Back to Inference

Topics:

1.) Estimation of parameters
2.) Confidence Intervals
3.) Hypothesis Testing
4.) Prediction

Estimation

Start with $X_1, \cdots, X_n \overset{iid}{\sim} f_X(x)$, where there is some parameter, say θ, associated with f_X. In Statistics, θ is unknown to us, so we must Estimate it from the data.

Def:

A statistic, $T(X_1, \cdots, X_n)$, that is used to learn about an unknown parameter θ is called an Estimator.

Notes: 1) The term Estimator is used when referring to a Statistic as a function of random variables

2) Typical notation is to put a "hat" over the parameter being estimated to denote an Estimator.

— $\hat{\theta}$ is an Estimator of θ

Def:

The observed value of a statistic used to learn about an unknown parameter is called an Estimate.

Notes: 1) An <u>Estimate</u> is a function of the Observed data values $x_1, \cdots x_n$

2) It is an Actual numeric value

<u>Ex</u>

I have a sample $X_1, \ldots, X_n \sim f_X(x)$ and want ~~to~~ an estimator for $\mathbb{E}(X) = \mu$. my <u>Estimator</u> will be $\hat{\mu} = \overline{X}$. I observe the Values 6, 7, 7, 8, 9, 10. my <u>estimate</u> of μ is

$$\overline{x} = \frac{6+7+7+8+9+10}{6} = 7.83$$

Since an <u>Estimator</u>, $\hat{\theta}$, is a function of R.V.'s, it is a R.V. also. Thus it has its own distribution called the <u>Sampling Distribution of $\hat{\theta}$.</u>

- $\mathbb{E}(\hat{\theta})$ is the mean of the sampling distribution
- The Standard Deviation of the sampling distribution is called the "Standard Error" and is denoted as $se(\hat{\theta})$
  - $se(\hat{\theta}) = \sqrt{Var(\hat{\theta})}$

We make use of the sampling distribution in confidence intervals and hypothesis testing

---

A natural question now is:

IS my Estimator any good?

There are some properties we can look at.

- unbiasedness
- consistency
- mean Square Error (MSE)
- Asymptotic Normality

in many cases, $\hat{\theta} \approx N(\theta, [se(\hat{\theta})]^2)$

Def

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

[on Average, we hit the target]

Def

An estimator, $\hat{\theta}$, is a Consistent estimator for θ if $\mathbb{P}(|\hat{\theta} - \theta| > \epsilon) \longrightarrow 0$ as $n \to \infty$

[as the sample size gets bigger, there is a high probability $\hat{\theta}$ will be close to θ]

Note: unbiasedness use to be a very desirable property. still is important. Consistency is a good thing to have in an Estimator

---

Earlier in Notes, we said that we should use

$\overline{X}$ as an Estimator for $\mathbb{E}(X) = \mu$ and

S^2 as an Estimator for $Var(X) = \sigma^2$

Thm

$\overline{X}$ and S^2 are both unbiased and consistent Estimators for their Respective Parameters

$\overline{X}$:

$$\mathbb{E}(\overline{X}) = \mathbb{E}\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n}\mathbb{E}\left(\sum X_i\right) = \frac{1}{n}\sum \mathbb{E}(X_i) = \frac{1}{\cancel{n}}\cancel{n}\mu = \mu$$

$\Rightarrow$ unbiased

$$\mathbb{P}(|\overline{X} - \mu| > \epsilon) \leq \frac{Var(\overline{X})}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

$\Rightarrow$ Consistent (using Chebychev's Inequality on the R.V. $\overline{X}$)

A popular metric to compare Estimators is the mean Square Error (MSE)

Def

The MSE of an Estimator $\hat{\theta}$ is $MSE(\hat{\theta}) = \mathbb{E}\left((\hat{\theta}-\theta)^2\right)$

It can be shown that $MSE(\hat{\theta}) = [Bias(\hat{\theta})]^2 + Var(\hat{\theta})$

we would like an Estimator with a good mix of Small bias and small variance.

