

## VIII. Interval Estimation

### A. A Few Important Definitions (Including Some Reminders)

1. Point Estimate - a single numerical value used as an estimate of a parameter.
2. Point Estimator - the sample statistic that provides the point estimate of a parameter.

Some point estimators (and the parameters they estimate) include

<u>Parameter</u>	<u>Point Estimator</u>
$\mu$	$\bar{x}$
$\sigma$	$s$
$p$	$\bar{p}$

---

---

---

---

---

---

---

---

3. Precision - the exactness of an estimator.

4. Accuracy - the correctness of an estimator.

Remember that

- Precision and Accuracy are inversely related
- Point estimators are i) perfectly precise and ii) almost certainly inaccurate

5. Sampling Error - the absolute difference between a parameter and its point estimator.

How do we address the existence of sampling error?

---

---

---

---

---

---

---

---

6. Confidence Interval - range of values, used as an estimate of a parameter, that will contain the true value of the parameter a given proportion of times over many independent, identical repeated trials - also referred to as an *interval estimate*

7. Confidence Level - the proportion of times a confidence interval can be expected to contain the true value of the parameter over many independent, identical repeated trials the correctness of an estimator OR

the probability that the interval estimation procedure will generate an interval that does contain the true value of the parameter - also referred to as the *confidence coefficient*

Note that  $(1 - \text{Confidence Level})$  is often referred to as  $\alpha$  or the *significance level* (so Confidence Level =  $1 - \alpha$ )

---

---

---

---

---

---

---

---

## B. Large-Sample Interval Estimation of the Population Mean $\mu$

Large Sample Confidence Interval for  $\mu$  - recall that we can find a symmetric interval that satisfies the formula:

$$P\left(\mu - z_{\alpha/2} \sigma_{\bar{x}} \leq \bar{x} \leq \mu + z_{\alpha/2} \sigma_{\bar{x}}\right) = \text{Tolerance Level} = 1 - \alpha$$

Note that, once we take a sample, we will have values for  $\bar{x}$ ,  $z$ , and  $\sigma_{\bar{x}}$  - but not for  $\mu$  (which is what we are trying to estimate) - how does this help us develop an estimate of  $\mu$ ?

---

---

---

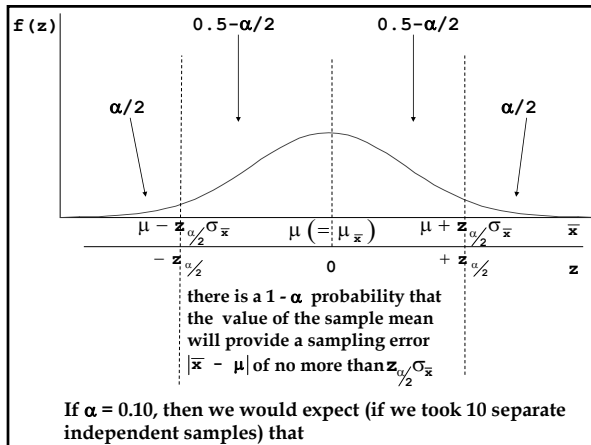
---

---

---

---

---




---

---

---

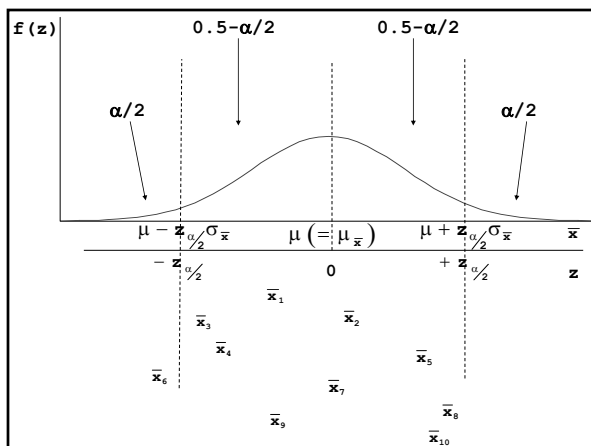
---

---

---

---

---




---

---

---

---

---

---

---

---

Problem: We are trying to estimate  $\mu$  - but the (tolerance) intervals are constructed around  $\mu$ ! How can we use  $\mu$  to estimate  $\mu$ ?

