

Review of Likelihoods

Pascal Bell

```
library(knitr)
library(magrittr)
library(kableExtra)
```

Instructions

- Write all narrative using full sentences. Write all interpretations and conclusions in the context of the data.
- Be sure all analysis code is displayed in the rendered pdf.
- If you are fitting a model, display the model output in a neatly formatted table. (The `tidy` and `kable` functions can help!)
- If you are creating a plot, use clear and informative labels and titles.
- Render and back up your work regularly, such as using Github.
- When you're done, we should be able to render the final version of the Rmd document to fully reproduce your pdf.
- Upload your pdf to Gradescope. Upload your Rmd, pdf (and any data) to Canvas.

Exercises

Exercise 1

Write out the likelihood for the Poisson distribution for $x_{1:n}$.

The probability mass function (PMF) of a Poisson-distributed random variable X with parameter λ is given by:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

where $\lambda > 0$ is the rate parameter, which represents both the mean and variance of the distribution.

$$L(\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$L(\lambda) = \left(\prod_{i=1}^n \frac{1}{x_i!} \right) \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}$$

$$\ell(\lambda) = \log \left[\left(\prod_{i=1}^n \frac{1}{x_i!} \right) \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \right]$$

$$\ell(\lambda) = \log(\lambda^{\sum_{i=1}^n x_i}) + \log(e^{-n\lambda}) - \log(\prod_{i=1}^n x_i!)$$

$$\ell(\lambda) = \sum_{i=1}^n x_i \log(\lambda) - \lambda n + \sum_{i=1}^n \log(x_i!)$$

$$\ell(\lambda) = \sum_{i=1}^n x_i \log(\lambda) - \lambda n + \sum_{i=1}^n \log(x_i!)$$

Exercise 2

Derive using calculus based methods the MLE of λ is $\sum_i x_i/n$ (sample mean) and show that it is in fact a maximum.

$$\frac{\partial \ell(\lambda)}{\partial \lambda} = \frac{\sum_{i=1}^n x_i}{\lambda} - n$$

Set the derivative equal to zero:

$$\frac{\sum_{i=1}^n x_i}{\lambda} - n = 0$$

$$\frac{\sum_{i=1}^n x_i}{\lambda} = n$$

$$\lambda = \frac{\sum_{i=1}^n x_i}{n}$$

The second derivative of $\ell(\lambda)$ is:

$$-\frac{\sum_{i=1}^n x_i}{\lambda^2}$$

The second derivative is always negative since lambda and x are always positive, so the the critical point we found above is a maximum.

Exercise 3

Verify using a grid-search that your solution matches to the calculus based one, where you may assume for simplicity that $\sum_i x_i = 500$. You may assume 100 observations. (Hint: show that the approximated MLE is 5.)

```
log_lik <- function(sum_xi, lambda, n) {  
  return(sum_xi * log(lambda) - (lambda * n) - sum(lfactorial(sum_xi)))  
}  
  
lambda <- seq(4.5, 5.5, length = 10000)  
lik <- log_lik(500, lambda, 100)  
lambda[lik == max(lik)]
```

```
## [1] 5.00005
```

Exercise 4 (Derived from Chapter 2 of BMLR).

The hot hand in basketball. @Gilovich1985 wrote a controversial but compelling article claiming that there is no such thing as “the hot hand” in basketball. That is, there is no empirical evidence that shooters have stretches where they are more likely to make consecutive shots, and basketball shots are essentially independent events. One of the many ways they tested for evidence of a “hot hand” was to record sequences of shots for players under game conditions and determine if players are more likely to make shots after made baskets than after misses. For instance, assume we recorded data from one player’s first 5 three-point attempts over a 5-game period. We can assume games are independent, but we’ll consider two models for shots within a game:

- No Hot Hand (1 parameter): p_B = probability of making a basket (thus $1 - p_B$ = probability of not making a basket).
 - Hot Hand (2 parameters): p_B = probability of making a basket after a miss (or the first shot of a game); $p_{B|B}$ = probability of making a basket after making the previous shot.
- a. Fill out Table @ref(tab:hothandchp2)—write out the contribution of each game to the likelihood for both models along with the total likelihood for each model.

Total likelihood for no hot hand model is $(p_B)^{10}(1 - p_B)^{15}$

Total likelihood for hot hand is $(p_B)^7(p_{B|B})^3(1 - p_{B|B})^4(1 - p_B)^{11}$

- b. Given that, for the No Hot Hand model, $\text{Lik}(p_B) = p_B^{10}(1 - p_B)^{15}$ for the 5 games where we collected data, how do we know that 0.40 (the maximum likelihood estimator (MLE) of p_B) is a better estimate than, say, 0.30?

