Introduction to Statistical Inference

Jianan Hui

3/5/2015

Background

- Populations and parameters
  - For a normal population
    population mean μ and s.d. σ
  - A binomial population
    population proportion p
- If parameters are unknown, we make statistical inferences about them using sample information.

What is statistical inference?

- Drawing conclusions based on data.
- **Estimation:**
  - Estimating the value of the parameter
  - "What is (are) the values of μ or *p*?"
- **Hypothesis Testing:**
  - Deciding about the value of a parameter based on some preconceived idea.
  - "Did the sample come from a population with $\mu = 5$ or *p* = .2?"

Example

- A consumer wants to estimate the average price of similar homes in her city before putting her home on the market.

Estimation: Estimate μ, the average home price.

- A manufacturer wants to know if a new type of steel is more resistant to high temperatures than the old type.

Hypothesis test: Is the new average resistance, μ_N greater to the old average resistance, μ_O?

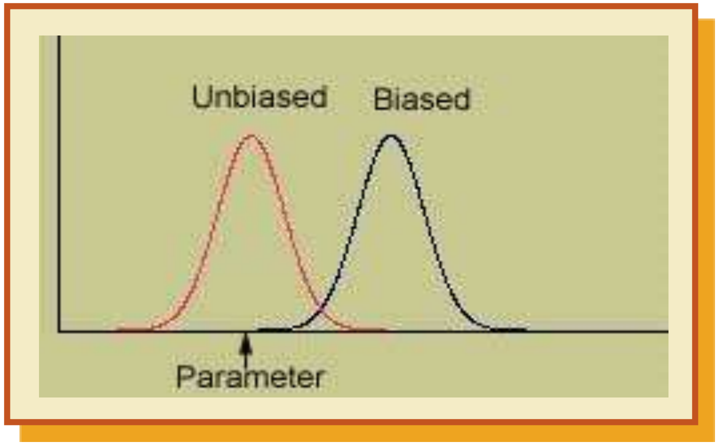
Part 1: Estimation

What is estimator?

- An **estimator** is a rule, usually a formula, that tells you how to calculate the estimate based on the sample.
- Estimators are calculated from sample observations, hence they are statistics.
  - **Point estimator:** A single number is calculated to estimate the parameter.
  - **Interval estimator:** Two numbers are calculated to create an interval within which the parameter is expected to lie.

"Good" Point Estimators

- An **estimator** is **unbiased** if its mean equals the parameter.
- It does not systematically overestimate or underestimate the target parameter.
- Sample mean($\bar{x}$)/proportion($\hat{p}$) is an unbiased estimator of population mean/proportion.



Example

- Suppose $X_1, X_2, \ldots X_n$ iid~ $N(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$.
- If $\hat{\boldsymbol{\mu}} = \text{Geometric Mean} = \sqrt[n]{X_1 X_2 \ldots X_n}$,

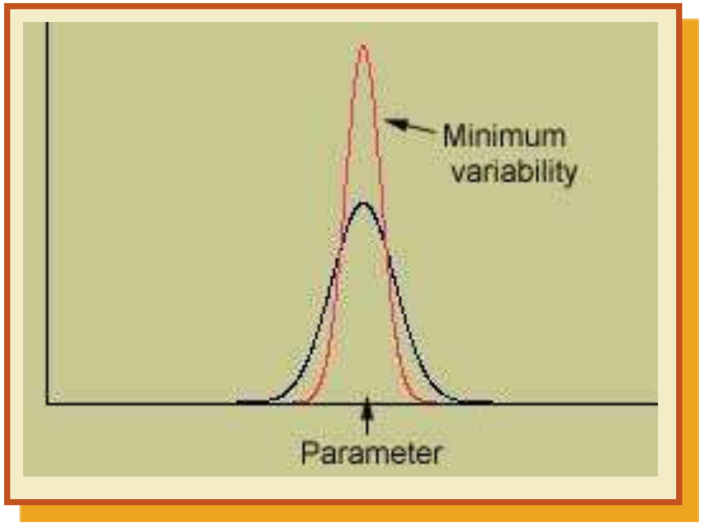
  then $E(\hat{\boldsymbol{\mu}}) \neq \boldsymbol{\mu}$.
- If $\hat{\boldsymbol{\mu}} = \text{Arithmetic Mean} = \overline{X} = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$,

