Lecture 4

Hypothesis testing and statistical inferences

The Law of Large Numbers and Central Limit Theorem

The Law of Large Numbers

- Assume X_1 , X_2 , … , X_n are random samples of X. $E(X) = \mu$ and *V*(X) exist. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, then for any given $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left\{\left|\overline{X} - \mu\right| < \varepsilon\right\} = 1$$

The Central Limit Theorem

Let $\overline{X}$ be the mean of a random sample X_1, X_2, …, X_n, of size *n* from a distribution with a finite mean μ and a finite positive variance σ^2. Then

$$Y = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \to N(0, 1)$$

Small probability event

- A **small probability event** is an event that has a low probability of occurring.
- The small probability event will hardly happen in one experiment. This principle is used for hypothesis and tests.
- An event is a small probability event, so it will hardly happen in theory. But if it happens actually, then we reject H_0.

Hypothesis and testing

Hypothesis testing: preliminaries

- A hypothesis is a statement that something is true.
- Null hypothesis: A hypothesis to be tested. We use the symbol H_0 to represent the null hypothesis
- Alternative hypothesis: A hypothesis to be considered as an alternative to the null hypothesis. We use the symbol H_a to represent the alternative hypothesis.
- The alternative hypothesis is the one believe to be true, or what you are trying to prove is true.

Hypothesis testing: Preliminaries

- In this course, we will always assume that the null hypothesis for a population parameter, Θ , always specifies a single value for that parameter. So, an equal sign always appears:

$$H_0 : \Theta = \Theta_0$$

- If the primary concern is deciding whether a population parameter is *different with* a specified value, the alternative hypothesis should be:

$$H_a : \Theta \neq \Theta_0$$

- This form of alternative hypothesis is called a two-tailed test.

Hypothesis testing: Preliminaries

- If the primary concern is whether a population parameter, Θ_0, is *less than* a specified value Θ , the alternative hypothesis should be:

$$H_a : \Theta < \Theta_0$$

- A hypothesis test whose alternative hypothesis has this form is called a left-tailed test.
- If the primary concern is whether a population parameter, Θ_0, is *greater than* a specified value Θ , the alternative hypothesis should be:

$$H_a : \Theta > \Theta_0$$

- A hypothesis test whose alternative hypothesis has this form is called a right-tailed test.
- A hypothesis test is called a one-tailed test if it is either right- or left-tailed, i.e., if it is not a two-tailed test.

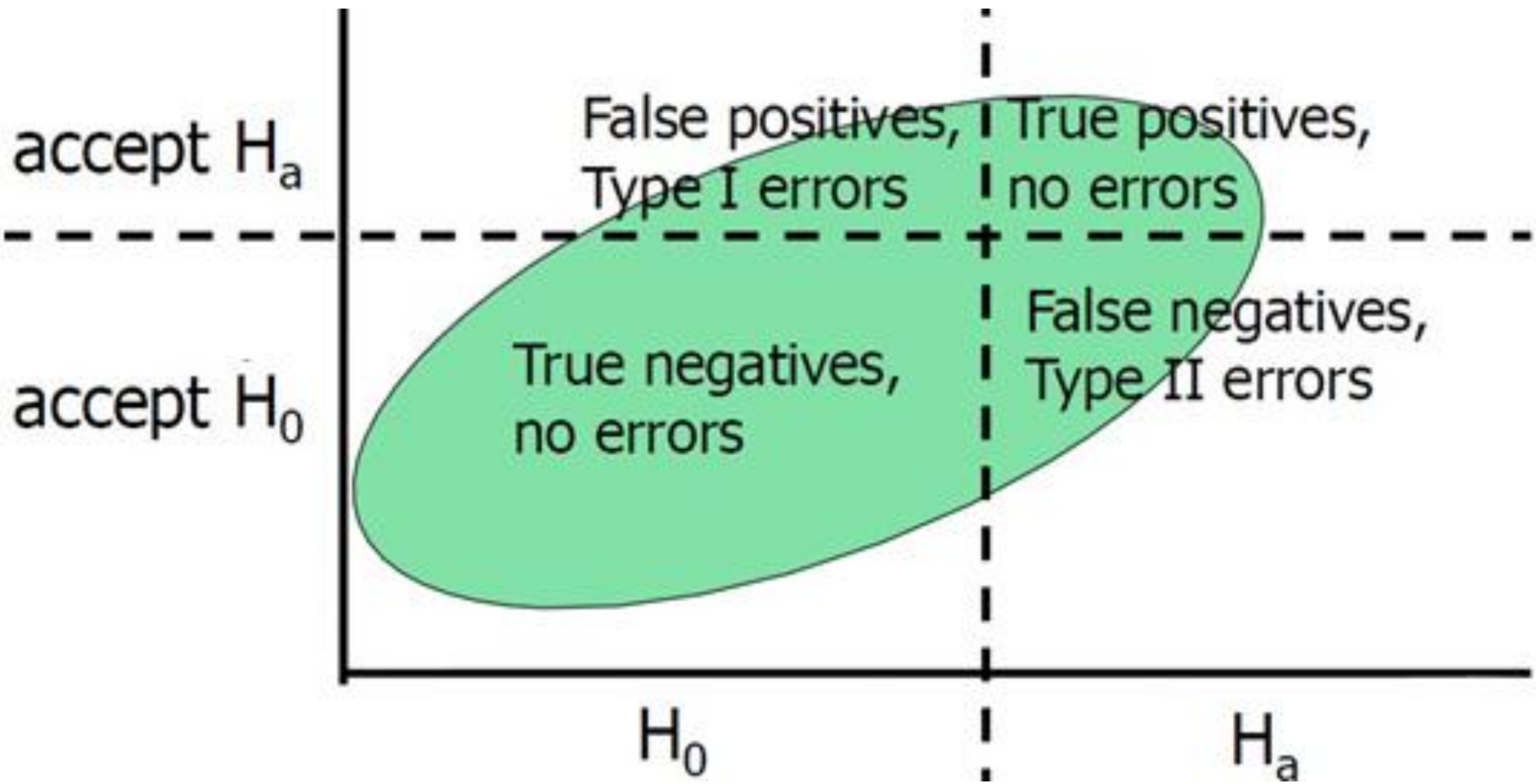
Hypothesis testing: Preliminaries

- After we have the null hypothesis, we have to determine whether to reject it or fail to reject it.
- The decision to reject or fail to reject is based on information contained in a sample drawn from the population of interest. The sample values are used to compute a single number, corresponding to a point on a line, which operates as a decision maker. This decision maker is called a ***test statistic***
- If test statistic falls in some interval which support alternative hypothesis, we reject the null hypothesis. This interval is called ***rejection region***
- It test statistic falls in some interval which support null hypothesis, we fail to reject the null hypothesis. This interval is called ***acceptance region***
- The value of the point, which divide the rejection region and acceptance one is called ***critical value***

