TOPIC: TWO GROUP TESTS FOR PROPORTIONS

Testing for the Difference Between Two Proportions

Using the normal approximation,

$$Z = \frac{P_{s1} - P_{s2}}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where,

P_{s1} = Sample proportion in population 1.

P_{s2} = Sample proportion in population 2.

$\overline{p}$ = pooled estimate of population proportion.

X_1 = # of "successes" in sample 1

X_2 = # of "successes" in sample 2

n_1 = sample size for group 1

n_2 = sample size for group 2

$P_{s1} = \frac{X_1}{n_1}$ $\qquad$ $P_{s2} = \frac{X_2}{n_2}$

and

$$\left[\overline{p} = \frac{X_1 + X_2}{n_1 + n_2}\right]$$

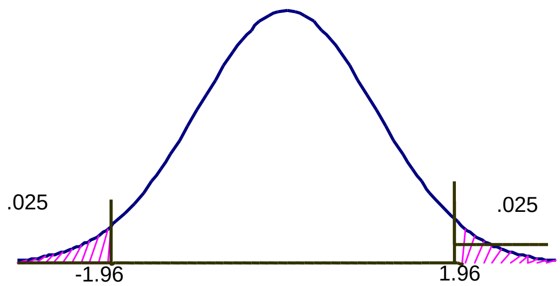
We use this test to compare two proportions. We are trying to determine whether the difference between two proportions is statistically significant or just sampling error.

EXAMPLE

Compare death rates of liver transplants at 2 hospitals. Test at α=.05

Hospital A 77/100 died within 6 months
Hospital B 120/200 died within 6 months

$$H_0 : P_1 = P_2$$
$$H_1 : P_1 \neq P_2$$



$$P_{S1} = .77$$
$$P_{s2} = .60$$

$$\overline{P} = \frac{77+120}{300} = \frac{197}{300} = .657$$

$$Z = \frac{.77-.60}{\sqrt{(.657)(.343)\left(\frac{1}{100}+\frac{1}{200}\right)}} = \frac{.17}{.058} = 2.93$$

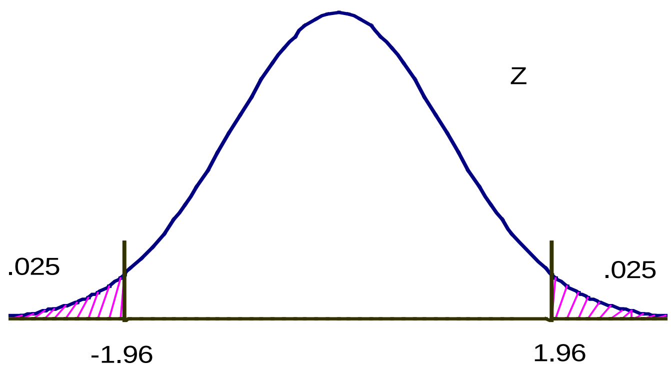
Reject H_0 $P < .05$.

EXAMPLE
Unemployment rates in 2 Counties. Is there a difference in the unemployment rates?
Test at $\alpha = .05$.

County A: 100/400 Unemployed **(25%)**
County B: 44/200 Unemployed **(22%)**

$$H_0 : P_1 = P_2$$
$$H_1 : P_1 \neq P_2$$

$$P_{s1} = .25$$
$$P_{s2} = .22$$
$$\bar{p} = \frac{144}{600} = .24$$



$$Z = \frac{.25 - .22}{\sqrt{(.24)(.76)\left(\frac{1}{400} + \frac{1}{200}\right)}} = \frac{.03}{.037} = .81$$

Do not reject H_0 **p > .05.**

EXAMPLE

Is there a difference between the two suppliers in proportion of defectives?
Teat at α=.05
Suppler A: 24/80 chips =defective
Suppler B: 30/60 chips =defective

