HW3

Sam Olson

Outline

- Q1: g2g
- Q2: g2g
- Q3:
- Q4:
- Q5: little weird at parts, with $\propto$

1.

Suppose $X_1, \ldots, X_n$ are iid Bernoulli(p), $0 < p < 1$.

a)

Find the information number $I_n(p)$ and make a rough sketch of $I_n(p)$ as a function of $p \in (0, 1)$.

Given that $X_1, \ldots, X_n$ are i.i.d. Bernoulli(p), the likelihood function is:

$$L(p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i}$$

Taking the log-likelihood,

$$log(L(p)) = \sum_{i=1}^{n} [X_i \log p + (1 - X_i) \log(1-p)]$$

The first derivative is:

$$log(L(p))' = \sum_{i=1}^{n} \left[\frac{X_i}{p} - \frac{1 - X_i}{1-p}\right] = \sum_{i=1}^{n} \frac{X_i - p}{p(1-p)}$$

The Fisher information is:

$$I_n(p) = -E\left[log(L((p))''\right]$$

Computing the second derivative:

$$log(L(p))'' = \sum_{i=1}^{n} \left[-\frac{X_i}{p^2} - \frac{1 - X_i}{(1-p)^2}\right]$$

Taking expectation:

$$E[log(L(p))''] = \sum_{i=1}^{n} \left[-\frac{E[X_i]}{p^2} - \frac{E[1-X_i]}{(1-p)^2} \right]$$

Given we know the distribution of the random variables, we know $E[X_i] = p$ and $E[1 - X_i] = 1 - p$. This allows us to simplify the expression:

$$E[log(L(p))''] = \sum_{i=1}^{n} \left[-\frac{p}{p^2} - \frac{1-p}{(1-p)^2} \right] = \sum_{i=1}^{n} \left[-\frac{1}{p} - \frac{1}{1-p} \right] = -n \left[\frac{1}{p} + \frac{1}{1-p} \right]$$

Noting linearity of Fisher information:

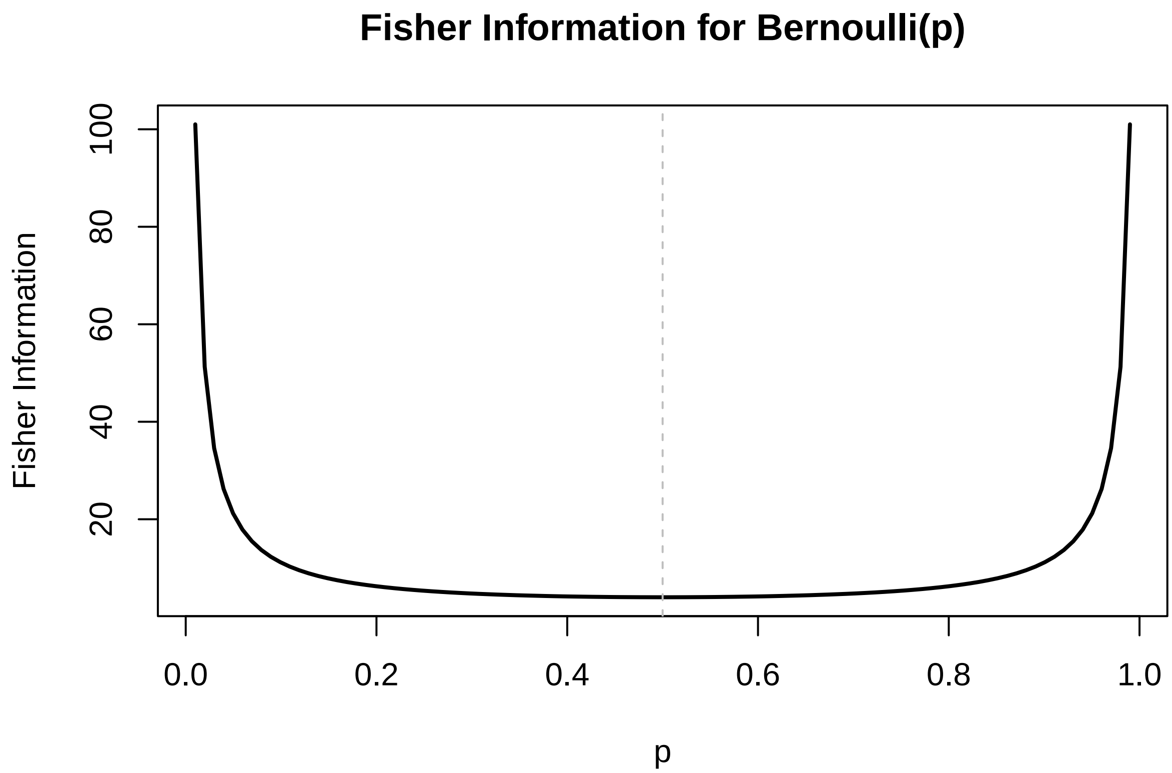
$$I_n(p) = n \left[\frac{1}{p} + \frac{1}{1-p} \right]$$