Hypothesis testing: Preliminaries

- We can make mistakes in the test.
- **Type I error:** reject the null hypothesis when it is true.
- Probability of type I error is denoted by α
- **Type II error:** accept the null hypothesis when it is wrong.
- Probability of type II error is denoted by β

Hypothesis testing: Preliminaries



Test of hypothesis for a population mean

- We are basically asking: What observed value of random variable X would be different enough from my null hypothesis value to convince me that my null is wrong
- We always talk in terms of type I errors, alpha, which are always small, for example, 0.1, 0.05, 0.01
- The smaller alpha gets the more tight your proof that the alternative is correct, because the probability of type I error is reduced, but the chances of type II error are increased
- Now we will introduce some test on populations which obey normal distribution

Test of hypothesis for a population mean

(two tailed and large sample, i.e. variance $\sigma^2 = \sigma_0^2$ is known)

1) Hypothesis: $H_0 : \mu = \mu_0$

$$H_a : \mu \neq \mu_0$$

2) Test statistic: large sample case

$$z_{obs} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

3) Critical value, rejection and acceptance region:

- The bigger the absolute value of z is, the more possible to reject null hypothesis.
- The critical value depend on the significance level α
- rejection region: $|z_{obs}| > z_{\alpha/2 \text{ or crit}}$

Example

- A sample of 60 students' grades is taken from a large class, the average grade in the sample is 80 with a sample standard deviation 10. Test the hypothesis that the average grade is 75 with 5% significance level (probability of making a type I error).
- Hypothesis: $H_0 : \mu = 75$

  $H_a : \mu \neq 75$
- Test statistic:

$$z_{obs} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{80 - 75}{10/\sqrt{60}} = 3.87 \sim N(0, 1)$$

$$|z_{obs}| > z_{0.05/2} = 1.96, \text{so we reject } H_0$$

Test of hypothesis for a population mean

(two tailed and small sample, i.e. variance σ^2 is unknown)

1) Hypothesis: $H_0 : \mu = \mu_0$

$H_a : \mu \neq \mu_0$

2) Test statistic: small sample case

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \sim t(n-1)$$

3) Critical value, rejection and acceptance region:

- The bigger the absolute value of *t* is, the more possible to reject null hypothesis.
- The critical value depends on significance level α
- rejection region: $|t| > t_{\alpha/2}$

An example

- Suppose you have a sample of 11 Econ midterm exam grades. The mean of that sample is 81 with a standard deviation of 9. Test hypothesis that average grade of the population is 75 under the 5% significance level.
- Hypothesis: $H_0: \mu = 75$

  $H_a: \mu \neq 75$

- Test statistic:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{81 - 75}{9/\sqrt{11}} = 2.21 \sim t(10)$$

$$\left| t_{obs} \right| < t_{0.05/2}(10) = 2.23, \text{ so we cannot reject } H_0$$

Test of hypothesis for a population variance

(mean $\mu = \mu_0$ is known)

1) Hypothesis: $H_0 : \sigma^2 = \sigma_0^2$

$H_a : \sigma^2 \neq \sigma_0^2$

2) Test statistic:

$$\chi^2 = \frac{\sum_{i=1}^{n} (x_i - \mu_0)^2}{\sigma_0^2} \sim \chi^2(n)$$

3) Critical value, rejection and acceptance region:

rejection region: $\chi^2 > \chi^2_{\alpha/2}(n)$ or $\chi^2 < \chi^2_{1-\alpha/2}(n)$

An example

- Suppose that X~N(1.45, σ^2). X1,…, X5 are random samples of X with values 1.32, 1.55, 1.36, 1.40 and 1.44. Test hypothesis that $H_0 : \sigma^2 = 0.048^2$ under the 5% significance level.
- Hypothesis: $H_0 : \sigma^2 = 0.048^2$

  $H_a : \sigma \neq 0.048^2$
- Test statistic:

$$\chi^2 = \frac{\sum_{i=1}^{5} (x_i - 1.45)^2}{0.048^2} = 16.32 \sim \chi^2(5)$$

$$\chi^2{}_{obs} > \chi^2_{0.05/2}(5) = 12.83, \text{ so we reject } H_0$$

Test of hypothesis for a population variance

(mean μ is unknown)

1) Hypothesis: $H_0 : \sigma^2 = \sigma_0^2$

$H_a : \sigma^2 \neq \sigma_0^2$

2) Test statistic:

$$\chi^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma_0^2} \sim \chi^2(n-1)$$

3) Critical value, rejection and acceptance region:

rejection region: $\chi^2 > \chi^2_{\alpha/2}(n-1)$ or $\chi^2 < \chi^2_{1-\alpha/2}(n-1)$

An example

- Suppose that variance of heading date for wheat is 2.781 which is caused by environmental effects. Now sample variance of 58 F_2 plants is s^2=6.458. Test if the variance is caused by environment under the 5% significance level.
- Hypothesis: $H_0: \sigma^2 = 2.781$

  $H_a: \sigma \neq 2.781$

- Test statistic:

$$\chi^2 = \frac{\sum_{i=1}^{5}(x_i - \overline{\mathrm{x}})^2}{2.781} = \frac{57s^2}{2.781} = 132.26 \sim \chi^2(57)$$

$$\chi^2{}_{obs} > \chi^2_{0.05/2}(57) = 79.75, \text{so we reject } H_0$$

Test of hypothesis for binomial proportion

1) Hypothesis: $H_0 : p = p_0$

Two-tailed: $H_a : p \neq p_0$

2) Test statistic: large sample case

$$\hat{p} = \frac{x}{n} \qquad z_{obs} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} \quad approximately \sim N(0,1)$$

3) Critical value, rejection and acceptance region:

rejection region (two-tailed): $|z_{obs}| > z_{\alpha/2}$

An example

- Suppose the rate of a sickness in one area is 2.7%. Now we surveyed 278 persons in this area and found 10 of them had this sickness. Test if the sickness rate normal or not?
- Hypothesis: $H_0 : p = 2.7\%$

$$H_a : p \neq 2.7\%$$

- Test statistic:

$$z_{obs} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{10/278 - 0.027}{\sqrt{0.027 \bullet (1 - 0.027)/278}} = 0.923 \sim N(0,1)$$

$$\left| z_{obs} \right| < z_{0.05/2} = 1.96, \text{ so we cannot reject } H_0$$

Test of difference between two population means

- Population 1 and population 2 are two populations
- μ_1 =mean of data in population 1
- μ_2 =mean of data in population 2
- Two sets of samples: one from population 1, the other from population 2

