Midterm Review Tutorial

MIPS 21.8

21.8 ⊞ (Exercises 19.7 and 20.7 continued) A certain type of plant can be divided into four types: starchy-green, starchy-white, sugary-green, and sugary-white. The following table lists the counts of the various types among 3839 leaves.

Type	Count
Starchy-green	1997
Sugary-white	32
Starchy-white	906
Sugary-green	904

Setting

$$X = \begin{cases} 1 & \text{if the observed leave is of type starchy-green} \\ 2 & \text{if the observed leave is of type sugary-white} \\ 3 & \text{if the observed leave is of type starchy-white} \\ 4 & \text{if the observed leave is of type sugary-green,} \end{cases}$$

the probability mass function p of X is given by

a	1	2	3	4
$p(a)$	$\frac{1}{4}(2+\theta)$	$\frac{1}{4}\theta$	$\frac{1}{4}(1-\theta)$	$\frac{1}{4}(1-\theta)$

and $p(a) = 0$ for all other a. Here $0 < \theta < 1$ is an unknown parameter, which was estimated in Exercise 19.7. We want to find a maximum likelihood estimate of θ.

a. Use the data to find the likelihood $L(\theta)$ and the loglikelihood $\ell(\theta)$.

b. What is the maximum likelihood estimate of θ using the data from the preceding table?

c. Suppose that we have the counts of n different leaves: n_1 of type starchy-green, n_2 of type sugary-white, n_3 of type starchy-white, and n_4 of type sugary-green (so $n = n_1 + n_2 + n_3 + n_4$). Determine the general formula for the maximum likelihood estimate of θ.

MIPS 21.8 a

Use the data to find the likelihood $L(\theta)$ and the loglikelihood $\ell(\theta)$.

MIPS 21.8 a

Use the data to find the likelihood $L(\theta)$ and the loglikelihood $\ell(\theta)$.

$$L(\theta) = \prod_{i=1}^{3839} p(x_i)$$

MIPS 21.8 a

Use the data to find the likelihood $L(\theta)$ and the loglikelihood $\ell(\theta)$.

$$L(\theta) = \prod_{i=1}^{3839} p(x_i)$$

$$L(\theta) = \left(\frac{2+\theta}{4}\right)^{1997} \cdot \left(\frac{\theta}{4}\right)^{32} \cdot \left(\frac{1-\theta}{4}\right)^{906} \cdot \left(\frac{1-\theta}{4}\right)^{904}$$

MIPS 21.8 a

Use the data to find the likelihood $L(\theta)$ and the loglikelihood $\ell(\theta)$.

$$L(\theta) = \prod_{i=1}^{3839} p(x_i)$$

$$L(\theta) = \left(\frac{2+\theta}{4}\right)^{1997} \cdot \left(\frac{\theta}{4}\right)^{32} \cdot \left(\frac{1-\theta}{4}\right)^{906} \cdot \left(\frac{1-\theta}{4}\right)^{904}$$

$$L(\theta) = \frac{1}{4^{3839}} \cdot (2+\theta)^{1997} \cdot \theta^{32} \cdot (1-\theta)^{1810}$$

MIPS 21.8 a

Use the data to find the likelihood $L(\theta)$ and the loglikelihood $\ell(\theta)$.

$$L(\theta) = \prod_{i=1}^{3839} p(x_i)$$

$$L(\theta) = \left(\frac{2+\theta}{4}\right)^{1997} \cdot \left(\frac{\theta}{4}\right)^{32} \cdot \left(\frac{1-\theta}{4}\right)^{906} \cdot \left(\frac{1-\theta}{4}\right)^{904}$$

$$L(\theta) = \frac{1}{4^{3839}} \cdot (2+\theta)^{1997} \cdot \theta^{32} \cdot (1-\theta)^{1810}$$

$$\ell(\theta) = -3839 \log(4) + 1997 \log(2+\theta) + 32 \log(\theta) + 1810 \log(1-\theta)$$

MIPS 21.8 b

What is the maximum likelihood estimate of θ using the data from the preceding table?

