STA255: Statistical Theory

Chapter 8: Estimation

Summer 2017

Introduction

- If $Y_1, Y_2, \ldots, Y_n \overset{i.i.d.}{\sim} N(\mu, 1)$, where μ is unknown, then how to estimate μ?
- An estimator is a rule, often expressed as a formula, that tells how to calculate the value of an estimate based on the measurements contained in a sample.
- Goal: Let $Y_1, Y_2, \ldots, Y_n \overset{i.i.d.}{\sim} f_\theta$, where the parameter θ is unknown. We want to find an estimate for the unknown parameter θ.

A Point Estimator

- A point estimator of the parameter θ is a function of the underlying random variables and so it is a random variable with a distribution function.
- Thus, a point estimate of θ is a function of the data.
- Notation: $\hat{\theta}$ is a point estimator of θ.
- **Example:** If a certain distribution has a mean μ and a variance σ^2, then:
  - $\overline{Y}$ is a point estimator of μ.
  - S^2 is a point estimator of σ^2.

Desirable properties of a point estimator

- Unbiased
- Consistent
- Minimum variance
- Has a known probability distribution

Definition

Let $\hat{\theta}$ be a point estimator for a parameter θ. Then $\hat{\theta}$ is an unbiased estimator if $E(\hat{\theta}) = \theta$. If $E(\hat{\theta}) \neq \theta$, then $\hat{\theta}$ is said to be biased.

Example

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample with $\mathrm{E}(\mathrm{Y}_i) = \mu$. Show that $\hat{\theta} = \bar{Y}$ is unbiased estimator of μ.

Bias and Mean Square Error

Definition

The bias of a point estimator $\hat{\theta}$ is given by $B(\hat{\theta}) = E(\hat{\theta}) - \theta$.

Definition

The mean square error of a point estimator $\hat{\theta}$ is $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$.

Note: $MSE(\hat{\theta}) = V(\hat{\theta}) + \left[B(\hat{\theta})\right]^2$.

Proof:

Example: # 8.6

Suppose that $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$, $V(\hat{\theta}_1) = \sigma_1^2$ and $V(\hat{\theta}_2) = \sigma_2^2$. Consider the estimator $\hat{\theta}_3 = a\hat{\theta}_1 + (1-a)\hat{\theta}_2$.

(a) Show that $\hat{\theta}_3$ is an unbiased estimator for θ.

(b) If $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent, how should the constant a be chosen in order to minimize the variance of $\hat{\theta}_3$?

Examples of Unbiased Estimators

- Let $Y_1, Y_2, \ldots, Y_n$ be a random sample with $E(Y_i) = \mu_1$ and $X_1, X_2, \ldots, X_n$ be a random sample with $E(X_i) = \mu_2$. Then:
  - $E(\bar{Y}) = \mu_1$ and $E(\bar{X}) = \mu_2$
  - $E(\bar{Y} - \bar{X}) = \mu_1 - \mu_2$
- Let $\bar{Y} \sim Bin(n, p_1)$ and $\bar{X} \sim Bin(n, p_2)$. Then $\hat{p}_1 = \frac{Y}{n}$ is the proportion of successes in the sample.
  - $E(\hat{p}_1) = E(\frac{Y}{n}) = \frac{1}{n}E(Y) = \frac{np_1}{n} = p_1$.
  - $E(\hat{p}_2) = E(\frac{Y}{n}) = \frac{1}{n}E(Y) = \frac{np_2}{n} = p_2$.
  - $E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$.

Bias and Mean Square Error

Definition

Let $\sigma^2_{\hat{\theta}}$ be the variance of the sampling distribution of the estimator $\hat{\theta}$ (i.e. $V(\hat{\theta}) = \sigma^2_{\hat{\theta}}$) then $\sqrt{V(\hat{\theta})} = \sqrt{\sigma^2_{\hat{\theta}}} = \sigma_{\hat{\theta}}$ is called the standard error of the estimator.

Examples:

- Let $Y_1, Y_2, \ldots, Y_n$ be a random sample with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$. $V(\bar{Y}) = \frac{\sigma^2}{n} \Rightarrow \sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}}$.
- Let $Y \sim Bin(n, p)$. Then $V(\hat{p}) = V(\frac{Y}{n}) = \frac{1}{n^2}V(Y) = \frac{npq}{n^2} = \frac{pq}{n}$. Thus, $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$.
- $V(\bar{Y} - \bar{X}) = V(\bar{Y}) + V(\bar{X}) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}$. Thus, $\sigma_{\bar{Y}-\bar{X}} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$

Example

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$. Show that

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

is a **biased** estimator for σ^2 and that

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

is **unbiased** estimator for σ^2.

Evaluating the Goodness of a Point Estimator

Definition

The error of estimation ϵ is the distance between an estimator and its target parameter. That is, $\epsilon = |\hat{\theta} - \theta|$.

Example: A sample of $n = 1000$ voters, randomly selected from a city, showed $y = 560$ in favour of candidate Jones. Estimate p, the fraction of voters in the population favouring Jones, and place a 2-standard-error bound on the error of estimation.

Solution:

Confidence Intervals

- An alternative to reporting a single value for the parameter being estimated is to calculate and report an entire interval of plausible values; i.e., a confidence interval (CI).
- Properties of the CI:
  - It contains true parameter θ.
  - It is relatively narrow.
- The lower and upper endpoints of a CI are called the lower and upper confidence limits.
- The probability that a CI will enclose θ is called the **confidence coefficient** or **confidence level**, denoted by $\alpha - 1$.