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

Test of difference between two population means for pairwise data

- Assume the population size of the two populations are both *n*. The sampling distribution of difference between population mean $\bar{x}_1 - \bar{x}_2$ is a normal distribution with mean $$\mu_{(\bar{x}_1 - \bar{x}_2)} = \bar{x}_1 - \bar{x}_2$$

and the standard deviation is

$$s_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sum_{i=1}^{n}[(x_{1i} - x_{2i}) - (\bar{x}_1 - \bar{x}_2)]^2}{n(n-1)}}$$

Test for difference of two population means

(pairwise data)

1) Hypothesis: D_0 is some specified difference that you wish to test. For many tests, you will wish to hypothesize that there is no difference between two means, that is $D_0=0$

$$H_0 : \mu_1 = \mu_2 \qquad H_a : \mu_1 \neq \mu_2$$

2) Test statistic: large sample case

$$t_{obs} = \frac{(\overline{x_1} - \overline{x_2}) - D_0}{s_{(\overline{x_1} - \overline{x_2})}} = \frac{(\overline{x_1} - \overline{x_2}) - D_0}{\sqrt{\dfrac{\sum_{i=1}^{n} [(x_{1i} - x_{2i}) - (\overline{x}_1 - \overline{x}_2)]^2}{n(n-1)}}} \sim t(n-1)$$

3) Critical value, rejection and acceptance region:

rejection region: $|t_{obs}| > t_{\alpha/2}$

An Example

- Ten silicosis patients were treated by a new medicine. Their HGB before and after treatment were

Patient i	1	2	3	4	5	6	7	8	9	10
Before (x1)	11.3	15.0	15.0	13.5	12.8	10.0	11.0	12.0	13.0	12.3
After (x2)	14.0	13.8	14.0	13.5	13.5	12.0	14.7	11.4	13.8	12.0
D=x1-x2	-2.7	1.2	1.0	0.0	-0.7	-2.0	-3.7	0.6	-0.8	0.3

- Test if the treatment may cause the change in HGB.

An Example

- Hypothesis: $H_0 : \mu_1 = \mu_2$

  $H_a : \mu_1 \neq \mu_2$

- Test statistic:

$$t_{obs} = \frac{\overline{x_1} - \overline{x_2}}{s_{(\overline{x_1} - \overline{x_2})}} = \frac{-0.68}{\sqrt{\dfrac{\sum_{i=1}^{n}[(x_{1i} - x_{2i}) + 0.68]^2}{10 \bullet 9}}} = -1.3067 \sim t(9)$$

$$\left|t_{obs}\right| < t_{0.05/2}(9) = 2.262, \text{so we cannot reject } H_0$$

Test of difference between two population means (large sample)

- In large sample case (i.e. σ_1^2 and σ_2^2 is known), the sampling distribution of $\overline{x}_1 - \overline{x}_2$ difference between population mean is a normal distribution with mean

$$\mu_{(\overline{x}_1 - \overline{x}_2)} = \mu_1 - \mu_2$$

and the standard deviation is

$$\sigma_{(\overline{x}_1 - \overline{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Test for difference of two population means

(two tailed and large sample)

1) Hypothesis: D_0 is some specified difference that you wish to test. For many tests, you will wish to hypothesize that there is no difference between two means, that is $D_0=0$

$$H_0 : \mu_1 - \mu_2 = D_0$$

$$H_a : \mu_1 - \mu_2 \neq D_0$$

2) Test statistic: large sample case

$$z_{obs} = \frac{(\overline{x_1} - \overline{x_2}) - D_0}{\sigma_{(\overline{x_1} - \overline{x_2})}} = \frac{(\overline{x_1} - \overline{x_2}) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

3) Critical value, rejection and acceptance region:

rejection region: $| z_{obs} | > z_{\alpha/2}$

Example: compare salary difference

- Population 1: faculty in public schools
- Population 2: faculty in private schools

- μ_1 =mean salary of faculty in public schools
- μ_2 =mean salary of faculty in private schools

- Sample 1: salaries of faculty members in public schools (n=30)
- Sample 2: salaries of faculty members in private schools (n=35)

$$\overline{x_1} = 57.48 \qquad \overline{x_2} = 66.39$$

$$s_1 = 9 \qquad s_2 = 9.5$$

- Test the hypothesis that the salaries are less for faculty in public school with 5% significance level

Example: compare salary difference

- They are large size populations.

$$z_{obs} = \frac{(\overline{x_1} - \overline{x_2})}{s_{(\overline{x_1} - \overline{x_2})}} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = -2.18$$

$$|z_{obs}| > z_{0.05/2} = 1.96, \text{so we reject } H_0$$

Test of difference between two population variances

1) Hypothesis: $H_0 : \sigma_1^2 = \sigma_2^2$

$H_a : \sigma_1^2 \neq \sigma_2^2$

2) Test statistic: small sample case

$$F_{obs} = \frac{s_1^2}{s_2^2} \sim F(n_1 - 1, n_2 - 1), \text{assume } s_1^2 > s_2^2$$

3) Critical value, rejection and acceptance region:

rejection region: $F_{obs} > F_{1-\alpha/2}$

An Example

- Two groups of mice with different feeds. The increased weights after eight weeks are (g):
- Gourp1: n1=12. Weights: 83, 146, 119, 104, 120, 161, 107, 134, 115, 129, 99, 123.
- Group 2: n2=7. Weights: 70, 118, 101, 85, 107, 132, 94.
- Test if the variances of the two groups are different.

An Example

- Hypothesis: $H_0: \sigma_1^2 = \sigma_2^2$

  $H_a: \sigma_1^2 \neq \sigma_2^2$

- Test statistic:

$$s_1^2 = 445.82; s_2^2 = 425.33$$

$$F_{obs} = \frac{s_1^2}{s_2^2} = 1.048 \sim F(11,6)$$

$$F_{obs} < F_{0.975}(11,6) = 5.43, \text{ so we cannot reject } H_0$$

Test of difference between two population means

- In small sample case, σ_1^2 and σ_2^2 is unknown, but $\sigma_1^2 = \sigma_2^2$, the sampling distribution of the difference between two means is the *t*-distribution with mean

$$\mu_{(\overline{x_1} - \overline{x_2})} = \mu_1 - \mu_2$$

- and standard deviation

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

- with $\boldsymbol{n_1 + n_2 - 2}$ degrees of freedom

Test for difference of two population means

(two tailed and small sample with the same variance)

1) Hypothesis:

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

2) Test statistic: small sample case

$$t_{obs} = \frac{\overline{x_1} - \overline{x_2}}{s} \sim t(n_1 + n_2 - 2)$$

3) Critical value, rejection and acceptance region:

rejection region $|t_{obs}| > t_{\alpha/2}$

Example: compare salary difference

- Population 1: faculty in public schools
- Population 2: faculty in private schools

- μ_1 =mean salary of faculty in public schools
- μ_2 =mean salary of faculty in private schools
- Sample 1: salaries of faculty members in public schools (n=10)
- Sample 2: salaries of faculty members in private schools (n=15)

$$\overline{x_1} = 57.48 \qquad \overline{x_2} = 66.39$$

$$s_1 = 9 \qquad s_2 = 9.5$$

- Test the hypothesis that the salaries are the same for faculty in public and private school with 5% significance level

Example: compare salary difference

- We have smaller sample sizes.

