Assignment 3 – Maximum Likelihood Estimator

Leonardo L Sanchez

2026-01-30

Problem 1 – In a college, p×100% of all art use a MacBook and p/2×100% of all business students use a MacBook.

The value of p is unknown. To estimate p, a random sample of 5 art students and 5 business students is selected. Each student is asked the question: "Do you use a MacBook?" The responses are summarized below. Art Students Students 1st 2nd 3rd 4th 5th Uses MacBook Yes Yes Yes Yes No

Business Students Students 1st 2nd 3rd 4th 5th Uses MacBook Yes Yes Yes No No

Questions:

a. Formulate the likelihood function L(p) based on the observed data.

```r
# Problem 1(a) - Likelihood function L(p)

# Data from the table
n_art <- 5;  x_art <- 4   # 4 "Yes" out of 5 art students
n_bus <- 5;  x_bus <- 3   # 3 "Yes" out of 5 business students

# Model (from statement):
#P(Mac | Art) = p,  P(Mac | Business) = p/2
# Likelihood (ignoring constants):
# L(p) = p^(x_art) (1-p)^(n_art-x_art) * (p/2)^(x_bus) * (1 - p/2)^(n_bus-x_bus)

L_noConst <- function(p){
  # valid parameter region: 0 < p < 1 (also ensures 1 - p/2 > 0)
  if(any(p <= 0 | p >= 1)) return(0)
  p^x_art * (1 - p)^(n_art - x_art) * (p/2)^x_bus * (1 - p/2)^(n_bus - x_bus)
}
```

b. Derive the log-likelihood function, l(p).

The Art component:

$$\ln(p^4(1-p)^1) = 4\ln(p) + 1\ln(1-p)$$

The Business component:

$$\ln((\frac{p}{2})^3(1-\frac{p}{2})^2) = 3\ln(\frac{p}{2}) + 2\ln(1-\frac{p}{2})$$

Note: We can simplify

$$3\ln(\frac{p}{2})$$

to

$$3\ln(p) - 3\ln(2)$$

.reduce:

$$l(p) = 4\ln(p) + \ln(1-p) + 3\ln(p) - 3\ln(2) + 2\ln(1 - \frac{p}{2})$$

$$l(p) = 7\ln(p) + \ln(1-p) + 2\ln(1 - \frac{p}{2}) - 3\ln(2)$$

```r
# Define the log-likelihood function in R
l_p <- function(p) {
  # Ensure p stays within the valid range (0, 1)
  if(any(p <= 0 | p >= 1)) return(-Inf)

  7*log(p) + log(1-p) + 2*log(1 - p/2) - 3*log(2)
}
```

Step-by-Step DifferentiationWe start with the simplified function from the previous step:

$$l(p) = 7\ln(p) + \ln(1-p) + 2\ln\left(1 - \frac{p}{2}\right) - 3\ln(2)$$

We apply the derivative rule for logarithms,

$$\frac{d}{dp}[\ln(u)] = \frac{1}{u} \cdot \frac{du}{dp}$$

:

First Term:

$$\frac{d}{dp}[7\ln(p)] = \frac{7}{p}$$

Second Term (Chain Rule):

$$\frac{d}{dp}[\ln(1-p)] = \frac{1}{1-p} \cdot \frac{d}{dp}(1-p) = \frac{1}{1-p} \cdot (-1) = -\frac{1}{1-p}$$

Third Term (Chain Rule):

$$\frac{d}{dp}\left[2\ln\left(1 - \frac{p}{2}\right)\right] = 2 \cdot \frac{1}{1 - \frac{p}{2}} \cdot \frac{d}{dp}\left(1 - \frac{p}{2}\right)$$

$$= 2 \cdot \frac{1}{1 - \frac{p}{2}} \cdot \left(-\frac{1}{2}\right) = -\frac{1}{1 - \frac{p}{2}}$$