  then

$$E(\hat{\boldsymbol{\mu}}) = \frac{1}{n} E(X_1 + X_2 + \ldots + X_n) = \frac{n}{n}\boldsymbol{\mu} = \boldsymbol{\mu}.$$

“Good” Point Estimators

- We also prefer the sampling distribution of the estimator has a **small spread** or **variability,** i.e. small standard deviation.



Example

- Suppose $X_1, X_2, \ldots X_n$ iid~ $N(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$.
- If $\hat{\boldsymbol{\mu}} = X_1$, then $\mathrm{var}(\hat{\boldsymbol{\mu}}) = \mathrm{var}(X_1) = \boldsymbol{\sigma}^2$.
- If $\hat{\boldsymbol{\mu}} = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$, then

$$\mathrm{var}(\hat{\boldsymbol{\mu}}) = \mathrm{var}(\frac{X_1 + X_2 + \ldots + X_n}{n}) = \frac{1}{n^2}\mathrm{var}(X_1 + X_2 + \ldots + X_n)$$

$$= \frac{1}{n^2} * n * \mathrm{var}(X_1) = \frac{\boldsymbol{\sigma}^2}{n}.$$

Measuring the Goodness of an Estimator

- A good estimator should have small bias as well as small variance.
- A common criterion could be Mean Square Error(MSE):

$$\mathrm{MSE}(\hat{\boldsymbol{\mu}}) = \mathrm{Bias}^2(\hat{\boldsymbol{\mu}}) + \mathrm{var}(\hat{\boldsymbol{\mu}}),$$
$$\text{where } \mathrm{Bias}(\hat{\boldsymbol{\mu}}) = \mathrm{E}(\hat{\boldsymbol{\mu}}) - \boldsymbol{\mu}.$$

Example

- Suppose $X_1, X_2, \ldots X_n$ iid~ $N(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$.
- If $\hat{\boldsymbol{\mu}} = X_1$, then

$$MSE(\hat{\boldsymbol{\mu}}) = Bias^2(\hat{\boldsymbol{\mu}}) + var(\hat{\boldsymbol{\mu}}) = 0 + \boldsymbol{\sigma}^2.$$

- If $\hat{\boldsymbol{\mu}} = \overline{X} = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$, then

$$MSE(\hat{\boldsymbol{\mu}}) = Bias^2(\hat{\boldsymbol{\mu}}) + var(\hat{\boldsymbol{\mu}}) = 0 + \frac{\boldsymbol{\sigma}^2}{n}.$$

Estimating Means and Proportions

- For a quantitative population,

$$\text{Point estimator of population mean } \mu : \bar{x}$$

- For a binomial population,

$$\text{Point estimator of population proportion } p : \hat{p} = x/n$$

Example

- A homeowner randomly samples 64 homes similar to her own and finds that the average selling price is $252,000 with a standard deviation of $15,000.
- Estimate the average selling price for all similar homes in the city.

$$\text{Point estimator of } \mu\text{: } \overline{x} = 252{,}000$$

Example

A quality control technician wants to estimate the proportion of soda cans that are underfilled. He randomly samples 200 cans of soda and finds 10 underfilled cans.

$n = 200$ $\quad$ $p =$ proportion of underfilled cans

Point estimator of p: $\hat{p} = x / n = 10 / 200 = .05$

Interval Estimator

- Create an interval (a, b) so that you are fairly sure that the parameter falls in (a, b).
- "Fairly sure" means "with high probability", measured by the confidence coefficient, $1-\alpha$.