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = 14.44$$

$$t_{obs} = \frac{\overline{x_1} - \overline{x_2}}{s} = \text{-}1.32$$

$$|t_{obs}| < t_{0.05/2}(23) = 2.07, \text{ we accept } H_0 \text{ or cannot reject } H_0$$

Test for difference of two population means

(two tails and small sample with different variances)

1) Hypothesis:

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

2) Test statistic: small sample case

$$t_{obs} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ approximately } \sim t(\upsilon)$$

Where

$$\upsilon = INTEGER\left[\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right) \Big/ \left(\left(\frac{s_1^2}{n_1}\right)^2 \Big/ (n_1 - 1) + \left(\frac{s_2^2}{n_2}\right)^2 \Big/ (n_2 - 1)\right)\right] (rounded)$$

3) Critical value, rejection and acceptance region:

rejection region $|t_{obs}| > t_{\alpha/2}$

Simplification

- Simplify the formula for degree of freedom in some special conditions: assume $n_1 = n_2 = n$, then

$$\upsilon = NINT\left[(n-1)\frac{\left(s_1^2 + s_2^2\right)^2}{s_1^4 + s_2^4}\right]$$

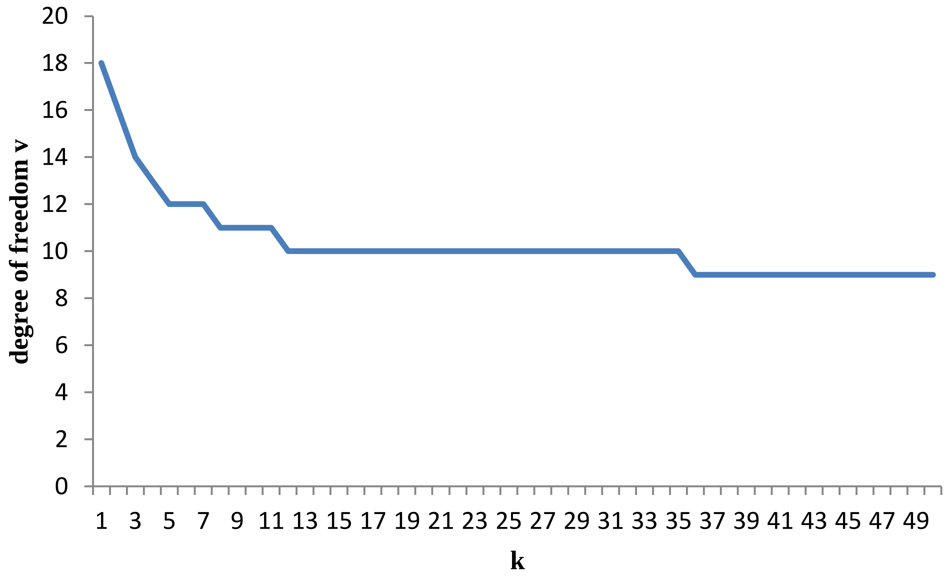
- Assume $s_1 = s_2$, $n_1 = n_2 = n$, then $\upsilon = 2(n-1)$

- In this case, $\upsilon = 2(n-1) = n_1 + n_2 - 2$

Degree of freedom

- Assume $s_1^2 = k s_2^2$ (k>=1), n_1=n_2=n, then

$$\upsilon = NINT\left[(n-1)\frac{(k+1)^2}{k^2+1}\right] \xrightarrow{k\to\infty} n-1$$



An example

- Suppose the blood pressures of 20 older people $\bar{x}_1 = 137mmHg, s_1^2 = 938$; 20 younger people $\bar{x}_2 = 128mmHg, s_1^2 = 193$. Test if the blood pressures of the two groups are different.
- First, test $H_0 : \sigma_1^2 = \sigma_2^2$

$$F_{obs} = \frac{s_1^2}{s_2^2} = \frac{938}{193} = 4.8601 \sim F(19,19)$$

$$F_{obs} > F_{0.975}(19,19) = 2.51, \text{so we reject } H_0$$

$$\text{So } \sigma_1^2 \neq \sigma_2^2$$

An example

- Then test $H_0 : \mu_1 = \mu_2$

$$\upsilon = INTEGER\left[\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2 \Big/ \left(\left(\frac{s_1^2}{n_1}\right)^2 \Big/ (n_1 - 1) + \left(\frac{s_2^2}{n_2}\right)^2 \Big/ (n_2 - 1)\right)\right] (rounded) \approx 27$$

- Test statistic:

$$t_{obs} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = \frac{137 - 128}{\sqrt{\frac{938}{20}+\frac{193}{20}}} = 1.1968 \text{ approximately } \sim t(27)$$

$$|t_{obs}| < t_{0.05/2}(27) = 2.052, \text{ we cannot reject } H_0$$

Test of hypothesis for difference in binomial proportions

1) Hypothesis: $H_0 : p_1 = p_2$

$H_a : p_1 \neq or > or < p_2$ one/two tail tests

2) Test statistic:

$$z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{s_{\hat{p}_1 - \hat{p}_2}} \text{ approximately} \sim N(0,1)$$

Where $s_{\hat{p}_1 - \hat{p}_2} = \sqrt{(\frac{1}{n_1} + \frac{1}{n_2})\hat{p}\hat{q}}$ $\hat{p} = \frac{n_1}{n_1 + n_2}\hat{p}_1 + \frac{n_2}{n_1 + n_2}\hat{p}_2, \hat{q} = 1 - \hat{p}$

3) Critical value, rejection and acceptance region:

rejection region $|z_{obs}| > z_{\alpha/2}$ or $z_{obs} > z_{\alpha}$ or $z_{obs} < -z_{\alpha}$

An example

- Two pesticides: one killed 460 of 700 pets; the other killed 364 of 500. Test if the two pesticides had similar effects.
- Hypothesis: $H_0 : p_1 = p_2$
- Test statistic:

$$z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{s_{\hat{p}_1 - \hat{p}_2}} = \frac{0.657 - 0.728}{0.02716} = -2.6141 \text{ approximately} \sim N(0,1)$$

$$\left| z_{obs} \right| > z_{0.05/2} = 1.96, \text{ so we reject } H_0$$

P-values

- The smallest value of alpha for which test results are statistically significant, or in other words, statistically different than the null hypothesis value.
- Smallest value at which you still reject the null.
- Example 1: You see a *p*-value of 0.025
- You would fail to reject at a 1% level of significance, but reject at 5%
- Example 2: 60 students are polled average of 72 observed with a standard deviation of 10, what is the p-value of the test whether the population average is 75?