Fourth Term:

$$\frac{d}{dp}[-3\ln(2)] = 0 \quad \text{(since it is a constant)}$$

The Resulting Derivative Combining these parts, the derivative

$$l'(p) is :' (p) = \frac{7}{p} - \frac{1}{1-p} - \frac{1}{1 - \frac{p}{2}}$$

To make this easier to solve for zero (the next step), we can simplify the last term by multiplying the numerator and denominator by

$$l'(p) = \frac{7}{p} - \frac{1}{1-p} - \frac{2}{2-p}$$

d. Let pˆ be the maximum likelihood estimate of p. Find pˆ.

Step 1: Set the Derivative to ZeroFrom the previous step, we have:

$$\frac{7}{p} - \frac{1}{1-p} - \frac{2}{2-p} = 0$$

Step 2: Clear the FractionsMultiply the entire equation by the common denominator

$$p(1-p)(2-p)$$

to eliminate the fractions:

$$7(1-p)(2-p) - 1(p)(2-p) - 2(p)(1-p) = 0$$

Step 3: Expand the Terms Term 1:

$$7(2 - 3p + p^2) = 14 - 21p + 7p^2$$

Term 2:

$$-1(2p - p^2) = -2p + p^2$$

Term 3:

$$-2(p - p^2) = -2p + 2p^2$$

Combine them into one equation:

$$(14 - 21p + 7p^2) + (-2p + p^2) + (-2p + 2p^2) = 0$$

Step 4: Simplify the Quadratic EquationGroup the like terms:

$$p^2$$

terms:

$$7p^2 + 1p^2 + 2p^2 = 10p^2$$

$p$$ terms:

$$-21p - 2p - 2p = -25p$$

Constant terms:

$$14$$

This gives us the quadratic equation:

$$10p^2 - 25p + 14 = 0$$

Step 5: Solve using the Quadratic Formula

$$p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where

$$a = 10, b = -25, c = 14$$

:

$$p = \frac{25 \pm \sqrt{(-25)^2 - 4(10)(14)}}{2(10)}$$

$$p = \frac{25 \pm \sqrt{625 - 560}}{20}$$

$$p = \frac{25 \pm \sqrt{65}}{20}$$

Step 6: Evaluate the SolutionsUsing

$$\sqrt{65} \approx 8.06$$

:

$$p_1 = \frac{25 + 8.06}{20} \approx 1.653$$

$$p_2 = \frac{25 - 8.06}{20} \approx 0.847$$

```
# Define the derivative function
deriv_l <- function(p) { 7/p - 1/(1-p) - 2/(2-p) }

# Solve for p in the range (0, 1)
mle_p <- uniroot(deriv_l, interval = c(0.01, 0.99))$root
print(mle_p) # Should be approx 0.84688
```

```
## [1] 0.8469056
```

e. Suppose you want to perf test whether p 0.8 I State the null and alternative hypotheses.

Since you want to test whether

$$p$$

is specifically different from 0.8, you are looking at a two-tailed test.I. State the null and alternative hypotheses Null Hypothesis (H_0): The probability p is equal to 0.8.

$$H_0 : p = 0.8$$

Alternative Hypothesis (H_a): The probability p is not equal to 0.8.

$$H_a : p \neq 0.8$$

Likelihood Ratio Test (LRT) statistic

$$\Lambda = -2 \ln \left(\frac{L(p_0)}{L(\hat{p})} \right) = 2 \left[l(\hat{p}) - l(p_0) \right]$$

Where:$p_0 = 0.8$ (the value under the null hypothesis).$\hat{p} \approx 0.84688$ (your calculated MLE).