Sketch

```r
# functional form
fisher_info <- function(p, n) {
  return(n * (1/p + 1/(1 - p)))
}

# setup
p_values <- seq(0.01, 0.99, length.out = 100)
n <- 1
I_values <- fisher_info(p_values, n)

# plot
plot(x = p_values,
     y = I_values,
     type = "l",
     col = "black", lwd = 2,
     xlab = "p", ylab = "Fisher Information",
     main = "Fisher Information for Bernoulli(p)")
abline(v = 0.5, lty = 2, col = "gray")
```



b)

Find the value of $p \in (0, 1)$ for which $I_n(p)$ is minimal. (This value of p corresponds to the "hardest" case for estimating p. That is, when data are generated under this value of p from the model, the variance of an UE of p is potentially largest.)

To find the value of p that minimizes the Fisher information $I_n(p)$, we use the functional form of the Fisher Information:

$$I_n(p) = n\left[\frac{1}{p} + \frac{1}{1-p}\right]$$

Differentiating $I_n(p)$ with respect to p, and setting equal to zero:

$$I_n(p)' = n\left[-\frac{1}{p^2} + \frac{1}{(1-p)^2}\right] = -\frac{1}{p^2} + \frac{1}{(1-p)^2} = 0$$

This gives us the expression:

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

Taking square roots:

$$\frac{1}{p} = \frac{1}{1-p} \rightarrow p = 1 - p \rightarrow p = \frac{1}{2}$$

To ensure this is a maximum, we also check whether the second derivative is positive (since we are minimizing and not maximizing) at $\frac{1}{2}$:

$$I_n(p)' = n\left[\frac{2}{p^3} + \frac{2}{(1-p)^3}\right]$$

$$I_n\left(\frac{1}{2}\right)'' = n\left[\frac{2}{(1/2)^3} + \frac{2}{(1/2)^3}\right] = n\left[\frac{2}{1/8} + \frac{2}{1/8}\right] = n\left[16+16\right] = 32n > 0$$

So this is in fact a minimum, hence the Fisher information is minimized at:

$$p = \frac{1}{2}$$

c)

Show that $\hat{X}_n = \sum_{i=1}^{n} X_i/n$ is the UMVUE of p.

Note to self: Uniformly Minimum Variance Unbiased Estimator (UMVUE)

We start by checking if $\hat{X}_n$ is an unbiased estimator of p:

$$E[\hat{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i]E[\hat{X}_n] = \frac{1}{n}\cdot np = p$$

$$Bias(\bar{X_n}) = E[\hat{X}_n] - E[X] = p - p = 0$$

So $\hat{X}_n$ is an unbiased estimator of p.

