CMD 305 - COURSE
(RESEARCH METHODOLOGY & BIOSTATISTICS)

TUTORIAL TOPIC : Sample size estimation(solutions)

Q1) We want to estimate the mean systolic blood pressure of Saudi females. The standard deviation is around 20 mmHg and we wish to estimate the true mean to within 10 mmHg with 95% confidence. What is the required sample size ?

Solution: You can get sample size in <u>two</u> ways:

(1) Sample size $n = Z^2\alpha s^2 / d^2$

Given :S=20 and d=10 Zα= 1.96 for 95% confidence level

Then $n = (1.96)^2 x (20)^2/(10)^2 = 15.37$

Since we cannot take 0.37 of a person, we round up to 16 women as our sample size.

(2) By using tables (Table.1A)

Q2) A public health department wanted an estimate of average retail price of twenty tablets of a commonly used tranquilizer sold in retail pharmacies of its region. How many sample of pharmacies is to be selected such that the estimate is required to be with in5 riyals of the true average price with95% confidence. (Based on a small pilot study , it was found the sd of price is 10 riyals).

Solution:

we can get the sample size by two ways 1) by formula and 2) By using ready made Table 1A

(1) Sample size $n = Z^2\alpha S^2 / d^2$

Given :s=10 and d=5 Zα= 1.96 for 95% level of confidence

Then $n = (1.96)^2 x (10)^2 / (5)^2 = 16$ pharmacies are to selected among the pharmacies of that region.

Q3) We wish to estimate the proportion of Saudi males who smoke. What sample size do we require to achieve a 95% confidence interval of width ± 5% (that is to be within 5% of the true value) ? A study some years ago found approximately 30% were smokers ?

Solution: You can get sample size in <u>two</u> ways:

(1) Sample size $n = Z^2_{\alpha}p(1-p)/d^2$

Given : $p=0.3$ $d=.05$ $Z_{\alpha}= 1.96$ for 95% confidence level

Then $n = (1.96)^2(0.3)(0.7)/(0.05)^2 = 322.7$ rounded to 323

We need 323 saudi males as our sample size to estimate the proportion of saudi males who smoke.

(2) By using tables (Table.1B)

Q4) An epidemiologist was asked to asked to estimate the proportionof children in a region receiving appropriate childhood vaccinations. How many children he should select, if the resulting estimate is to fall within 10 percentage points of the true proportion with 95% confidence?.(From a pilot study he came to know that this rate is 20%)

Solution: You can get sample size in <u>two</u> ways:1) by formula and 2) by using readymade Table 1B

(1) Sample size $n = Z^2_{\alpha}P(1-P)/d^2$

Given : $P=0.2$ $d=0.10$ $Z_{\alpha}= 1.96$ for 95% confidence level

Then $n = (1.96)^2(0.2)(0.8)/(0.1)^2 = 62$

He needs 62children as his sample to estimate the proportion of children in a region receiving appropriate childhood vaccinations.

Q5) An epidemiologist wants to test whether a iron supplement for pregnant women will increase the increase their Hb level. One group of women will receive new supplement and the other group the usual supplement. From a pilot study the sd of Hb is 4 g/dl and is assumed to be same for both groups.. what is the sample size required to test the hypothesis of no difference in mean Hb level at 99% level of confidence and 90% power of detecting an increase of 2 g/dl.

Solution: You can get sample size in two ways: 1) by using formula and(2) By using ready madetables (Table.2A)

(1) Sample size $n = 2S^2 (Z_\alpha + Z_\beta)^2 / d^2$ per arm

Given d=2, S=4,
$Z\alpha$ = 2.58 for 99% level of confidenceand
Z_β = 1.28 for 90% power

Then $n = 2 \times 4^2 (2.58+1.28)^2 / (2)^2 = 119$
We require at least 119 per group (Total 238 women)

(2) By using tables (Table.2A)

Q6.Suppose it has been estimated that the rate of caries is 800 per 1000 school children in one district and 600 per 1000 in another district. What is the sample size required from each district to determine whether the difference is significant at the 95% level if we wish to have an 90% of chance of detecting the difference if it is real?