$l(p)$ is the log-likelihood function we derived: $l(p) = 7\ln(p) + \ln(1-p) + 2\ln(1-p/2) - 3\ln(2)$.

```
# 1. Define p0 and retrieve your MLE
p0 <- 0.8
p_hat <- mle_p  # From your uniroot calculation (approx 0.84688)



# 2. Calculate the Likelihood Ratio R
# Using the log-likelihood function l_p already defined
log_R <- l_p(p0) - l_p(p_hat)
R <- exp(log_R)

# 3. Calculate the Deviance Statistic D (LRT statistic)
d_stat <- -2 * log(R)
```

```r
# 4. Calculate the P-value [cite: 89, 90, 96]
# area.left.tail calculates area to the left
area.left.tail <- pchisq(d_stat, df = 1)
# pvalue is the area in the right tail
pvalue <- 1 - area.left.tail

# Output results
print(paste("Likelihood Ratio (R):", round(R, 4)))
```

```
## [1] "Likelihood Ratio (R): 0.9495"
```

```r
print(paste("LRT Statistic (D):", round(d_stat, 4)))
```

```
## [1] "LRT Statistic (D): 0.1037"
```

```r
print(paste("P-value:", round(pvalue, 4)))
```

```
## [1] "P-value: 0.7475"
```

The z-statistic is calculated using the formula:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

In your specific problem:$\hat{p}$ (Sample Proportion): ≈ 0.84688 (calculated from your MLE).p_0 (Null Hypothesis): 0.8.n (Total Sample Size): 10 (5 art students + 5 business students).

```r
# 1. Define inputs based on your problem
n <- 10           # Total sample size

# 2. Calculate Z-statistic
# se utilizes p0 under the null hypothesis as shown in teacher's slides
se <- sqrt((p0 * (1 - p0)) / n)
z <- (p_hat - p0) / se

# 3. Calculate P-value (Two-tailed)
# pnorm finds the area to the left; we multiply the tail area by 2
p_value <- 2 * (1 - pnorm(abs(z)))

# Print results to match the 'RR Console' style in slides
z
```

```
## [1] 0.3708213
```

```r
p_value
```

```
## [1] 0.7107707
```

Summary of Results

The results of the Likelihood Ratio Test and the Z-Test are summarized below:Testp-value

Likelihood Ratio Test0.7475 Z-Test0.7108

The p-values obtained from the two tests are very similar, indicating a strong relation between the Likelihood Ratio Test and the Z-Test.

Since the p-value (0.7108) is much greater than the significance level (0.05), we fail to reject the null hypothesis. There is not sufficient evidence to conclude that the proportion is different from 0.8.

Problem 2 - A statistics instructor models the probability that a student passes an exam as a function of the number of hours studied, given by p=0.2x where: p is the probability that a student passes the exam, and x is the number of hours the student studies.

For example, if a student studies for 2 hours, then the probability of passing the exam is p= 0.2×2=0.4. Suppose a random sample of four students (A, B, C, and D) is selected. The number of hours each student studies is shown below. Student A B C D Hours 2 2 3 4

Assuming that students' exam outcomes are independent, calculate the probability that Students A, C, and D pass the exam, while Student B does not pass the exam.

Student	Hours (x)	Outcome	Probability
A	2	Pass	0.4
B	2	Fail	1-0.4
C	3	Pass	0.6
D	4	Pass	0.8

Student A (2 hours): $0.2 \times 2 = \mathbf{0.4}$ (Chance of passing)

Student B (2 hours): $0.2 \times 2 = \mathbf{0.4}$ (Chance of passing). BUT, the question asks for the probability that Student B fails. Since the chance of passing is 0.4, the chance of failing is $1 - 0.4 = \mathbf{0.6}$.