MIPS 21.8 b

What is the maximum likelihood estimate of θ using the data from the preceding table?

$$\ell(\theta) = -3839 \log(4) + 1997 \log(2+\theta) + 32 \log(\theta) + 1810 \log(1-\theta)$$

MIPS 21.8 b

What is the maximum likelihood estimate of θ using the data from the preceding table?

$$\ell(\theta) = -3839 \log(4) + 1997 \log(2+\theta) + 32 \log(\theta) + 1810 \log(1-\theta)$$

$$\ell'(\theta) = \frac{1997}{2+\theta} + \frac{32}{\theta} - \frac{1810}{1-\theta}$$

MIPS 21.8 b

$$0 = \frac{1997}{2 + \hat{\theta}} + \frac{32}{\hat{\theta}} - \frac{1810}{1 - \hat{\theta}}$$

MIPS 21.8 b

$$0 = \frac{1997}{2+\hat{\theta}} + \frac{32}{\hat{\theta}} - \frac{1810}{1-\hat{\theta}}$$

$$0 = 1997\hat{\theta}\left(1-\hat{\theta}\right) + 32\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - 1810\hat{\theta}\left(2+\hat{\theta}\right)$$

MIPS 21.8 b

$$0 = \frac{1997}{2+\hat{\theta}} + \frac{32}{\hat{\theta}} - \frac{1810}{1-\hat{\theta}}$$

$$0 = 1997\hat{\theta}\left(1-\hat{\theta}\right) + 32\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - 1810\hat{\theta}\left(2+\hat{\theta}\right)$$

$$0 = -3839\hat{\theta}^2 - 1655\hat{\theta} + 64$$

MIPS 21.8 b

$$
\begin{aligned}
0 &= \frac{1997}{2+\hat{\theta}} + \frac{32}{\hat{\theta}} - \frac{1810}{1-\hat{\theta}} \\
0 &= 1997\hat{\theta}\left(1-\hat{\theta}\right) + 32\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - 1810\hat{\theta}\left(2+\hat{\theta}\right) \\
0 &= -3839\hat{\theta}^2 - 1655\hat{\theta} + 64 \\
&\dots
\end{aligned}
$$

MIPS 21.8 b

$$
\begin{aligned}
0 &= \frac{1997}{2+\hat{\theta}} + \frac{32}{\hat{\theta}} - \frac{1810}{1-\hat{\theta}} \\
0 &= 1997\hat{\theta}\left(1-\hat{\theta}\right) + 32\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - 1810\hat{\theta}\left(2+\hat{\theta}\right) \\
0 &= -3839\hat{\theta}^2 - 1655\hat{\theta} + 64 \\
&\dots \\
\hat{\theta} &= \frac{1655 - \sqrt{3721809}}{-7678}
\end{aligned}
$$

MIPS 21.8 b

$$
\begin{aligned}
0 &= \frac{1997}{2+\hat{\theta}} + \frac{32}{\hat{\theta}} - \frac{1810}{1-\hat{\theta}} \\
0 &= 1997\hat{\theta}\left(1-\hat{\theta}\right) + 32\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - 1810\hat{\theta}\left(2+\hat{\theta}\right) \\
0 &= -3839\hat{\theta}^2 - 1655\hat{\theta} + 64 \\
&\cdots \\
\hat{\theta} &= \frac{1655-\sqrt{3721809}}{-7678} \\
\hat{\theta} &\approx 0.0357
\end{aligned}
$$

MIPS 21.8 c

Suppose that we have the counts of n different leaves: n_1 of type starchy-green, n_2 of type sugary-white, n_3 of type starchy-white, and n_4 of type sugary-green (so $n = n_1 + n_2 + n_3 + n_4$). Determine the general formula for the maximum likelihood estimate of θ.

MIPS 21.8 c

$$L(\theta) = \prod_{i=1}^{n} p(x_i)$$

MIPS 21.8 c

$$L(\theta) = \prod_{i=1}^{n} p(x_i)$$

$$L(\theta) = \left(\frac{2+\theta}{4}\right)^{n_1} \cdot \left(\frac{\theta}{4}\right)^{n_2} \cdot \left(\frac{1-\theta}{4}\right)^{n_3} \cdot \left(\frac{1-\theta}{4}\right)^{n_4}$$