What would happen if we built intervals of half-width  $z_{\alpha/2}\sigma_{\bar{x}}$  around each sample mean?

Could we even do so?

---

---

---

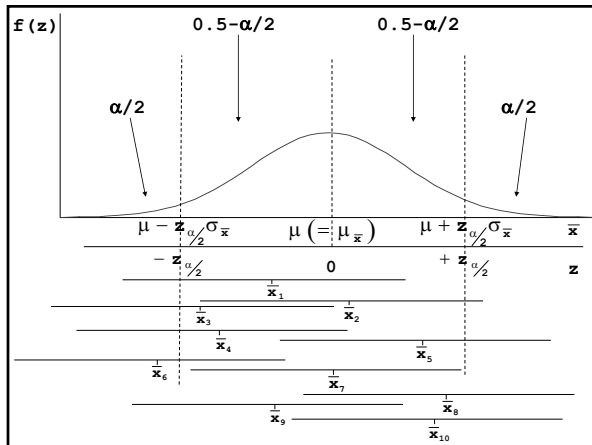
---

---

---

---

---




---

---

---

---

---

---

---

---

The resulting interval of half-width  $z_{\alpha/2}\sigma_{\bar{x}}$  around the sample mean is given by

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

which by substitution can be rewritten as

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

if the parent population's standard deviation  $\sigma$  is known and either i) the parent population is normal or ii) the sample is sufficiently large ( $n \geq 30$ ). Note that

- this is often referred to as an interval estimate or a confidence interval.
- $(1 - \alpha)100\%$  is often referred to as the confidence level or confidence coefficient.
- $z_{\alpha/2}\sigma_{\bar{x}}$  is often referred to as the margin of error.

---

---

---

---

---

---

---

---

Another Problem: We need to know the parent population's standard deviation  $\sigma$  - this (usually) necessitates a census! This implies a knowledge about the population mean  $\mu$ , which is what we are trying to estimate!

It can be shown that the sample standard deviation  $s$  is an *extremely* reliable estimator of the parent population's standard deviation  $\sigma$  if we take a large sample ( $n \geq 30$ ).

---

---

---

---

---

---

---

---

The resulting interval of half-width  $z_{\alpha/2} s_{\bar{x}}$  around the sample mean is given by

$$\bar{x} \pm z_{\alpha/2} s_{\bar{x}}$$

which by substitution can be rewritten as

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

if the parent population's standard deviation  $\sigma$  is not known and the sample is sufficiently large ( $n \geq 30$ ). Note that

- this is often referred to as an interval estimate or a confidence interval.
- $(1 - \alpha)100\%$  is often referred to as the confidence level or confidence coefficient.
- $z_{\alpha/2} s_{\bar{x}}$  is often referred to as the margin of error.

---

---

---

---

---

---

---

---

Example: Suppose we take a sample of the ages of forty-five customers from a population whose standard deviation is 6.5 and collect the following observations:

Observation #	X	Observation #	X	Observation #	X
1	55	16	32	31	27
2	42	17	23	32	33
3	37	18	41	33	33
4	24	19	34	34	33
5	28	20	38	35	34
6	33	21	32	36	18
7	36	22	29	37	34
8	24	23	36	38	37
9	40	24	33	39	21
10	38	25	26	40	26
11	36	26	36	41	37
12	33	27	38	42	31
13	35	28	37	43	35
14	39	29	25	44	31
15	28	30	39	45	37

---

---

---

---

---

---

---

---

How would we construct a 95% confidence interval for these data?

Since the random variable  $X$  (age) is quantitative (interval or ratio), we recognize the parameter we wish to estimate is the population mean  $\mu$ .

We also recognize that we have a large sample ( $n = 45 \geq 30$ ) and we know the parent population's standard deviation is  $\sigma = 6.5$ .

