Math 3070, Applied Statistics

Section 1

October 23, 2019

Lecture Outline, 10/23

Section 6.1

- Estimators

Parameters and Estimators

A **parameter** is a constant describing a distribution:

- In a normal distribution, the mean μ and variance σ^2 are parameters.
- In an exponential distribution, the rate λ is a parameter.
- We will often use the symbol θ to represent a parameter generically.

An **estimator** is a random variable which is used to estimate a parameter.

- Given a random sample $X_1, \ldots, X_n$, the sample mean $\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$ is an estimator for the mean μ.
- The sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is an estimator for the variance σ^2.
- We will often use the symbol $\hat{\theta}$ to represent an estimator generically.

Unbiased Estimators

An estimator $\hat{\theta}$ is an **unbiased** estimator for θ if $E(\hat{\theta}) = \theta$.

- If X is a binomial random variable, then the *sample proportion* $\hat{p} = \frac{X}{n}$ is an unbiased estimator of p:

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n} \cdot np = p$$

- If $X_1, \ldots, X_n$ is a random sample from a distribution with mean μ, the sample mean $\overline{X}$ is an unbiased estimator for μ:

$$\begin{aligned} E(\overline{X}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} \cdot n\mu = \mu \end{aligned}$$

Example: Uniform Distribution

Given a random sample $X_1, \dots, X_n$ from a uniform distribution on $[0, \theta]$, how do we estimate θ?

- The mean μ is the midpoint of the interval, $\mu = \theta/2$. Therefore, $\theta = 2\mu$. Since we can estimate μ with $\overline{X}$, we can estimate θ with $\hat{\theta} = 2\overline{X}$.
- Each of the observations $X_1, \dots, X_n$ will be less than θ, and if n is large we expect one of them to be close to θ. So we may estimate θ using the maximum: $\hat{\theta} = \max\{X_1, \dots, X_n\}$.

Example: Uniform Distribution

We gave two possible estimators for the parameter θ of a uniform distribution on $[0, \theta]$:

$$\hat{\theta}_1 = 2\overline{X}$$
$$\hat{\theta}_2 = \max\{X_1, \ldots, X_n\}$$

Question: Are these estimators unbiased?

We may calculate the expected value of $\hat{\theta}_1$:

$$E(\hat{\theta}_1) = E(2\overline{X}) = 2E(\overline{X}) = 2\mu = \theta$$

Since $E(\hat{\theta}_1) = \theta$, this means that $\hat{\theta}_1$ is an unbiased estimator for θ.

Example: Uniform Distribution

Since $\hat{\theta}_2 = \max\{X_1, \ldots, X_n\}$ is always less than θ, intuition suggests that $\hat{\theta}_2$ will underestimate θ and hence must be biased.

Now we will calculate $E(\hat{\theta}_2)$. For $0 \le x \le \theta$, the cdf $F(x)$ of $\hat{\theta}_2$ is

$$\begin{aligned} F(x) &= P(\hat{\theta}_2 \le x) = P(\max\{X_1, \ldots, X_n\} \le x) \\ &= P(X_1 \le x, X_2 \le x, \ldots, X_n \le x) \\ &= P(X_1 \le x)P(X_2 \le x) \cdots P(X_n \le x) = (x/\theta)^n \end{aligned}$$

Therefore the pdf of $\hat{\theta}_2$ is, for $0 \le x \le \theta$,

$$f(x) = F'(x) = \frac{d}{dx}(x/\theta)^n = nx^{n-1}/\theta^n$$

So the expected value of $\hat{\theta}_2$ is

$$E(\hat{\theta}_2) = \int_{-\infty}^{\infty} xf(x)\ dx = \int_0^{\theta} x \cdot \frac{nx^{n-1}}{\theta^n}\ dx = \frac{n}{n+1}\theta$$

Example: Uniform Distribution

We considered two estimators for the parameter θ of a uniform distribution on $[0, \theta]$,

$$\hat{\theta}_1 = 2\overline{X}$$
$$\hat{\theta}_2 = \max\{X_1, \ldots, X_n\}$$

We found that $\hat{\theta}_1$ was unbiased but that $\hat{\theta}_2$ was biased. However, it is easy to modify $\hat{\theta}_2$ to produce an unbiased estimator $\hat{\theta}_3$:

$$\hat{\theta}_3 = \frac{n+1}{n}\hat{\theta}_2 = \frac{n+1}{n}\max\{X_1, \ldots, X_n\}$$

Since $E(\hat{\theta}_2) = \frac{n}{n+1}\theta$, it follows that

$$E(\hat{\theta}_3) = E\left(\frac{n+1}{n}\hat{\theta}_2\right) = \frac{n+1}{n}E(\hat{\theta}_2) = \frac{n+1}{n} \cdot \frac{n}{n+1}\theta = \theta$$

so $\hat{\theta}_3$ is in fact an unbiased estimator for θ.

