Math 10B with Professor Stankova
Worksheet, Discussion #23; Thursday, 4/18/2019
GSI name: Roy Zhao

Maximum Likelihood Estimation (MLE)

Concepts

1. The **likelihood function** $L(\theta)$ which is the probability that we see the data we see if we set the parameter equal to θ. Namely, $L(\theta|x_1, \ldots, x_n) = P(x_1, \ldots, x_n|\theta)$, the probability we see $x_1, \ldots, x_n$ if our parameter is equal to θ. Then we choose the value of θ that maximizes this function by taking the derivative and setting it equal to 0.

Distribution	PMF/PDF	$E(X)$	Variance
Uniform	If $\#R(X) = n$, then $f(x) = \frac{1}{n}$ for all $x \in R(X)$.	$\sum_{i=1}^n \frac{x_i}{n}$	$\sum_{i=1}^n \frac{(x_i-\mu)^2}{n}$
Bernoulli Trial	$f(0) = 1-p$, $f(1) = p$	p	$Var(X) = p(1-p)$
Binomial	$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Geometric	$f(k) = (1-p)^k p$	$\frac{1-p}{p}$	$Var(X) = \frac{1-p}{p^2}$
Hyper-Geometric	$f(k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$	$\frac{nm}{N}$	$\frac{nm(N-m)(N-n)}{N^2(N-1)}$
Poisson	$f(k) = \frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ
Pareto	$f(x) = \frac{\alpha-1}{x^\alpha}, x \geq 1$	$\frac{\alpha-1}{\alpha-2}$	$\frac{\alpha-1}{(\alpha-2)^2(\alpha-3)}$
Uniform	$f(x) = \frac{1}{b-a}$ for $a \leq x \leq b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2
Exponential	$f(x) = ce^{-cx}$	$\frac{1}{c}$	$\frac{1}{c^2}$
Laplacian	$f(x) = \frac{1}{2}e^{-\|x\|}$	0	2

Examples

2. The number of threes made during an NBA game is Poisson distributed. Last Saturday, the number of threes made were $14, 26, 25,$ and 13. Calculate the maximum likelihood estimate for the parameter λ.

Solution: The likelihood function is $L(\lambda|14, 26, 25, 13) = P(14, 26, 25, 13|\lambda) = \frac{\lambda^{14}e^{-\lambda}}{14!} \cdot \frac{\lambda^{26}e^{-\lambda}}{26!} \cdot \frac{\lambda^{25}e^{-\lambda}}{25!} \cdot \frac{\lambda^{13}e^{-\lambda}}{13!} = \frac{\lambda^{78}e^{-4\lambda}}{14!26!25!13!}$. Taking the derivative and setting it equal to 0 gives

$$\lambda^{78}(-4e^{-4\lambda}) + 78\lambda^{77}e^{-4\lambda} = \lambda^{77}e^{-4\lambda}(-4\lambda + 78) = 0.$$

Thus $\lambda = 0$ or $\lambda = \frac{78}{4} = 19.5$. Since $\lambda = 0$ doesn't make sense and 19.5 is a local maximum, we have $\lambda = 19.5$.

Problems

3. True **FALSE** The maximum likelihood estimate for the standard deviation of a normal distribution is the sample standard deviation ($\hat{\sigma} = s$).

Solution: We showed in class that the maximum likelihood is actually the biased estimator s_*.

4. True **FALSE** The maximum likelihood estimate is always unbiased.

Solution: The estimator for σ of a normal distribution is biased.

5. 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 $3p^2 - 4p^3 \implies p = \frac{3}{4}$.

6. During Cal Day, my two friends and I asked prospective students where they were from until we found someone who wasn't from California. We had to ask $23, 18, 46$ people respectively before finding someone not from California. What is the maximum likelihood estimate for the percentage of students from California?

Solution: This is a geometric distribution. Let p be the probability of success which in this case is finding someone not from California. Then, the likelihood function is

$$L(p|23, 18, 46) = P(23, 18, 46|p) = (1-p)^{23}p \cdot (1-p)^{18}p \cdot (1-p)^{46}p = (1-p)^{87}p^3.$$

We want to maximize this and hence take its derivative and set it equal to 0. Since this is a product, we take its log derivative. So we first take the log which gives $87\ln(1-p) + 3\ln p$. Then we take the derivative and set it equal to 0 to get

$$\frac{-87}{1-p} + \frac{3}{p} = 0.$$

So $\frac{3}{p} = \frac{87}{1-p}$ and $3 - 3p = 87p$ so $p = \frac{1}{30}$ is the maximum. Thus, the maximum likelihood for the percentage of students from California is 1 minus this or $\frac{29}{30}$.

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

$$\frac{1}{\sigma^3\sqrt{2\pi}^3}e^{-(7-\mu)^2/2\sigma^2-(8-\mu)^2/2\sigma^2-(9-\mu)^2/2\sigma^2}$$

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

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

Hypothesis Testing

Concepts

8. In general, statistics does not allow you to prove anything is true, but instead allows you to show that things are probably false. So when we do hypothesis testing, the **null hypothesis** H_0 is something that we want to show is false and the **alternative hypothesis** H_1 is something that you want to show is true. For example, to show that a drug cures cancer, the null hypothesis would be that the drug does nothing and the alternative hypothesis would be that the drug does help cure cancer.

