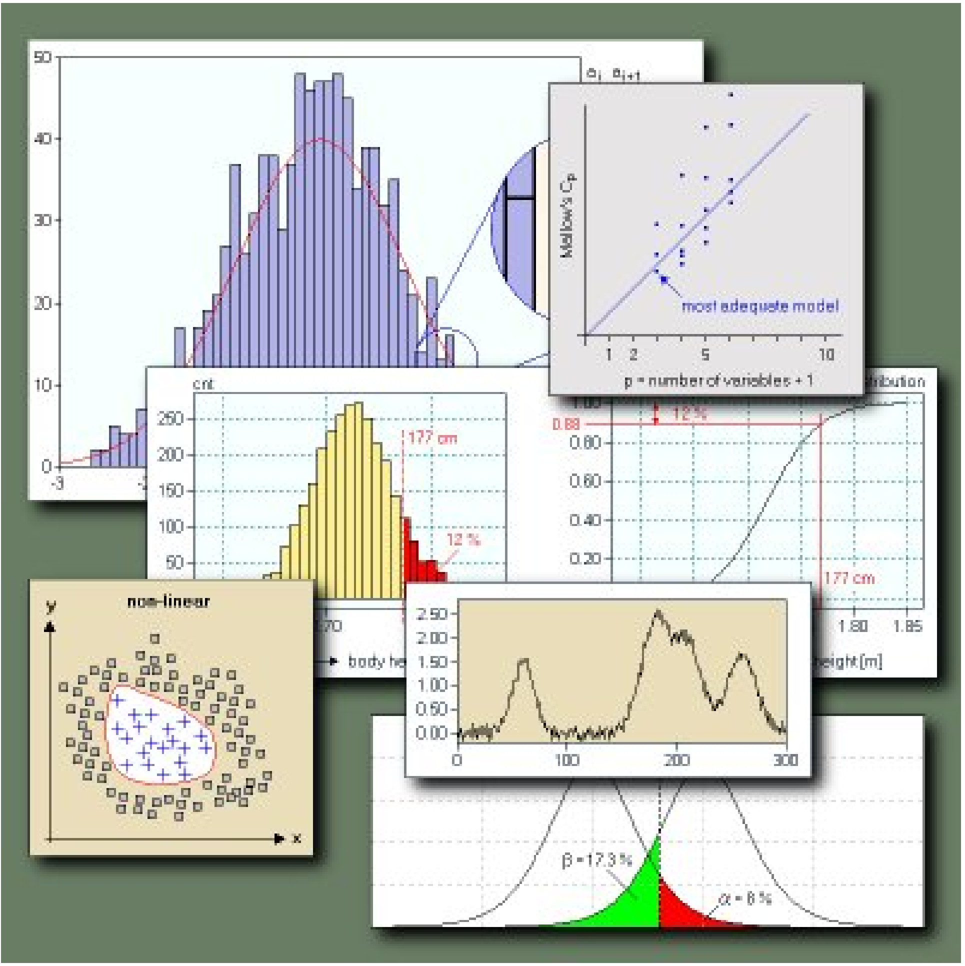


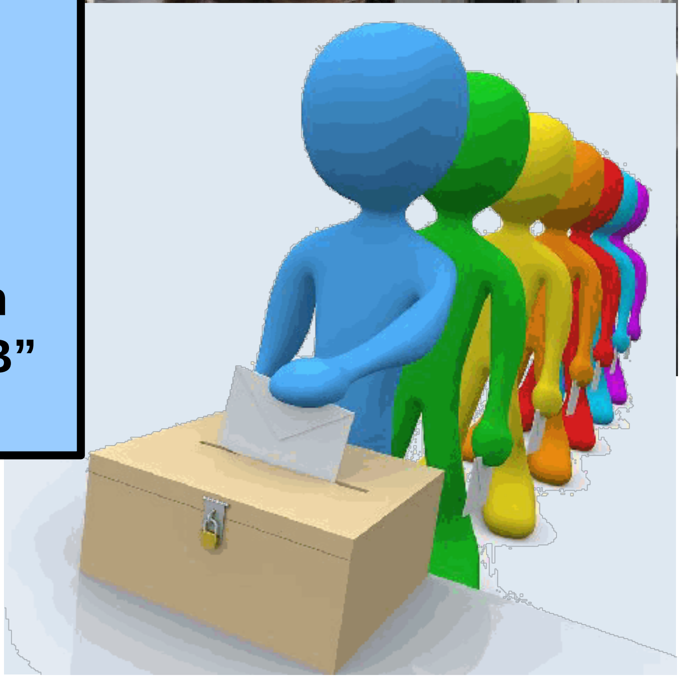
Confidence Interval Estimation

Learning Objectives

You will learn:

- To construct and interpret ***confidence interval estimates*** for the population mean and the population proportion (**μ** and **π**) .
- How to determine the sample size necessary to develop a ***confidence interval*** for the population mean or population proportion

Confidence Interval Estimate

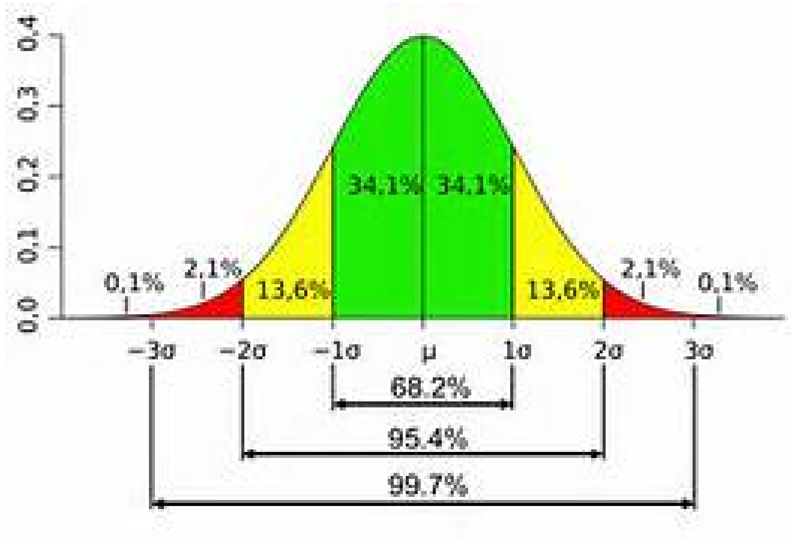


- I am 95% confident that the true average income of U.S. appliance factory workers is between $51,000.00 and $56,000.00 per year.

- I am 90% confident that the true proportion of potential voters supporting candidate “B” is between 35% and 38%.

EXAMPLES

Outline



1. **Confidence Intervals for the Population Mean, μ**
   - when Population Standard Deviation **σ** is <u>known</u>
   - when Population Standard Deviation **σ** is <u>unknown</u>
2. **Confidence Intervals for the Population Proportion, π**
3. **Determining the Required Sample Size, n**

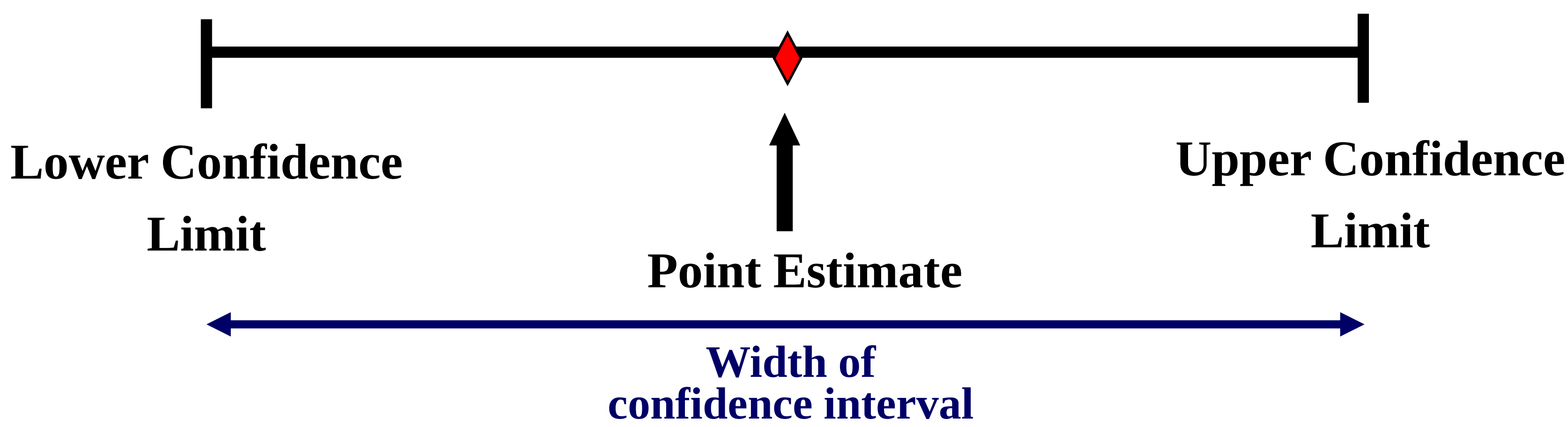
Point Estimates

The starting point for developing the confidence interval for μ or π

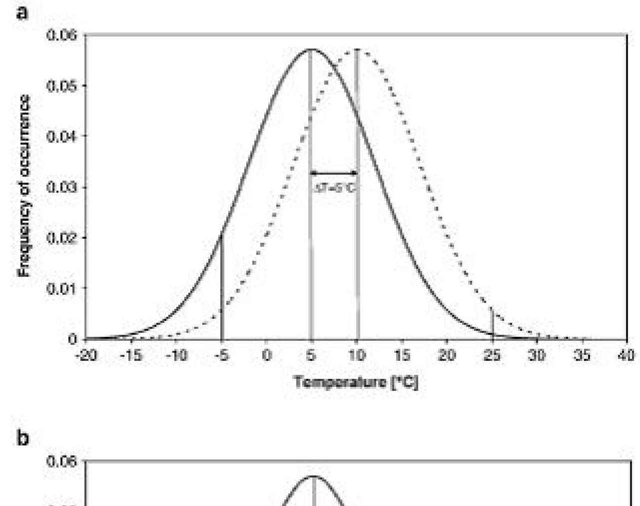
The value of a single sample statistic: mean or proportion

A range of numbers constructed around the point estimate

- A **point estimate** is a single number. For the population mean and population standard deviation, a point estimate is the sample mean and sample standard deviation.
- A **confidence interval** provides additional information about variability



Confidence Interval Estimates



- **A confidence interval gives a range estimate of values:**
  - Takes into consideration variation in sample statistics from sample to sample*
  - Based on all the observations from <u>one</u> sample
  - Gives information about closeness to unknown population parameters
  - Stated in terms of level of confidence
    - Example: **95%** confidence, **99%** confidence
    - Can **never** be **100%** confident

