PS 2010: 10. Estimation and Inference 2

Qing Chang

Department of Political Science
University of Pittsburgh

Fall 2023

Agenda

- Bias and consistency
- Confidence interval
- Hypothesis Testing

Bias

- To measure a goodness of an estimator we can use bias measure
- An unbiased point estimator produces a sample statistic with sampling distribution that has a mean equal to the population parameter the statistic is intended to estimate.

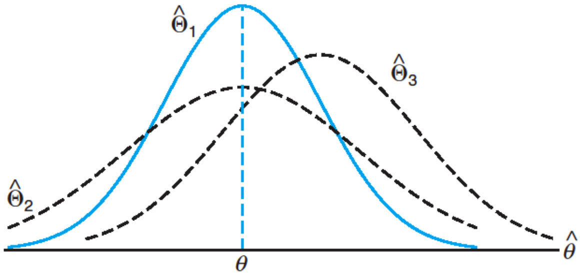
  Let $\hat{\theta}$ be a point estimator for a parameter θ. Then $\hat{\theta}$ is an unbiased estimator if $E(\hat{\theta}) = \theta$. If $E(\hat{\theta}) \neq \theta, \hat{\theta}$ is said to be biased.
- The bias of a point estimator $\hat{\theta}$ is given by $B(\hat{\theta}) = E(\hat{\theta}) - \theta$

Mean Square Error

- Another way to measure the goodness of an estimator in mean square error (MSE). The MSE for $\hat{\theta}$ is

$$\mathrm{MSE}(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right].$$

$$\begin{aligned}
\mathrm{MSE} &= \mathbb{E}\left\{(\hat{\theta}-\theta)^2\right\} \\
&= \mathbb{E}\left[\{(\hat{\theta}-\mathbb{E}(\hat{\theta}))+(\mathbb{E}(\hat{\theta})-\theta)\}^2\right] \\
&= \mathbb{E}\left[\{\hat{\theta}-\mathbb{E}(\hat{\theta})\}^2\right]+\{\mathbb{E}(\hat{\theta})-\theta\}^2 \\
&= \text{ variance } + \text{ bias }^2.
\end{aligned}$$



$\hat{\Theta}_1$
$\hat{\Theta}_3$
$\hat{\Theta}_2$
θ
$\hat{\theta}$

Property of Point Estimator: Efficiency and Consistency

- MSE cares about both variance and bias
- **Efficiency**, we want an estimator is unbias and has the smallest variance among all unbiased estimator.
- RMSE is just the root of MSE
- A consistent point estimator converges in probability to the true parameter value.

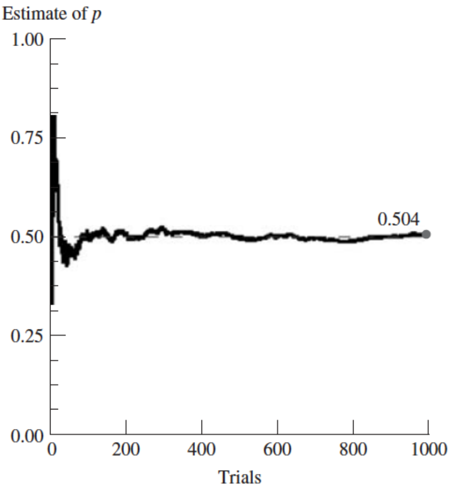
$$\lim_{n \to \infty} P\left(\left|\hat{\theta}_n - \theta\right| \leq \varepsilon\right) = 1$$

or, equivalently,

$$\lim_{n \to \infty} P\left(\left|\hat{\theta}_n - \theta\right| > \varepsilon\right) = 0$$

Understanding Consistency

- Flip a coin 5, 10, 20, 100, 1000 times and calculate $P(Head)$



Unbiasedness and Consistency

An unbiased estimator $\hat{\theta}_n$ for θ is a consistent estimator of θ if

$$\lim_{n\to\infty} V\left(\hat{\theta}_n\right) = 0$$

- Example: Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a distribution with mean μ and variance $\sigma^2 < \infty$. Show that $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ is a consistent estimator of μ.