MIPS 21.8 c

$$L(\theta) = \prod_{i=1}^{n} p(x_i)$$
$$L(\theta) = \left(\frac{2+\theta}{4}\right)^{n_1} \cdot \left(\frac{\theta}{4}\right)^{n_2} \cdot \left(\frac{1-\theta}{4}\right)^{n_3} \cdot \left(\frac{1-\theta}{4}\right)^{n_4}$$
$$L(\theta) = \frac{1}{4^n} \cdot (2+\theta)^{n_1} \cdot \theta^{n_2} \cdot (1-\theta)^{n_3+n_4}$$

MIPS 21.8 c

$$L(\theta) = \prod_{i=1}^{n} p(x_i)$$

$$L(\theta) = \left(\frac{2+\theta}{4}\right)^{n_1} \cdot \left(\frac{\theta}{4}\right)^{n_2} \cdot \left(\frac{1-\theta}{4}\right)^{n_3} \cdot \left(\frac{1-\theta}{4}\right)^{n_4}$$

$$L(\theta) = \frac{1}{4^n} \cdot (2+\theta)^{n_1} \cdot \theta^{n_2} \cdot (1-\theta)^{n_3+n_4}$$

$$\ell(\theta) = -n\log(4) + n_1\log(2+\theta) + n_2\log(\theta) + (n_3+n_4)\log(1-\theta)$$

MIPS 21.8 c

$$L(\theta) = \prod_{i=1}^{n} p(x_i)$$

$$L(\theta) = \left(\frac{2+\theta}{4}\right)^{n_1} \cdot \left(\frac{\theta}{4}\right)^{n_2} \cdot \left(\frac{1-\theta}{4}\right)^{n_3} \cdot \left(\frac{1-\theta}{4}\right)^{n_4}$$

$$L(\theta) = \frac{1}{4^n} \cdot (2+\theta)^{n_1} \cdot \theta^{n_2} \cdot (1-\theta)^{n_3+n_4}$$

$$\ell(\theta) = -n\log(4) + n_1 \log(2+\theta) + n_2 \log(\theta) + (n_3+n_4)\log(1-\theta)$$

$$\ell'(\theta) = \frac{n_1}{2+\theta} + \frac{n_2}{\theta} - \frac{n_3+n_4}{1-\theta}$$

MIPS 21.8 c

$$0 = \frac{n_1}{2 + \hat{\theta}} + \frac{n_2}{\hat{\theta}} - \frac{n_3 + n_4}{1 - \hat{\theta}}$$

MIPS 21.8 c

$$
\begin{aligned}
0 &= \frac{n_1}{2+\hat{\theta}} + \frac{n_2}{\hat{\theta}} - \frac{n_3+n_4}{1-\hat{\theta}} \\
0 &= n_1\hat{\theta}\left(1-\hat{\theta}\right) + n_2\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - (n_3+n_4)\hat{\theta}\left(2+\hat{\theta}\right)
\end{aligned}
$$

MIPS 21.8 c

$$
\begin{aligned}
0 &= \frac{n_1}{2+\hat{\theta}} + \frac{n_2}{\hat{\theta}} - \frac{n_3+n_4}{1-\hat{\theta}} \\
0 &= n_1\hat{\theta}\left(1-\hat{\theta}\right) + n_2\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - (n_3+n_4)\hat{\theta}\left(2+\hat{\theta}\right) \\
0 &= -n\hat{\theta}^2 + (n_1 - n_2 - 2n_3 - 2n_4)\hat{\theta} + 2n_2
\end{aligned}
$$

MIPS 21.8 c

$$0 = \frac{n_1}{2+\hat{\theta}} + \frac{n_2}{\hat{\theta}} - \frac{n_3+n_4}{1-\hat{\theta}}$$

$$0 = n_1\hat{\theta}\left(1-\hat{\theta}\right) + n_2\left(2+\hat{\theta}\right)\left(1-\hat{\theta}\right) - (n_3+n_4)\hat{\theta}\left(2+\hat{\theta}\right)$$

$$0 = -n\hat{\theta}^2 + (n_1 - n_2 - 2n_3 - 2n_4)\hat{\theta} + 2n_2$$

$$\hat{\theta} = \frac{-n_1 + n_2 + 2n_3 + 2n_4 - \sqrt{(n_1 - n_2 - 2n_3 - 2n_4)^2 + 8nn_2}}{-2n}$$

Additional Problems

1. Determine the method of moments estimator for θ using the first moment.
2. Describe the empirical bootstrapping process for $B = 500$.
   a. How many plants should you resample at each iteration?
3. Describe the parametric bootstrapping process for $B = 500$.
   a. How can you transform $U \sim \text{Unif}(0, 1)$ to sample from $p(a)$?
4. When would you choose empirical vs. parametric bootstrapping?