Recall that the sample mean will vary from sample to sample

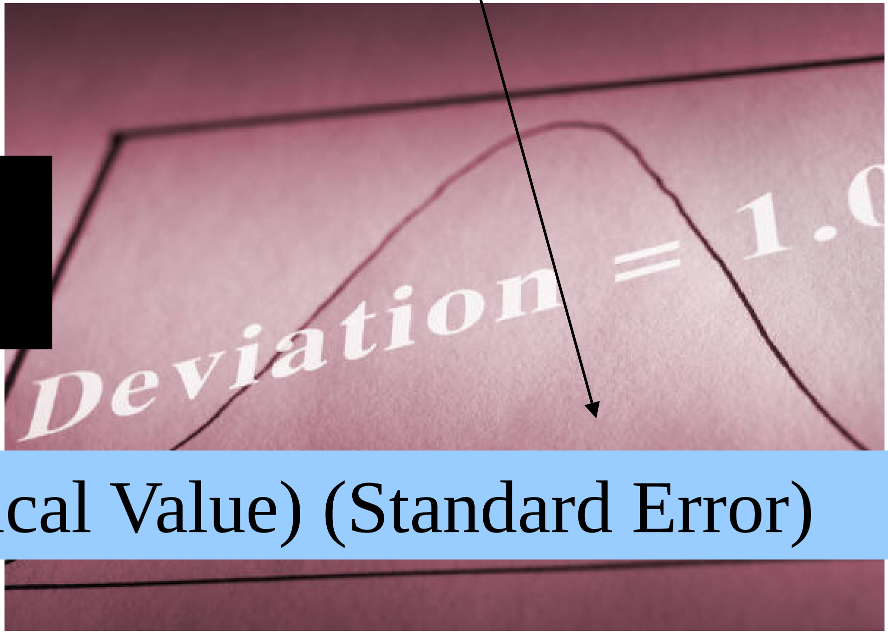
Confidence Interval Estimates

The general formula for
all confidence intervals is:

σ / √ n or s / √ n

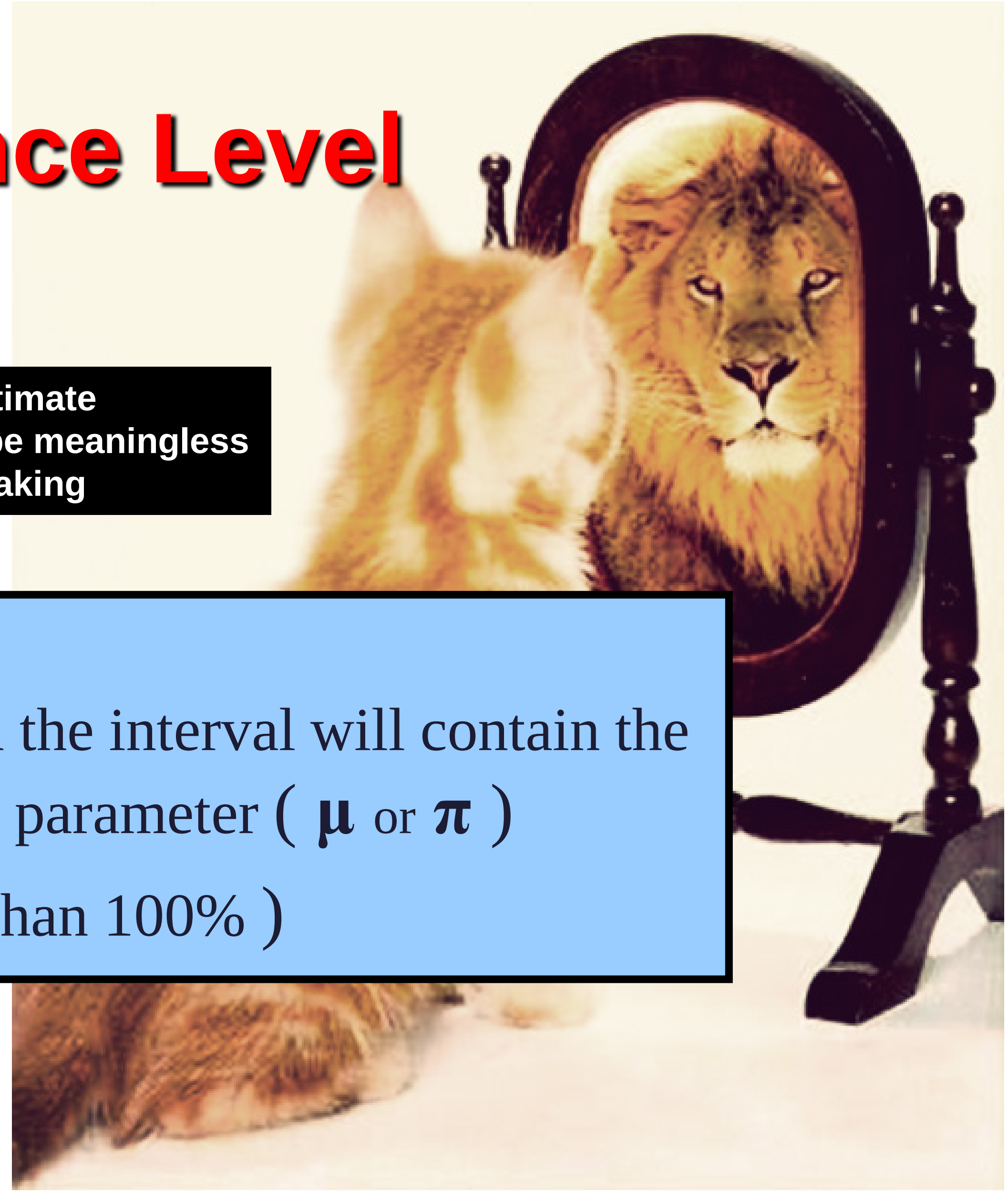
Sample Mean
or
Sample Proportion

The “z” or “t”
Critical Value



Point Estimate ± (Critical Value) (Standard Error)

Confidence Level



A confidence interval estimate of 100% would be so wide as to be meaningless for practical decision making

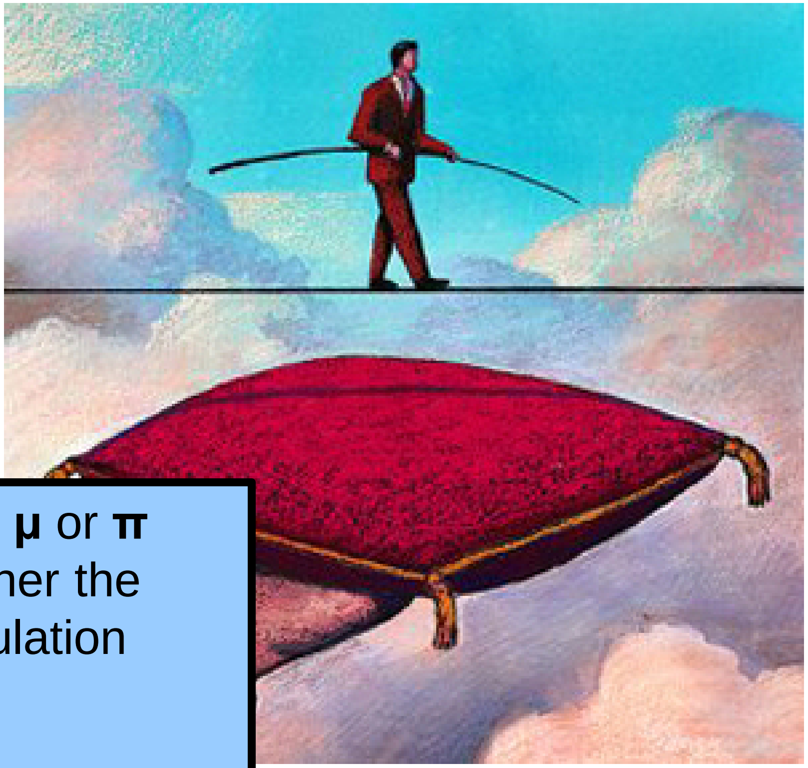
Confidence Level

- Confidence in which the interval will contain the unknown population parameter (μ or π)
- A percentage (less than 100%)

Confidence Level

- Suppose confidence level = **95%**
- Also written $(1 - \alpha) =$ **.95**
- A relative frequency interpretation:
  - In the long run, **95%** of all the confidence intervals that **<u>could be</u>** constructed will contain the unknown true parameter (μ , π)
- A specific interval either will contain or will not contain the true parameter

The Level of Significance (α)



- Because we only select one sample, and **μ** or **π** are unknown, we never really know whether the confidence interval includes the true population mean or proportion, or not.

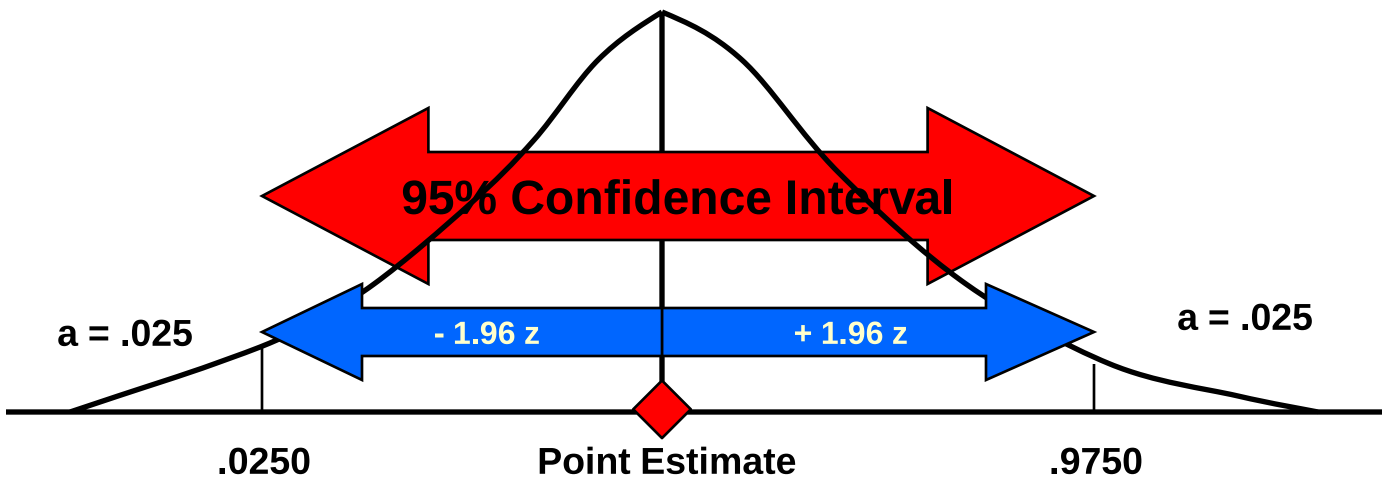
- The ***level of significance***, or "**α**" **risk** is the chance we take that the true population parameter is not contained in the confidence interval.

