3. Another Example of Maximum Likelihood Estimator

MLE for a Loaded Die: Likelihood

1/1 point (graded)

You have a loaded (i.e. possibly unfair) six-sided die with the probability that it shows a "3" equal to η and the probability that it shows any other number equal to $(1-\eta)/5$.

Let X be a random variable representing a roll of this die. You roll this die n times, and record your data set, consisting of the values of the faces as $X_1, X_2, X_3, \ldots X_n$.

Let the outcome of a set of n rolls of the die be modeled by the i.i.d. random variable sequence $(X_1, \ldots, X_n)$. We model the i'th roll as X_i where $X_i = j$ if the top face of the die shows a "j".

You roll the die n times and you see the outcome $X_i = 3$ exactly k times. What is the likelihood function $L_n(x_1, \ldots, x_n, \eta)$?

(Enter **eta** for η.)

eta^k*((1-eta)/5)^(n-k)

✔ **Answer:** eta^k*((1-eta)/5)^(n-k)

$\eta^k \cdot \left(\frac{1-\eta}{5}\right)^{n-k}$

STANDARD NOTATION

Solution:

Denote by $p_\eta(x)$ the pmf of X_i. Then, the likelihood function is

$$\begin{aligned} L_n(x_1, \ldots, x_n, \eta) &= \prod_{i=1}^{n} p_\eta(x_i) \\ &= \eta^k \left(\frac{1-\eta}{5}\right)^{n-k}. \end{aligned}$$

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

MLE for a Loaded Die: MLE

1/1 point (graded)

Find the ML estimator $\hat{\eta}_n^{\mathrm{MLE}}$.

k/n

✔ **Answer:** k/n

$\frac{k}{n}$

STANDARD NOTATION

Solution:

Since we are looking for the $\operatorname{argmax}_{\eta\in[0,1]} L_n(x_1,\ldots,x_n,\eta)$, we can ignoring any scaling constant in $L_n(x_1,\ldots,x_n,\eta)$. Hence, we will maximize $\widetilde{L}_n(x_1,\ldots,x_n,\eta) = \eta^k(1-\eta)^{n-k}$.

Taking the derivative of $\widetilde{L}_n(x_1,\ldots,x_n,\eta)$ with respect to η and setting it to 0, we get

$$\begin{aligned} k(1-\eta) &= (n-k)\eta \\ \implies \hat{\eta}_n^{\mathrm{MLE}} &= \frac{k}{n}. \end{aligned}$$

Remark: The function $\widetilde{L}_n(x_1,\ldots,x_n,\eta) = \eta^k(1-\eta)^{n-k}$ whose maximizer is $\hat{\eta}_n^{\mathrm{MLE}}$ is the same as the likelihood function for a Bernoulli experiment with parameter η, even though each roll of a die has 6 potential outcomes.

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

(Optional) Generalization of the Loaded Die Problem

Show

Discussion

Hide Discussion

Topic: Unit 3 Methods of Estimation:Lecture 10: Consistency of MLE, Covariance Matrices, and Multivariate Statistics / 3. Another Example of Maximum Likelihood Estimator

Add a Post

All Posts

loaded die

question posted 5 days ago by **nbourbon**

I tried using a simple representation of the problem as if it is a binomial variable given that I'm only interested in prob x =3.. but the grader gave me a red mark although I understand pmf and likelihood are very similar removing the constant only

Any hint?

This post is visible to everyone.

Add a Response

4 responses

denizstij
5 days ago

if you follow the method (take the derivative of likelihood function according to η), u will get there (result might be similar to bernouli distrubution or fair diced) ...

I was doing it on paper and mixing up my eta and n :P . got it right after two mistakes

posted 4 days ago by **SuhailWali**

This problem tried to force us to have respectable handwriting. I opted for θ instead.

posted 4 days ago by **derekgriffing**

denizstij... I'm still not able to pass the likelihood function that is my problem at this point. I see it as a binomial but either factorials are not allowed or even if I remove the factorials I can't get it right.

posted 4 days ago by **nbourbon**

Add a comment

Gaylyn
4 days ago

I also see the likelihood function as a binomial, however I can't get past the first question since the grader doesn't accept factorials. Any helpful hints to help me move along?