So we decide to estimate the interval by using

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

---

---

---

---

---

---

---

---

We find the appropriate level of  $z_{\alpha/2}$  from the Standard Normal Distribution table for the 95% level of confidence

$$z_{\alpha/2} = 1.96$$

and calculate the sample mean

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{55 + 42 + 37 + 24 + \dots + 37}{45} = 33.2$$

and plug the results into the formula

$$\begin{aligned} \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} &= 33.2 \pm 1.96 \frac{6.5}{\sqrt{45}} \\ &= 33.2 \pm 1.899 = (31.301, 35.099) \end{aligned}$$

---

---

---

---

---

---

---

---

**Example:** If we ask thirty-three randomly selected individuals to rate a product on a 100 point (1 - 100) scale, and our responses are

Observation #	X	Observation #	X	Observation #	X
1	72	12	76	23	72
2	78	13	65	24	88
3	79	14	73	25	84
4	66	15	79	26	74
5	83	16	82	27	78
6	74	17	86	28	80
7	77	18	76	29	70
8	79	19	74	30	74
9	69	20	86	31	89
10	82	21	84	32	94
11	91	22	76	33	83

---

---

---

---

---

---

---

---

How would we construct a 90% confidence interval for these data?

Since the random variable  $X$  (product rating) is quantitative (interval or ratio), we recognize the parameter we wish to estimate is the population mean  $\mu$ .

We also recognize that we have a large sample ( $n = 33 \geq 30$ ) but we don't know the parent population's standard deviation.

So we decide to estimate the interval by using

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

---

---

---

---

---

---

---

---

We find the appropriate level of  $z_{\alpha/2}$  from the Standard Normal Distribution table for the 90% level of confidence

$$z_{\alpha/2} = 1.64 \text{ or } 1.65 \text{ (or } 1.645)$$

and calculate the sample mean

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{72 + 78 + 79 + 66 + \dots + 83}{33} = 78.6$$

---

---

---

---

---

---

---

---

calculate the sample standard deviation

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$
$$= \sqrt{\frac{(72-78.6)^2 + (78-78.6)^2 + (79-78.6)^2 + (66-78.6)^2 + \dots + (83-78.6)^2}{33-1}}$$
$$= 6.97$$

and plug the results into the formula

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} = 78.6 \pm 1.645 \frac{6.97}{\sqrt{33}}$$
$$= 78.6 \pm 1.997$$
$$= (76.603, 80.597)$$

---

---

---

---

---

---

---

---

Example: If we ask thirty-three randomly selected individuals to rate a product on a 100 point (1 - 100) scale, and calculate from our results a sample mean of 78.6 and a sample variance of 48.63, what would be the appropriate 86% confidence interval?

We have a large sample ( $n = 33 \geq 30$ ), we do not know the parent population's standard deviation  $\sigma$ , and  $z_{\alpha/2} = 1.48$ , so

$$\begin{aligned} \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} &= 78.6 \pm 1.48 \frac{\sqrt{48.63}}{\sqrt{33}} \\ &= 78.6 \pm 1.797 \\ &= (76.803, 80.397) \end{aligned}$$

Note that we were provided the *sample variance* and not the *sample standard deviation* in this problem!

Notice that as the confidence level decreases, the confidence interval narrows - why?

---

---

---

---

---

---

---

---

### C. Small-Sample Interval Estimation of the Population Mean $\mu$

If the parent population's standard deviation  $\sigma$  is not known and the parent population is normally distributed, but the sample is not sufficiently large ( $n \geq 30$ ) we must use a special probability distribution.

The t-Distribution (or the Student-t Distribution) - family of probability distributions that are used to construct interval estimates of the population mean when the population standard deviation  $s$  is unknown and the parent population is normally or near-normally distributed.

---

---

---

---

---

---

---

---

- the t-distribution is similar to the normal distribution except that it is shorter and wider (to allow for the fact that we are using an estimate of the parent population's standard deviation  $\sigma$ ).