Ex

Let $X_1 \cdots X_n \sim N(\mu, \sigma^2)$ and we want an Estimator for μ. Two choices we will compare:

$$\left.\begin{array}{l} 1.)\ \hat{\mu}_1 = X_1 \\ 2.)\ \hat{\mu}_2 = \overline{X} \end{array}\right\} \text{compare MSEs}$$

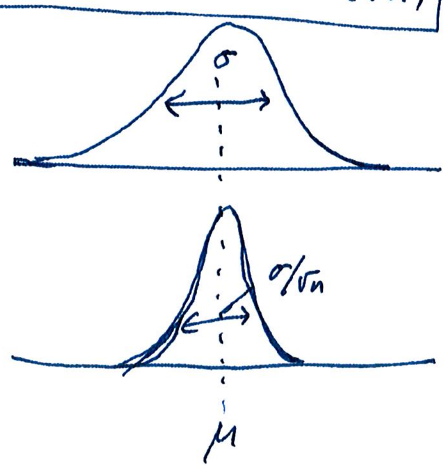
First, both Estimators have sampling distributions that are the Normal Distribution.

Second, $\mathbb{E}(X_1) = \mu$ and $\mathbb{E}(\overline{X}) = \mu$ so both are unbiased Estimators

Thus, to compare MSEs we just have to compare variances.

$Var(X_1) = \sigma^2$ and $var(\overline{X}) = \sigma^2/n$

Thus, $\boxed{MSE(\overline{X}) < MSE(X_1)}$ and should be preferred



Sampling distribution of X_1

Sampling Distribution of $\overline{X}$

Statistical models

We want a model for our entire sample that we can use for Inference on parameters etc.

Def

A Statistical model is the joint distribution of our sample.

Recall: we saw joint distributions for two discrete R.V.'s earlier in the semester $\Rightarrow$ $P_{XY}(x,y) = \mathbb{P}(X=x, Y=y)$

"and"

Also, if X & Y were Independent, then the joint distribution could be written as $P_{XY}(x,y) = \mathbb{P}(X=x, Y=y) = P_X(x)P_Y(y)$

Let $X_1, \dots, X_n \overset{iid}{\sim} f_X(x)$. The joint distribution of our sample is

$$\boxed{f(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f_X(x_i)}$$

The joint distribution will involve unknown parameter(s) that we need to estimate To have a "working" model.

The first thing we can do as in the previous pages is use our Statistical model to come up with a single estimate (Point Estimate) of the parameter(s) which we can then plug in so the model is usable

- In statistics this is called "fitting the model"
- In machine Learning this is called "Learning the model" with training data

EX I have a sample $X_1, \ldots, X_n \overset{iid}{\sim} f_X(x)$

where X_i = # goals scored by ISU women's soccer team in game i.

1.) come up with a model for the sample

2.) Estimate the parameter

3.) use fitted model to estimate probability the score more than 2 goals in next game.

Each X_i is a discrete R.V. that is the # of occurences (goals) in some time frame (1 game) {maybe poisson Distribution?}

Our assumption will be:

$X_1 \cdots X_n \overset{iid}{\sim} pois(\lambda)$ [Recall: $P_X(x_i) = \dfrac{e^{-\lambda}\lambda^{x_i}}{x_i!}$ for $X_i \sim pois(\lambda)$]

So our joint model for our sample will be:

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} \underbrace{f_X(x_i)}_{\text{distribution where the } X\text{'s are coming from}} = \boxed{\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}}$$

There is one parameter, λ, that needs to be Estimated from the sample

Since $\lambda = \mathbb{E}(X)$ I will use the plug in Estimator $\bar{X}$.

suppose the observed values were

0, 0, 1, 0, 1, 2, 2, 0, 1, 1

my estimate of λ is: $\hat{\lambda} = \bar{x} = \boxed{.8}$. Thus I assume that a poisson distribution with $\lambda = .8$ is the distribution that generated my data.

So, for the next game I can use my fitted model to answer $\mathbb{P}(X > 2) = 1 - \mathbb{P}(X \leq 2)$ where $X \sim pois(.8)$

$$= \boxed{.047}$$