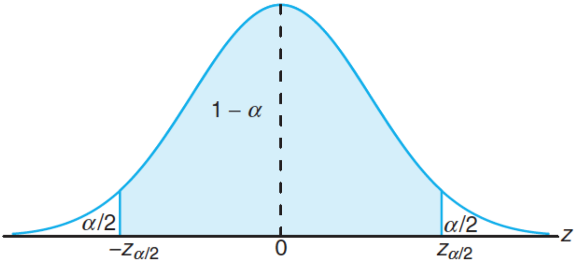
Confidence Interval

- We have learned how to measure a goodness and properties of an estimator
- Another way of measuring uncertainty of an estimator is confidence interval
- It involves of characterizing and describing the sampling distribution of statistics
- Confidence intervals give a range of values that are likely to include the true value of the parameter.
    - over a hypothetically repeated data-generating process
    - contain the true value of the parameter
    - with the probability specified by the confidence level

Confidence Interval

- Also called confidence bounds, error bands, interval estimators
- Two sided CI have lower and upper bounds. One sided CI is also allowed
- Research need to decide the confidence level by $1 - \alpha$

$$P\left(\hat{\theta}_L \leq \theta \leq \hat{\theta}_U\right) = 1 - \alpha,$$



Constructing CI for Large Sample

- Under CLT, the sampling distribution of an estimator is normally distributed
- We could then transform it to standard normal
- That is $Z = \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}$
- If we want CI as $P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1-\alpha$
- Then, $P\left(-z_{\alpha/2} < \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} < z_{\alpha/2}\right) = 1-\alpha$
- Finally

$$P\left(\hat{\theta} - z_{\alpha/2}\sigma_{\hat{\theta}} < \theta < \hat{\theta} + z_{\alpha/2}\sigma_{\hat{\theta}}\right) = 1-\alpha$$

- The CI is

$$\hat{\theta} - z_{\alpha/2}\sigma_{\hat{\theta}} < \theta < \hat{\theta} + z_{\alpha/2}\sigma_{\hat{\theta}}$$

CI for μ in Large Sample

If $\bar{x}$ is the mean of a random sample of size n from a population with known variance σ^2, a $100(1-\alpha)\%$ confidence interval for μ is given by

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}},$$

where $z_{\alpha/2}$ is the z-value leaving an area of $\alpha/2$ to the right.

- If σ is unknown, use s sample standard deviation to replace it.
- However, this is not ok under small sample size

Examples

- The above are general procedures, let's look at some examples
- SAT mathematics scores of a random sample of 500 high school seniors in the state of Texas are collected, and the sample mean and standard deviation are found to be 501 and 112, respectively. Find a 99% confidence interval on the mean SAT mathematics score for seniors in the state of Texas.
- The shopping times of n $=$ 64 randomly selected customers at a local supermarket were recorded. The average and variance of the 64 shopping times were 33 minutes and 256, respectively. Estimate μ, the true average shopping time per customer, with a confidence level of $1 - \alpha = .90$.

Summary

- Under large sample, we have no assumption on the population distribution
- CLT makes sure sampling distribution of estimators have normal distribution
- We can also safely replace σ using s and using standard normal for the critical value
- However, using s to replace σ is student t in small sample size if sample comes from normal distribution

Constructing CI for Small Sample

- Notice, in the small sample and the variance of population is unknown, the sampling distribution is not normally distributed any more
- Under the assumption that sample is randomly draw from normal distribution
- We need to use student t distribution to construct the CI for population mean μ
- The rest of the procedures are the same

CI for μ in Small Sample

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal population with unknown variance σ^2, a $100(1-\alpha)\%$ confidence interval for μ is

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}},$$

where $t_{\alpha/2}$ is the t-value with $v = n - 1$ degrees of freedom, leaving an area of $\alpha/2$ to the right.

Example

A meat inspector has randomly selected 30 packs of 95% lean beef. The sample resulted in a mean of 96.2% with a sample standard deviation of 0.8%. Find a 99% prediction interval for the leanness of a new pack. Assume normality.

CI for Proportion

- In a binomial experiment, the proportion of success can be represented by $\hat{p} = \frac{x}{n}$
- For large sample, under CLT, the variance of sampling distribution is $\sigma^2_{\hat{P}} = \frac{\hat{p}\hat{q}}{n}$
- Therefore, we could have CI for proportion as

$$P\left(\widehat{P} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \widehat{P} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) \approx 1 - \alpha$$

- This is large sample result

Example

In a random sample of $n = 500$ families owning television sets in the city of Hamilton, Canada, it is found that $x = 340$ subscribe to HBO. Find a 95% confidence interval for the actual proportion of families with television sets in this city that subscribe to HBO.