Example: Uniform Distribution

We now have two unbiased estimators for the parameter θ of a uniform distribution on $[0, \theta]$:

$$\hat{\theta}_1 = 2\overline{X}$$

$$\hat{\theta}_3 = \frac{n+1}{n} \max\{X_1, \ldots, X_n\}$$

Question: Which of these estimators is better?

To answer this, we need a measure of how good an estimator is. One commonly used such measure is the *variance* of the estimator.

We can calculate the variance of $\hat{\theta}_1$:

$$V(\hat{\theta}_1) = V(2\overline{X}) = 4V(\overline{X}) = \frac{4}{n} V(X_1) = \frac{4}{n} \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

Example: Uniform Distribution

Recall the pdf of $\hat{\theta}_2$ is $f(x) = nx^{n-1}/\theta^n$. Therefore,

$$E(\hat{\theta}_2^2) = \int_0^\infty x^2 f(x)\ dx = \int_0^\infty nx^{n+1}/\theta^n\ dx = \frac{n}{n+2}\theta^2$$

$$V(\hat{\theta}_2) = E(\hat{\theta}_2^2) - [E(\hat{\theta}_2)]^2 = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2$$

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

$$V(\hat{\theta}_3) = V\left(\frac{n+1}{n}\hat{\theta}_2\right) = \left(\frac{n+1}{n}\right)^2 V(\hat{\theta}_2)$$

$$= \left(\frac{n+1}{n}\right)^2 \cdot \frac{n}{(n+1)^2(n+2)}\theta^2 = \frac{\theta^2}{n(n+2)}$$

Example: Uniform Distribution

We considered two unbiased estimators for the parameter θ of a uniform distribution on $[0, \theta]$:

$$\hat{\theta}_1 = 2\overline{X}$$
$$\hat{\theta}_3 = \frac{n+1}{n} \max\{X_1, \ldots, X_n\}$$

We calculated that their variances were $V(\hat{\theta}_1) = \frac{\theta^2}{3n}$ and $V(\hat{\theta}_3) = \frac{\theta^2}{n(n+2)}$. Therefore, for $n > 1$, $\hat{\theta}_3$ has a smaller variance.

More advanced statistical theory can be used to show that in fact $\hat{\theta}_3$ is a *minimum variance unbiased estimator*: it has a smaller variance than any other unbiased estimator.

Given a distribution with variance σ^2, the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ is an unbiased estimator for σ^2.

$$
\begin{aligned}
E(S^2) &= E\left[\frac{1}{n-1}\left(\sum_{i=1}^n X_i^2 - n\overline{X}^2\right)\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^n E(X_i^2) - nE(\overline{X}^2)\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^n (V(X_i) + [E(X_i)]^2) - n(V(\overline{X}) + [E(\overline{X})]^2)\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^n (\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2)\right] = \sigma^2
\end{aligned}
$$

Minimum of Exponential Random Variables

Suppose $X_1, \ldots, X_n$ are iid exponential random variables with mean μ. Find an unbiased estimator for μ based on $\min\{X_1, \ldots, X_n\}$.

Solution: First we need to identify the distribution of $T = \min\{X_1, \ldots, X_n\}$. Letting $\lambda = \frac{1}{\mu}$, the cdf of T is

$$\begin{aligned} F(t) = P(T \leq t) &= 1 - P(T > t) \\ &= 1 - P(\min\{X_1, \ldots, X_n\} > t) \\ &= 1 - P(X_1 > t, X_2 > t, \ldots, X_n > t) \\ &= 1 - (e^{-\lambda t})^n = 1 - e^{-n\lambda t} \end{aligned}$$

This is the cdf of an exponential random variable with rate $n\lambda$. Now

$$E(T) = \frac{1}{n\lambda} = \frac{\mu}{n}$$

So $\hat{\mu} = nT = n\min\{X_1, \ldots, X_n\}$ is an unbiased estimator of μ.

Minimum of Exponential Random Variables

Given a random sample $X_1, \ldots, X_n$ from an exponential distribution with mean μ, we found an unbiased estimator for μ:

$$\hat{\mu} = nT = n \min\{X_1, \ldots, X_n\}$$

What is the variance of this estimator?

Recalling that T is exponential with rate $n\lambda$, we calculate

$$V(nT) = n^2 V(T) = n^2 \cdot \frac{1}{(n\lambda)^2} = \frac{1}{\lambda^2} = \mu^2$$

We note that as the sample size n increases, the variance does not decrease but remains a constant μ^2. This suggests that $\hat{\mu}$ is a poor estimator of μ. Compare this to the sample mean:

$$V(\overline{X}) = \frac{V(X_1)}{n} = \frac{1/\lambda^2}{n} = \frac{\mu^2}{n}$$