Definition

A $100(1-\alpha)\%$ **confidence interval** for a parameter θ is a random interval $[\hat{\theta}_L, \hat{\theta}_U]$ such that $P(\hat{\theta}_L \leq \theta \leq \hat{\theta}_U) = 1 - \alpha$, regardless of the value of θ.

Confidence Intervals

- Two-sided CI: $[\hat{\theta}_L, \hat{\theta}_U]$
- Lower one-sided CI: $[\hat{\theta}_L, \infty)$
- Upper one-sided CI: $(-\infty, \hat{\theta}_U]$
- The confidence level is denoted by $100(1-\alpha)\%$. The most common confidence levels are 90%, 95% and 99%.
- The higher the confidence level, the more strongly we believe that the true value of the parameter being estimated lies within the interval.

Finding CI: Pivotal Method

Let $Y_1, \ldots, Y_n$ be a random sample and we observed the data $y_1, \ldots, y_n$ which are the realization of these random variables. We want a CI for some parameter θ.

- **Pivotal Method:**
- We need to find a random variable, called a pivot, that is typically a function of the estimator of θ satisfying the following:
- It depends on $Y_1, \ldots, Y_n$ and θ.
- Its probability distribution does not depend on θ or any other unknown parameter.

Example

Suppose that we are to obtain a single observation Y from an exponential distribution with mean θ. Use Y to form a confidence interval for with confidence coefficient 0.90 or 90% confidence level.

Example: # 8.47

Assume that $Y_1, Y_2, \ldots, Y_n$ is a sample of size n from an exponential distribution with mean θ.

(a) Use the method of moment-generating functions to show that $2\sum_{i=1}^{n} Y_i/\theta$ is a pivotal quantity with a χ^2 distribution with $2n$ df.

(b) Use the pivotal quantity $2\sum_{i=1}^{n} Y_i/\theta$ to derive a 95% confidence interval for θ.

(c) If a sample of size $n = 7$ yields $\bar{y} = 4.77$, use the result from part (b) to give a 95% confidence interval for θ.

Large-Sample Confidence Intervals

- If the target parameter θ is $\mu, p, \mu_1 - \mu_2$ or $p_1 - p_2$, then for large samples (CLT),
- $Z = \frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}} \sim N(0, 1)$
- The distribution of Z doesn't depend on θ. So it is pivotal.
- The pivotal method can be employed to develop confidence intervals for the target parameter θ.

Example

Let $\hat{\theta}$ be a statistic that is normally distributed with mean θ and standard error $\sigma_{\hat{\theta}}$. Find a confidence interval for θ that possesses a confidence coefficient equal to $1 - \alpha$.

Example: # 8.54

Is Americas romance with movies on the wane? In a Gallup Poll of $n = 800$ randomly chosen adults, 45% indicated that movies were getting better whereas 43% indicated that movies were getting worse.

(a) Find a 98% confidence interval for p, the overall proportion of adults who say that movies are getting better.

(b) Does the interval include the value $p = .50$? Do you think that a majority of adults say that movies are getting better?

Confidence Intervals Based on t-distribution

CI for μ when σ is unknown:

- Pivotal: $T = \frac{\bar{Y}-\mu}{s/\sqrt{n}} \sim t_{n-1}$.
- $100(1-\alpha)\%$ CI for μ is:

$$\left[\bar{Y} - t_{\alpha/2}\frac{S}{\sqrt{n}}, \bar{Y} + t_{\alpha/2}\frac{S}{\sqrt{n}}\right]$$

Proof

Example

A manufacturer of gunpowder has developed a new powder, which was tested in eight shells. The resulting muzzle velocities, in feet per second, were as follows:

$$3005, 2925, 2935, 2965, 2995, 3005, 2937, 2905$$

Find a 95% confidence interval for the true average velocity μ for shells of this type. Assume that muzzle velocities are approximately normally distributed.

Confidence Intervals Based on t-distribution

CI for $\mu_1 - \mu_2$:

- Let $Y_1, Y_2, \dots, Y_n \overset{i.i.d.}{\sim} N(\mu_1, \sigma_1^2)$ and $X_1, X_2, \dots, X_n \overset{i.i.d.}{\sim} N(\mu_2, \sigma_2^2)$
- Assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$, where σ is unknown.
- $\bar{Y} - \bar{X} \sim N(\mu_1 - \mu_2, \sigma^2(\frac{1}{n_1} + \frac{1}{n_2}))$.
- Let $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1-1)+(n_2-2)}$, where S_1^2 and S_2^2 are the sample variances from the first and the second samples, respectively.
- We know that $W = \frac{n_1+n_2-2}{\sigma^2} S_p^2 \sim \chi^2_{n_1+n_2-2}$ and $Z = \frac{\bar{Y}-\bar{X}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim N(0,1)$.
- Pivotal: $T = \frac{Z}{\sqrt{W/(n_1+n_2-2)}} = \frac{\bar{Y}-\bar{X}-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$
- $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ is:

$$\left[\bar{Y} - \bar{X} - t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \bar{Y} - \bar{X} + t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$$

Confidence Interval for σ^2

- Pivotal: $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$.
- $100(1-\alpha)\%$ CI for σ^2 is:

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}\right]$$

Proof

Example

An experimenter wanted to check the variability of measurements obtained by using equipment designed to measure the volume of an audio source. Three independent measurements recorded by this equipment for the same sound were 4.1, 5.2, and 10.2. Estimate σ^2 with confidence coefficient 0.90.