- the height and width of the t-distribution is determined by a parameter called the degrees of freedom.

for a confidence interval of the population mean, the degrees of freedom are  $n - 1$ .

as the degrees of freedom increase, the t-distribution becomes more normal (taller and narrower).

values for the t-distribution for various combinations of  $\alpha/2$  and degrees of freedom increase are given in Appendix A, Table 2 of ASW.

---

---

---

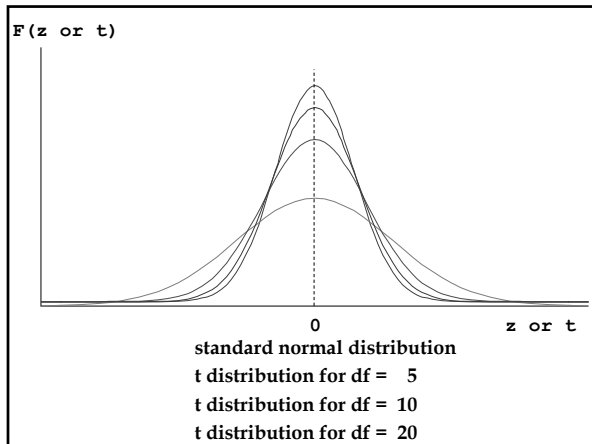
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

The resulting interval of half-width  $t_{\alpha/2} s_{\bar{x}}$  around the sample mean is given by

$$\bar{x} \pm t_{\alpha/2} s_{\bar{x}}$$

which by substitution can be rewritten as

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

if the parent population's standard deviation  $\sigma$  is not known and the sample is not sufficiently large ( $n < 30$ ), but the parent population is relatively normally distributed. Note that

- this is often referred to as an interval estimate or a confidence interval.
- $(1 - \alpha)100\%$  is often referred to as the confidence level or confidence coefficient.
- $t_{\alpha/2} s_{\bar{x}}$  is often referred to as the margin of error.

---

---

---

---

---

---

---

---

---

---

degrees of freedom	Area in Upper Tail					
	0.10	0.05	0.025	0.01	0.005	0.001
1	3.078	6.314	12.706	31.821	63.656	318.289
2	1.886	2.920	4.303	6.965	9.925	22.328
3	1.638	2.353	3.182	4.541	5.841	10.214
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.894
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.238	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.644
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
60	1.296	1.671	2.000	2.390	2.660	3.232
120	1.289	1.658	1.980	2.358	2.617	3.160
$\infty$	1.282	1.645	1.960	2.326	2.576	3.090

---

---

---

---

---

---

---

---

---

---

If we ask 17 randomly selected individuals to rate a product on a 100 point (1 - 100) scale, and we wish to calculate the 90% confidence interval, what would be the appropriate value of t?

We have a sample of  $n = 17$  observations, so the degrees of freedom are  $n - 1 = 17 - 1 = 16$ . Thus  $t_{\alpha/2} = 1.746$ .

degrees of freedom	Area in Upper Tail				
	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.656
⋮	⋮	⋮	⋮	⋮	⋮
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878

---

---

---

---

---

---

---

---

---

---

---

---

Example: Suppose we ask seventeen randomly selected individuals to rate a product on a 100 point (1 - 100) scale, and collect the following sample results:

Observation #	X	Observation #	X
1	74	10	85
2	79	11	91
3	85	12	76
4	69	13	69
5	84	14	70
6	77	15	79
7	77	16	85
8	79	17	88
9	69		

Also suppose that product ratings have historically been relatively normally distributed.

---

---

---

---

---

---

---

---

---

---

---

---

How would we construct a 90% confidence interval for these data?

Since the random variable X (product rating) is quantitative (interval or ratio), we recognize the parameter we wish to estimate is the population mean  $\mu$ .

We also recognize that we have a small sample ( $n = 17 < 30$ ) from a normal population, and we don't know the parent population's standard deviation.