Student C (3 hours): $0.2 \times 3 = \mathbf{0.6}$ (Chance of passing)

Student D (4 hours): $0.2 \times 4 = \mathbf{0.8}$ (Chance of passing)

Event of interest:

$$A = P,\ B = F,\ C = P,\ D = P$$

Step 3: Use independence (multiply probabilities)

Because the students' exam outcomes are independent, the joint probability is:

$$P(A = P, B = F, C = P, D = P) = P(A = P) \cdot P(B = F) \cdot P(C = P) \cdot P(D = P)$$

Substitute the values:

$$= (0.4)(0.6)(0.6)(0.8)$$

$$\boxed{0.1152}$$

Problem 3 - A statistics instructor models the probability that a student passes an exam as a function of the number of hours studied, given by p= x where: p is the probability that a student passes the exam, and x is the number of hours the student studies.   is unknown parameter to be estimate

To estimate , a random sample of four students (A, B, C, D) is selected. The number of hours studied and the exam outcomes for these students are shown below. Student A B C D Hours Studied (x) 2 2 3 4 Exam Outcome Pass Fail Pass Pass Assume that students' exam outcomes are independent. Questions: Formulate the likelihood function L( ) based on the observed data.

If a student Passes: The probability is $p = \frac{e^{\beta x}}{1+e^{\beta x}}$.If a student Fails: The probability is $1-p$. Using the math from your notes, this simplifies to $\frac{1}{1+e^{\beta x}}$.

Apply the Data to the ModelWe plug in the hours (x) for each student:StudentHours (x)Outcome Probability Contribution A2Pass$\frac{e^{2\beta}}{1+e^{2\beta}}$

B2Fail$\frac{1}{1+e^{2\beta}}$ C3Pass$\frac{e^{3\beta}}{1+e^{3\beta}}$ D4Pass$\frac{e^{4\beta}}{1+e^{4\beta}}$

3. The Likelihood Function $L(\beta)$

Because the outcomes are independent, the likelihood is the product of all these individual probabilities:

$$L(\beta) = \left(\frac{e^{2\beta}}{1+e^{2\beta}}\right) \cdot \left(\frac{1}{1+e^{2\beta}}\right) \cdot \left(\frac{e^{3\beta}}{1+e^{3\beta}}\right) \cdot \left(\frac{e^{4\beta}}{1+e^{4\beta}}\right)$$

4. Simplified VersionTo make it look like the "teacher way," we combine the terms:Top (Numerator): $e^{2\beta} \cdot 1 \cdot e^{3\beta} \cdot e^{4\beta} = e^{(2+3+4)\beta} = e^{9\beta}$Bottom (Denominator): $(1+e^{2\beta}) \cdot (1+e^{2\beta}) \cdot (1+e^{3\beta}) \cdot (1+e^{4\beta}) = (1+e^{2\beta})^2(1+e^{3\beta})(1+e^{4\beta})$

Final Answer for (a):

$$L(\beta) = \frac{e^{9\beta}}{(1+e^{2\beta})^2(1+e^{3\beta})(1+e^{4\beta})}$$

b. Derive the log-likelihood function, l( ).

The log-likelihood is defined as $l(\beta) = \ln[L(\beta)]$. Using the laws of logarithms:

$\ln(\frac{a}{b}) = \ln(a) - \ln(b)$

$\ln(a \cdot b) = \ln(a) + \ln(b)$

$\ln(a^n) = n \cdot \ln(a)$

$\ln(e^x) = x$

3. Step-by-Step DerivationFirst, apply the log to the whole fraction:

$$l(\beta) = \ln(e^{9\beta}) - \ln[(1+e^{2\beta})^2(1+e^{3\beta})(1+e^{4\beta})]$$

Next, simplify the first term:

$$\ln(e^{9\beta}) = 9\beta$$

Then, break apart the second term (the denominator):

$$\ln[(1+e^{2\beta})^2(1+e^{3\beta})(1+e^{4\beta})] = \ln(1+e^{2\beta})^2 + \ln(1+e^{3\beta}) + \ln(1+e^{4\beta})$$

Finally, move the exponent in the first part of the denominator to the front:

$$\ln(1+e^{2\beta})^2 = 2\ln(1+e^{2\beta})$$

4. Final Log-Likelihood Function Combining everything together, we get the final expression for $l(\beta)$:

$$l(\beta) = 9\beta - 2\ln(1+e^{2\beta}) - \ln(1+e^{3\beta}) - \ln(1+e^{4\beta})$$

C. Compute the derivative of l( ) with respect to .

. Differentiating Term-by-Step

We apply the derivative $\frac{d}{d\beta}$ to each part of the equation:

First Term:

$$\frac{d}{d\beta}(9\beta) = 9$$

Second Term: Use the chain rule on $2\ln(1+e^{2\beta})$.