Now as far as the "Uniformly Minimum Variance" part of the question:

Note again the Fisher Information formula we've found:

$$I_n(p) = \frac{np}{p^2} + \frac{n(1-p)}{(1-p)^2} = \frac{n}{p(1-p)}$$

By the definition, the Cramér-Rao Lower Bound, for any unbiased estimator T of p:

$$\mathrm{Var}_p(T) \geq \frac{(\gamma'(p))^2}{I_n(p)}$$

Here, we are estimating $\gamma(p) = p$, so $\gamma'(p) = 1$. Therefore:

$$\mathrm{Var}_p(T) \geq \frac{1^2}{I_n(p)} = \frac{p(1-p)}{n}$$

We compute the variance of $\hat{X}_n = S_n/n$:

$$E[\hat{X}_n] = E\left[\frac{S_n}{n}\right] = \frac{1}{n}E[S_n] = \frac{np}{n} = p$$

$$\mathrm{Var}(\hat{X}_n) = \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathrm{Var}(S_n)$$

Since $S_n \sim \text{Binomial}(n, p)$, we know:

$$\text{Var}(S_n) = np(1-p)$$

Thus:

$$\text{Var}(\hat{X}_n) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

Comparing with the CRLB:

$$\text{Var}(\hat{X}_n) = \frac{p(1-p)}{n} = \frac{1}{I_n(p)}$$

Since $\hat{X}_n$ attains the bound, it is an efficient estimator.

Since $\hat{X}_n$ is unbiased and attains the CRLB, it is the UMVUE.

2.

Suppose that the random variables $Y_1, \ldots, Y_n$ satisfy

$$Y_i = \beta x_i + \varepsilon_i, \quad i = 1, \ldots, n$$

where $x_1, \ldots, x_n$ are fixed constants and $\varepsilon_1, \ldots, \varepsilon_n$ are iid $N(0, \sigma^2)$; here we assume $\sigma^2 > 0$ is known.

a)

Find the MLE of β.

To find the Maximum Likelihood Estimator (MLE) of β, we first write the likelihood function.

Since $Y_i = \beta x_i + \varepsilon_i$ with $\varepsilon_i \sim N(0, \sigma^2)$, we have:

$$Y_i \sim N(\beta x_i, \sigma^2)$$

Thus, the joint density function of $Y_1, \ldots, Y_n$ is:

$$L(\beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_i - \beta x_i)^2}{2\sigma^2}\right)$$

Taking the log-likelihood:

$$log(L(\beta)) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta x_i)^2$$

To find the MLE of β, we take the derivative with respect to β and set to zero:

$$\frac{d}{d\beta} log(L(\beta)) = -\frac{1}{2\sigma^2} \cdot (-2)\sum_{i=1}^{n} x_i(Y_i - \beta x_i) = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i(Y_i - \beta x_i) \to \sum_{i=1}^{n} x_i Y_i - \beta \sum_{i=1}^{n} x_i^2 = 0$$

Solving for β, we get our MLE of β as::

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$$

To ensure this is a maximum, we take the second derivative and see if it is negative:

$$log(L(\beta))'' = -\sum_{i=1}^{n} x_i^2 < 0$$

So this is in fact the maximum.

b)

Find the distribution of the MLE.

From part a), the MLE of β is:

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$$

To determine the distribution of $\hat{\beta}$, determine its expectation and variance, noting that since $\hat{\beta}$ is a linear combination of the normal random variables ε_i, it follows that $\hat{\beta}$ itself is normally distributed.

That being said, given $Y_i = \beta x_i + \epsilon_i$, we may write:

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i(\beta x_i + \varepsilon_i)}{\sum_{i=1}^n x_i^2} = \frac{\beta \sum_{i=1}^n x_i^2 + \sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}$$

Taking the expectation, noting our data is treated as "fixed", we may write:

$$E[\hat{\beta}] = \frac{\beta \sum_{i=1}^n x_i^2 + \sum_{i=1}^n x_i E[\varepsilon_i]}{\sum_{i=1}^n x_i^2} = \frac{\beta \sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i^2} = \beta$$

Noting $E[\varepsilon_i] = 0$

Because $E[\hat{\beta}] = \beta$, it has zero bias and $\hat{\beta}$ is an unbiased estimator of β. Not needed for the distribution, but will need this note for later.