Solution: You can get sample size in two ways:1) by using formula and (2) By using ready madetable (Table.2B)

(1) Sample size $n = \dfrac{(Z_\alpha + Z_\beta)^2 \left((p_1 q_1) + (p_2 q_2)\right)}{(p_1 - p_2)^2}$, *per arm, where* $q_1 = (1 - p_1), q_2 = (1 - p_2)$.

Given p1=80% and p2=60% ;
$Z\alpha$= 1.96 for 95% confidence level ;
Z_β= 1.282 for 90% power

$n = (1.96+1.282)^2((0.8 \times 0.2)+(0.6 \times 0.4)) / (0.8-0.6)^2$
= 105 patients for each group Total sample size=210 children

.

Sample size formulae

A) If outcome variable is quantitative/continuous then outcome measure is **mean**
Example: Height, weight, BMI, HB, BP etc.,

For **a <u>single mean</u>** Sample size : $n = Z^2_\alpha S^2 / d^2$

Where , **<u>S</u>** (=sd, <u>get from the literature review or from the pilot study)</u>

For<u>two means</u>Sample size **:**$n = 2S^2 (Z_\alpha + Z_\beta)^2 / d^2$, per arm

Where , **<u>S</u>** (=sd, <u>get from the literature review or from the pilot study)</u>

B) If outcome variable is qualitative/categorical then outcome measure is **proportion**
Example: proportion of smokers, diabetes, anemia etc.,

For a single proportion Sample size **:**$n = Z^2_\alpha P(1-P)/d^2$

Where , **P** (=proportion, <u>get from the literature review or from the pilot study)</u>

For two proportions Sample size **:**

$$n = \frac{(Z_\alpha + Z_\beta)^2 ((p_1 q_1) + (p_2 q_2))}{(p_1 - p_2)^2} \quad per\ arm\ , where \quad q_1 = (1 - p_1), q_2 = (1 - p_2)$$

Where , P_1 and P_2(are proportions for group1 and group2 we are studying ,for example , obese-non obese, smokers-non smokers etc., , <u>get from the literature review or from the pilot study)</u>

d = precision (the researcher has to decide)
Z_α = 1.96 for 95% confidence level, usually
Z_β = 1.282 for 90% power, usually