0.4 is a better estimate than 0.3 because when p_B is 0.4, the likelihood of seeing the sampled data is maximized, with a probability of 4.930247e-08 compared to a probability of 2.803388e-08 when p_B is 0.3

- c. Find the MLEs for the parameters in each model, and then use those MLEs to determine if there's significant evidence that the hot hand exists using a likelihood ratio test (LRT). Be sure to specify the test and provide all details of your approach, including reproducible code used.

We can create a hypothesis test between the two models, where the null and alternate hypothesis are:

$H_0 : p_{B(\text{conditional})} = p_{B|B} = p_{B(\text{unconditional})}$ $H_a : \text{At least one of } p_{B(\text{conditional})}, p_{B|B}, p_{B(\text{unconditional})} \text{ differs from the others.}$

Let's test the null hypotheses:

MLE for no hot hand model: $\hat{p}_B = \frac{10}{25} = 0.4$

MLEs for hot hand model: $\hat{p}_B = \frac{7}{18} = 0.39$ $\hat{p}_{B|B} = \frac{3}{7} = 0.43$

The maximum value for the log likelihood for the unconditional model is -16.825 and the maximum value for the log likelihood for the conditional model is approximately -16.809.

The differences in the maximum likelihood is approximately 0.0164.

The LRT is 0.033.

The LRT statistic follows a chi squared distribution with 1 degree of freedom. Therefore, the p-value is $P(X^2 > LRT)$ and equals 0.856. The p-value is very large, so we fail to reject the null hypothesis. We do not find evidence that the conditional model is an improvement over the unconditional model, and we do not find significant evidence that hot hand exists.

Below is the code that calculates all the variables used above, with the p-value printed below.

```
log_lik_no_hot <- function(p_b) {
  return(log((p_b)^10 * (1-p_b)^15))
}

log_lik_hot_pb <- function(pb) {
  return(7*log(pb) + 11*log(1-pb))
}

log_lik_hot_pbb <- function(pbb) {
  return(3*log(pbb) + 4*log(1-pbb))
}

# calculate mle's for each parameter
mle_no_hot_pb <- unname(unlist(optimize(log_lik_no_hot, interval = c(0, 1), maximum = TRUE)[1]))
mle_hot_pb <- unname(unlist(optimize(log_lik_hot_pb, interval = c(0, 1), maximum = TRUE)[1]))
mle_hot_pbb <- unname(unlist(optimize(log_lik_hot_pbb, interval = c(0, 1), maximum = TRUE)[1]))

# plug in MLEs to find max log likelihood
max_log_lik_no_hot <- log_lik_no_hot(mle_no_hot_pb)
max_log_lik_hot <- log_lik_hot_pb(mle_hot_pb) + log_lik_hot_pbb(mle_hot_pbb)
```

```
# LRT test
diff <- max_log_lik_hot - max_log_lik_no_hot
LRT <- 2 * diff
pchisq(LRT, 1, lower.tail = FALSE)
```

```
## [1] 0.8560132
```

Game	First five shots	Likelihood (no hot hand)	Likelihood (hot hand)
1	BMMBB	$(p_B)^3(1 - p_B)^2$	$(p_B)^2(p_{B B})^1(1 - p_{B B})^1(1 - p_B)^1$
2	MBMBM	$(p_B)^2(1 - p_B)^3$	$(p_B)^2(p_{B B})^0(1 - p_{B B})^2(1 - p_B)^1$
3	MMBBB	$(p_B)^3(1 - p_B)^2$	$(p_B)^1(p_{B B})^2(1 - p_{B B})^0(1 - p_B)^2$
4	BMMMB	$(p_B)^2(1 - p_B)^3$	$(p_B)^2(p_{B B})^0(1 - p_{B B})^1(1 - p_B)^2$
5	MMMMM	$(p_B)^0(1 - p_B)^5$	$(p_B)^0(p_{B B})^0(1 - p_{B B})^0(1 - p_B)^5$

Grading

Total	15
Ex 1	2
Ex 2	5
Ex 3	5
Ex 4	8
Workflow & formatting	3

The “Workflow & formatting” grade is to based on the organization of the assignment write up along with the reproducible workflow. This includes having an organized write up with neat and readable headers, code, and narrative, including properly rendered mathematical notation. It also includes having a reproducible Rmd or Quarto document that can be rendered to reproduce the submitted PDF.