So we decide to estimate the interval by using

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

---

---

---

---

---

---

---

---

---

---

---

---

We find the appropriate level of  $t_{\alpha/2}$  from the Student's t Distribution table for the 90% level of confidence and  $n - 1 = 16$  degrees of freedom

$$t_{\alpha/2} = 1.746$$

and calculate the sample mean

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{74 + 79 + 85 + 69 + \dots + 88}{17} = 78.6$$

---

---

---

---

---

---

---

---

calculate the sample standard deviation

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(74-78.6)^2 + (79-78.6)^2 + (85-78.6)^2 + (69-78.6)^2 + \dots + (88-78.6)^2}{17-1}}$$

$$= 6.97$$

and plug the results into the formula

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 78.6 \pm 1.746 \frac{6.97}{\sqrt{17}}$$

$$= 78.6 \pm 2.953$$

$$= (75.647, 81.553)$$

Notice that as this confidence level is wider than the confidence interval of (76.603, 80.597) derived using the standard normal distribution - why?

---

---

---

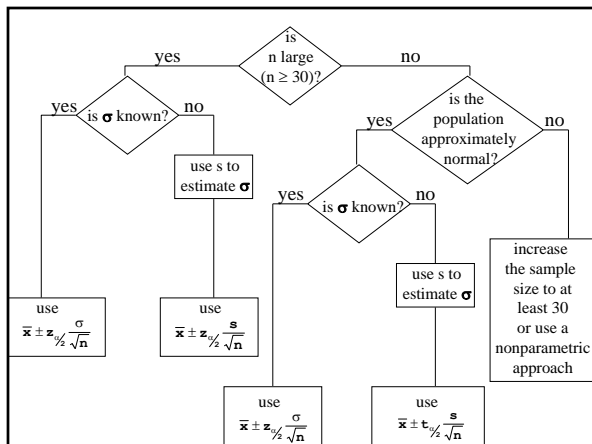
---

---

---

---

---




---

---

---

---

---

---

---

---

Can we determine the sample size necessary to ensure a desired confidence level and margin of error?

If we assume a given population standard deviation  $\sigma$  and state the desired margin of error (denoted as E), sample size necessary to ensure a desired confidence level and margin of error is

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{E^2}$$

This is nothing but an algebraic manipulation of the definition of the margin of error

$$E = \frac{z_{\alpha/2} \sigma}{\sqrt{n}}$$

---

---

---

---

---

---

---

---

Example: Suppose we wish to determine the minimum sample size necessary to ensure a 95% confidence level and a margin of error of 2.5. If we believe that the population standard deviation is 21, how large must the sample size be?

We have have that  $\sigma = 21.0$ ,  $E = 2.5$ , and  $z_{\alpha/2} = 1.96$ , so

$$n = \frac{(1.96)^2 (21.0)^2}{2.5^2} = 271.063$$

or 272 (why?).

---

---

---

---

---

---

---

---

#### D. Interval Estimation of the Population Proportion p

Recall that if the sample size is sufficiently large ( $np \geq 5$  and  $n(1-p) \geq 5$ ), that  $\bar{p}$  is approximately normally distributed with a mean of p and standard deviation of

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

so the resulting  $(1 - \alpha)100\%$  confidence interval is

$$\bar{p} \pm z_{\alpha/2} \sigma_{\bar{p}}$$

which through substitution can be rewritten as

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

---

---

---

---

---

---

---

---

Another Problem: How do we use the population proportion  $p$  to estimate the population proportion  $p$ ?

We simply substitute  $\bar{p}$  for  $p$ , i.e.

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

and

$$n\bar{p} \geq 5 \text{ and } n(1-\bar{p}) \geq 5$$

Another alternative is to use  $p = 0.50$  in

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.50(1-0.50)}{n}} = \sqrt{\frac{0.25}{n}} \text{ (Why?)}$$

---

---

---

---

---

---

---

---

---

---

Example: Suppose we poll one-hundred randomly selected registered voters individuals and ask if they intend to vote for Candidate Jones, and our sample results are as follows:

---

---

---

---

---

---

---

---

---

---

Observation #	Vote For Jones?	Observation #	Vote For Jones?	Observation #	Vote For Jones?	Observation #	Vote For Jones?
1	YES	26	NO	50	NO	76	NO
2	NO	27	NO	51	YES	77	NO
3	NO	28	YES	52	NO	78	YES
4	NO	29	YES	53	NO	79	YES
5	YES	30	NO	54	YES	80	NO
6	NO	31	YES	55	NO	81	NO
7	NO	32	YES	56	YES	82	YES
8	YES	33	NO	57	NO	83	NO
9	YES	34	YES	58	YES	84	YES
10	NO	35	YES	59	YES	85	YES
11	NO	36	YES	60	NO	86	NO
12	YES	37	NO	61	YES	87	YES
13	NO	38	NO	62	NO	88	YES
14	YES	39	YES	63	NO	89	NO
15	NO	40	NO	64	YES	90	NO
16	NO	41	YES	65	NO	91	NO
17	YES	42	NO	66	YES	92	YES
18	NO	43	NO	67	NO	93	NO
19	YES	44	YES	68	YES	94	NO
20	NO	45	NO	69	NO	95	YES
21	NO	46	YES	70	YES	96	NO
22	NO	47	NO	71	NO	97	NO
23	YES	48	NO	72	YES	98	NO
24	NO	49	YES	73	YES	99	YES
25	YES	50	YES	74	NO	100	NO

---

---

---

---

---

---

---

---

---

---

Since the random variable of interest (whether a respondent indicated s/he would vote for Jones) is nominal, we recognize the parameter we wish to estimate is the population proportion  $p$ .

We also recognize that the sample proportion is

$$\bar{p} = \frac{45}{100} = 0.45$$

so

$$n\bar{p} = 100(0.45) = 45 \geq 5$$

and

$$n(1 - \bar{p}) = 100(1 - 0.45) = 55 \geq 5$$

...Cochran's rules are satisfied and we can estimate the confidence interval using

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

---

---

---

---

---

---

---

---

---

---

For a 92% level of confidence we have that

$$z_{\alpha/2} = 1.75$$

so the 92% confidence interval is

$$\begin{aligned} \bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} &= 0.45 \pm 1.75 \sqrt{\frac{0.45(1 - 0.45)}{100}} \\ &= 0.45 \pm 0.087 \\ &= (0.363, 0.537) \end{aligned}$$

What do you think of the likelihood that Jones will win?

---

---

---

---

---

---

---

---

---

---

Can we determine the sample size necessary to ensure a desired confidence level and margin of error?

If we assume a given population standard deviation  $s$  and state the desired margin of error (denoted as  $E$ ), sample size necessary to ensure a desired confidence level and margin of error is

$$n = \frac{z_{\alpha/2}^2 p(1 - p)}{E^2}$$

Again, this is nothing but an algebraic manipulation of the definition of the margin of error

$$E = z_{\alpha/2} \sqrt{\frac{p(1 - p)}{n}}$$

---

---

---

---

---

---

---

---

---

---

Example: Suppose we wish to determine the minimum sample size necessary to ensure a 95% confidence level and a margin of error of .025 for Candidate Jones. If we believe that Jones is currently favored by 50% of the population, how large must the sample size be?

We have have that  $p = 0.50$ ,  $E = 0.025$ , and  $z_{\alpha/2} = 1.96$ , so

$$n = \frac{(1.96)^2 0.50(1 - 0.50)^2}{0.025^2} = 1536.64$$

or 1537 (why?).

Again, if we don't have a good estimate (guess) for the population proportion  $p$ , we should always use  $p = 0.50$  (why?).

---

---

---

---

---

---

---

---

---

---

Example: Raymond Kenneth, Brand manager for Imani Diot perfume, wants to assess the market potential of a new fragrance *Eau de Commodé*. Historically a consumer that rates a Diot perfume 80 or better on a 100 point scale after sampling the product is usually willing to consider purchasing the scent. Mr. Kenneth commissions Tommie Krook Research to collect ratings of *Eau de Commodé* from forty-eight frequent perfume purchasers. The results of Krook's survey follow.