Power of a statistical test

- P(reject the null hypothesis when it is false)=1-β
- (1-α) is the probability we accept the null when it was in fact true
- (1-β) is the probability we reject when the null is in fact false - this is the power of the test.
- The power changes depending on what the actual population parameter is.

Impact factors of power

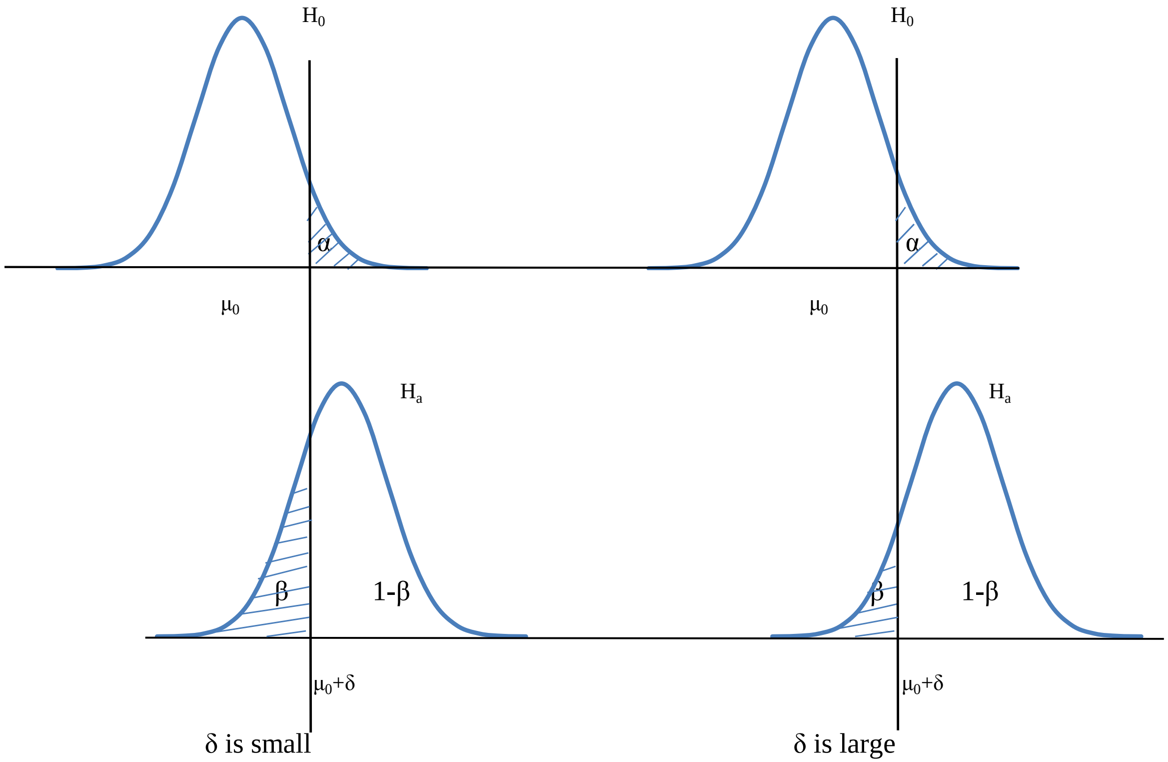
- For example: H_0: $\mu = \mu_0$, H_a: $\mu > \mu_0$
- Test statistic $$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$
- If we want to reject H_0, we need

$$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \geq Z_\alpha$$

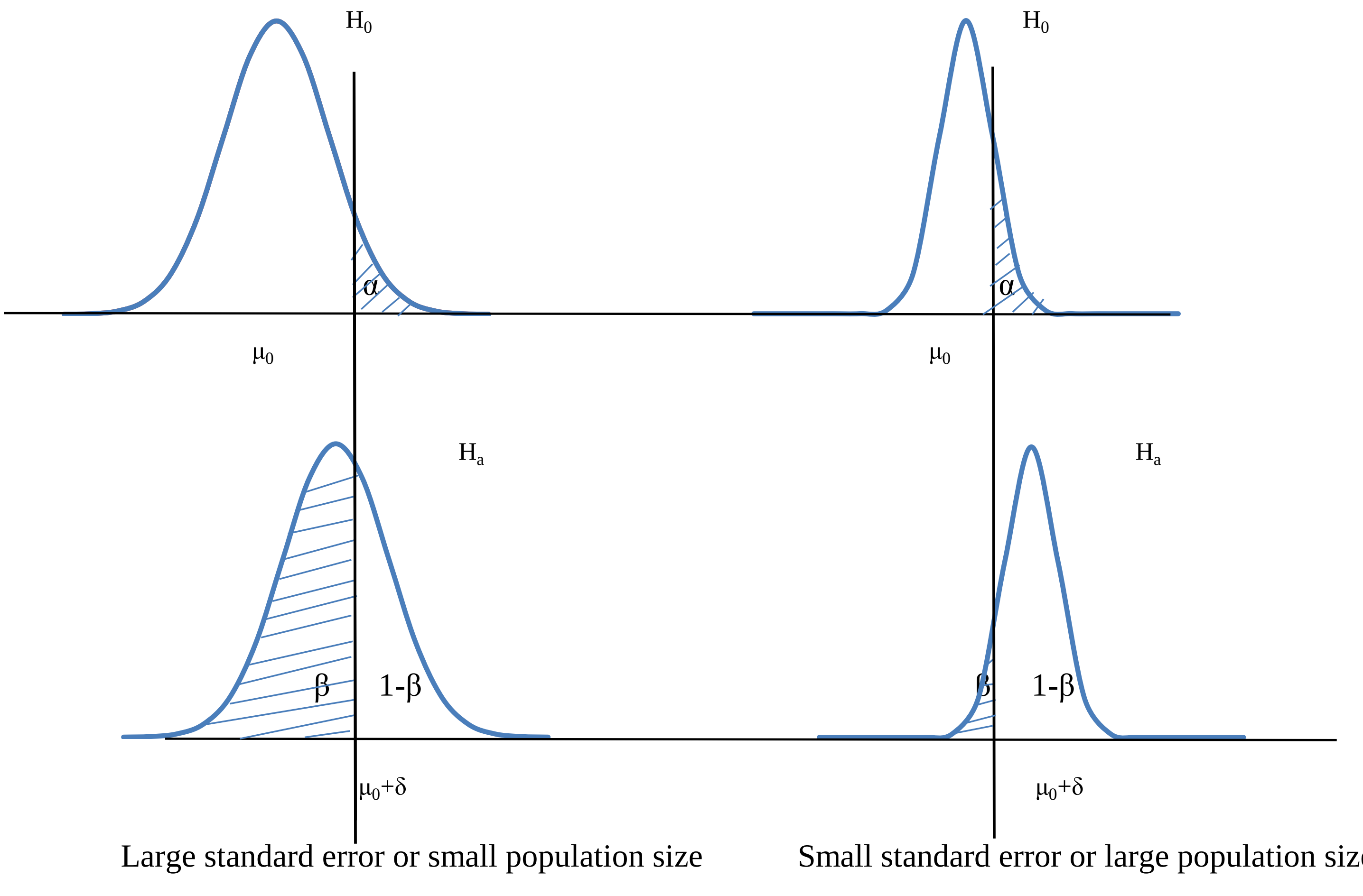
- So the power depends on $\delta = \bar{x} - \mu_0$, σ, n, and α