The derivative of the inside $(1 + e^{2\beta})$ is $2e^{2\beta}$.

$$\frac{d}{d\beta}[-2\ln(1 + e^{2\beta})] = -2 \cdot \frac{1}{1 + e^{2\beta}} \cdot (2e^{2\beta}) = -\frac{4e^{2\beta}}{1 + e^{2\beta}}$$

Third Term: Similar logic for $\ln(1 + e^{3\beta})$.

The derivative of the inside is $3e^{3\beta}$.

$$\frac{d}{d\beta}[-\ln(1 + e^{3\beta})] = -\frac{3e^{3\beta}}{1 + e^{3\beta}}$$

Fourth Term: Similar logic for $\ln(1 + e^{4\beta})$.

The derivative of the inside is $4e^{4\beta}$.

$$\frac{d}{d\beta}[-\ln(1 + e^{4\beta})] = -\frac{4e^{4\beta}}{1 + e^{4\beta}}$$

3. Final Derivative $l'(\beta)$ Combining these parts, the derivative of the log-likelihood function is:

$$l'(\beta) = 9 - \frac{4e^{2\beta}}{1 + e^{2\beta}} - \frac{3e^{3\beta}}{1 + e^{3\beta}} - \frac{4e^{4\beta}}{1 + e^{4\beta}}$$

Let ˆ be the maximum likelihood estimate of . Find ˆ.

```
# Definimos la derivada de la función de log-verosimilitud
dl_db <- function(b) {
  9 - ((4 * exp(2 * b)) / (1 + exp(2 * b))) -
      ((3 * exp(3 * b)) / (1 + exp(3 * b))) -
      ((4 * exp(4 * b)) / (1 + exp(4 * b)))
}

# uniroot busca el valor de b que hace que la derivada sea 0
mle_beta_result <- uniroot(dl_db, interval = c(0, 5))
mle_beta <- mle_beta_result$root

# Mostramos el resultado
mle_beta
```

```
## [1] 0.5221865
```

Suppose you want to test whether 0.2 State the null and alternative hypotheses. Compute the likelihood ratio test statistic. Provide the R code to calculate the p-value based on the likelihood ratio test.

I. HypothesesNull Hypothesis (H_0): $\beta = 0.2$ Alternative Hypothesis (H_a): $\beta \neq 0.2$

II. Compute the LRT Statistic (D)The formula for the test statistic is:

$$D = 2 \times [l(\hat{\beta}) - l(0.2)]$$

Using $\hat{\beta} = 0.5222$ and the $l(\beta)$ function you derived in part b.

```
# 1. Define the log-likelihood function from part (b)
l_beta <- function(b) {
  9*b - 2*log(1 + exp(2*b)) - log(1 + exp(3*b)) - log(1 + exp(4*b))
}
```

```r
# 2. Assign values
beta_hat <- 0.5222167
beta_0 <- 0.2

# 3. Calculate D and P-value
D_stat <- 2 * (l_beta(beta_hat) - l_beta(beta_0))
p_val_lrt <- 1 - pchisq(D_stat, df = 1)

cat("LRT Statistic (D):", D_stat, "\n")
```

```
## LRT Statistic (D): 0.5617109
```

```r
cat("P-value:", p_val_lrt)
```

```
## P-value: 0.4535717
```

Using a linear equation such as $p = 0.2x$ to model probability is like to model a straight line when, in fact, you are looking at a curve. The first problem with this model is when you consider a student who prepares for 20 hours.