- Therefore, a 95% confidence interval would have an "α" of 5%

The Level of Significance (α)

If α = .05, then each tail has .025 area

The critical values of "z" that define the "α" areas are -1.96 and + 1.96

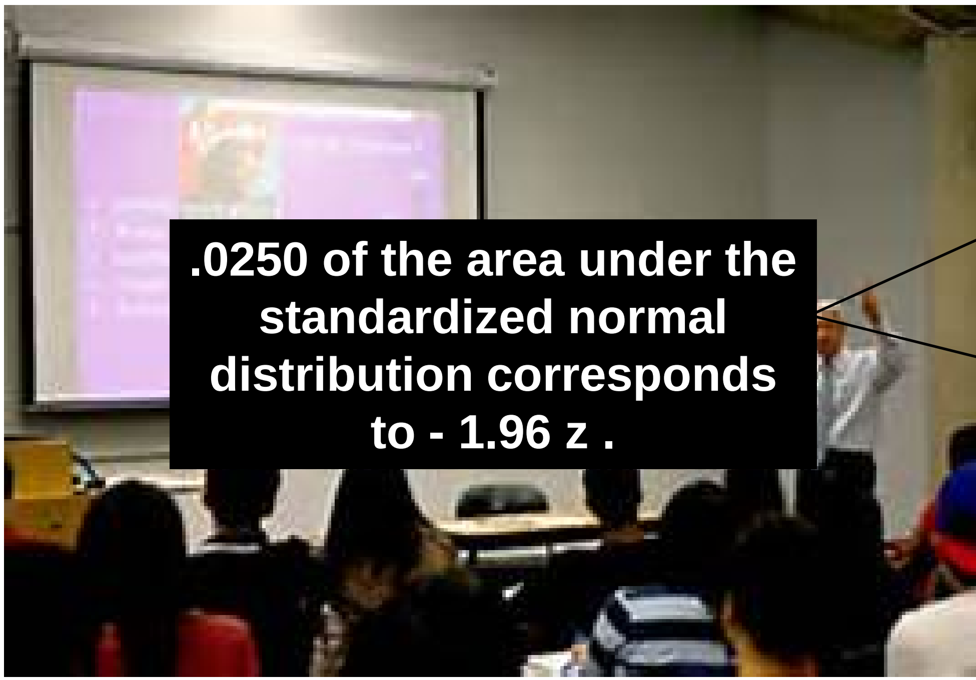


Z	.06
- 1.9	.0250

Z	.06
+ 1.9	.9750

" α " is the proportion in the tails of the sampling distribution that is outside the established confidence interval.

The ‘z’ Table



.0250 of the area under the standardized normal distribution corresponds to - 1.96 z .

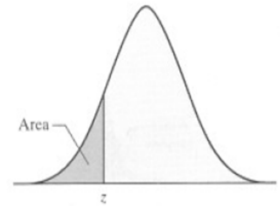


TABLE II

Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	[illegible]	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	[illegible]	[illegible]	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	[illegible]	[illegible]	[illegible]	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	[illegible]	0.0197	[illegible]	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	[illegible]	0.0250	[illegible]	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	[illegible]	0.0314	[illegible]	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	[illegible]	[illegible]	[illegible]	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	[illegible]	[illegible]	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4027	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

The ‘z’ Table

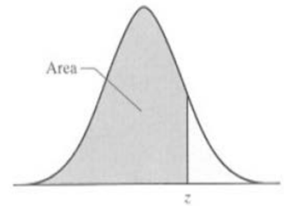
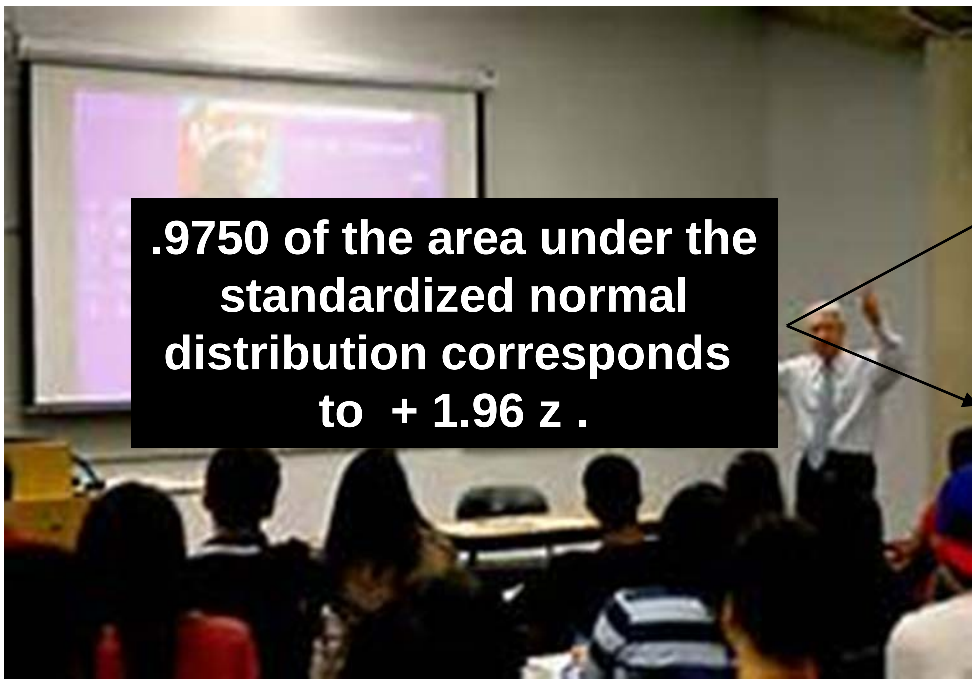


TABLE II *(continued)*

Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	[illegible]	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	[illegible]	[illegible]	[illegible]	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	[illegible]	[illegible]	[illegible]	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	[illegible]	0.9686	[illegible]	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	[illegible]	0.9750	[illegible]	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	[illegible]	0.9803	[illegible]	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	[illegible]	[illegible]	[illegible]	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	[illegible]	[illegible]	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	[illegible]	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998



.9750 of the area under the standardized normal distribution corresponds to + 1.96 z .

Confidence Interval for μ (when σ is Known)

Assumptions

- Population standard deviation **σ** is **known**
- Population is **normally** distributed
- If population is not normal, use large sample (n > 30)

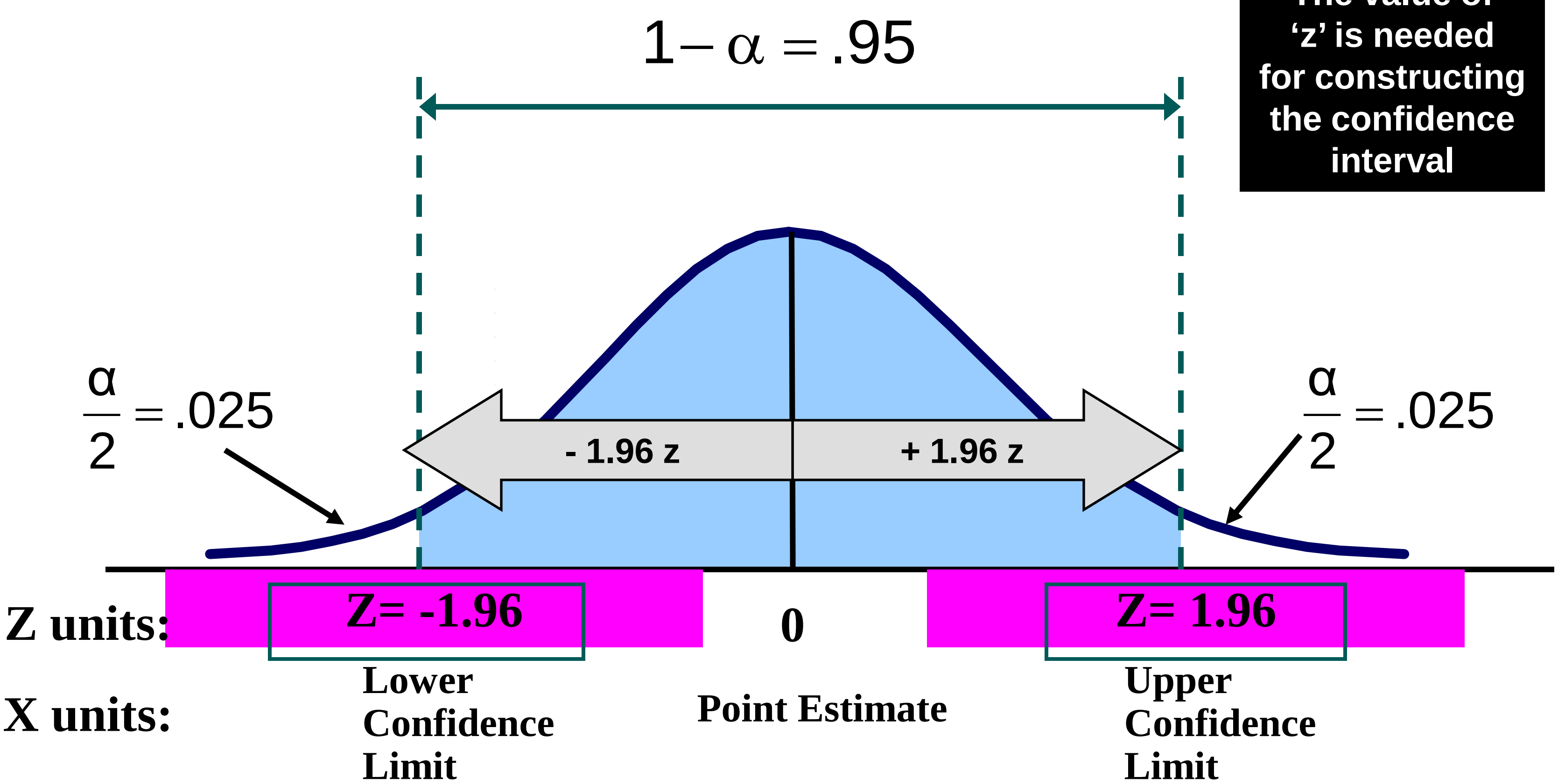
Confidence interval estimate:

$$\overline{X} \pm Z \frac{\sigma}{\sqrt{n}}$$

(where Z is the standardized normal distribution *critical value* for a probability of α/2 in each tail)

Finding the Critical Value, Z

Consider a 95% confidence interval:



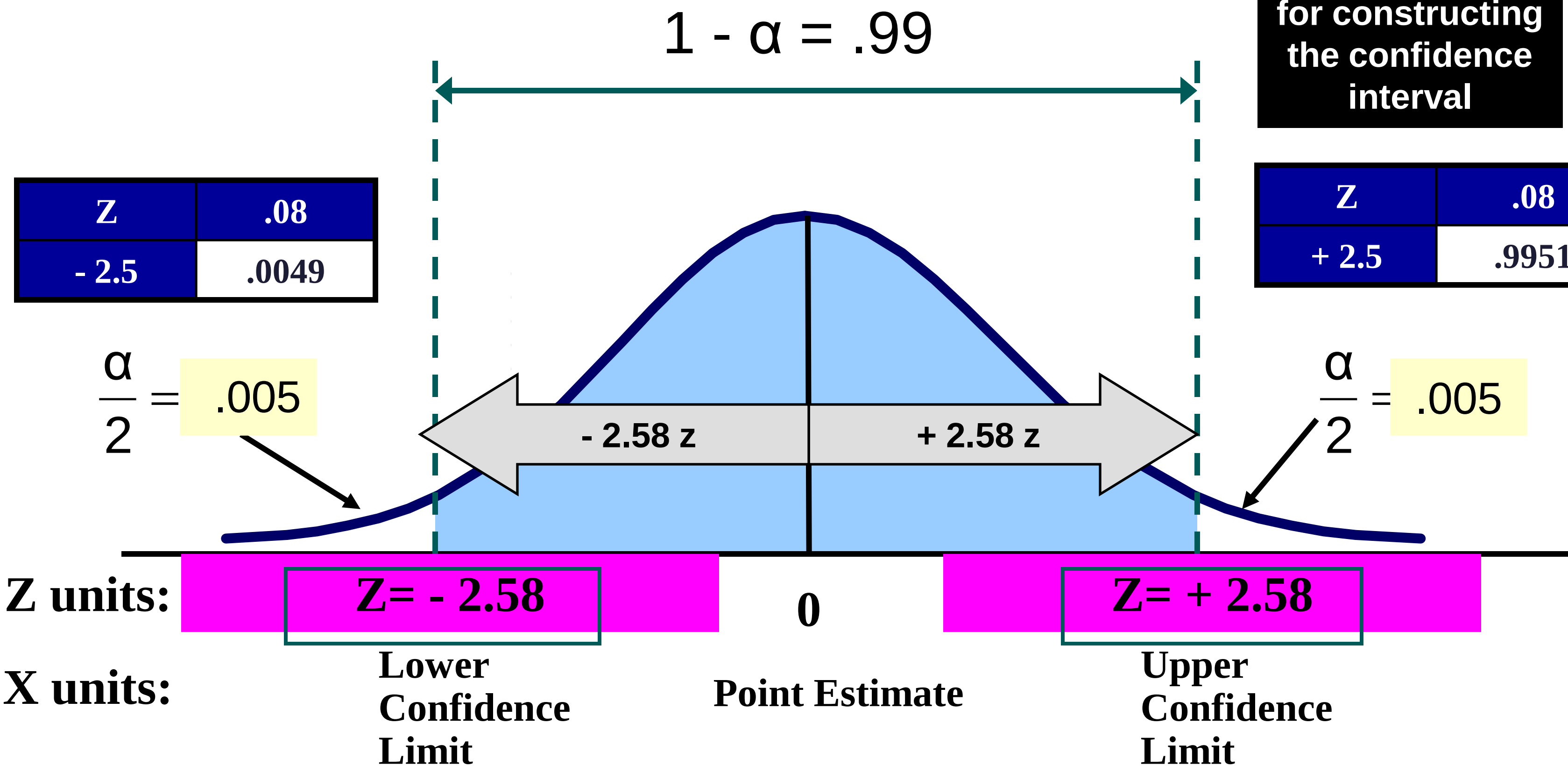
Finding the Critical Value, Z

Consider a 99% confidence interval:

The value of 'z' is needed for constructing the confidence interval

Z	.08
- 2.5	.0049

Z	.08
+ 2.5	.9951



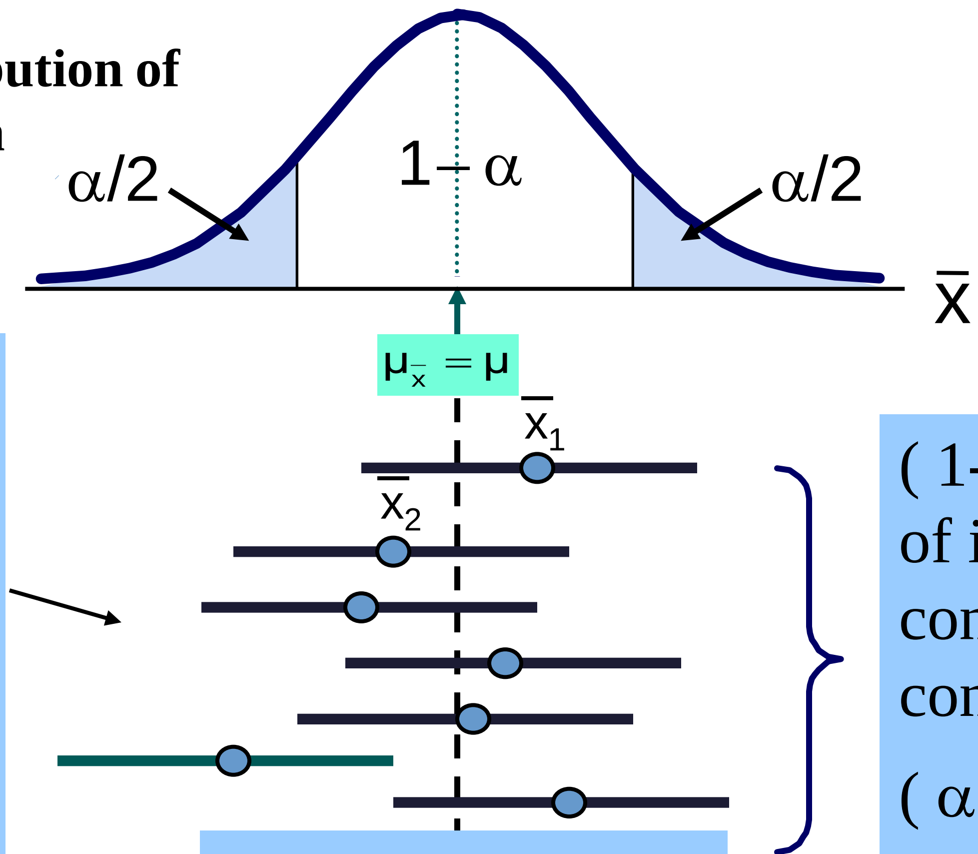
Finding the Critical Value, Z

Commonly used confidence levels are 90%, 95%, and 99%

Confidence Level	Confidence Coefficient	‘z’ Value	‘a’ / 2 Value
80%	.80	1.28	.1000
90%	.90	1.645	.0500
95%	.95	1.96	.0250
98%	.98	2.33	.0100
99%	.99	2.58	.0050
99.8%	.998	3.08	.0010
99.9%	.999	3.27	.0005

Intervals and Level of Confidence

Sampling Distribution of the Mean



Confidence Intervals

Intervals extend from

$$\bar{X} - Z\frac{\sigma}{\sqrt{n}}$$

to

$$\bar{X} + Z\frac{\sigma}{\sqrt{n}}$$

(1-α) x 100% of intervals constructed contain μ;

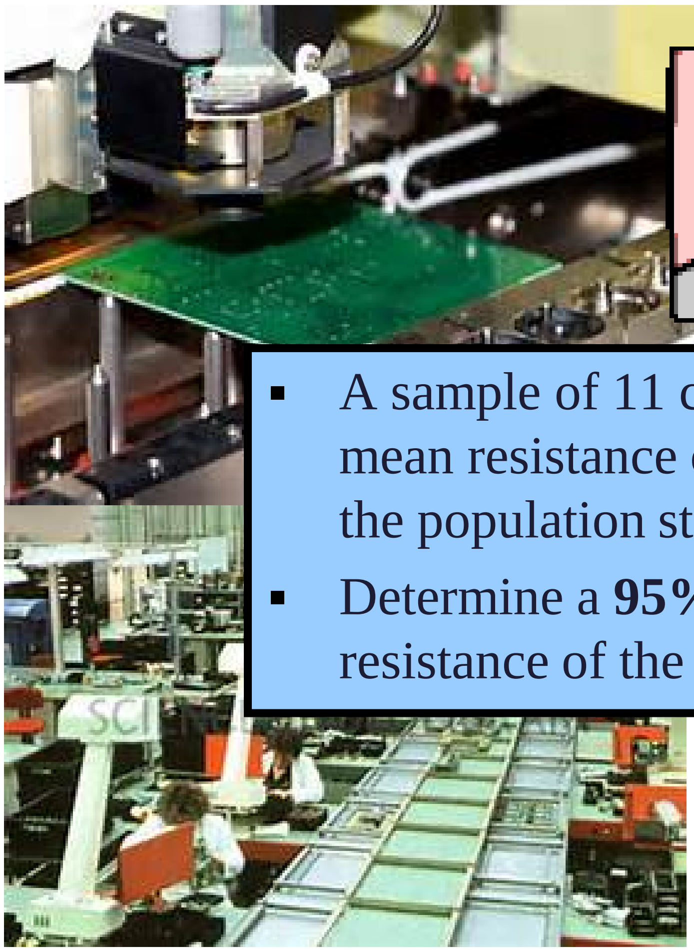
(α) x 100% do not.

Means vary from sample to sample

Accordingly, confidence intervals of the same percentage will have different lower and upper limits

The true population mean may not be in them at all !

Confidence Interval for μ (σ Known) Example



- A sample of 11 circuits from a large normal population has a mean resistance of **2.20** ohms. We know from past testing that the population standard deviation is **.35** ohms.
- Determine a **95%** confidence interval for the true mean resistance of the population.

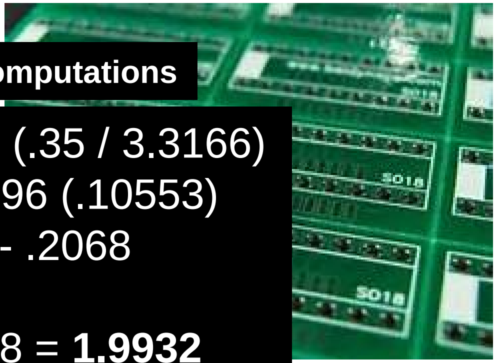
Confidence Interval for μ (σ Known) Example

$$\bar{X} \pm Z \frac{\sigma}{\sqrt{n}}$$

$$= 2.20 \pm 1.96\,(.35/\sqrt{11})$$

$$= 2.20 \pm .2068$$

$$(1.9932\,,\ 2.4068)$$



Detailed Computations

2.20 + / - 1.96 (.35 / 3.3166)
2.20 + / - 1.96 (.10553)
2.20 + / - .2068

2.20 - .2068 = **1.9932**

2.20 + .2068 = **2.4068**

- We are **95%** confident that the true mean resistance is between **1.9932** and **2.4068** ohms
- Although the true mean may or may not be in this interval, **95%** of intervals formed in this manner will contain the true mean

Confidence Interval for μ (σ Unknown)

- If the population standard deviation **σ** is unknown, we can substitute the **sample standard deviation**: (s) .
- This introduces extra uncertainty, since '**s**' is variable from sample to sample
- So we use the '***t***' ***distribution*** instead of the normal distribution

"S" is variable from sample to sample, and even more so, in very small samples

Confidence Interval for μ
(σ Unknown)

Actually, we rarely know the true population standard deviation. We must instead develop a confidence interval using the sample standard deviation

Assumptions

- Population standard deviation is unknown
- Population is normally distributed
- If population is not normal, use large sample (n > 30)