The larger the difference δ is, the higher the power is

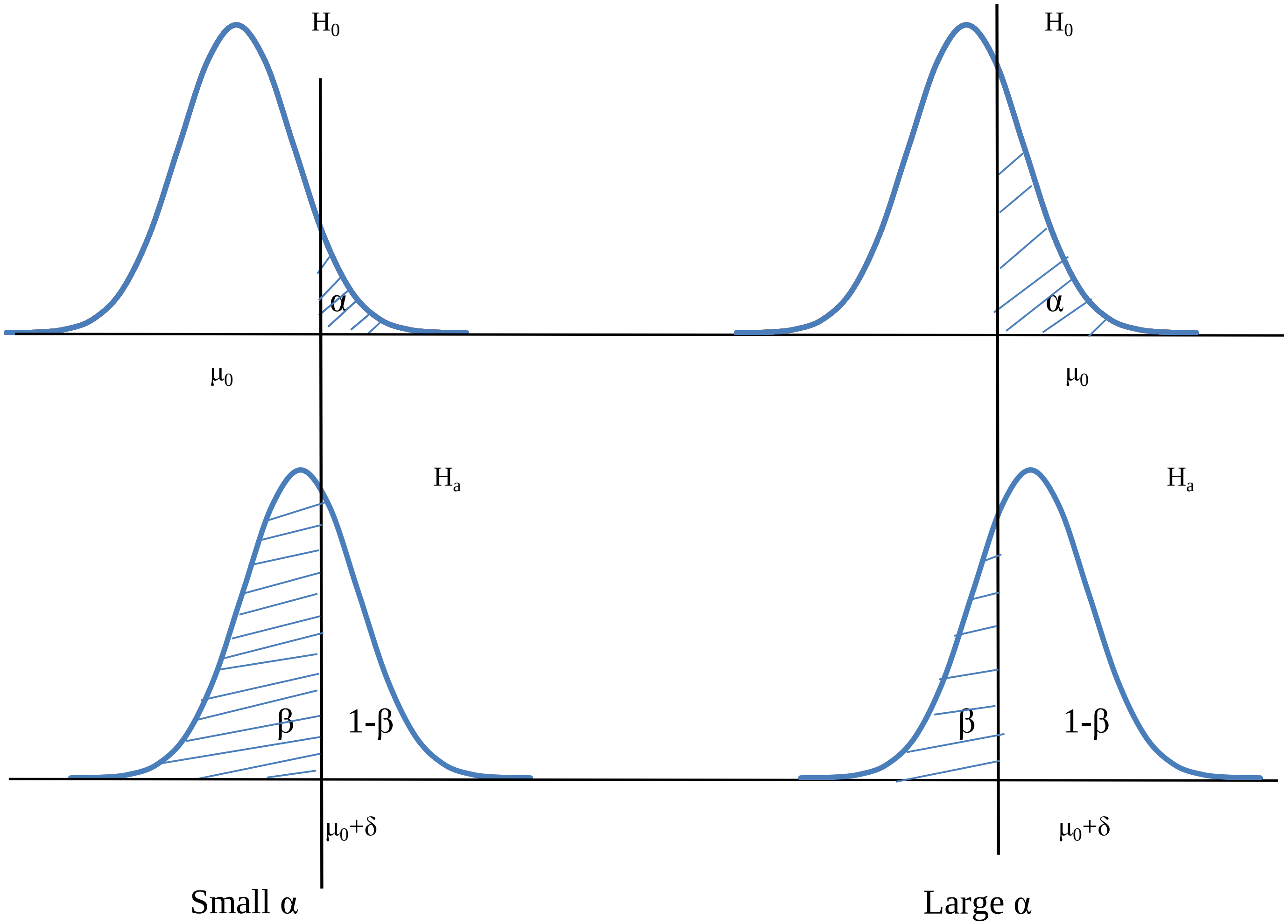
- $\bar{X} \sim N(\mu, \sigma^2/n)$
- If H_0 is true, $\bar{X} \sim N(\mu_0, \sigma^2/n)$
- If H_a is true, $\bar{X} \sim N(\mu_0+\delta, \sigma^2/n)$



The smaller the standard error or the larger the population size is, the higher the power is



The larger α is, the higher the power is



How many replications?
Or what is a suitable sample size?

- The number of replications in a research study affects the precision of estimates for treatment means and the power of statistical tests to detect differences among the means of treatment groups.
- The method for determining the number of replications is often based on a test of a hypothesis about differences among treatment group means.

How many replications?

- The required number of replications is affected primarily by four factors that are required for calculations:
  - The variance (σ^2) or the percent coefficient of variation (%CV)
  - The size of difference (that has physical significance) between two means (δ)
  - The significance level of the test (α), or the probability of Type I error
  - The power of test $1-\beta$, or the probability of detecting δ, where β is the probability of a Type II error

The required replication number for each treatment group

- The required replication number for each treatment group, r, for two-sided alternatives is estimated with

$$r \geq 2(z_{\alpha/2} + z_{\beta})^2 \left(\frac{\sigma}{\delta} \right)^2$$

- Where $z_{\alpha/2}$ is the standard normal variable exceeded with probability α/2 and z_{β} is exceeded with probability β.