Additional Problem 1

a	1	2	3	4
$p(a)$	$\frac{1}{4}(2+\theta)$	$\frac{1}{4}\theta$	$\frac{1}{4}(1-\theta)$	$\frac{1}{4}(1-\theta)$

Determine the method of moments estimator for θ using the first moment.

Additional Problem 1

a	1	2	3	4
$p(a)$	$\frac{1}{4}(2+\theta)$	$\frac{1}{4}\theta$	$\frac{1}{4}(1-\theta)$	$\frac{1}{4}(1-\theta)$

Determine the method of moments estimator for θ using the first moment.

$$E[X] = \sum_{x \in \{1,2,3,4\}} x \cdot P(X = x)$$

Additional Problem 1

a	1	2	3	4
$p(a)$	$\frac{1}{4}(2+\theta)$	$\frac{1}{4}\theta$	$\frac{1}{4}(1-\theta)$	$\frac{1}{4}(1-\theta)$

Determine the method of moments estimator for θ using the first moment.

$$E[X] = \sum_{x \in \{1,2,3,4\}} x \cdot P(X = x)$$

$$E[X] = \frac{2+\theta}{4} + \frac{2\theta}{4} + \frac{3(1-\theta)}{4} + \frac{4(1-\theta)}{4}$$

Additional Problem 1

a	1	2	3	4
$p(a)$	$\frac{1}{4}(2+\theta)$	$\frac{1}{4}\theta$	$\frac{1}{4}(1-\theta)$	$\frac{1}{4}(1-\theta)$

Determine the method of moments estimator for θ using the first moment.

$$E[X] = \sum_{x \in \{1,2,3,4\}} x \cdot P(X = x)$$

$$E[X] = \frac{2+\theta}{4} + \frac{2\theta}{4} + \frac{3(1-\theta)}{4} + \frac{4(1-\theta)}{4}$$

$$E[X] = \frac{9}{4} - \theta$$

Additional Problem 1

$$\bar{X} = \frac{9}{4} - \hat{\theta}$$

Additional Problem 1

$$\bar{X} = \frac{9}{4} - \hat{\theta}$$
$$\hat{\theta} = \frac{9}{4} - \bar{X}$$

Additional Problem 2

Describe the empirical bootstrapping process for $B = 500$.

Additional Problem 2

Describe the empirical bootstrapping process for $B = 500$.

Given a dataset $x_1, x_2, \ldots, x_n$, determine its empirical distribution function F_n as an estimate of F.

Additional Problem 2

Describe the empirical bootstrapping process for $B = 500$.

Given a dataset $x_1, x_2, \ldots, x_n$, determine its empirical distribution function F_n as an estimate of F.

Compute the parameter

$$\theta^* = \hat{\theta}$$

(e.g., using MLE or MOM) corresponding to F_n.

Additional Problem 2

Describe the empirical bootstrapping process for $B = 500$.

1. Generate a bootstrap dataset $x_1^*, x_2^*, \ldots, x_n^*$ from F_n (i.e., sample from the dataset with replacement).

Additional Problem 2

Describe the empirical bootstrapping process for $B = 500$.

1. Generate a bootstrap dataset $x_1^*, x_2^*, \ldots, x_n^*$ from F_n (i.e., sample from the dataset with replacement).
2. Compute the estimate $\hat{\theta}^*$ of $\theta^* = \hat{\theta}$ (e.g., using MLE or MOM) for the bootstrap dataset.

Additional Problem 2

Describe the empirical bootstrapping process for $B = 500$.

1. Generate a bootstrap dataset $x_1^*, x_2^*, \ldots, x_n^*$ from F_n (i.e., sample from the dataset with replacement).
2. Compute the estimate $\hat{\theta}^*$ of $\theta^* = \hat{\theta}$ (e.g., using MLE or MOM) for the bootstrap dataset.
3. Repeat steps 1 and 2 many times until you generate $B = 500$ bootstrapped estimates.

Additional Problem 2a

How many plants should you resample at each iteration?

Additional Problem 2a

How many plants should you resample at each iteration?