Use **Student's '*t*' Distribution**

The Confidence Interval Estimate: $$\overline{X} \pm t_{n\text{-}1} \frac{S}{\sqrt{n}}$$

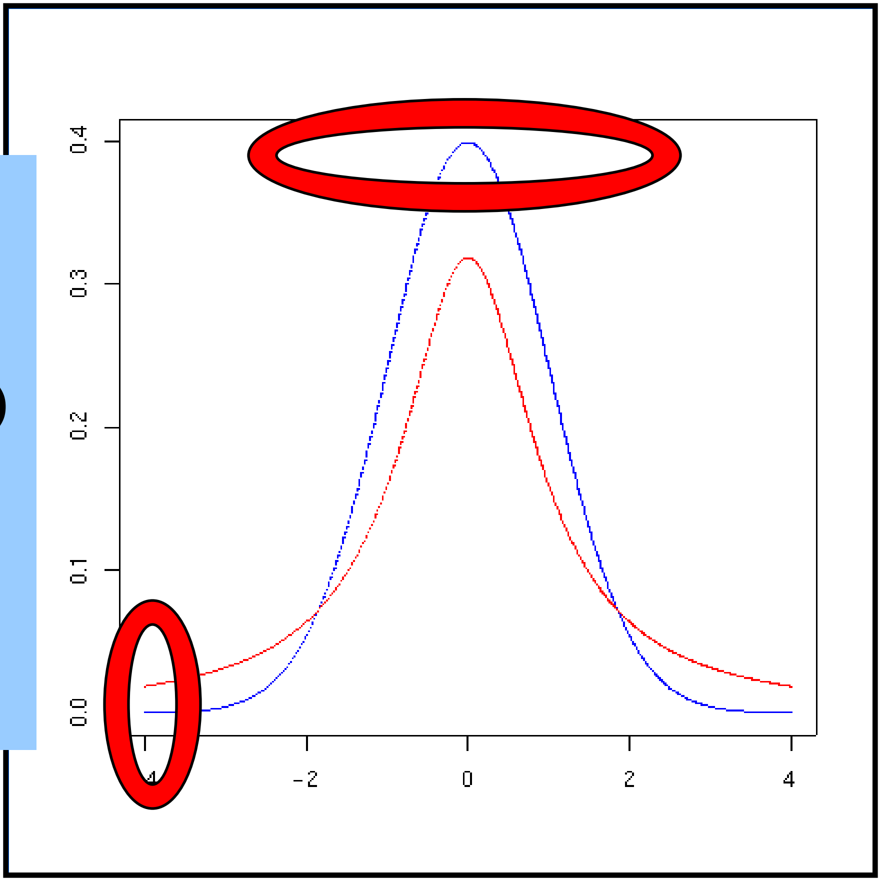
Sample Standard Deviation

Degrees of Freedom

(where t is the critical value of the t distribution with 'n-1' degrees of freedom and an area of α/2 in each tail)

Student ' *t* ' Distribution

- it is bell-shaped
- has fatter tails
- has flatter center (peak or "kurtosis")
- the sampling distributions of small samples follow the '*t* ' distribution
- the smaller the sample size taken, the more variable the '*t* ' distribution



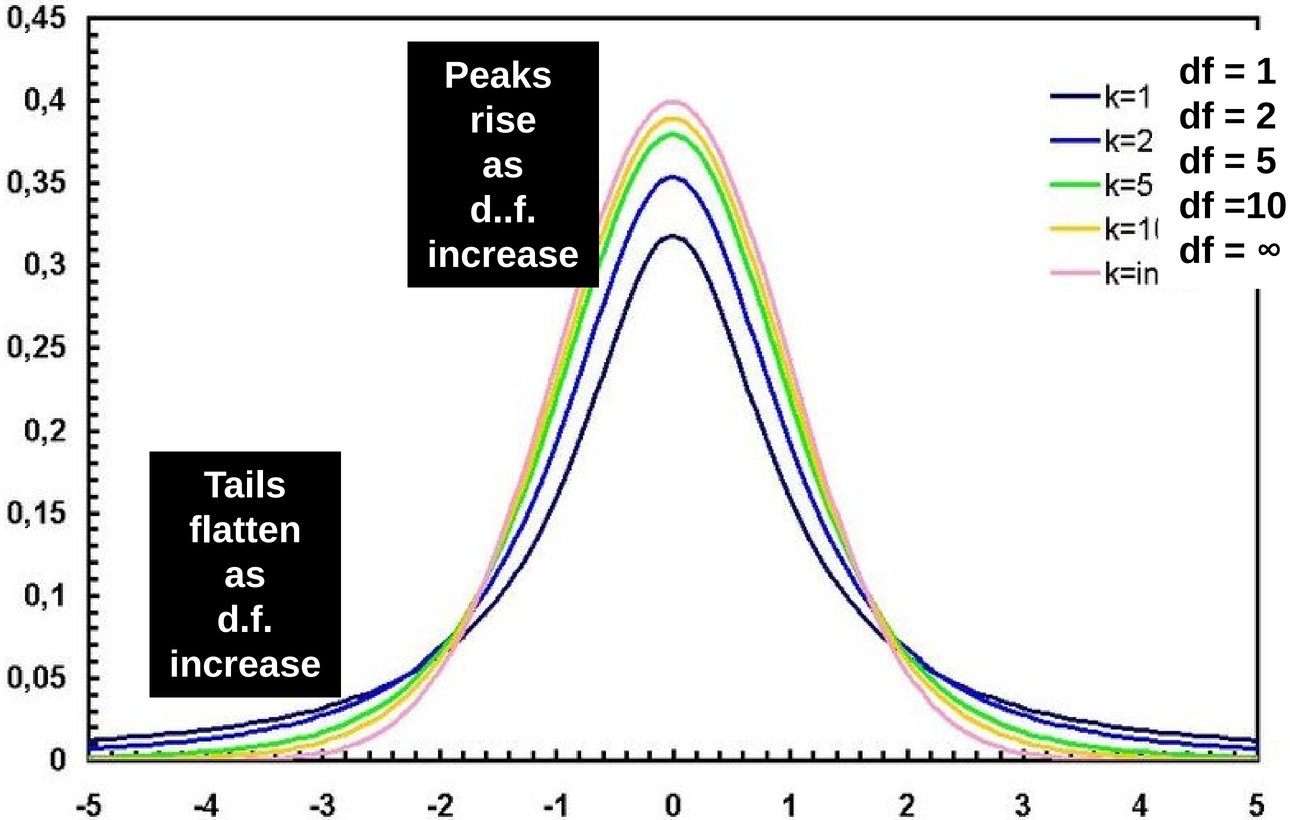
Student's ' *t* ' Distribution

d.f. = n - 1

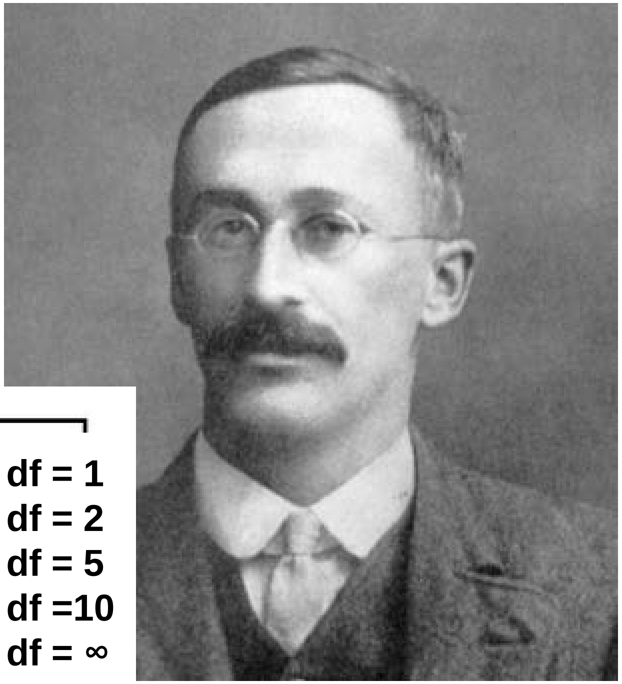
Sample size

- **The ' *t* ' value depends on *<u>degrees of freedom</u>* (d.f.)**
- **The number of observations in the sample that are free to vary after the sample mean has been calculated.**
- **The ' *t* ' distribution <u>changes</u> as the number of degrees of freedom changes.**
- **small samples have greater variability and the ' *t* ' reflects that variability when we are finding areas under the sampling distribution curve.**

Student's t Distribution



df = 1
df = 2
df = 5
df =10
df = ∞



William S. Gosset of the Guiness Brewery who published under the pen name "student" (1913)

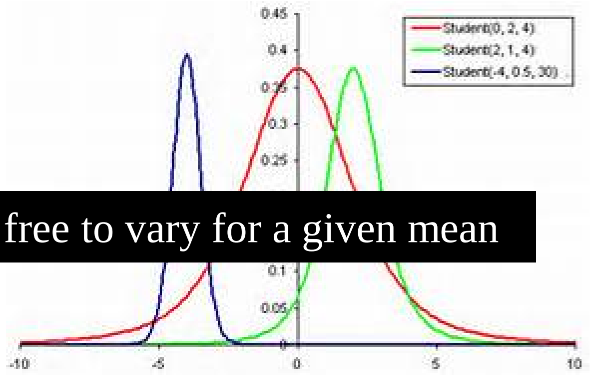
Degrees of Freedom

Concept: The number of observations that are free to vary after the sample mean has been calculated