Usually, $1-\alpha = .90, .95, .98, .99$

How to find an interval estimator?

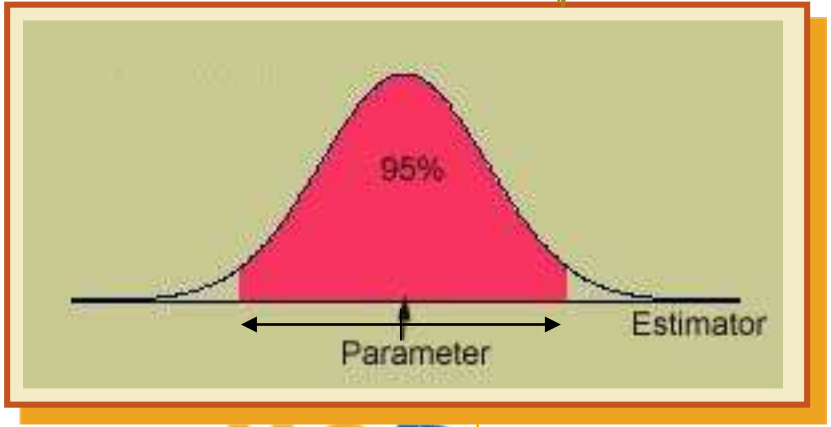
- Suppose 1-α = .95 and that the point estimator has a normal distribution.

$$P(\mu - 1.96SE < \overline{X} < \mu + 1.96SE) = .95$$

$$\Leftrightarrow P(\overline{X} - 1.96SE < \mu < \overline{X} + 1.96SE) = .95$$

$$a = \overline{X} - 1.96SE; \quad b = \overline{X} + 1.96SE$$

Empirical Rule



95% C.I. of μ is:

Estimator ± *1.96*SE

In general, 100(1-α)% C.I. of a parameter is:

Estimator ± $z_{\alpha/2}$SE

How to obtain the z score?

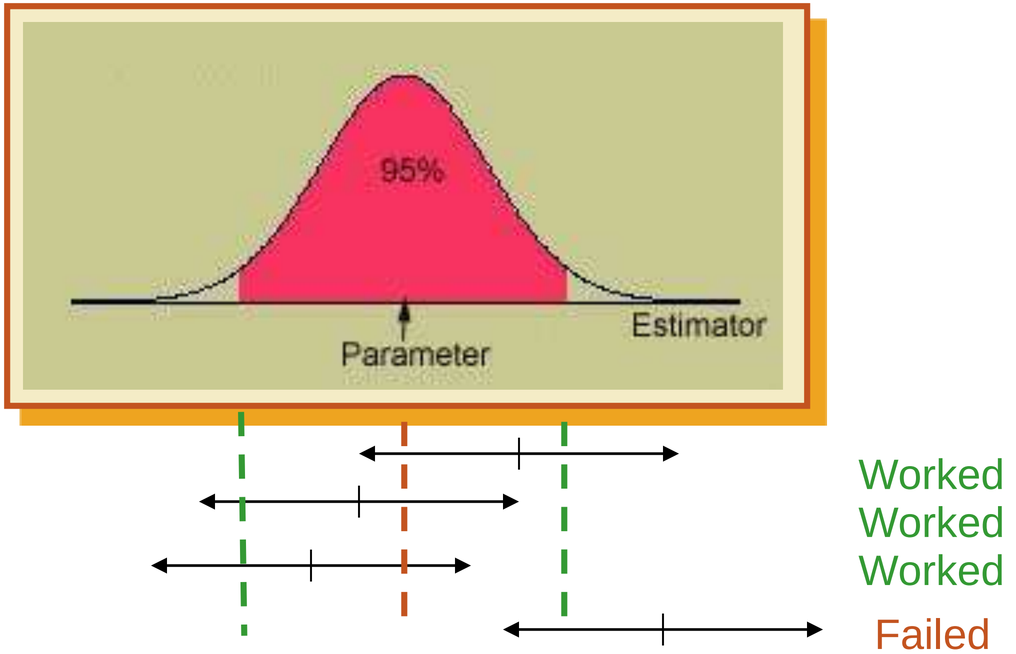
- We can find z score based on the z table of standard normal distribution.

