STAT 4202

Chapter 11 – Interval Estimation

1

Interval Estimation

Goal: Come up with an interval, $[\hat{\theta}_1, \hat{\theta}_2]$ such that we can confidently claim that θ is inside this interval and at the same time the length of this interval is as small as possible.

This is called the interval estimation problem.

2

Definition

Let $\widehat{\Theta}_1$ and $\widehat{\Theta}_2$ be two statistics (functions of $X_1, X_2, \ldots, X_n$). Let $\hat{\theta}_1 = \widehat{\Theta}_1(x_1, x_2, \ldots, x_n)$ and $\hat{\theta}_2 = \widehat{\Theta}_2(x_1, x_2, \ldots, x_n)$ be the values of $\widehat{\Theta}_1$ and $\widehat{\Theta}_2$ when $x_1, \ldots, x_n$ are observed.

If

$$P\left(\widehat{\Theta}_1 < \theta < \widehat{\Theta}_2\right) = 1 - \alpha$$

For some $0 < \alpha < 1$, we refer to the interval $\left[\hat{\theta}_1, \hat{\theta}_2\right]$ as a $100(1-\alpha)\%$ confidence interval for θ.

The value $100(1-\alpha)\%$ is called the **degree of confidence**, and $\hat{\theta}_1$ and $\hat{\theta}_2$ are called the **upper** and **lower confidence limits**.

3

Some Examples

When $\alpha =$ ________, we have a________% confidence interval

When $\alpha =$ ________, we have a________% confidence interval

Given a value of α, we want to design a confidence interval with $E\left(\left|\widehat{\Theta}_2 - \widehat{\Theta}_1\right|\right)$ as small as possible.

4

STAT 4202

Section 11.2 – Estimation of Means

5

Estimation of Means

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. How do we construct a confidence interval for μ?

We will use the fact that

$$Z \coloneqq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

6

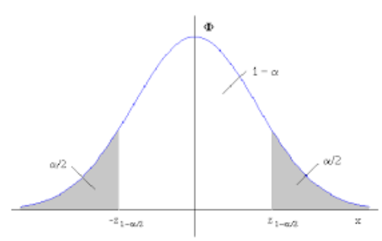
Estimation of Means

7

Estimation of Means

We want to pick $z_{\alpha/2}$ such that

$$P\left(|Z| < z_{\alpha/2}\right) = 1 - \alpha$$



8

Theorem 11.2

If the value of σ is known, then

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

Is a $100(1-\alpha)\%$ confidence interval for μ.

9

Theorem 11.1

$$P\left(|\bar{X} - \mu| < z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

How does the length of the interval change as α changes?

10

Example 1

A random sample of size $n = 20$ is taken from a normal distribution. The variance from the distribution is $\sigma^2 = 225$, and the sample mean is $\bar{x} = 64.3$. Construct a 95% confidence interval for the population mean μ.

11

Example 1

12

Example 1

The table lists the area between the median value (0) and the z value constructed by adding the value in the first column of the given cell to the value in the first row of the column of the given cell

Table III: Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4988
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Also, for $z = 4.0$, 5.0, and 6.0, the probabilities are 0.49997, 0.4999997, and 0.499999999.

13

Example 1

14

A Good Applet for Confidence Intervals

https://digitalfirst.bfwpub.com/stats_applet/stats_applet_4_ci.html

15

When σ is not known

If we do not know σ, but our sample size is at least 30 ($n \geq 30$), then

$$\left[\bar{X} - z_{\alpha/2}\frac{s}{\sqrt{n}}\ , \bar{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right]$$

Where s is the sample standard deviation

Is an appropriate $100(1-\alpha)\%$ confidence interval for μ.

16

Example 2

Given the following summary statistics:

$$n = 36 \quad \bar{x} = 19.92 \quad s = 5.73$$

Construct a 95% confidence interval for μ.

17

What about when n < 30 and σ is unknown

In this case, we would need to assume that the data sampled are drawn from a normal distribution. Then,

$$T \coloneqq \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

Follows a T-distribution with $n - 1$ degrees of freedom. We also write

$$T \sim t_{n-1}$$

18

Theorem 11.3

The confidence interval for a population mean μ of a normal distribution (when σ is unknown) is given by

$$\left[\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right]$$

Where $P\left(T \leq t_{\frac{\alpha}{2},n-1}\right) = 1 - \alpha$

19

Example 3

Given the following summary statistics

$$n = 12 \quad \bar{x} = 66.3 \quad s = 8.4$$

Calculate a 95% confidence interval for μ, assuming the data were drawn from a normal distribution.

20

Example 3 – Finding the t-statistic

Table IV: Values of $t_{\alpha,\nu}$ †

ν	$\alpha = .10$	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$	ν
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
inf.	1.282	1.645	1.960	2.326	2.576	inf.

† Based on Richard A. Johnson and Dean W. Wichern, *Applied Multivariate Statistical Analysis*, 2nd ed., © 1988, Table 2, p. 592. By permission of Prentice Hall, Upper Saddle River, N.J.

21

Example 3

Given the following summary statistics

$$n = 12 \quad \bar{x} = 66.3 \quad s = 8.4$$

Calculate a 95% confidence interval for μ, assuming the data were drawn from a normal distribution.

22

Summary of Confidence Intervals for μ

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

If σ is known

$$\left[\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right]$$

If σ is unknown

These intervals are exact, but require the data to come from a normal distribution

$$\left[\bar{X} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right]$$

If σ is unknown, $n \geq 30$

This is an approximate interval, but applies to any population distribution

23