TABLE 1A	SAMPLE SIZES FOR A SINGLE MEAN FOR VARIOUS d and sd for 95% level, Za=1.96																			
	d																			
sd	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	16	4	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	35	9	4	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	62	16	7	4	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
5	97	25	11	7	4	3	2	2	2	1	1	1	1	1	1	1	1	1	1	1
6	139	35	16	9	6	4	3	3	2	2	2	1	1	1	1	1	1	1	1	1
7	189	48	21	12	8	6	4	3	3	2	2	2	2	1	1	1	1	1	1	1
8	246	62	28	16	10	7	6	4	4	3	3	2	2	2	2	1	1	1	1	1
9	312	78	35	20	13	9	7	5	4	4	3	3	2	2	2	2	2	1	1	1
10	385	97	43	25	16	11	8	7	5	4	4	3	3	2	2	2	2	2	2	1
11	465	117	52	30	19	13	10	8	6	5	4	4	3	3	3	2	2	2	2	2
12	554	139	62	35	23	16	12	9	7	6	5	4	4	3	3	3	2	2	2	2
13	650	163	73	41	26	19	14	11	9	7	6	5	4	4	3	3	3	3	2	2
14	753	189	84	48	31	21	16	12	10	8	7	6	5	4	4	3	3	3	3	2
15	865	217	97	55	35	25	18	14	11	9	8	7	6	5	4	4	3	3	3	3
16	984	246	110	62	40	28	21	16	13	10	9	7	6	6	5	4	4	4	3	3
17	1111	278	124	70	45	31	23	18	14	12	10	8	7	6	5	5	4	4	4	3
18	1245	312	139	78	50	35	26	20	16	13	11	9	8	7	6	5	5	4	4	4
19	1387	347	155	87	56	39	29	22	18	14	12	10	9	8	7	6	5	5	4	4
20	1537	385	171	97	62	43	32	25	19	16	13	11	10	8	7	7	6	5	5	4
21	1695	424	189	106	68	48	35	27	21	17	15	12	11	9	8	7	6	6	5	5
22	1860	465	207	117	75	52	38	30	23	19	16	13	12	10	9	8	7	6	6	5
23	2033	509	226	128	82	57	42	32	26	21	17	15	13	11	10	8	8	7	6	6
24	2213	554	246	139	89	62	46	35	28	23	19	16	14	12	10	9	8	7	7	6
25	2401	601	267	151	97	67	49	38	30	25	20	17	15	13	11	10	9	8	7	7
26	2597	650	289	163	104	73	53	41	33	26	22	19	16	14	12	11	9	9	8	7
27	2801	701	312	176	113	78	58	44	35	29	24	20	17	15	13	11	10	9	8	8
28	3012	753	335	189	121	84	62	48	38	31	25	21	18	16	14	12	11	10	9	8
29	3231	808	359	202	130	90	66	51	40	33	27	23	20	17	15	13	12	10	9	9
30	3458	865	385	217	139	97	71	55	43	35	29	25	21	18	16	14	12	11	10	9
31	3692	923	411	231	148	103	76	58	46	37	31	26	22	19	17	15	13	12	11	10
32	3934	984	438	246	158	110	81	62	49	40	33	28	24	21	18	16	14	13	11	10
33	4184	1046	465	262	168	117	86	66	52	42	35	30	25	22	19	17	15	13	12	11
34	4441	1111	494	278	178	124	91	70	55	45	37	31	27	23	20	18	16	14	13	12
35	4706	1177	523	295	189	131	97	74	59	48	39	33	28	25	21	19	17	15	14	12
36	4979	1245	554	312	200	139	102	78	62	50	42	35	30	26	23	20	18	16	14	13
37	5260	1315	585	329	211	147	108	83	65	53	44	37	32	27	24	21	19	17	15	14
38	5548	1387	617	347	222	155	114	87	69	56	46	39	33	29	25	22	20	18	16	14
39	5844	1461	650	366	234	163	120	92	73	59	49	41	35	30	26	23	21	19	17	15
40	6147	1537	683	385	246	171	126	97	76	62	51	43	37	32	28	25	22	19	18	16