$z_{\alpha/2}$	$1-\alpha$
1.645	.90
1.96	.95
2.33	.98
2.58	.99

100(1-α)% Confidence Interval:

Estimator $\pm\ z_{\alpha/2}$SE

What does 1-α stand for?



- 1-α is the proportion of intervals that capture the parameter in repeated sampling.
- More intuitively, it stands for the probability of the interval will capture the parameter.

Confidence Intervals for Means and Proportions

- For a Quantitative Population

Confidence Interval for a Population Mean μ :

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

- For a Binomial Population

Confidence Interval for Population Proportion p :

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Example

- A random sample of $n = 50$ males showed a mean average daily intake of dairy products equal to 756 grams with a standard deviation of 35 grams. Find a 95% confidence interval for the population average μ.

$$\bar{x} \pm 1.96\frac{s}{\sqrt{n}} \Rightarrow 756 \pm 1.96\frac{35}{\sqrt{50}} \Rightarrow 756 \pm 9.70$$

$$\text{or } 746.30 < \mu < 765.70 \text{ grams.}$$

Example

UCR

- Find a 99% confidence interval for μ, the population average daily intake of dairy products for men.

$$\bar{x} \pm 2.58\frac{s}{\sqrt{n}} \Rightarrow 756 \pm 2.58\frac{35}{\sqrt{50}} \Rightarrow 756 \pm 12.77$$

$$\text{or } 743.23 < \mu < 768.77 \text{ grams.}$$

The interval must be wider to provide for the increased confidence that it does indeed enclose the true value of μ.

Summary

I. Types of Estimators

1. **Point estimator**: a single number is calculated to estimate the population parameter.
2. **Interval estimator**: two numbers are calculated to form an interval that contains the parameter.

II. Properties of Good Point Estimators

1. **Unbiased**: the average value of the estimator equals the parameter to be estimated.
2. **Minimum variance**: of all the unbiased estimators, the best estimator has a sampling distribution with the smallest standard error.

Summary

Estimator for normal mean and binomial proportion

Parameter	Point Estimator	Margin of Error
μ	$\bar{x}$	$\pm 1.96\left(\frac{s}{\sqrt{n}}\right)$
p	$\hat{p} = \frac{x}{n}$	$\pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}}$
$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	$\pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
$p_1 - p_2$	$(\hat{p}_1 - \hat{p}_2) = \left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right)$	$\pm 1.96\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$

Part 2: Hypothesis Testing

Introduction

- Suppose that a pharmaceutical company is concerned that the mean potency μ of an antibiotic meet the minimum government potency standards. They need to decide between two possibilities:
  - **The mean potency μ does not exceed the mean allowable potency.**
  - **The mean potency μ exceeds the mean allowable potency.**

- This is an example of **hypothesis testing.**

Hypothesis Testing

- Hypothesis testing is to make a choice between two hypotheses based on the sample information.
- We will work out hypothesis test in a simple case but the ideas are all universal to more complicated cases.

Hypothesis Testing Framework

1. Set up null and alternative hypothesis.
2. Calculate test statistic (often using common descriptive statistics).
3. Calculate P-value based on the test statistic.
4. Make rejection decision based on P-value and draw conclusion accordingly.

1 Set up Null and Alternative Hypothesis

- One wants to test if the average height of UCR students is greater than 5.75 feet or not. The hypothesis are:

  - $H_0: \mu = 5.75$
  - $H_a: \mu > 5.75$

- Null hypothesis is H_0 and alternative is H_a

Structure of Null and Alternative

- H_0 always has the equality sign and H_a never has an equality sign.