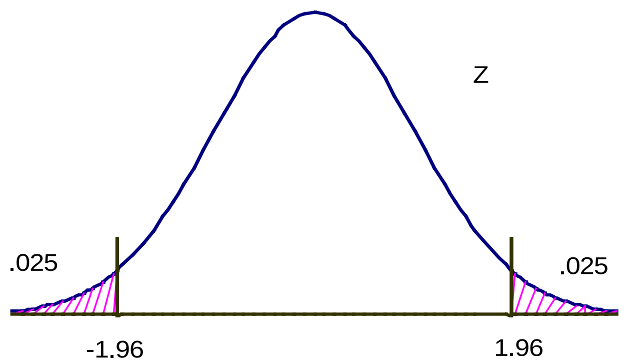
$$\mathrm{H}_0 : P_1 = P_2$$

$$\mathrm{H}_1 : P_1 \neq P_2$$

$$P_{S1} = .30$$

$$P_{s2} = .50$$

$$\overline{p} = \frac{54}{140} = .386$$



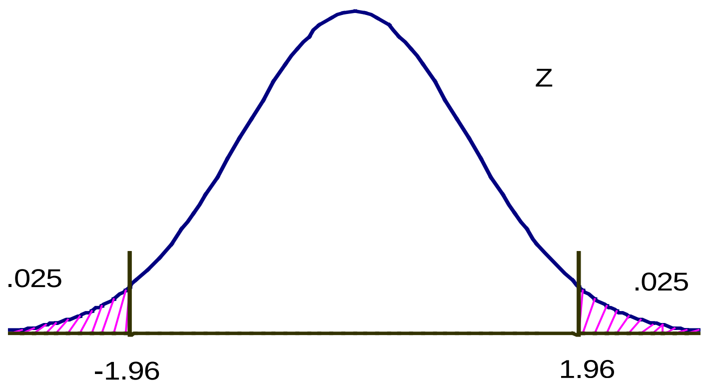
$$Z = \frac{.30 - .50}{\sqrt{(.386)(.614)\left(\frac{1}{80} + \frac{1}{60}\right)}} = \frac{-.20}{.083} = -2.41$$

Reject H_0 $\quad$ $P <. 05.$

EXAMPLE
Effect of estrogen on Alzheimer's disease
Test at α=.05

Women receiving estrogen: 7/100 developed Alzheimer
Women not receiving estrogen: 27/150 developed Alzheimer

$$P_{S1} = .07$$
$$P_{s2} = .18$$
$$\mathrm{H}_0 : P_1 = P_2$$
$$\mathrm{H}_1 : P_1 \neq P_2$$



$$\bar{p} = \frac{34}{250} = .136$$

$$Z = \frac{.07 - .18}{\sqrt{(.136)(.864)\left(\frac{1}{100} + \frac{1}{150}\right)}} = \frac{-.11}{.044} = -2.5$$

Reject H_0 $\quad$ $P < .05$.

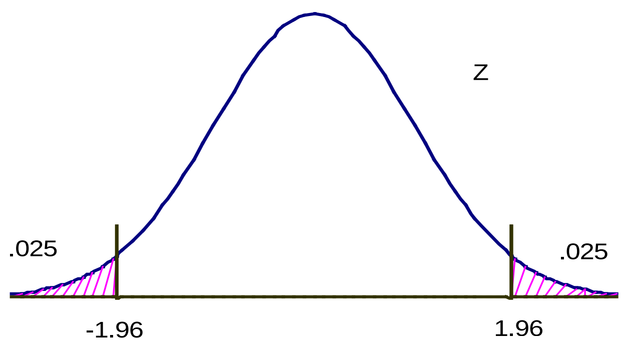
EXAMPLE

Direct Mail –Should Company use Sweepstakes. Test at α=.05

	Sweepstakes	No Sweepstakes
Mailed Out	5,000	4,000
# of Orders	100	60

$$H_0 : P_1 = P_2$$

$$H_1 : P_1 \neq P_2$$



$$P_{S1} = .020$$

$$P_{s2} = .015$$

$$\overline{p} = 160/9000 = .018$$

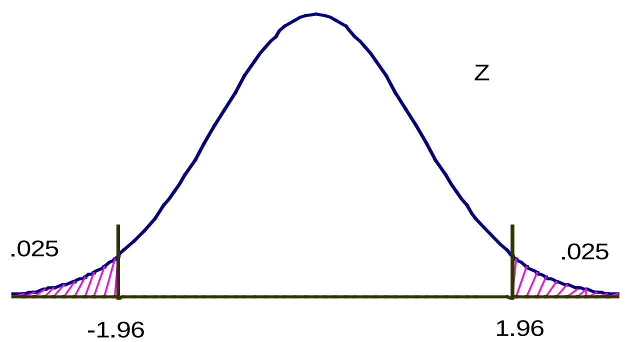
$$Z = \frac{.020 - .015}{\sqrt{(.018)(.982)\left(\frac{1}{5,000} + \frac{1}{4,000}\right)}} = \frac{.005}{.0028} = 1.79$$

Do not reject H_0. P > .05.

EXAMPLE
Who does better under adverse conditions, men or women?
Donner Party: Caught in blizzard and stranded in mountains for 6 months without food.

Women: 10/34 died
Men: 30/53 died

$$\mathrm{H}_0 : P_1 = P_2$$
$$\mathrm{H}_1 : P_1 \neq P_2$$



$$P_{S1} = .294$$
$$P_{s2} = .566$$
$$\overline{p} = 40/87 = .46$$

$$Z = \frac{.294 - .566}{\sqrt{(.46)(.54)\left(\frac{1}{34} + \frac{1}{53}\right)}} = \frac{-.272}{.1095} = -2.48$$

Reject H_0. P < .05 The difference in death rates between men and women was statistically significant. This study was published by Donald K. Grayson in a journal. His theory was that women have more fat on them (this is necessary so that the fetus will survive even if there is an insufficient amount of food) than men. This enables them to do better in times of famine.

(Another Real Example)

Interesting study published in *Family Practice*: Two groups were compared: One group of subjects were exposed to the cold by placing their feet in ice water; the other group (control group) placed their feet in an empty pail. 26/90 in first group developed cold within five days; 8/90 in control group developed colds. Is difference statistically significant? (Yes, $Z = 3.45$ $p < .05$).

Researchers concluded that mothers are right and not wearing enough clothing in the winter and thus becoming chilled can lead to colds.