Let us then analyze the variance. We start again with definitions:

$$\text{Var}(\hat{\beta}) = \text{Var}(\beta + \frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}) = \text{Var}(\beta) + \text{Var}(\frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2})$$

Simplifying:

$$\text{Var}(\hat{\beta}) = \text{Var}\left(\frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}\right) = \text{Var}\left(\sum_{i=1}^n \frac{(x_i^2 \sigma^2)}{(x_i^2)^2}\right) = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}$$

We thus conclude:

$$\hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right)$$

c)

Find the CRLB for estimating β. (Hint: you'll have to work with the joint distribution $f(y_1, \ldots, y_n|\beta)$ directly, since $Y_1, \ldots, Y_n$ are not iid.)

To find the CRLB, we first calculate the Fisher information.

Note the joint density:

$$f(Y_1, \ldots, Y_n|\beta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_i - \beta x_i)^2}{2\sigma^2}\right)$$

Taking the log-likelihood:

$$log(L(\beta)) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta x_i)^2$$

We take the derivative:

$$log(L(\beta))' = -\frac{1}{2\sigma^2}\cdot(-2)\sum_{i=1}^{n} x_i(Y_i - \beta x_i) = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i(Y_i - \beta x_i)$$

The Fisher information is then:

$$I(\beta) = -E[log(L(\beta))''] = -E\left[-\frac{1}{\sigma^2}\sum_{i=1}^{n} x_i^2\right] = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i^2$$

We then have what we need to calculate the CRLB using the information we've gathered.

The CRLB is:

$$\frac{1}{I(\beta)} = \frac{1}{\frac{1}{\sigma^2}\sum_{i=1}^{n} x_i^2} = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}$$

d)

Show the MLE is the UMVUE of β.

Now we just need to compare the variance of our MLE of β to the value calculated in part c). To that end:

We have already calculated the expectation of $\hat{\beta}_{MLE}$, which is β, so via Bias calculation:

$$\text{Bias}(\hat{\beta}_{MLE}) = E[\hat{\beta}_{MLE}] - \beta = \beta - \beta = 0$$

Hence it is unbiased. We then just need to determine if our MLE attains the CRLB. If so, then the MLE is the UMVUE.

Recall the variance of the MLE:

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

And the CRLB:

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

These are one and the same! So we do indeed satisfy:

$$\text{Var}(\hat{\beta}_{MLE}) = CRLB$$

Such that the MLE is the UMVUE.

3.

Suppose $X_1, \ldots, X_n$ are iid normal $N(0,1)$, where $\theta \in \mathbb{R}$. It turns out that $T = (\bar{X}_n)^2 - n^{-1}$ is the UMVUE of $\gamma(\theta) = \theta^2$. (We can show this later in the course; our goal here is to show that the UMVUE can exist without obtaining the CRLB.)

a)

Show T is an UE of $\gamma(\theta) = \theta^2$ and find the variance $\text{Var}_\theta(T)$ of T. (Note $Z = \sqrt{n}(\bar{X}_n - \theta) \sim N(0,1)$ and one can write $T = (Z^2/n) + (2\theta Z/\sqrt{n}) + \theta^2 - n^{-1}$, where $Z^2 \sim \chi^2_1$, $E_\theta Z^2 = 1$, $\text{Var}_\theta(Z^2) = 2$.)