Bonus

a. Why the Logistic Function resolves linear model problems

A linear probability model ($p = \beta x$) often results in predicted probabilities that are less than 0 or greater than 1, which are mathematically invalid. The logistic function:

$$p = \frac{e^{\beta x}}{1 + e^{\beta x}}$$

is a sigmoid (S-shaped) curve that is mathematically constrained between 0 and 1. Regardless of how many hours a student studies (x), the probability will always remain within a realistic range.

b. Probability of Failing

The probability of failing is the complement of passing $(1 - p)$:

$$P(\text{fail}) = 1 - \frac{e^{\beta x}}{1 + e^{\beta x}} = \frac{1 + e^{\beta x} - e^{\beta x}}{1 + e^{\beta x}} = \frac{1}{1 + e^{\beta x}}$$

Maximum Likelihood Estimation (MLE)

I. Formulating the Likelihood Function $L(\beta)$

Given the independent outcomes for students A, B, C, and D: * **Student A** ($x = 2$, Pass): $P = \frac{e^{2\beta}}{1+e^{2\beta}}$ * **Student B** ($x = 2$, Fail): $P = \frac{1}{1+e^{2\beta}}$ * **Student C** ($x = 3$, Pass): $P = \frac{e^{3\beta}}{1+e^{3\beta}}$ * **Student D** ($x = 4$, Pass): $P = \frac{e^{4\beta}}{1+e^{4\beta}}$

The Likelihood Function is the product:

$$L(\beta) = \left(\frac{e^{2\beta}}{1 + e^{2\beta}}\right)\left(\frac{1}{1 + e^{2\beta}}\right)\left(\frac{e^{3\beta}}{1 + e^{3\beta}}\right)\left(\frac{e^{4\beta}}{1 + e^{4\beta}}\right)$$

$$L(\beta) = \frac{e^{(2+3+4)\beta}}{(1 + e^{2\beta})^2(1 + e^{3\beta})(1 + e^{4\beta})} = \frac{e^{9\beta}}{(1 + e^{2\beta})^2(1 + e^{3\beta})(1 + e^{4\beta})}$$

II. Deriving the Log-likelihood Function $\ell(\beta)$

$$\ell(\beta) = \ln(e^{9\beta}) - \ln[(1 + e^{2\beta})^2(1 + e^{3\beta})(1 + e^{4\beta})]$$
$$\ell(\beta) = 9\beta - [2\ln(1 + e^{2\beta}) + \ln(1 + e^{3\beta}) + \ln(1 + e^{4\beta})]$$

III. Computing the Derivative of $\ell(\beta)$

Using the chain rule $\frac{d}{d\beta}\ln(1 + e^{kx}) = \frac{ke^{kx}}{1+e^{kx}}$:

$$\ell'(\beta) = 9 - \left[\frac{4e^{2\beta}}{1 + e^{2\beta}} + \frac{3e^{3\beta}}{1 + e^{3\beta}} + \frac{4e^{4\beta}}{1 + e^{4\beta}}\right]$$

IV. Finding the Maximum Likelihood Estimate $\hat{\beta}$

Since $\ell'(\beta) = 0$ does not have a simple analytical solution, we use numerical optimization in R.

```r
# Define the negative log-likelihood function
neg_log_lik <- function(beta) {
  -(9*beta - (2*log(1 + exp(2*beta)) + log(1 + exp(3*beta)) + log(1 + exp(4*beta))))
}

# Numerical optimization
mle_result <- optim(par = 0.5, fn = neg_log_lik, method = "BFGS")
hat_beta <- mle_result$par

cat("The Maximum Likelihood Estimate (MLE) for beta is:", round(hat_beta, 4))
```

```
## The Maximum Likelihood Estimate (MLE) for beta is: 0.5222
```