In the scenario from Exercise 21.8, you should resample 3839 plants at each iteration since the sample size of the original dataset is $n = 3839$.

Additional Problem 3

Describe the parametric bootstrapping process for $B = 500$.

Additional Problem 3

Describe the parametric bootstrapping process for $B = 500$.

Given a dataset $x_1, x_2, \ldots, x_n$, compute an estimate $\hat{\theta}$ for θ (e.g., using MLE or MOM).

Additional Problem 3

Describe the parametric bootstrapping process for $B = 500$.

Given a dataset $x_1, x_2, \ldots, x_n$, compute an estimate $\hat{\theta}$ for θ (e.g., using MLE or MOM).

Determine $F_{\hat{\theta}}$ as an estimate for F_θ, and use the parameter $\theta^* = \hat{\theta}$ corresponding to $F_{\hat{\theta}}$.

Additional Problem 3

Describe the parametric bootstrapping process for $B = 500$.

1. Generate a bootstrap dataset $x_1^*, x_2^*, \ldots, x_n^*$ from $F_{\hat{\theta}}$.

Additional Problem 3

Describe the parametric bootstrapping process for $B = 500$.

1. Generate a bootstrap dataset $x_1^*, x_2^*, \ldots, x_n^*$ from $F_{\hat{\theta}}$.
2. Compute the estimate $\hat{\theta}^*$ of $\theta^* = \hat{\theta}$ (e.g., using MLE or MOM) for the bootstrap dataset.

Additional Problem 3

Describe the parametric bootstrapping process for $B = 500$.

1. Generate a bootstrap dataset $x_1^*, x_2^*, \ldots, x_n^*$ from $F_{\hat{\theta}}$.
2. Compute the estimate $\hat{\theta}^*$ of $\theta^* = \hat{\theta}$ (e.g., using MLE or MOM) for the bootstrap dataset.
3. Repeat steps 1 and 2 many times until you generate $B = 500$ bootstrapped estimates.

Additional Problem 3a

How can you transform $U \sim \text{Unif}(0, 1)$ to sample from $p(a)$?

Additional Problem 3a

How can you transform $U \sim \text{Unif}(0, 1)$ to sample from $p(a)$?

Suppose that the CDF of X from Exercise 21.8 is F. We can use it to map probabilities back to the true percentiles of the distribution.

Additional Problem 3a

How can you transform $U \sim \text{Unif}(0,1)$ to sample from $p(a)$?

Suppose that the CDF of X from Exercise 21.8 is F. We can use it to map probabilities back to the true percentiles of the distribution.

1. Simulate a random sample of probabilities (i.e., randomly sample from the interval $[0,1]$, which is the same as sampling from $\text{Unif}(0,1)$).

Additional Problem 3a

How can you transform $U \sim \text{Unif}(0, 1)$ to sample from $p(a)$?

Suppose that the CDF of X from Exercise 21.8 is F. We can use it to map probabilities back to the true percentiles of the distribution.

1. Simulate a random sample of probabilities (i.e., randomly sample from the interval $[0, 1]$, which is the same as sampling from $\text{Unif}(0, 1)$).
2. Transform the probabilities using F^{-1}.

Additional Problem 3a

How can you transform $U \sim \text{Unif}(0, 1)$ to sample from $p(a)$?

Suppose that the CDF of X from Exercise 21.8 is F. We can use it to map probabilities back to the true percentiles of the distribution.

1. Simulate a random sample of probabilities (i.e., randomly sample from the interval $[0, 1]$, which is the same as sampling from $\text{Unif}(0, 1)$).
2. Transform the probabilities using F^{-1}.
3. The resulting values represent a random sample from the distribution $p(a)$ of X.

Additional Problem 4

When would you choose empirical vs. parametric bootstrapping?

Additional Problem 4

When would you choose empirical vs. parametric bootstrapping?

- Use empirical bootstrapping when you are not certain about the model and want to test assumptions around a parametric model, or when you don't have any parametric model.

Additional Problem 4

When would you choose empirical vs. parametric bootstrapping?

- Use empirical bootstrapping when you are not certain about the model and want to test assumptions around a parametric model, or when you don't have any parametric model.
- Use parametric bootstrapping when you are confident about the model, especially when the dataset is small and empirical resampling may not result in distributions representative of the population distribution.