I'm still with the problem...

posted 4 days ago by **nbourbon**

You're given the pdf explicitly. You don't need to interpret it as any particular distribution.

posted 2 days ago by **synnfusion**

Add a comment

rickytyagi
4 days ago

I recall from the lecture the definition of $L_n\left(x_1, \ldots, x_n, \eta\right)$
$= P_\eta\left(X_1 = x_1, .. X_n = x_n\right)$
$= P_\eta\left(X_1 = x_1\right) * \ldots * P_\eta\left(X_n = x_n\right)$...independence property

then, if you think about observing any one instance of such sequence given the criteria of k number of 3's and $n-k$ number of remaining observation, it should lead to the the right answer.

Are you saying that this should not be modeled as Binomial? To me Binomial seems to be perfect here. Since the grader does not accept factorial as part of the answer, I simply removed the constant which should be OK for a likelihood function. And using the result as the base for MLE does get to the right answer. Yet the grader disagrees.

posted 3 days ago by **sean_s_wang**

Think carefully whether the binomial coefficient is required here...

posted 3 days ago by **sudarsanvsr_mit** (Staff)

To comply with the requirement of being a probability value, it should have the coefficient. But the grader does not accept factorial operation. So maybe there is another way to enter the coefficient which is what I am missing.

Technically the coefficient does not contain η. The term will be removed after log and gradient operations. So it would not make any difference for the MLE.

posted 3 days ago by **sean_s_wang**

I cannot stress again that you need to think carefully why the binomial coefficient is not required. It has nothing to do with making it a valid probability value. In fact, including it is *wrong* in this case.

posted 3 days ago by **sudarsanvsr_mit** (Staff)

Yes you are right. After reading the description of the question carefully I realized that you are right. We are counting one specific sequence. Including the coefficient here is indeed wrong.

Thanks a lot! The staff of the course are super helpful.

posted 3 days ago by **sean_s_wang**

It's also worth noticing that the pdf is given explicitly here. There's no need to try to cast it as any specific distribution,

posted 2 days ago by **synnfusion**

...

Ah, because we're looking at $(X_1 = x_1, \dots, X_n = x_n)$, not just K=k.

posted 2 days ago by **Gaylyn**

Add a comment

nbourbon
2 days ago

+
...

I finally got the green tick and here is the tip If we get any probability p you would be tempted to use 1-p as the complement but think carefully.. as the complement is given in the exercise ..

...

People are overthinking the problem. The pmf is specified exactly. Only it is written in words.

posted about 21 hours ago by **Analesdey**

...

I got the tick, but I don't get the solution. The die probabilities don't add to 1, shouldn't a likelihood function contain all possible outcomes?

posted about 13 hours ago by **Frank_Coumans**

...

@Frank_coumans The likelihood isn't a probability density function. The values that the likelihood function takes on are not probabilities. It's just a function that our goal is to maximize. Remember in its formulation we were dropping constants all over the place because they didn't change the value of θ that maximizes it. So we don't expect the values of the likelihood to necessarily sum to 1.

posted about 13 hours ago by **synnfusion**

As I can conclude from the definitions in this course and in several books, likelihood **is** a joint PDF / PMF. In this particular problem, likelihood is a legitimate PMF by construction.

One has to think carefully how this PMF is constructed. I will just reiterate what has already been mentioned here: this is neither Bernoulli nor Binomial distribution! Try to write down explicitly the PMF of a single random variable; then, by independence, construct a joint PMF.

IMO this was one of the coolest problems in the course.

posted about 3 hours ago by **Hryhorchuk**

I'll go back and take another look. Maybe I was wrong in my explanation.

posted about an hour ago by **synnfusion**

Add a comment

Showing all responses

Add a response:

Preview

Submit

Learn About Verified Certificates