   A **type 1 error** is rejecting a true null which means that in our example, saying a drug cures cancer when it doesn't. A **type 2 error** is failing to reject a false null which means in our case as saying that the drug doesn't do anything when it does. The **significance level** is the probability of making a type 1 error. The **power** is 1 minus the probability of making a type 2 error.

Examples

9. Chip bag manufacturers claim the weights are normally distributed with a mean of 14 ounces and a standard deviation of 0.5 ounces. Your bag is 13 ounces. What can you say with significance level $\alpha = 0.05$?

Solution: The null hypothesis is that the chips have a weight of 14 ounces. We calculate the z score as $P(X \leq 13) = P(Z \leq \frac{13-14}{0.5}) = P(Z \leq 2) = \frac{1}{2} - z(2) = 0.0228 < \alpha/2$. Thus, we can say that the chip bag makers are lying.

Problems

10. **TRUE** False The null hypothesis is something we want to be false.

11. True **FALSE** If we get a p value that is not smaller than α, then we have shown that the null hypothesis is true.

Solution: We simply do not have enough evidence to show that it is false, not proven that it is true.

12. True **FALSE** We want our test to have a high significance level and high power.

Solution: The significance level is the probability of making a type 1 error so we want that low and the power is 1 minus the probability of making a type 2 error so we want that large. So small significance level and high power.

13. **TRUE** False A type-2 error made by a road patrol may result in letting drunken drivers continue driving.

Solution: A type-2 error is when you fail to reject the null hypothesis. We are asking if someone is drunk driving so the null hypothesis, the thing we want to disprove is that he is not drunk driving. So we fail to reject this so that means that they are drunk driving but we do not think so.

14. You think a die is rigged so that it will roll 5 less often. You continually roll the die until you get a 5 and you have to roll 10 times before this happens. Is your suspicion correct with $\alpha = 0.05$?

Solution: The null hypothesis is that the die is fair and the probability of rolling 5 is $p = \frac{1}{6}$. This is a geometric distribution and let X be the number of times I roll until I get a 5. Our p value is the probability we have to roll at least 10 times which is $P(X \geq 10) = (1-p)^{10} = \frac{5^{10}}{6^{10}} = 0.16 > \alpha$. Therefore we cannot reject the null hypothesis.

15. (True story) A woman claims that she can smell when someone has Parkinson's disease. She is given 10 people's shirts and correctly said whether the person had the disease in 9 of the 10 cases. Does she have this ability with $\alpha = 0.05$? (The 10th person who she said had Parkinson's actually developed it months later so she was really 10 for 10).

> **Solution:** The null hypothesis would be that she guessed randomly and the probability that she was successful once is $p = \frac{1}{2}$. We want to calculate the probability that she did at least this well so $P(X \geq 9) = P(X = 9) + P(X = 10) = \binom{10}{9}\frac{1}{2^9}(1-\frac{1}{2}) + \binom{10}{10}\frac{1}{2^{10}} = \frac{11}{2^{10}} = 0.01 < \alpha$. Therefore, we can reject the null hypothesis and say that she does have this ability.

16. (True story) A woman claimed that she could tell whether milk or tea was added first to a cup. She was given 4 cups with milk added first then tea and 4 cups of the opposite. She guessed all 8 correctly. Can we say that she has this ability with $alpha = 0.05$?

> **Solution:** The null hypothesis would be that she guessed randomly and the probability that she was successful once is $p = \frac{1}{2}$. We want to calculate the probability that she did at least this well so $P(X \geq 8) = P(X = 8) = \binom{8}{8}\frac{1}{2^8} = \frac{1}{2^8} = 0.004 < \alpha$. Therefore, we can reject the null hypothesis and say that she does have this ability.

17. If the woman in the previous problem had only been given 4 total cups to test, explain why we would never be able to reject the null hypothesis.

> **Solution:** The most extreme case would be if she got all 4 right and even then the p value would be $P(X \geq 4) = \binom{4}{4}\frac{1}{2^4} = \frac{1}{16} = 0.0625 > \alpha$. So we wouldn't been able to reject the null hypothesis.

18. Some scientists publish a study that says that the average height of men is normally distributed with mean 64 inches and standard deviation 2 inches. If I am 67 inches tall, can I say that the scientists are wrong with $\alpha = 0.05$?

> **Solution:** The null hypothesis is that the scientists are right and we want to calculate $P(X \geq 67)$ and check if it is less than $\alpha/2$ (two sided). Since it is normally distributed, we can use z scores and $P(X \geq 67) = P(Z \geq \frac{67-64}{2}) = P(Z \geq 1.5) = 0.5 - z(1.5) = 0.0668 > \alpha/2$. So we cannot reject the null hypothesis.