{S}_n = \frac{1}{n} \sum (x_i - \bar{X})^2 \)
Properties of MLE (estimators)

1) Asymptotically unbiased

2) MLE of \( f(\theta) = f(\text{MLE of } \theta) \)

   Example: \( X \sim \text{Exp}(\theta) \)
   \[ X: 15, 20, 15, 20 \]
   \[ \Theta: \text{MLE of } \Pr(X > 25) \]
   \[ A: \text{MLE of } \Theta = \hat{\Theta} = 20 \ (\approx \bar{X}) \]
   \[ \therefore \ Pr(X > 25) \approx e^{-\frac{25}{20}} e^{-25 \cdot 20} = 0.865... \]

3) If \( X \) \& \( X' \) are in a one-to-one relationship, then
   the MLE of parameters are the same whether you use
   observations of one or corresponding observations of
   the other.

   Example: \( X \sim \text{LN}(\mu, \sigma^2) \) \iff \( X = e^{X'}, X' \sim N(\mu, \sigma^2) \)

   Given \( X: x_1, x_2, \ldots, x_n \)

   then \( X': \ln(x_1), \ln(x_2), \ldots, \ln(x_n) \)