Example: Suppose the **mean** of 3 numbers is **8.0**

- **Let $X_1 = 7$**
- **Let $X_2 = 8$**
- **What is X_3?**

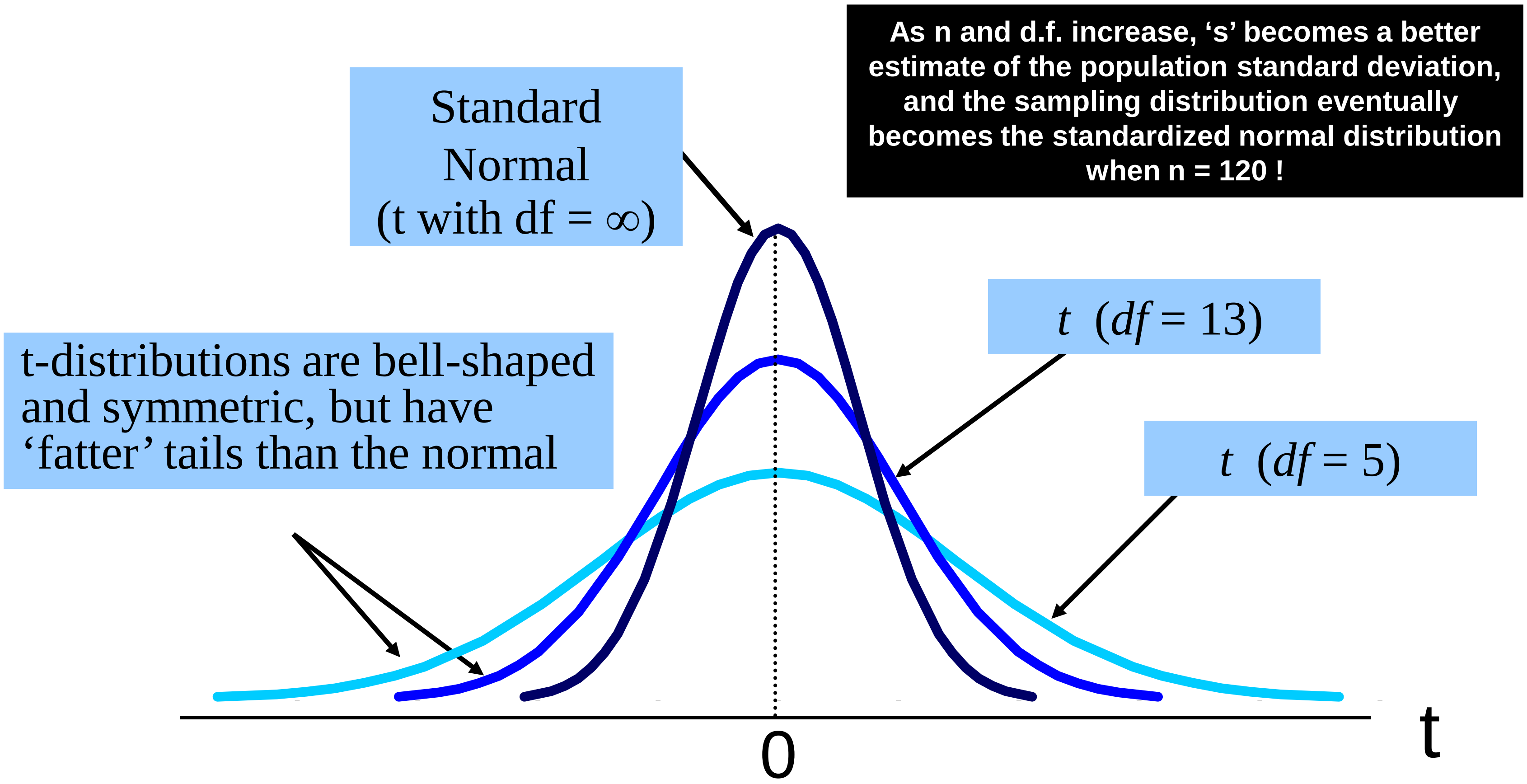
If the mean of these three values is 8.0,
then X_3 must be 9
(i.e., X_3 is not free to vary)



2 values can be any numbers, that is, free to vary, but the third is not free to vary for a given mean

Student's t Distribution

Note: 't' approaches 'Z' as 'n' increases



Student's t Table

It's set up for the upper tail area only !

	Upper Tail Area		
df	.25	.10	**.05**
1	1.000	3.078	6.314
2	0.817	1.886	**2.920**
3	0.765	1.638	2.353

For 90% Confidence Interval

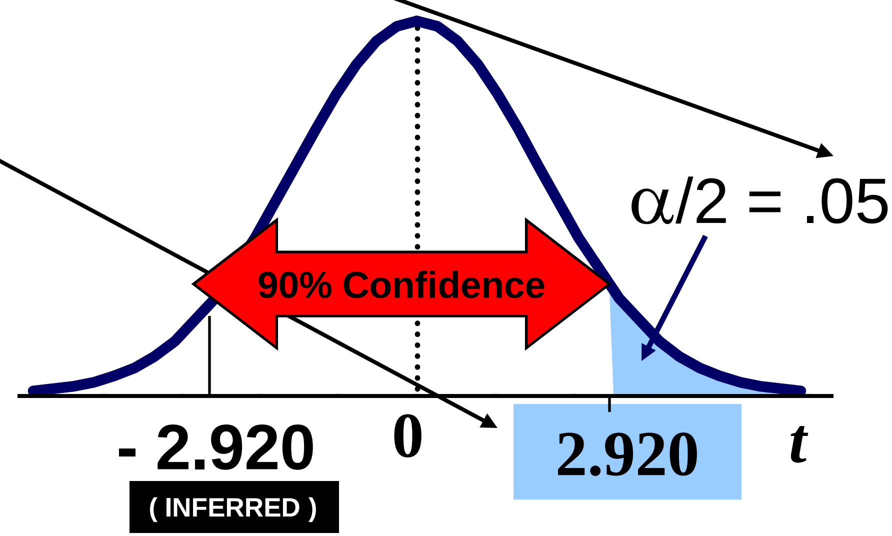
Let: n = **3**

df = n - 1 = **2**

α = **.10**

$\alpha/2$ =**.05**

The body of the table contains t values, **<u>not</u>** probabilities !

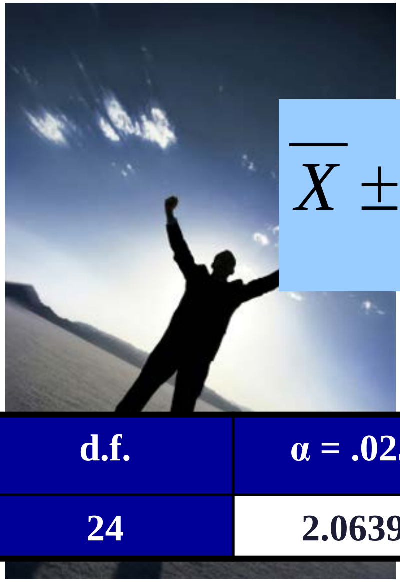


Confidence Interval for μ (σ Unknown) Example

A random sample of $n = 25$ has $\bar{X} = 50$ and $S = 8$. Form a 95% confidence interval for μ

- d.f. = $n - 1 = 24$, so

The confidence interval is



$$\bar{X} \pm t_{\alpha/2,\,\text{n-1}} \frac{S}{\sqrt{n}} = 50 \pm (2.0639) \frac{8}{\sqrt{25}}$$

d.f.	α = .025
24	2.0639

(46.698 , 53.302)

Detailed Calculations

50 + / - (2.0639)(8/5)
50 + / - (2.0639)(1.6)
50 + / - 3.30224

50 - 3.30224 = 46.698

50 + 3.30224 = 53.302

Using Confidence Intervals

Before calculating a confidence interval for μ or p there are three important **conditions** that you should check.

1) **Random**: The data should come from a well-designed random sample or randomized experiment.

2) **Normal:** The sampling distribution of the statistic is approximately Normal.

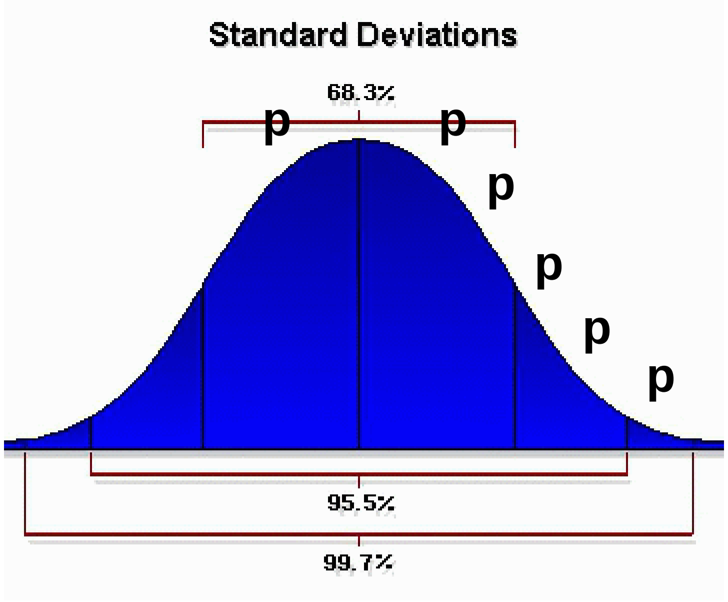
For means: The sampling distribution is exactly Normal if the population distribution is Normal. When the population distribution is not Normal, then the central limit theorem tells us the sampling distribution will be approximately Normal if n is sufficiently large ($n \geq 30$).

For proportions: We can use the Normal approximation to the sampling distribution as long as $np \geq 10$ and $n(1 - p) \geq 10$.

3) **Independent**: Individual observations are independent. When sampling without replacement, the sample size n should be no more than 10% of the population size N (the *10% condition*) to use our formula for the standard deviation of the statistic.

Confidence Intervals for the Population Proportion, π

- An confidence interval estimate for the population proportion (π) can be calculated.



'p' is the sample proportion

Recall, that if np => 5 and if n (1 - p) => 5, that is, if the sample size is large enough, we can approximate the proportion's *binomial distribution* with a *normally distributed* sampling distribution.

Confidence Intervals for the Population Proportion, π

Again, recall that the distribution of the sample proportion is approximately normal if the sample size is large, with standard deviation:

$$\sigma_{\mathrm{p}} = \sqrt{\frac{\pi(1-\pi)}{\mathrm{n}}}$$

We will estimate this with sample data:

(‘p’ is the sample proportion)

$$\left.\sqrt{\frac{p(1-p)}{n}}\;\right\}$$

Estimated σp

Confidence Intervals for the Population Proportion, π

Upper and lower confidence limits for the population proportion are calculated with the formula:

$$p \pm Z\sqrt{\frac{p(1-p)}{n}}$$

Alternately,

***p* +/- *Z* (estimated σp)**

where

- **Z** is the standardized normal value for the level of confidence desired
- **p** is the sample proportion
- **n** is the sample size

Confidence Intervals for the Population Proportion: Example

A random sample of 100 people shows that 25 have opened IRA's this year. Form a 95% confidence interval for the true proportion of the population who have opened IRA's.