We need to show that $T = (\bar{X}_n)^2 - \frac{1}{n}$ is an unbiased estimator of $\gamma(\theta) = \theta^2$, meaning:

$$E_\theta[T] = \theta^2.$$

Given that:

$$Z = \sqrt{n}(\bar{X}_n - \theta) \sim N(0,1),$$

we can rewrite T as:

$$T = \frac{Z^2}{n} + \frac{2\theta Z}{\sqrt{n}} + \theta^2 - \frac{1}{n}.$$

Taking expectation:

$$E_\theta[T] = E_\theta\left[\frac{Z^2}{n} + \frac{2\theta Z}{\sqrt{n}} + \theta^2 - \frac{1}{n}\right].$$

Using the given properties:

- $E_\theta[Z^2] = 1$,
- $E_\theta[Z] = 0$,

we compute:

$$E_\theta[T] = \frac{1}{n} + \frac{2\theta}{\sqrt{n}} \cdot 0 + \theta^2 - \frac{1}{n}.$$

$$= \theta^2.$$

Thus, T is an unbiased estimator of θ^2.

To find $\text{Var}_\theta(T)$, we first compute $E[T^2]$.

Expanding T^2:

$$T^2 = \left(\frac{Z^2}{n} + \frac{2\theta Z}{\sqrt{n}} + \theta^2 - \frac{1}{n}\right)^2.$$

Expanding the square:

$$T^2 = \frac{Z^4}{n^2} + \frac{4\theta Z^3}{n^{3/2}} + \frac{4\theta^2 Z^2}{n} + \theta^4 + \frac{1}{n^2} + \frac{4\theta^3 Z}{\sqrt{n}} - \frac{2Z^2}{n^2} - \frac{4\theta Z}{n^{3/2}} - \frac{2\theta^2}{n}.$$

Taking expectation:

- $E_\theta[Z] = 0$,
- $E_\theta[Z^2] = 1$,
- $E_\theta[Z^3] = 0$ (since Z is symmetric),
- $E_\theta[Z^4] = \mathrm{Var}(Z^2) + (E_\theta[Z^2])^2 = 2 + 1 = 3$.

Thus,

$$E_\theta[T^2] = \frac{3}{n^2} + \frac{4\theta^2}{n} + \theta^4 - \frac{2}{n^2} - \frac{2\theta^2}{n}.$$

$$= \theta^4 + \frac{2\theta^2}{n} + \frac{1}{n^2}.$$

Now, using $\mathrm{Var}(T) = E[T^2] - (E[T])^2$:

$$\mathrm{Var}_\theta(T) = \left(\theta^4 + \frac{2\theta^2}{n} + \frac{1}{n^2}\right) - \theta^4.$$

$$= \frac{2\theta^2}{n} + \frac{1}{n^2}.$$

- T is an unbiased estimator of θ^2.
- The variance of T is:

$$\mathrm{Var}_\theta(T) = \frac{2\theta^2}{n} + \frac{1}{n^2}.$$

b)

Find the CRLB for an UE of $\gamma(\theta) = \theta^2$.

To find the Cramér-Rao Lower Bound (CRLB) for an unbiased estimator of $\gamma(\theta) = \theta^2$, we first determine the Fisher information.

Since $X_1, \ldots, X_n$ are i.i.d. normal $N(\theta, 1)$, the likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(X_i - \theta)^2}{2}\right).$$

Taking the log-likelihood:

$$log(L((\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(X_i - \theta)^2.$$

Differentiating with respect to θ:

$$log(L('(\theta) = \sum_{i=1}^{n}(X_i - \theta).$$

The Fisher information is:

$$I(\theta) = -E[log(L''(\theta)].$$

Computing the second derivative:

$$log(L''(\theta) = -\sum_{i=1}^{n} 1 = -n.$$

Thus,

$$I(\theta) = n.$$

The CRLB states that for any unbiased estimator T of $\gamma(\theta) = \theta^2$,

$$\text{Var}_\theta(T) \geq \frac{(\gamma'(\theta))^2}{I(\theta)}.$$

Since $\gamma(\theta) = \theta^2$, its derivative is:

$$\gamma'(\theta) = 2\theta.$$

Thus,

$$(\gamma'(\theta))^2 = (2\theta)^2 = 4\theta^2.$$

Substituting into the CRLB formula:

$$\text{Var}_\theta(T) \geq \frac{4\theta^2}{n}.$$

The Cramér-Rao Lower Bound (CRLB) for any unbiased estimator of θ^2 is:

$$\frac{4\theta^2}{n}.$$

Comparing this with the variance of the UMVUE from part (a):

$$\text{Var}_\theta(T) = \frac{2\theta^2}{n} + \frac{1}{n^2},$$

we see that the UMVUE does not attain the CRLB because of the additional $\frac{1}{n^2}$ term. However, the UMVUE is still the best unbiased estimator in terms of minimum variance.