Test for One Mean

$$H_0: \mu = \mu_0 \qquad H_a: \mu = \mu_1 (= \mu_0 + \delta)$$

$$\text{Under } H_0, Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{r}} \sim N(0,1)$$

$$\text{Under } H_a, Z = \frac{\bar{X} - \mu_1}{\sigma/\sqrt{r}} \sim N(0,1)$$

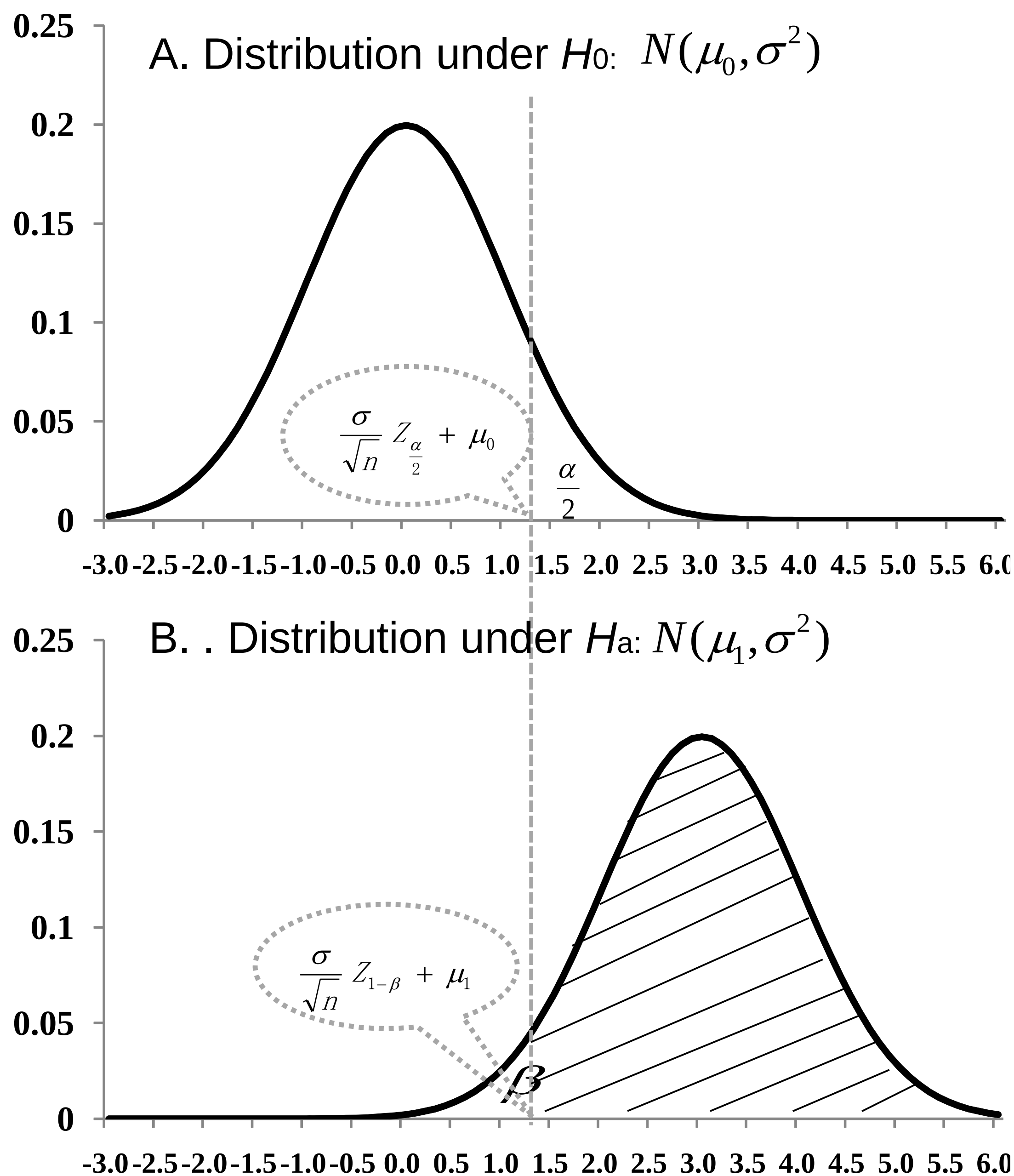
$$x = \frac{\sigma}{\sqrt{r}} Z_{\frac{\alpha}{2}} + \mu_0$$

$$= \frac{\sigma}{\sqrt{r}} Z_{1-\beta} + \mu_1$$

$$r = \left[\frac{\left(Z_{1-\beta} - Z_{\frac{\alpha}{2}} \right) \sigma}{\mu_0 - \mu_1} \right]^2$$

$$Z_{1-\beta} = Z_\beta$$

$$r = (z_{\alpha/2} + z_\beta)^2 \left(\frac{\sigma}{\delta} \right)^2$$



Test for Two Means

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 = \delta, \delta \neq 0$$

$$\text{Under } \mathrm{H}_0, Z = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{2/n}} \sim N(0,1)$$

$$\text{Under } \mathrm{H}_a, Z = \frac{\overline{X}_1 - \overline{X}_2 - \delta}{\sigma\sqrt{2/n}} \sim N(0,1)$$