- H_a can be 1 of 3 types(for this example):
  - $H_a: \mu < 5.75$; $H_a: \mu \neq 5.75$; $H_a: \mu > 5.75$
  - H_a reflects the question being asked

Are these correct?

- $H_0: \mu > 5.75$
  $H_a: \mu = 5.75$

- $H_0: \mu = 5.75$
  $H_a: \mu \geq 5.75$

- $H_0: \bar{X} = 5.75$
  $H_a: \bar{X} > 5.75$

2 Calculating a Test Statistic

- Let's say that we collected a sample of 25 UCR students heights and $\overline{\mathrm{X}} = 5.9$ and $S = .75$

- Our test statistic would be: $\mathrm{T}^*_{\mathrm{n}-1} = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} = \frac{\bar{X}-5.75}{\frac{S}{\sqrt{n}}}$

- How is this test statistic formed and why do we use it?

Test Statistic

- We are using this test statistic because:
  - $\mathrm{T}_{\mathrm{n}-1}^{*}$ is expected small when H_0 is true, and large when H_a is true.
  - $\mathrm{T}_{\mathrm{n}-1}^{*}$ follows a known distribution after standardization.
- When the data are from normal distribution, the test statistics follows T distribution.

3 Calculating P-value

- Our T test statistic is calculated to be:

$$T^*_{24} = \frac{5.9 - 5.75}{\frac{0.75}{\sqrt{25}}} = \frac{0.15}{0.15} = 1$$

  - Therefore, P-value = $P(T > 1)$

- A p-value is the chance of observing a value of test statistic that is at least as bizarre as 1 under H_0.
- A small p-value indicates that 1 is bizarre under H_0.

P-value based on T table

	PROPORTION IN ONE TAIL				
	0.25	0.10	0.05	0.025	0.01
	PROPORTION IN TWO TAILS COMBINED				
df	0.50	0.20	0.10	0.05	0.02
1	1.000	3.078	6.314	12.706	31.821
2	0.816	1.886	2.920	4.303	6.965
3	0.765	1.638	2.353	3.182	4.541
23	0.685	1.319	1.714	2.069	2.500
24	0.685	1.318	1.711	2.064	2.492
25	0.684	1.316	1.708	2.060	2.485

- Since we have a one tail test, our T-value = 1 is between 0.685 and 1.318. This implies that P-value is between 0.1 and 0.25.

4 Make rejection decision

- If our p-value is less than α, then we say that 1 is not likely under H_0 and therefore, we reject H_0.
- If our p-value is no less than α, we say that we do not have enough evidence to reject H_0.
- α is threshold to determine whether p-value is small or not. The default is 0.05. In statistics, it's called significance level.

Decision and Conclusion

- *Rejection decision:* we would say we fail to reject H_0, since p-value is between .1 and .25 which is greater than .05.
- *Conclusion:* there is insufficient evidence to indicate that $\mu > 5.75$.
- Does this mean we support that $\mu = 5.75$?

Conclusions

- While we did not have enough evidence to indicate $\mu > 5.75$; we are not stating that $\mu = 5.75$
- There could be a number of reasons why we did not have enough evidence
  - sample is not representative
  - not having a large enough sample size
  - incorrect assumptions
- While it is a possibility that $\mu = 5.75$, our conclusion does not reflect that possibility.

Discussions

- We can test many other hypothesis under the same framework.

$$H_0 : \mu_1 - \mu_2 = 0 \;\; v.s. \; H_a : \mu_1 - \mu_2 > 0$$

$$H_0 : \sigma^2 = \sigma_0^2 \;\; v.s. \; H_a : \sigma^2 \neq \sigma_0^2$$

- Different test statistics can follow different distributions under H_0.
- Since T-test require the data to be normally distributed, we need a new test for non-normal data.

- The End!
- Thank you!