TABLE 1B	SAMPLE SIZESFOR A SINGLE PROPORTION FOR VARIOUS P and d for 95% level, Za=1.96																			
	d																			
P	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
0.01	16	4	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.02	31	8	4	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.03	45	12	5	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.04	60	15	7	4	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
0.05	73	19	9	5	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1
0.06	87	22	10	6	4	3	2	2	2	1	1	1	1	1	1	1	1	1	1	1
0.07	101	26	12	7	5	3	3	2	2	2	1	1	1	1	1	1	1	1	1	1
0.08	114	29	13	8	5	4	3	2	2	2	1	1	1	1	1	1	1	1	1	1
0.09	126	32	14	8	6	4	3	2	2	2	2	1	1	1	1	1	1	1	1	1
0.1	139	35	16	9	6	4	3	3	2	2	2	1	1	1	1	1	1	1	1	1
0.11	151	38	17	10	7	5	4	3	2	2	2	2	1	1	1	1	1	1	1	1
0.12	163	41	19	11	7	5	4	3	3	2	2	2	1	1	1	1	1	1	1	1
0.13	174	44	20	11	7	5	4	3	3	2	2	2	2	1	1	1	1	1	1	1
0.14	186	47	21	12	8	6	4	3	3	2	2	2	2	1	1	1	1	1	1	1
0.15	196	49	22	13	8	6	4	4	3	2	2	2	2	1	1	1	1	1	1	1
0.16	207	52	23	13	9	6	5	4	3	3	2	2	2	2	1	1	1	1	1	1
0.17	217	55	25	14	9	7	5	4	3	3	2	2	2	2	1	1	1	1	1	1
0.18	227	57	26	15	10	7	5	4	3	3	2	2	2	2	2	1	1	1	1	1
0.19	237	60	27	15	10	7	5	4	3	3	2	2	2	2	2	1	1	1	1	1
0.2	246	62	28	16	10	7	6	4	4	3	3	2	2	2	2	1	1	1	1	1
0.21	255	64	29	16	11	8	6	4	4	3	3	2	2	2	2	1	1	1	1	1
0.22	264	66	30	17	11	8	6	5	4	3	3	2	2	2	2	2	1	1	1	1
0.23	273	69	31	18	11	8	6	5	4	3	3	2	2	2	2	2	1	1	1	1
0.24	281	71	32	18	12	8	6	5	4	3	3	2	2	2	2	2	1	1	1	1
0.25	289	73	33	19	12	9	6	5	4	3	3	3	2	2	2	2	1	1	1	1
0.26	296	74	33	19	12	9	7	5	4	3	3	3	2	2	2	2	2	1	1	1
0.27	303	76	34	19	13	9	7	5	4	4	3	3	2	2	2	2	2	1	1	1
0.28	310	78	35	20	13	9	7	5	4	4	3	3	2	2	2	2	2	1	1	1
0.29	317	80	36	20	13	9	7	5	4	4	3	3	2	2	2	2	2	1	1	1
0.3	323	81	36	21	13	9	7	6	4	4	3	3	2	2	2	2	2	1	1	1
0.31	329	83	37	21	14	10	7	6	5	4	3	3	2	2	2	2	2	2	1	1
0.32	335	84	38	21	14	10	7	6	5	4	3	3	2	2	2	2	2	2	1	1
0.33	340	85	38	22	14	10	7	6	5	4	3	3	3	2	2	2	2	2	1	1
0.34	345	87	39	22	14	10	8	6	5	4	3	3	3	2	2	2	2	2	1	1
0.35	350	88	39	22	14	10	8	6	5	4	3	3	3	2	2	2	2	2	1	1
0.36	355	89	40	23	15	10	8	6	5	4	3	3	3	2	2	2	2	2	1	1
0.37	359	90	40	23	15	10	8	6	5	4	3	3	3	2	2	2	2	2	1	1
0.38	363	91	41	23	15	11	8	6	5	4	3	3	3	2	2	2	2	2	2	1
0.39	366	92	41	23	15	11	8	6	5	4	4	3	3	2	2	2	2	2	2	1
0.4	369	93	41	24	15	11	8	6	5	4	4	3	3	2	2	2	2	2	2	1