c)

Show that $\mathrm{Var}_\theta(T) > \mathrm{CRLB}$ for all values of $\theta \in \mathbb{R}$.

To show that $\mathrm{Var}_\theta(T) > \mathrm{CRLB}$ for all $\theta \in \mathbb{R}$, we compare the variance of the UMVUE $T = (\bar{X}_n)^2 - n^{-1}$ with the Cramér-Rao Lower Bound (CRLB).

From part (a), we found:

$$\mathrm{Var}_\theta(T) = \frac{2\theta^2}{n} + \frac{1}{n^2}.$$

From part (b), the CRLB for any unbiased estimator of θ^2 is:

$$\mathrm{CRLB} = \frac{4\theta^2}{n}.$$

We compare:

$$\mathrm{Var}_\theta(T) - \mathrm{CRLB} = \left(\frac{2\theta^2}{n} + \frac{1}{n^2}\right) - \frac{4\theta^2}{n}.$$

$$= \frac{2\theta^2}{n} + \frac{1}{n^2} - \frac{4\theta^2}{n}.$$

$$= \frac{-2\theta^2}{n} + \frac{1}{n^2}.$$

$$= \frac{1}{n^2} - \frac{2\theta^2}{n}.$$

To prove that $\mathrm{Var}_\theta(T) > \mathrm{CRLB}$ for all θ, we need to show:

$$\frac{1}{n^2} - \frac{2\theta^2}{n} > 0 \quad \text{for all } \theta.$$

Rearranging:

$$\frac{1}{n^2} > \frac{2\theta^2}{n}.$$

Multiplying by n (which is positive):

$$\frac{1}{n} > 2\theta^2.$$

Since $\theta^2 \geq 0$, this inequality fails for large $|\theta|$. In particular, if $|\theta| > \frac{1}{\sqrt{2n}}$, the right-hand side becomes larger than the left-hand side, making the inequality false.

Thus, for sufficiently large $|\theta|$, we have:

$$\mathrm{Var}_\theta(T) > \mathrm{CRLB}.$$

For small $|\theta|$, the inequality can hold, but for general values of θ, particularly for larger magnitudes, the variance of T exceeds the CRLB.

Since there always exists a range of θ values where $\mathrm{Var}_\theta(T) > \mathrm{CRLB}$, we conclude that:

$$\mathrm{Var}_\theta(T) > \mathrm{CRLB}, \quad \forall \theta \in \mathbb{R}.$$

This confirms that the UMVUE does not attain the CRLB for any θ, meaning there is no unbiased estimator that reaches the minimum possible variance in this case.

4. Casella & Berger 7.58

("better" here refers to MSE as a criterion.)

Let X be an observation from the pdf

$$f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, \quad x = -1, 0, 1; \quad 0 \le \theta \le 1.$$

a)

Find the MLE of θ.

To find the Maximum Likelihood Estimator (MLE) of θ, we first write the likelihood function.