Hypothesis Testing

- A statistical hypothesis is an assertion or conjecture concerning one or more populations.
- We use data from random sample to test hypothesis
- This is a probability issue
  - Rejection of a hypothesis implies that the sample evidence refutes it
  - We could be wrong, but probability is very small
- Support for one theory is obtained by showing lack of support for its converse (in a sense, a proof by contradiction)
- Example: Vaccine, election prediction

Elements of a Statistical Test

- **Null Hypothesis (H_0):**
  - The hypothesis to be tested
  - The opposite of the H_1
  - We **reject or fail to reject** H_0
- **Alternative Hypothesis (H_1):**
  - The hypothesis to be accepted in case H_0 is rejected
  - Represents the question to be answered or the theory to be tested
- **Test statistic**: a function of sample which the statistical decision will be based
- **Rejection region**: specifies the values of the test statistic for which the null hypothesis is to be rejected in favor of the alternative hypothesis.
- Example: Jury trial, high school example again

Rejection Region

- Choosing RR is very important
- Two errors we may make
- **Type I error:** Rejection of the null hypothesis when it is true
  - P(Type I error) $= \alpha$ called the level of the test
- **Type II error:** Nonrejection of the null hypothesis when it is false
  - P(Type I error) $= \beta$

	H_0 is true	H_0 is false
Do not reject H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

Relation between Type I and II error

- A decrease in the probability of one generally results in an increase in the probability of the other.
- Adjust the RR is to adjust the type I error
- Increase sample size can decrease both errors

Test Concerning a Single Mean for Large Sample

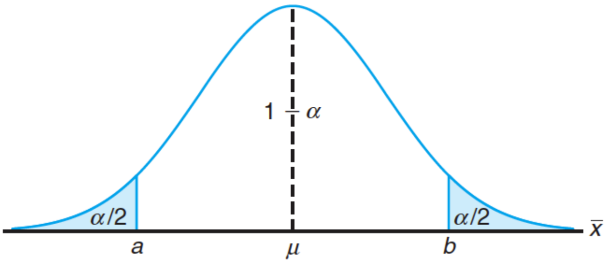
- Suppose X_1, X_2, X_3 ... X_n are randomly sample from a distribution with mean μ and variance σ^2
- We want to test if population $\mu = \mu_0$

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

- To estimate population mean, we could use the mean of sampling distribution based data
- Under large sample size (CLT), we know $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$
- Then, we could calculate $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$, given the null hypothesis

Test Concerning a Single Mean for Large Sample

- We also need to decide α
- For two tail test, we need $\frac{\alpha}{2}$ to determine RR
- For one tail test, we use α to determine RR
- Once we have RR, we also have critical values
- Reject H_0 falls into RR region



Examples

- A random sample of 100 recorded deaths in the United States during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.
- A vice president in charge of sales claims that salespeople are averaging no more than 15 sales contacts per week. As a check on his claim, n $=$ 36 salespeople are selected at random, and the number of contacts made by each is recorded for a single randomly selected week. The mean and variance of the 36 measurements were 17 and 9, respectively. Does the evidence contradict the vice president's claim? Use a test with level $\alpha = .05$.

Example

- A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms. Test the hypothesis that $\mu = 8$ kilograms against the alternative that $\mu \neq 8$ kilograms if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kilograms with a standard deviation of 0.5 kilogram. Use a 0.01 level of significance.

Test Concerning a Single Mean for Small Sample

- The logic and procedures are the same as in the large sample
- Core difference is when replace population σ^2 with sample s^2, the test statistic becomes to t distribution
- That is: $T = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$
- Notice, we also need to make sure the population distribution is a normal

Example

- A study claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes included in a planned study indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a standard deviation of 11.9 kilowatt hours, does this suggest at the 0.05 level of significance that vacuum cleaners use, on average, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal.

P-value

Definition: If W is a test statistic, the p-value, or attained significance level, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

- In all the above hypothesis testing, we pre-select α
- Instead, we can calculate the test statistic first and then calculate the P-value
- We then make decisions
- Previous examples

Examples

- Just prior to an important election, in a random sample of 749 voters, 397 preferred Candidate Y over Candidate Z.
  - Construct a 90% confidence interval for the overall proportion of voters who prefer Candidate Y over Candidate Z.
  - Test if the candidate is going to win less than 58% of vote, and calculate P value.
- In a random sample of 25 direct flights from New York to Boston by Delta Airline, the sample mean flight time was 56 minutes and the sample standard deviation was 8 minutes.
  - Construct a 99% confidence interval for the overall mean flight time on this route. (Assume the flight times are approximately normally distributed.)
  - Test the claim that the mean flight time is more than 50 mins, and calculate P value