	Table 2A SAMPLE SZES FOR TWO MEANS FOR VARIOUS VALUES OF D AND SD **Za for 99% level=2.58** **Zb for 90% power = 1.28** **d**																			
Sd	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	30	7	3	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	119	30	13	7	5	3	2	2	1	1	1	1	1	1	1	0	0	0	0	0
3	268	67	30	17	11	7	5	4	3	3	2	2	2	1	1	1	1	1	1	1
4	477	119	53	30	19	13	10	7	6	5	4	3	3	2	2	2	2	1	1	1
5	745	186	83	47	30	21	15	12	9	7	6	5	4	4	3	3	3	2	2	2
6	1073	268	119	67	43	30	22	17	13	11	9	7	6	5	5	4	4	3	3	3
7	1460	365	162	91	58	41	30	23	18	15	12	10	9	7	6	6	5	5	4	4
8	1907	477	212	119	76	53	39	30	24	19	16	13	11	10	8	7	7	6	5	5
9	2414	603	268	151	97	67	49	38	30	24	20	17	14	12	11	9	8	7	7	6
10	2980	745	331	186	119	83	61	47	37	30	25	21	18	15	13	12	10	9	8	7
11	3606	901	401	225	144	100	74	56	45	36	30	25	21	18	16	14	12	11	10	9
12	4291	1073	477	268	172	119	88	67	53	43	35	30	25	22	19	17	15	13	12	11
13	5036	1259	560	315	201	140	103	79	62	50	42	35	30	26	22	20	17	16	14	13
14	5841	1460	649	365	234	162	119	91	72	58	48	41	35	30	26	23	20	18	16	15
15	6705	1676	745	419	268	186	137	105	83	67	55	47	40	34	30	26	23	21	19	17
16	7629	1907	848	477	305	212	156	119	94	76	63	53	45	39	34	30	26	24	21	19
17	8612	2153	957	538	344	239	176	135	106	86	71	60	51	44	38	34	30	27	24	22
18	9655	2414	1073	603	386	268	197	151	119	97	80	67	57	49	43	38	33	30	27	24
19	10758	2689	1195	672	430	299	220	168	133	108	89	75	64	55	48	42	37	33	30	27
20	11920	2980	1324	745	477	331	243	186	147	119	99	83	71	61	53	47	41	37	33	30
21	13141	3285	1460	821	526	365	268	205	162	131	109	91	78	67	58	51	45	41	36	33
22	14423	3606	1603	901	577	401	294	225	178	144	119	100	85	74	64	56	50	45	40	36
23	15764	3941	1752	985	631	438	322	246	195	158	130	109	93	80	70	62	55	49	44	39
24	17164	4291	1907	1073	687	477	350	268	212	172	142	119	102	88	76	67	59	53	48	43
25	18625	4656	2069	1164	745	517	380	291	230	186	154	129	110	95	83	73	64	57	52	47

Table 2B Sample sizes for two proportions at 95% Confidence level, (Zα=1.96) and for 90% Power(Zβ= 1.282)

	P1(Larger proportion)																		
P2	***10%***	***15%***	***20%***	***25%***	***30%***	***35%***	***40%***	***45%***	***50%***	***55%***	***60%***	***65%***	***70%***	***75%***	***80%***	***85%***	***90%***	***95%***	***100%***
10%		**914**	**263**	**130**	**79**	**53**	**39**	**29**	**22**	**18**	**14**	**11**	**9**	**7**	**5**	**4**	**3**	**2**	**1**
15%			**1209**	**331**	**158**	**93**	**62**	**44**	**32**	**25**	**19**	**15**	**12**	**9**	**7**	**5**	**4**	**3**	**2**
20%				**1461**	**389**	**181**	**105**	**69**	**48**	**35**	**26**	**20**	**16**	**12**	**9**	**7**	**5**	**4**	**3**
25%					**1671**	**436**	**200**	**114**	**74**	**51**	**37**	**27**	**21**	**16**	**12**	**9**	**7**	**5**	**4**
30%						**1839**	**473**	**214**	**121**	**77**	**53**	**38**	**28**	**21**	**16**	**12**	**9**	**6**	**5**
35%							**1965**	**499**	**223**	**125**	**79**	**53**	**38**	**27**	**20**	**15**	**11**	**8**	**6**
40%								**2050**	**515**	**228**	**126**	**79**	**53**	**37**	**26**	**19**	**14**	**10**	**7**
45%									**2092**	**520**	**228**	**125**	**77**	**51**	**35**	**25**	**18**	**12**	**9**
50%										**2092**	**515**	**223**	**121**	**74**	**48**	**32**	**22**	**15**	**11**
55%											**2050**	**499**	**214**	**114**	**69**	**44**	**29**	**19**	**13**
60%												**1965**	**473**	**200**	**105**	**62**	**39**	**25**	**16**
65%													**1839**	**436**	**181**	**93**	**53**	**32**	**20**
70%														**1671**	**389**	**158**	**79**	**43**	**25**
75%															**1461**	**331**	**130**	**62**	**32**
80%																**1209**	**263**	**97**	**42**
85%																	**914**	**184**	**60**
90%																		**578**	**95**
95%																			**200**
100%																			