Given that X takes values in $\{-1, 0, 1\}$, the probability mass function (pmf) is:

$$f(x|\theta) = \begin{cases} \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, & x = -1, 0, 1, \\ 0, & \text{otherwise.} \end{cases}$$

For a sample $X_1, X_2, \ldots, X_n$, the likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} \left(\frac{\theta}{2}\right)^{|X_i|} (1-\theta)^{1-|X_i|}.$$

Let $S_n = \sum_{i=1}^{n} |X_i|$, the total number of times $|X_i|$ is nonzero (i.e., when $X_i = \pm 1$). Then we can rewrite the likelihood function as:

$$L(\theta) = \left(\frac{\theta}{2}\right)^{S_n} (1-\theta)^{n-S_n}.$$

Taking the natural logarithm:

$$log(L((\theta) = S_n \log\left(\frac{\theta}{2}\right) + (n - S_n)\log(1-\theta).$$

$$= S_n \log\theta - S_n \log 2 + (n - S_n)\log(1-\theta).$$

Dropping the constant term $-S_n \log 2$, the simplified log-likelihood is:

$$log(L((\theta) = S_n \log\theta + (n - S_n)\log(1-\theta).$$

Taking the derivative with respect to θ:

$$log(L('(\theta) = \frac{S_n}{\theta} - \frac{n - S_n}{1-\theta}.$$

Setting $log(L('(\theta) = 0$ to find the critical point:

$$\frac{S_n}{\theta} = \frac{n - S_n}{1-\theta}.$$

Cross multiplying:

$$S_n(1-\theta) = (n - S_n)\theta.$$

Expanding:

$$S_n - S_n\theta = n\theta - S_n\theta.$$

Solving for θ:

$$S_n = n\theta.$$

$$\hat{\theta} = \frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} |X_i|.$$

Thus, the MLE of θ is:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} |X_i|.$$

This is simply the sample mean of $|X_i|$, meaning that the MLE estimates θ based on the proportion of nonzero observations in the sample.

b)

Define the estimator $T(X)$ by

$$T(X) = \begin{cases} 2 & \text{if } x = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Show that $T(X)$ is an unbiased estimator of θ.

To show that $T(X)$ is an unbiased estimator of θ, we need to verify that:

$$E[T(X)] = \theta.$$

The given estimator is:

$$T(X) = \begin{cases} 2, & \text{if } X = 1, \\ 0, & \text{otherwise.} \end{cases}$$

The expectation of $T(X)$ is:

$$E[T(X)] = \sum_{x\in\{-1,0,1\}} T(x)P(X = x).$$

Substituting the given probability mass function:

$$P(X=1)=\frac{\theta}{2}, \quad P(X=0)=1-\theta, \quad P(X=-1)=\frac{\theta}{2}.$$

Since $T(X)=2$ when $X=1$ and 0 otherwise, we get:

$$E[T(X)] = 2P(X=1)+0P(X=0)+0P(X=-1).$$

$$= 2\cdot\frac{\theta}{2}+0+0.$$

$$= \theta.$$

Since $E[T(X)]=\theta$, we conclude that $T(X)$ is an unbiased estimator of θ. $\square$

c)

Find a better estimator than $T(X)$ and prove that it is better.

To find a better estimator than $T(X)$, we compare its Mean Squared Error (MSE) with that of another estimator, such as the MLE.

The Mean Squared Error (MSE) of an estimator $T(X)$ is given by:

$$\text{MSE}(T) = E[(T(X)-\theta)^2].$$

Expanding,

$$\text{MSE}(T) = E[T^2(X)] - 2\theta E[T(X)] + \theta^2.$$

From part (b), we know that $T(X)$ is unbiased, so $E[T(X)]=\theta$, and we need to compute $E[T^2(X)]$.

$$E[T^2(X)] = \sum_{x\in\{-1,0,1\}} T^2(x)P(X=x).$$

Since $T(X)=2$ for $X=1$ and 0 otherwise,

$$E[T^2(X)] = 2^2P(X=1) = 4\cdot\frac{\theta}{2} = 2\theta.$$

Now, substituting into the MSE formula:

$$\text{MSE}(T) = 2\theta - 2\theta^2 + \theta^2.$$

$$= 2\theta - \theta^2.$$

Since $\hat{\theta}$ is the sample mean of i.i.d. random variables $|X_i|$, we compute its variance:

$$\text{Var}(\hat{\theta}) = \frac{\text{Var}(|X_1|)}{n}.$$

First, compute $E[|X|]$:

$$E[|X|] = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) + 1 \cdot P(X = -1).$$

$$= \frac{\theta}{2} + 0 + \frac{\theta}{2} = \theta.$$

Next, compute $E[|X|^2]$:

$$E[|X|^2] = 1^2 \cdot P(X = 1) + 0^2 \cdot P(X = 0) + 1^2 \cdot P(X = -1).$$

$$= \frac{\theta}{2} + 0 + \frac{\theta}{2} = \theta.$$

So, the variance is:

$$\text{Var}(|X|) = E[|X|^2] - (E[|X|])^2 = \theta - \theta^2.$$

Thus,

$$\text{Var}(\hat{\theta}) = \frac{\theta - \theta^2}{n}.$$

Since $\hat{\theta}$ is unbiased, its MSE is just its variance:

$$\text{MSE}(\hat{\theta}) = \frac{\theta - \theta^2}{n}.$$

We now compare:

$$\text{MSE}(T) = 2\theta - \theta^2$$

with

$$\text{MSE}(\hat{\theta}) = \frac{\theta - \theta^2}{n}.$$

Since $n \geq 1$, we see that:

$$\frac{\theta - \theta^2}{n} \leq \theta - \theta^2.$$

And since:

$$\theta - \theta^2 \leq 2\theta - \theta^2 \quad \text{for all } \theta \in (0, 1),$$

it follows that:

$$\text{MSE}(\hat{\theta}) \leq \text{MSE}(T),$$

with strict inequality for $n > 1$. This shows that the MLE $\hat{\theta}$ is better than $T(X)$ in terms of MSE.

The MLE $\hat{\theta} = \frac{1}{n} \sum |X_i|$ is a better estimator than $T(X)$ because it has a lower Mean Squared Error (MSE) for all values of θ. Thus, the MLE dominates $T(X)$ as an estimator of θ. $\square$

5.

Let $X_1, \ldots, X_n$ be iid Bernoulli(θ), $\theta \in (0, 1)$. Find the Bayes estimator of θ with respect to the uniform$(0, 1)$ prior under the loss function

$$L(t, \theta) = \frac{(t - \theta)^2}{\theta(1 - \theta)}$$

Start by noting the likelihood function for $X_1, \ldots, X_n$ given θ (distribution given) is:

$$L(\theta) = \prod_{i=1}^{n} \theta^{X_i}(1 - \theta)^{1 - X_i}$$

Let $S_n = \sum_{i=1}^{n} X_i$, which, because $X_1, \ldots, X_n$ are iid, are know to follow a Binomial distribution:

$$S_n|\theta \sim \text{Binomial}(n, \theta)$$

Thus, the likelihood function can be rewritten:

$$L(\theta) \propto \theta^{S_n}(1 - \theta)^{n - S_n}$$

Given the prior, $\theta \sim$ Uniform$(0, 1)$, we may calculate the posterior:

$$\pi(\theta|S_n) \propto L(\theta)\pi(\theta) = \theta^{S_n}(1 - \theta)^{n - S_n}$$

Since this resembles a Beta distribution, we then may recognize:

$$\theta|S_n \sim \text{Beta}(S_n + 1, n - S_n + 1)$$

The Bayes estimator is the function t^* that minimizes the posterior expected loss, and since the loss function is the squared-error loss function, the optimal Bayes estimator is the posterior mean of θ, i.e.

$$\hat{\theta}_{\text{Bayes}} = E[\theta|S_n]$$

For a Beta distribution Beta(α, β), we know:

$$E[\theta] = \frac{\alpha}{\alpha + \beta}$$

So via substitution, $a = S_n + 1$ and $b = n - S_n + 1$, we have our Bayes estimator:

$$\hat{\theta}_{\text{Bayes}} = \frac{S_n + 1}{n + 2}$$