Math 10A with Professor Stankova
Worksheet, Discussion #39; Friday, 12/1/2017
GSI name: Roy Zhao

Maximum Likelihood Estimation (MLE)

Examples

1. You assume that the lifespan of lightbulbs are exponentially distributed (PDF is $\lambda e^{-\lambda t}$ for $t \geq 0$) and notice that your three light bulbs go out in 1, 2, and 3 years. What is the maximum likelihood estimator for λ?

Solution: We want to find the maximum likely λ given our sample of $1, 2, 3$. So, we want to maximize $L(\lambda | x_1, x_2, x_3)$ where $x_i = i$. By definition, we have that $L(\lambda | x_1, x_2, x_3) = P(x_1, x_2, x_3 | \lambda) = P(x_1|\lambda)P(x_2|\lambda)P(x_3|\lambda)$ by independence. We calculate that as $\lambda^3 e^{-6\lambda}$. In order to find the maximum, we take the derivative and set it equal to 0 to get

$$3\lambda^2 e^{-6\lambda} - 6\lambda^3 e^{-6\lambda} = 0 \implies \lambda = 0, 2.$$

The solution $\lambda = 0$ doesn't make sense and hence $\lambda = 2$.

2. I have a bag with 5 red and blue balls. I pull out a ball and it is red. I put it back and I add 3 blue balls and pull out another ball, which is blue. What is the maximum likelihood for the original number of blue balls.

Solution: Let n be the given number of blue balls. Then the probability for picking a red is $1 - n/5$. After, the probability of getting a blue ball after adding 3 blue balls is $(n+3)/8$. So, we want to maximize

$$\frac{5-n}{5} \cdot \frac{n+3}{8} = \frac{-n^2 + 2n + 15}{40}.$$

Taking the derivative and setting it equal to 0, we get that $2n = 2$ or $n = 1$ is the maximum likelihood.

Problems

3. True **FALSE** In MLE, you calculate the probability that your parameter θ is a given value.

Solution: You calculate the likelihood that the parameter is a given value by calculating the probability of the outcomes given the outcome.

4. True **FALSE** In the ball example from above, the solutions is always found by setting the derivative to 0.

Solution: If I solve and get a noninteger number, like 2.2, then I have to plug in $n = 2$ and $n = 3$ and take the higher one.

5. There is a bag with 12 balls colored red and blue. You pull out three balls (with replacement) and get BBR. What is the maximum likelihood for the number of blue balls in bag?

Solution: Let n be the number of blue balls. Then the likelihood is given by $\frac{n}{12} \cdot \frac{n}{12} \cdot \frac{12-n}{12}$. Taking the derivative and setting equal to 0 gives $n = 8$.

6. You have a coin that you think is biased. you flip it 4 times and get the sequence $HHHT$. What is the maximum likelihood estimate for the probability of getting heads?

Solution: Let p be the probability of getting heads. Then the probability of getting $HHHT$ is $p^3(1-p)$. Taking the derivative and setting equal to zero gives $3p^2 - 4p^3 \implies p = \frac{3}{4}$.

7. You know that baby weights are normally distributed with mean μ and standard deviation σ. You have three babies weighing $7, 8, 9$ ounces. What is the maximum likelihood for μ?

Solution: The PDF for a normal distribution is $\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$. So, we want to maximize

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-(7-\mu)^2/2\sigma^2-(8-\mu)^2/2\sigma^2-(9-\mu)^2/2\sigma^2}$$

for μ. We can take the logarithmic derivative and set that equal to 0. Doing so gives us

$$-2(7-\mu) - 2(8-\mu) - 2(9-\mu) = 0 \implies \mu = 8.$$

8. There is a bag with 10 balls colored red and blue. You pull out two balls (with replacement) and get BR. What is the maximum likelihood for the number of blue balls in bag?

Solution: Let n be the number of blue balls. Then the likelihood is given by $\frac{n}{10} \cdot \frac{10-n}{10}$. Taking the derivative and setting equal to 0 gives $n = 5$.

9. You have a coin that you think is biased. you flip it 3 times and get the sequence HTH. What is the maximum likelihood estimate for the probability of getting heads?

Solution: Let p be the probability of getting heads. Then the probability of getting HTH is $p^2(1-p)$. Taking the derivative and setting equal to zero gives $2p^2 - 3p^2 \implies p = \frac{2}{3}$.

10. You know that baby weights are normally distributed with mean μ and standard deviation σ. You have three babies weighing $6, 8, 10$ ounces. What is the maximum likelihood for μ?

Solution: The PDF for a normal distribution is $\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$. So, we want to maximize

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-(6-\mu)^2/2\sigma^2 - (8-\mu)^2/2\sigma^2 - (10-\mu)^2/2\sigma^2}$$

for μ. We can take the logarithmic derivative and set that equal to 0. Doing so gives us

$$-2(6-\mu) - 2(8-\mu) - 2(10-\mu) = 0 \implies \mu = 8.$$

Hypothesis Testing

Example

11. In patients with lung cancer, about 90% of them die within three years. They undergo an experimental treatment that claims to reduce their likelihood of dying. Out of 400 patients who underwent this treatment, 348 of them still die. Can we say that this treatment is successful? Use a significance level of 0.05.

Solution: The null hypothesis is something that we want to disprove. So, we set H_0 to be the hypothesis that this treatment doesn't work. This is one sided because we want to know if this treatment reduces the likelihood. Assuming the null hypothesis, our sample of 400 patients will be normally distributed by CLT and have a mean of 0.9 and a sample standard error of $\sigma/\sqrt{400} = \sigma/20$, where σ is the standard deviation of a given person. This standard deviation is given by $\sqrt{p(1-p)} = 0.3$. So by CLT, our sample average percentage of people dying should be normally distributed with $\mu = 0.9$ and $\sigma = 0.3/20$.

Under these conditions, the probability that we get a sample with mean less than or equal to $348/400 = 0.87$ is $\frac{1}{2} - z(|0.87 - 0.9|/(0.3/20)) = \frac{1}{2} - z(2) = 0.0228 < 0.05$. Therefore, we can reject the null hypothesis and this treatment is successful.

Problems

12. **TRUE** False Accusing a student of cheating when they didn't is a Type I error.

Solution: In this, the null hypothesis would be that students are not cheating. So, we are rejecting this hypothesis incorrectly, which is type I.

13. True **FALSE** We usually want to show that the null hypothesis is true.

Solution: We usually want to reject the null hypothesis.