Detailed Computations

.25 +/- 1.96 √ .1875 / 100

.25 +/- 1.96 √ .001875

.25 +/- 1.96 (.0433)

.25 +/- .084868

.25 + .084868 = .3349

.25 - .084868 = .1651

$$p \pm Z\sqrt{p(1-p)/n}$$

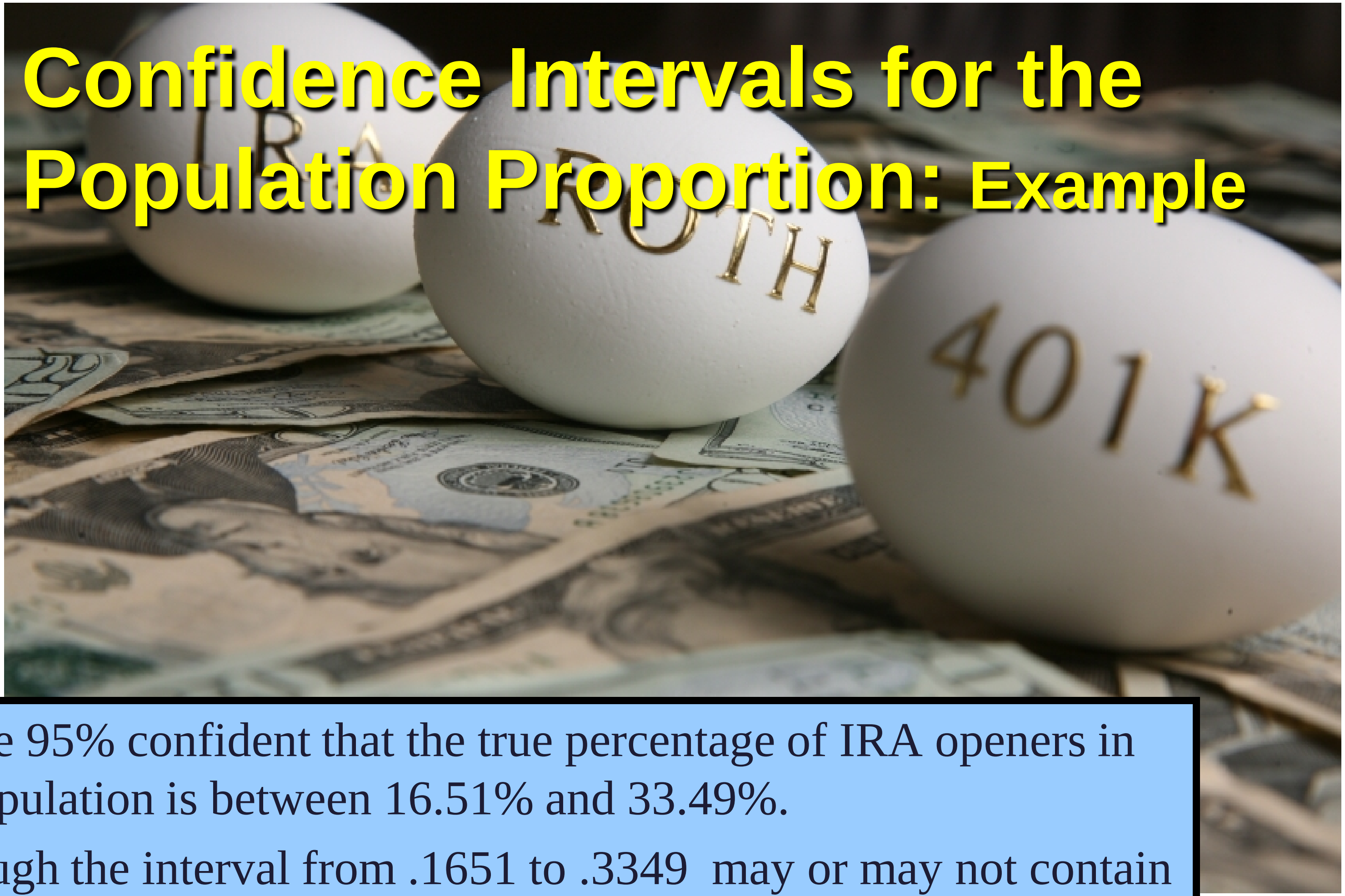
$$= 25/100 \pm 1.96\sqrt{.25(.75)/100}$$

$$= .25 \pm 1.96\,(.0433)$$

(0.1651 , 0.3349)

alternately, $0.1651 \leq \pi \leq 0.3349$

Confidence Intervals for the Population Proportion: Example



We are 95% confident that the true percentage of IRA openers in the population is between 16.51% and 33.49%.

Although the interval from .1651 to .3349 may or may not contain the true proportion, 95% of intervals formed from samples of size 100 in this manner will contain the true proportion.

Determining Sample Size

- The required sample size can be found to reach a desired margin of error (**e**) with a specified level of confidence (**1 - α**)
- The margin of error is also called ***sampling*** error or ***tolerated*** or ***allowable*** error.
  - the amount of imprecision in the estimate of the population parameter
  - the amount added and subtracted to the point estimate to form the confidence interval

We can determine sample size for the mean and the proportion

Determining Sample Size

- **To determine the required sample size for the mean, you must know:**
  - The desired level of confidence (1 - α), which determines the critical ‘***Z***’ value
  - The acceptable sampling error (margin of error), ‘***e***’
  - The standard deviation, ‘***σ***’

$$e = Z\frac{\sigma}{\sqrt{n}}$$

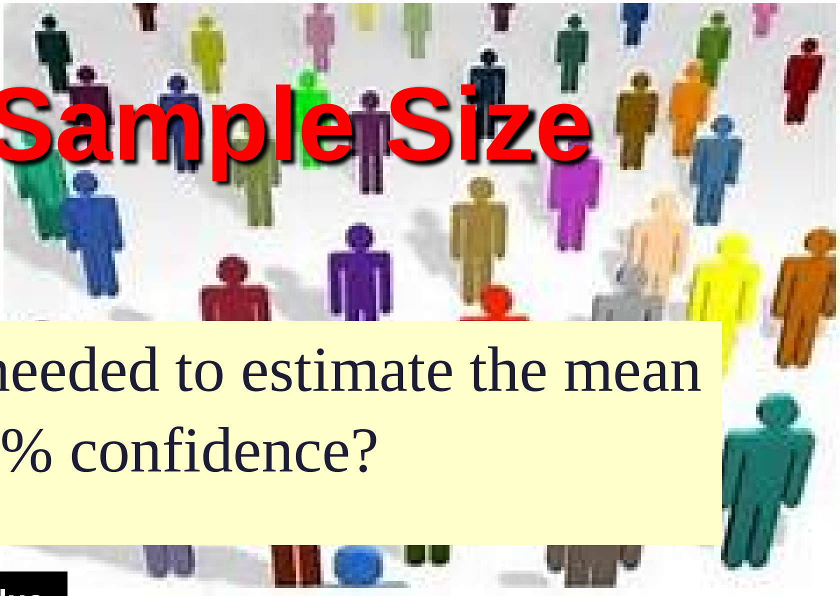
Now solve for n to get

$$n = \frac{Z^2\ \sigma^2}{e^2}$$

If unknown the **standard deviation**, σ, **estimate by:**

- past data
- educated guess
- estimate σ: [σ = range/4] This estimate is derived from the empirical rule stating that approximately 95% of the values in a normal distribution are within +/- 2σ of the mean, giving a range within which most of the values are located.

Determining Sample Size



If σ = 45, what sample size is needed to estimate the mean within ± 5 with 90% confidence?

interpolated Z value

$$n = \frac{Z^2\ \sigma^2}{e^2} = \frac{(1.645)^2(45)^2}{5^2} = 219.19$$

So the required sample size is ***n* = 220**

We always round up to the meet or exceed the requirements for the confidence interval and the margin of error.

Determining Sample Size

$$n = \frac{Z^2\ \sigma^2}{e^2} = \frac{(1.645)^2(45)^2}{5^2} = 219.19$$

$$\boldsymbol{n} = \frac{(\ 2.706025\ )\ (\ 2025\ )}{25}$$

$$\boldsymbol{n} = \frac{5479.7006}{25} = 219.18802$$

Detailed Computations

The ‘z’ Table

There is no ‘z’ value for ‘.9500’ in the table.
We need to interpolate:

Z	.04
1.6	.9495

Z	.05
1.6	.9505

Z becomes 1.645

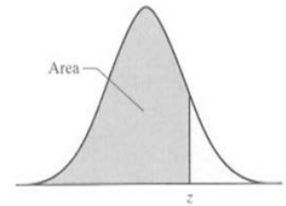


TABLE II (continued)

Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	[illegible]	[illegible]	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
[illegible]	0.9332	0.9345	0.9357	[illegible]	[illegible]	[illegible]	[illegible]	0.9418	0.9429	0.9441
1.6	[illegible]	0.9463	0.9474	0.9484	0.9495	0.9505	[illegible]	0.9525	0.9535	0.9545
[illegible]	0.9554	0.9564	0.9573	[illegible]	[illegible]	[illegible]	[illegible]	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

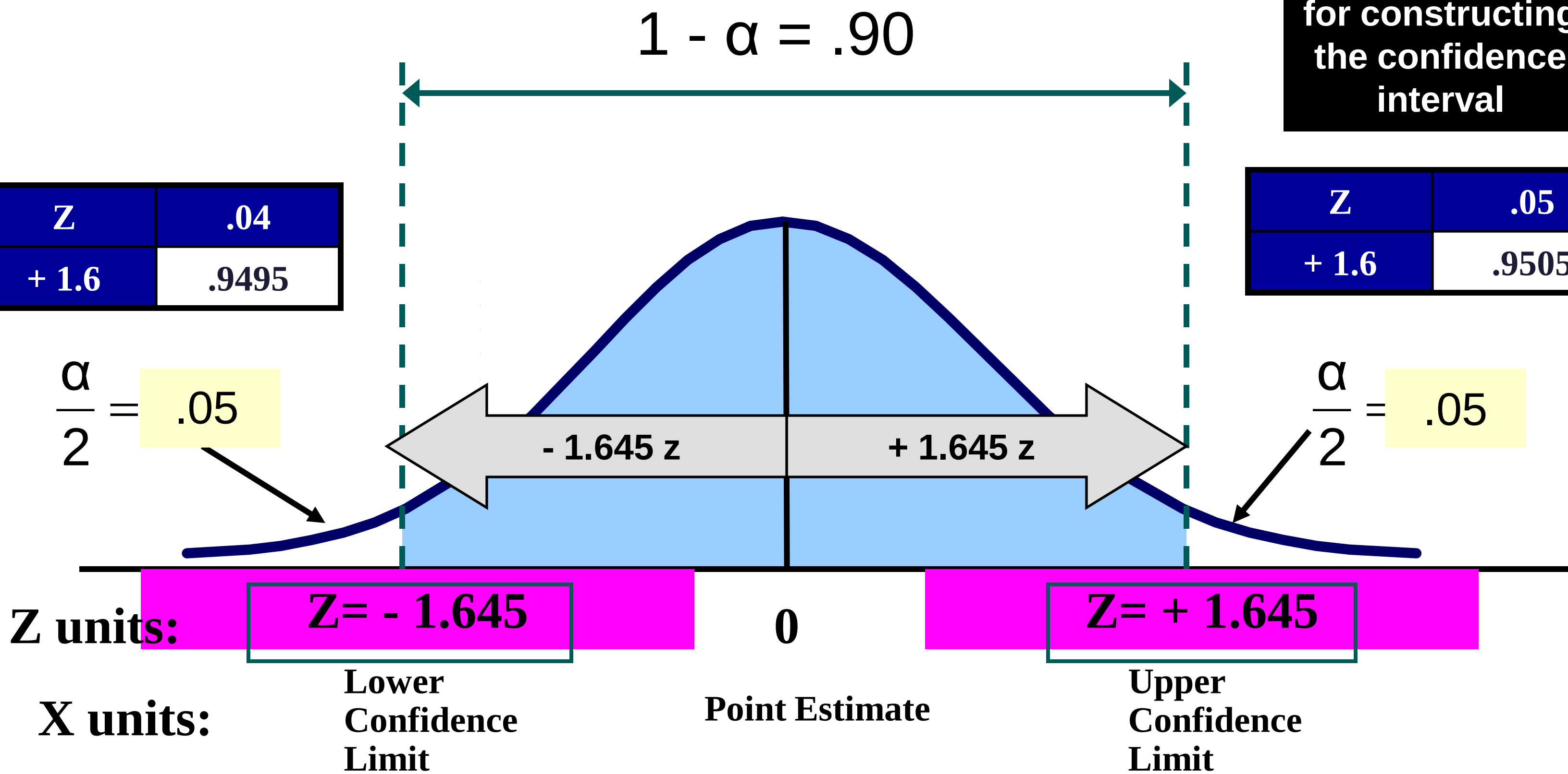
Finding the Critical Value, Z

Consider a 90% confidence interval:

The value of 'z' is needed for constructing the confidence interval

Z	.04
+ 1.6	.9495

Z	.05
+ 1.6	.9505



Determining Sample Size for the Proportion

To determine the required sample size for the proportion, you must know 3 factors:

- The desired level of confidence (**1 -** α), which determines the critical ***Z*** value
- The acceptable sampling error (margin of error), ***e***
- The true proportion of "successes", $\boldsymbol{\pi}$
  - $\boldsymbol{\pi}$ can be estimated with a pilot sample, if necessary (or conservatively use $\pi = .50$)

$$e = Z\sqrt{\frac{\pi(1-\pi)}{n}}$$

Now solve for n to get

$$n = \frac{Z^2\,\pi\,(1-\pi)}{e^2}$$

Determining Sample Size for the Proportion

Ironically, π is actually the population parameter that we are trying to estimate !

