STATISTICS IS THE GRAMMAR OF SCIENCE

PROBABILITY AND STATISTICS

LECTURE – 25 & 26

INFERENTIAL STATISTICS

CONFIDENCE INTERVALS

PREPARED BY
HAZBER SAMSON
FAST NUCES ISLAMABAD

CONFIDENCE INTERVAL

We estimate population parameters by using either point estimates or interval estimates.

POINT ESTIMATE The statistic, computed from sample information, that estimates a population parameter. Sample measures (statistics) are used to estimate population measures (parameters). These statistics are called **estimators**.

INTERVAL ESTIMATE An interval estimate of a parameter is an interval or a range of values used to estimate the parameter. This estimate may or may not contain the value of the parameter being estimated.

CONFIDENCE LEVEL The confidence level of an interval estimate of a parameter is the probability that the interval estimate will contain the parameter, assuming that a large number of samples are selected and that the estimation process on the same parameter is repeated. It is usually denoted by $1-\alpha$. The probability that the interval does not contain the parameter is denoted by α.

CONFIDENCE INTERVAL A confidence interval is a specific interval estimate of a parameter determined by usingndata obtained from a sample and by using the specific confidence level of the estimate.

A **confidence interval** consists of a range of values together with a percentage that specifies how confident we are that the parameter lies in the interval.

STANDARD ERROR A value that measures the spread of the sample means around the population mean. The standard error is reduced when the sample size is increased.

CONFIDENCE INTERVAL A range of values constructed from sample data so that the population parameter is likely to occur within that range at a specified probability. The specified probability is called the *level of confidence*.

Confidence Interval General Format

$$\text{Point estimate} \pm (\text{Critical value})(\text{Standard error})$$

A range of values used to estimate a population parameter is known as interval estimation or *estimation by confidence interval* and the interval (a,b) that will include the population parameter with a high probability $(eg\ 0.90, 0.95\ or\ 0.99)$ is known as *confidence interval.*

The $90\%, 95\%\ or\ 99\%$ confidence interval shows that our computed interval does in fact contain the unknown *population parameter*. The limit $'a'\ and\ 'b'$ are called lower and upper *confidence limits* of the interval, the probability $0.90, 0.95\ or\ 0.99$ is called the *confidence coefficient or level of confidence* and is denoted by $1-\alpha$.

CONFIFENCE INTERVALS FOR MEAN

CONFIDENCE INTERVAL FOR MEAN WHEN SIGMA IS KNOWN

When σ is **known** then we use z-stats

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

- The $(1-\alpha)$ 100% **confidence interval** for μ is

$$\bar{x} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

- The $(1-\alpha)$ 100% **confidence interval upper bound** for μ is

$$\mu < \bar{x} + Z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}$$

- The $(1-\alpha)$ 100% **confidence interval lower bound** for μ is

$$\bar{x} - Z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}} < \mu$$

CONFIDENCE INTERVAL FOR MEAN WHEN SIGMA IS UNKNOWN

- When σ **is unknown** then we use t-statistic

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \quad with \;\; d.f = n-1$$

- The $(1-\alpha)$ 100% **confidence interval** for μ is

$$\bar{x} - t_{\frac{\alpha}{2}(n-1)} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}(n-1)} \cdot \frac{S}{\sqrt{n}}$$

- The $(1-\alpha)$ 100% **confidence interval upper bound** for μ is

$$\mu < \bar{x} + t_{\alpha(n-1)} \cdot \frac{S}{\sqrt{n}}$$

- The $(1-\alpha)$ 100% **confidence interval lower bound** for μ is

$$\bar{x} - t_{\alpha(n-1)} \cdot \frac{S}{\sqrt{n}} < \mu$$

CRITICAL VALUES OF Z

Significance Level (α)	Confidence Level $(1-\alpha)$	One-Tailed Test	Two-Tailed Test
$\alpha = 0.01$	99%	$Z_{\alpha} = 2.33$	$Z_{\alpha/2} = 2.58$
$\alpha = 0.02$	98%	$Z_{\alpha} = 2.05$	$Z_{\alpha/2} = 2.33$
$\alpha = 0.03$	97%	$Z_{\alpha} = 1.88$	$Z_{\alpha/2} = 2.17$
$\alpha = 0.04$	96%	$Z_{\alpha} = 1.75$	$Z_{\alpha/2} = 2.05$
$\alpha = 0.05$	95%	$Z_{\alpha} = 1.65$	$Z_{\alpha/2} = 1.96$
$\alpha = 0.06$	94%	$Z_{\alpha} = 1.55$	$Z_{\alpha/2} = 1.88$
$\alpha = 0.07$	93%	$Z_{\alpha} = 1.48$	$Z_{\alpha/2} = 1.81$
$\alpha = 0.08$	92%	$Z_{\alpha} = 1.41$	$Z_{\alpha/2} = 1.75$
$\alpha = 0.09$	91%	$Z_{\alpha} = 1.34$	$Z_{\alpha/2} = 1.70$
$\alpha = 0.10$	90%	$Z_{\alpha} = 1.28$	$Z_{\alpha/2} = 1.65$

STUDENTS T-DISTRIBUTION TABLE

df	α = 0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.706	31.821	63.656	318.289	636.578
2	1.886	2.92	4.303	6.965	9.925	22.328	31.6
3	1.638	2.353	3.182	4.541	5.841	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.61
5	1.476	2.015	2.571	3.365	4.032	5.894	6.869
6	1.44	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.86	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.25	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.93	4.318
13	1.35	1.771	2.16	2.65	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.14
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.12	2.583	2.921	3.686	4.015
17	1.333	1.74	2.11	2.567	2.898	3.646	3.965
18	1.33	1.734	2.101	2.552	2.878	3.61	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.85
21	1.323	1.721	2.08	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.5	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.06	2.485	2.787	3.45	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.689
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.66
30	1.31	1.697	2.042	2.457	2.75	3.385	3.646
60	1.296	1.671	2	2.39	2.66	3.232	3.46
120	1.289	1.658	1.98	2.358	2.617	3.16	3.373
∞	1.282	1.645	1.96	2.326	2.576	3.091	3.291