12.1

Means → Proportions

- What % of seniors are planning to attend an IN school?
- Does the majority of LN students have a Facebook account?

$\hat{p} \longrightarrow p$

One Proportion Z Test

Conditions

1) SRS

2) Normal Sampling Distribution

HypTest: $np_0 \geq 10$, $n(1-p_0) \geq 10$

CI: $n\hat{p} \geq 10$, $n(1-\hat{p}) \geq 10$

3) $N > 10n$

Hypothesis Tests

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Compare to critical z^* or Use normalcdf

Confidence Intervals

$$CI = \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

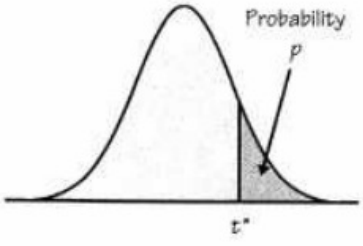

Table entry for p and C is the critical value t^* with probability p lying to its right and probability C lying between $-t^*$ and t^*.

TABLE C t distribution critical values

	Upper tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

1- PROPORTION Z TEST

This test is used to compare a sample proportion ($\hat{p}$) to a population proportion (p) or to determine a confidence interval for a population proportion.

In 1995, 7,741 students identified themselves as binge drinkers (from an SRS of 140 colleges and 17, 592 students).

Does this constitute strong evidence that more than 40% of college students were binge drinkers in 1995?

P) IDENTIFY POPULATION PARAMETER:

p = proportion of college students who were binge drinkers in 1995

H) STATE HYPOTHESES:

$H_0: p = .40$ $H_a: p > .40$

A) VERIFY CONDITIONS REQUIRED FOR TEST:

a) ✓SRS

Say so in problem

b) $N > 10n$

$N > 10(17,592) > 175,920$ ... probably

c) $n p_0 \geq 10$ $n(1-p_0) \geq 10$ – Norm Samp Dist

$(17592)(.40) \geq 10$? $(17592)(.60) \geq 10$?

$7036.8 \geq 10$ ✓ $10,555.2 \geq 10$ ✓

T) PERFORM TEST USING

$\hat{p} = \frac{7741}{17592} = .44$

a) TABLE A:

Calculate z test statistic and check Table:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{.44 - .40}{\sqrt{\frac{(.40)(.60)}{17592}}} = 10.84$$

$\downarrow$

$p = 0$

$(p = 1 \times 10^{-27})$

b) CALCULATOR:

STAT → TESTS → 1-Prop Z Test → $z = 10.84$, $p = 1 \times 10^{-27}$

S) STATE CONCLUSION:

There is very strong evidence ($p = 0$) to reject H_0 and conclude that more than 40% of college students were binge drinkers in 1995

CONFIDENCE INTERVAL (Use PAIS):

Norm Samp Dist

After checking for normal distribution [$n\hat{p} > 10$ $n(1-\hat{p}) > 10$], a 95% confidence interval for the proportion of college students who have engaged in binge drinking can be found using:

STAT ---> TEST ---> A: 1-Prop Z Int = (.433, .447)

We are 95% confident that between 43% and 45% of college students were binge drinkers in 1995.

Sample Size vs Desired Margin of Error

$$CI = \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

MOE

Ex How large of a sample would be needed to construct a 95% CI with a margin of error $\leq 3\%$

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq .03$$

$$1.960 \sqrt{\frac{(.50)(.50)}{n}} \leq .03$$

$$\frac{(1.960)(.50)}{\sqrt{n}} \leq \frac{.03}{1}$$

$$.03\sqrt{n} \geq (1.960)(.50)$$

$$n \geq 1067.1$$

Sec 12.2

2-Proportion Z Test

Conditions (For Both Samples)

1) Independent Samples

2) SRS

3) Normal Sampling Distributions

$n_1\hat{p}_1 \geq 5$ $\quad$ $n_2\hat{p}_2 \geq 5$

$n_1(1-\hat{p}) \geq 5$ $\quad$ $n_2(1-\hat{p}_2) \geq 5$

4) $N > 10n$

Hypothesis Test

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\,(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$\hat{p} = \frac{\text{total \# successes in both samples}}{\text{total \# observations in both samples}}$$

Pooled Sample Proportion

Confidence Interval

$$CI = (\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

2-PROPORTION Z TEST

This test is used to compare proportions from 2 independent samples.

In s study done in Michigan, it was determined 38 (out of 62) poor children who attended pre-school needed social services later in life compared to 49 (out of 61) poor children who did not attend preschool.

Does this study provide significant evidence that preschool reduces the need for social services later in life?

P) IDENTIFY POPULATION PARAMETERS:

p_1 = proportion of preschooled children needing social services

p_2 = proportion of nonpreschooled children needing social services

H) STATE HYPOTHESES:

$H_0: p_1 = p_2$ $H_a: p_1 < p_2$

A) VERIFY CONDITIONS REQUIRED FOR TEST:

a) SRS ?

Results may invalid...

b) Normal Sampling Distribution:

$n_1 \hat{p}_1 \geq 5$ $n_1 (1 - \hat{p}_1) \geq 5$ $n_2 \hat{p}_2 \geq 5$ $n_2 (1 - \hat{p}_2) \geq 5$

$38 \geq 5$ ✓ $24 \geq 5$ ✓ $49 \geq 5$ ✓ $12 \geq 5$ ✓

c) $N > 10n$

$N_1 > 10(62) > 620$ preschooled ... probably

$N_2 > 10(61) > 610$ nonpreschooled ... probably

T) PERFORM TEST USING

a) **TABLE C**

Calculate z-statistic and compare to critical values from Table C:

$\hat{p}_1 = \frac{38}{62} = .61$ $\hat{p}_2 = \frac{49}{61} = .80$

$\hat{p} = \frac{\text{total number of success in both samples}}{\text{total number of observations in both samples}} = \frac{87}{123} = .71$

$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{.61 - .80}{\sqrt{(.71)(.29)\left(\frac{1}{62} + \frac{1}{61}\right)}} = -2.32$

P-value < .05 Since $-2.32 < -1.645$

[normalcdf $(-10, -2.32) \approx .011$]

b) **CALCULATOR:**

STAT → TESTS → 2-PropZTest → $z = -2.32$, $p = .0102$

S) STATE CONCLUSION:

There is evidence ($p < .05$) to reject the H_0 and conclude that the proportion of preschooled poor children needed social services less than the proportion of nonpreschooled poor children in Michigan

CONFIDENCE INTERVAL (Use PAIS):

Construct a 95% confidence interval for the difference in proportions of people needing social services after attending preschool:

STAT ---> TESTS ---> B: 2-Prop Z Int = (-.35, -.03)

We are 95% confident that the percentage of people needing social service after attending preschool was between 3% and 35% less than those who did not attend preschool. (The interval is wide because the samples are quite small.)