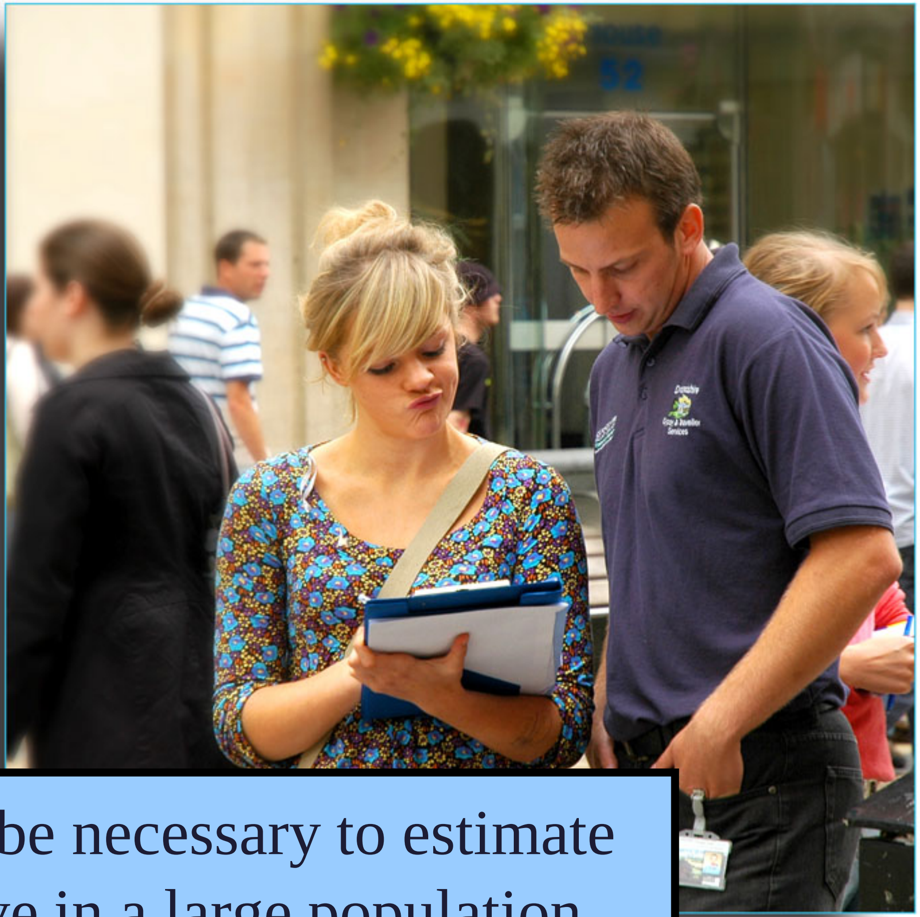
1. Consider using an "**educated**" **guess** for π .

2. Consider using π = **0.50** which will generate the maximum sample size needed.

3. Be aware the maximum sample size results in the highest sampling costs.

4. That said, there is then increased precision in estimating π, because the confidence interval is narrowed !

Determining Sample Size



How large a sample would be necessary to estimate the true proportion defective in a large population within ± **3%,** with **95%** confidence ?

(Assume a pilot sample yields p = .12)

Determining Sample Size

Solution:

For **95%** confidence, use Z = **1.96**

e = **.03**

p = **.12**, so use this to estimate π

Z	.06
+ 1.9	.9750

Z	.06
- 1.9	.0250

$$n = \frac{Z^2 \, \pi(1-\pi)}{e^2} = \frac{(1.96)^2(.12)(1-.12)}{(.03)^2} = 450.74$$

Detailed Computations
(3.8416) (.12) (.88) / .0009 = 0.4056729 / .0009 = 450.74766 ≈ 451

So use n = **451**

Determining Sample Size

Solution:

For **95%** confidence, use Z = **1.96**

e = **.03**

p = **.50**, if using this to estimate π (max 'n')

Z	.06
+ 1.9	.9750

Z	.06
- 1.9	.0250

$$n = \frac{Z^2\ \pi(1-\pi)}{e^2} = \frac{(1.96)^2\ (.50)(1 - .50)}{(.03)^2} = \boxed{1{,}067}$$

Detailed Computations
(3.8416) (.50) (.50) / .0009 = (3.8416) (.25) / .0009 = .9604 / .0009 = 1,067.11 ≈ 1,067

So use n = **1,067**

Example

An investor is trying to estimate the return on investment in companies that won quality awards last year.

A random sample of 41 such companies is selected, and the return on investment is recorded for each company. The data for the 41 companies have

$$\bar{x} = 14.75 \quad s = 8.18$$

Construct a 95% confidence interval for the mean return.

$$\bar{x} \pm t\frac{s}{\sqrt{n}} \qquad d.f. = n-1$$

$$\bar{x} = 14.75 \quad s = 8.18$$

$$\text{degrees of freedom} = 41 - 1 = 40$$

$$\text{from t-table, } t = 2.0211$$

$$\bar{x} \pm t\frac{s}{\sqrt{n}} = 14.75 \pm 2.0211\frac{8.18}{\sqrt{41}}$$

$$= 14.75 \pm 2.61 = [12.14, 17.36]$$

We are 95% confident that the interval (12.14, 17.36) contains the population mean return on investment for companies that win quality awards.

Example

Because cardiac deaths increase after heavy snowfalls, a study was conducted to measure the cardiac demands of shoveling snow by hand

The maximum heart rates for 10 adult males were recorded while shoveling snow. The sample mean and sample standard deviation were $\bar{x} = 175,\ s = 15$

Find a 90% CI for the population mean max. heart rate for those who shovel snow.

Solution

$$\bar{x} \pm t \frac{s}{\sqrt{n}} \qquad d.f. = n - 1$$

$\bar{x} = 175,\ s = 15\ n = 10$

From the t - table, t = 1.8331

$$175 \pm 1.8331 \frac{15}{\sqrt{10}} = 175 \pm 8.70$$

$= (166.30,\ 183.70)$

We are 90% confident that the interval (166.30, 183.70) contains the mean maximum heart rate for snow shovelers

EXAMPLE: Consumer Protection Agency

Selected random sample of 16 packages of a product whose packages are marked as weighing 1 pound.

From the 16 packages: $\bar{x} = 1.10$ pounds, $s = .36$ pound

a. find a 95% CI for the mean weight μ of the 1-pound packages

b. should the company's claim that the mean weight μ is 1 pound be challenged ?

Solution

$$\bar{x} \pm t\frac{s}{\sqrt{n}} \qquad d.f. = n-1$$

95% CI, n=16, df=15, $\bar{x}$=1.10

s=.36

critical value of t is $t = 2.1315$

$\bar{x} \pm t\frac{s}{\sqrt{n}}$ becomes

$$1.10 \pm (2.1315)\left(\frac{.36}{\sqrt{16}}\right) = 1.10 \pm .19 = (.91, 1.29)$$

Since 1 pound is in the interval, the company's claim appears reasonable.

Example:

Suppose the marketing manager wishes to estimate the population mean annual usage of home heating oil to **within +/- 50 gallons** of the true value, and he wants to be **95% confident** of correctly estimating the true mean.

On the basis of a study taken the previous year, he believes that the **standard deviation** can be estimated as **325 gallons**.

Find the **sample size** needed.

Solution

With e =50, σ = 325, and 95% confidence (Z = 1.96)

$$n = \frac{Z^2\ \sigma^2}{e^2} = (1.96)^2\,(325)^2 / (50)^2$$

n = 162.31

Therefore, n = 163. As a general rule for determining sample size, always round up to the next integer value in order to slightly **over satisfy** the criteria desired.

Example:

Suppose you want to estimate the average age of all Boeing 727 airplanes now in active domestic U.S. service.

You want to be 95% confident, and you want your estimate to be within 2 years of the actual figure.

The 727 was first placed in service about 30 years ago, but you believe that no active 727s in the U.S. domestic fleet are more than 25 years old.

How large a sample should you take?

Solution

With e = 2 years,

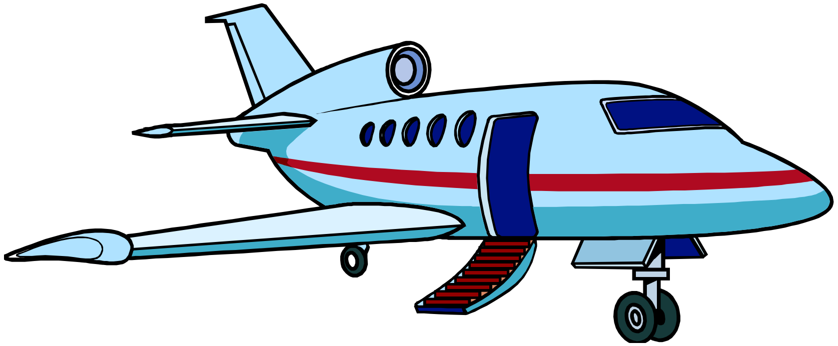
and Z value for 95% = 1.96,

and σ unknown,

it must be estimated by using σ ≈ range ÷ 4.

As the range of ages is 0 to 25 years,

σ = 25 ÷ 4 = 6.25.



$$n = \frac{Z^2 \ \sigma^2}{e^2}$$ $= (1.96)^2\ (6.25)^2\ /\ (2)^2 = 37.52$ airplanes.

Because you cannot sample 37.52 units, the required sample size is 38.

If you randomly sample 38 planes, you can estimate the average age of active 727s within 2 years and be 95% confident of the results.