---

---

---

---

---

---

---

---

---

---

Respondent #	Rating	Respondent #	Rating	Respondent #	Rating	Respondent #	Rating
1	84	13	84	25	48	37	81
2	65	14	85	26	66	38	44
3	99	15	82	27	89	39	29
4	91	16	69	28	94	40	81
5	82	17	88	29	65	41	84
6	88	18	83	30	81	42	88
7	82	19	83	31	90	43	66
8	81	20	100	32	81	44	67
9	90	21	36	33	84	45	80
10	82	22	84	34	81	46	81
11	80	23	80	35	40	47	53
12	84	24	94	36	82	48	84

Use Krook's survey results to build an appropriate confidence interval at the 90% confidence level for Mr. Kenneth.

---

---

---

---

---

---

---

---

---

---

Should you build a confidence interval of the

population mean or population proportion

of *Eau de Commodé*?

---

---

---

---

---

---

---

---

How would we construct a 90% confidence interval for the population mean rating  $\mu_{\text{Eau de Commodé}}$  of *Eau de Commodé* with these data?

We recognize that we have a large sample ( $n = 48 \geq 30$ ) but we don't know the parent population's standard deviation.

So we decide to estimate the interval by using

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

---

---

---

---

---

---

---

---

We find the appropriate level of  $z_{\alpha/2}$  from the Standard Normal Distribution table for the 90% level of confidence

$$z_{\alpha/2} = 1.64 \text{ or } 1.65 \text{ (or } 1.645)$$

and calculate the sample mean

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{84 + 29 + 99 + 91 + \dots + 82}{48} = 74.77$$

---

---

---

---

---

---

---

---

and calculate the sample standard deviation

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(84-74.77)^2 + (29-74.77)^2 + (99-74.77)^2 + (91-74.77)^2 + \dots + (82-74.77)^2}{48-1}}$$

$$= 20.64$$

and plug the results into the formula

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} = 74.77 \pm 1.645 \frac{20.64}{\sqrt{48}}$$

$$= 74.77 \pm 4.902$$

$$= (69.868, 79.672)$$

What does this mean for Mr. Kenneth?

---

---

---

---

---

---

---

---

How would we construct a 90% confidence interval for the population proportion of consumers  $p_{Eau\ de\ Commodé}$  who will consider purchasing *Eau de Commodé* with these data?

We recognize that the sample proportion is 0.75 (why?)

---

---

---

---

---

---

---

---

Respondent #	Rating	Respondent #	Rating	Respondent #	Rating	Respondent #	Rating
1	84	13	84	25	48	37	81
2	65	14	85	26	66	38	44
3	99	15	82	27	89	39	29
4	91	16	69	28	94	40	81
5	82	17	88	29	65	41	84
6	88	18	83	30	81	42	88
7	82	19	83	31	90	43	66
8	81	20	100	32	81	44	67
9	90	21	36	33	84	45	80
10	82	22	84	34	81	46	81
11	80	23	80	35	40	47	53
12	84	24	94	36	84	48	82

$$\bar{p} = \frac{36}{48} = 0.75$$

---

---

---

---

---

---

---

---

so

$$np = 48(0.75) = 36 \geq 5$$

and

$$n(1 - p) = 48(1 - 0.75) = 12 \geq 5$$

...Cochran's rules are satisfied and we can estimate the confidence interval using

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

---

---

---

---

---

---

---

---

For a 90% level of confidence we have that

$$z_{\alpha/2} = 1.645$$

so the 90% confidence interval is

$$\begin{aligned} \bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} &= 0.75 \pm 1.645 \sqrt{\frac{0.75(1 - 0.75)}{48}} \\ &= 0.75 \pm 0.103 \\ &= (0.647, 0.853) \end{aligned}$$

What does this mean for Mr. Kenneth?

---

---

---

---

---

---

---

---

Which is more appropriate - the 90% confidence interval of the population mean or the 90% confidence interval of the population proportion?

In which random variable should Mr. Kenneth be interested:

- The individual rating consumers give *Eau de Commodé*?
- Whether individual consumers give *Eau de Commodé* a rating of 80 or higher?

Answer this question and you have determined whether the confidence interval of the *population mean* or the confidence interval of the *population proportion* is more appropriate for this problem.

---

---

---

---

---

---

---

---