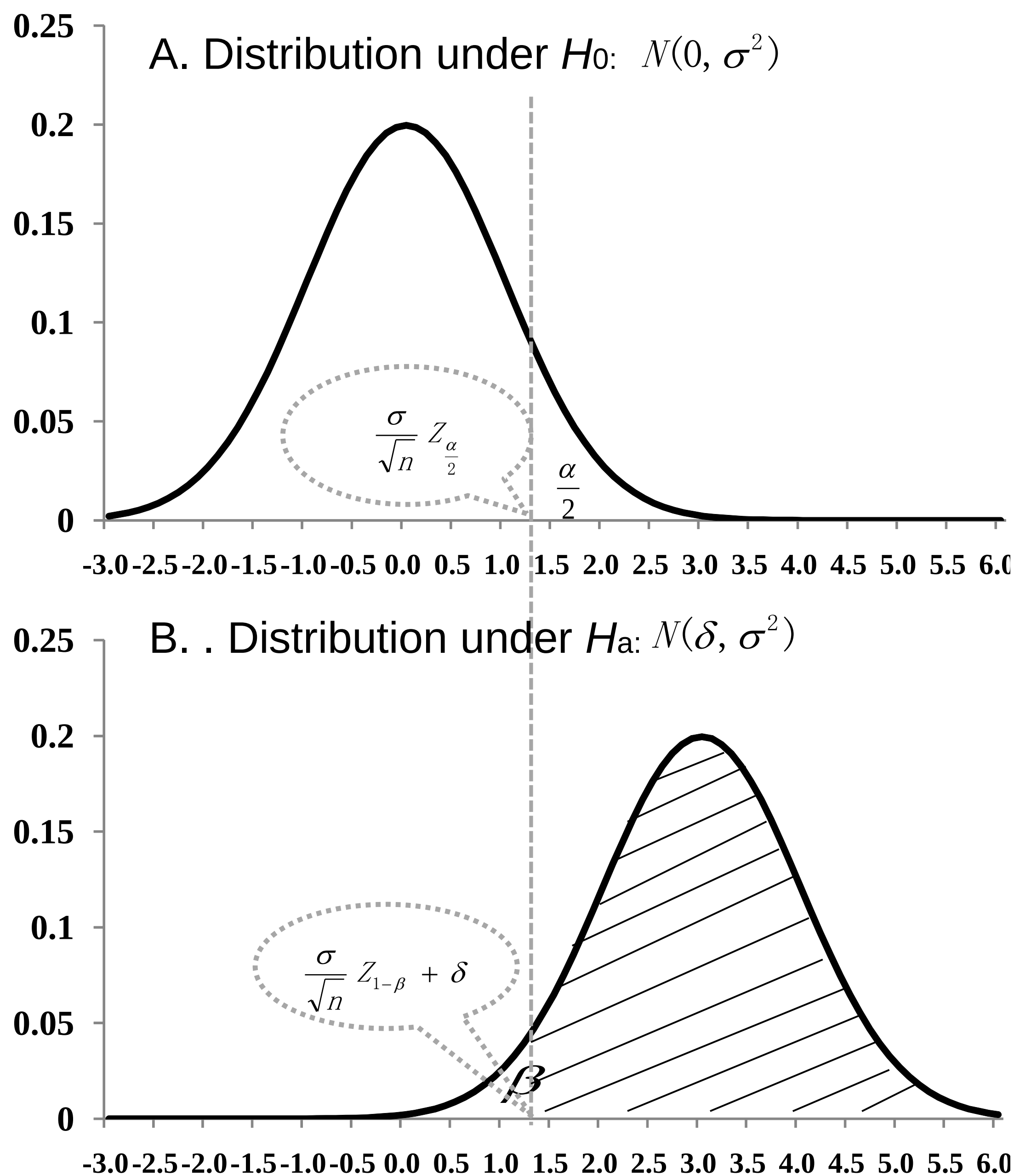
$$\overline{X}_1 - \overline{X}_2 = \frac{\sigma}{\sqrt{r/2}} Z_{\frac{\alpha}{2}}$$

$$= \frac{\sigma}{\sqrt{r/2}} Z_{1-\beta} + \delta$$

$$r = 2\left[\frac{\left(Z_{1-\beta} - Z_{\frac{\alpha}{2}}\right)\sigma}{\mu_0 - \mu_1}\right]^2$$

$$Z_{1-\beta} = Z_\beta$$

$$r = 2(z_{\alpha/2} + z_\beta)\left(\frac{\sigma}{\delta}\right)^2$$



The required replication number for each treatment group

- The replication number can be estimated with knowledge of the percent coefficient of variation, %CV. The %CV is substituted for σ where %CV =100(σ/μ). The difference δ must be expressed as a percentage of the overall expected mean of the experiment, %δ=100(δ/ μ).

$$r \geq 2(z_{\alpha/2} + z_{\beta})^2 \left(\frac{\%CV}{\%\delta} \right)^2$$

The required replications

		α=0.05		**α=0.01**	
		%δ		%δ	
%CV	1-β	10	20	10	20
5	0.80	4	1	6	2
	0.95	7	2	9	3
10	0.80	16	4	24	6
	0.95	26	7	36	9

Let's find Type I and II errors together

- Binomial distribution
- H_0: p=0.5; H_a: p≠0.5; n=6, X~B(n=6, p)
- Reject H_0 when X=0, or 6
- Accept H_0 when X=1, ..., 5
- α =P(X=0|p=0.5) + P(X=6|p=0.5) =0.0156+0.0156=0.0312
- Given p=0.75, find the Type II error β, and the statistical power
  - β= P(1<=X<=5|p=0.75)=0.8218; Power=0.1782
- Given p=0.90, find the Type II error β, and the statistical power
  - β= P(1<=X<=5|p=0.90)=0.4686; Power=0.5314

	X~Binomial (n, p), n=6		
	{0, 5} is the rejection region, i.e. EstP=0.0 or 1.0		
	{1,2,3,4} is the acceptance region		
X	H0: p=0.5	p=0.75	p=0.9
0	0.015625	0.000244	0.000001
1	0.093750	0.004395	0.000054
2	0.234375	0.032959	0.001215
3	0.312500	0.131836	0.014580
4	0.234375	0.296631	0.098415
5	0.093750	0.355957	0.354294
6	0.015625	0.177979	0.531441
	Type I error	Power	Power
	0.031250	0.178223	0.531442
		Type II error	Type II error
		0.821777	0.468558

Let's find Type I and II errors together

- Binomial distribution
- H_0: p=0.5; H_a: $p \neq 0.5$; n=30, X~B(30, p)
- Reject H_0 when X<=9, or >=21
- Accept H_0 when X=10, ..., 20
- Find α
- Given p=0.75, find the Type II error β, and the statistical power
- Given p=0.90, find the Type II error β, and the statistical power

X	H0: p=0.5	p=0.75	p=0.9
0	0.000000	0.000000	0.000000
1	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000
3	0.000004	0.000000	0.000000
4	0.000026	0.000000	0.000000
5	0.000133	0.000000	0.000000
6	0.000553	0.000000	0.000000
7	0.001896	0.000000	0.000000
8	0.005451	0.000000	0.000000
9	0.013325	0.000000	0.000000
10	0.027982	0.000002	0.000000
[illegible]	[illegible]	[illegible]	0.000000
[illegible]	[illegible]	[illegible]	0.000000
[illegible]	[illegible]	[illegible]	0.000000
[illegible]	[illegible]	[illegible]	0.000000
15	0.144464	0.001931	[illegible]
16	0.135435	0.005430	[illegible]
17	0.111535	0.013414	[illegible]
18	0.080553	0.029065	[illegible]
19	0.050876	0.055070	[illegible]
20	0.027982	0.090865	0.000365
21	0.013325	0.129807	0.001565
22	0.005451	0.159309	0.005764
23	0.001896	0.166236	0.018043
24	0.000553	0.145456	0.047363
25	0.000133	0.104728	0.102305
26	0.000026	0.060420	0.177066
27	0.000004	0.026853	0.236088
28	0.000000	0.008631	0.227656
29	0.000000	0.001786	0.141304
30	0.000000	0.000179	0.042391

Type I error	Power	Power
0.042774	0.803407	0.999546
	Type II error	Type II error
	0.196593	0.000454