# Statistics Ch 10 Hypothesis Test for 2 samples

## Ch 10.3 Hypothesis Test for 2 proportions.

Main objectives in chapter 10
- Compare the same population parameter (p or μ) from two different populations by using sample from each of the two populations.
- Use the samples to tell if the population parameters are significantly different, are the same, or one is higher (or lower) than the other.
 - Use the samples to estimate the difference in the two population parameters if there is significant difference.

Terms:
Significant different – means the parameters are not the same statistically, taking into considering of sampling error.

Independent vs dependent samples
Samples are independent if values are not naturally paired.

Inference about proportions from two populations.
1) Test a claim about p1 and p2 where p1 and p2 are population proportion (of the same category) from population one and two.
2) Estimate the confidence interval of the difference of p1 – p2.

Notations:
$\widehat{p_1} = \frac{x_1}{n_1}$ , $\widehat{p_2} = \frac{x_2}{n_2}$ . These are sample proportions from the two populations. (x1= n1·$\widehat{p_1}$, x2= n2· $\widehat{p_2}$)
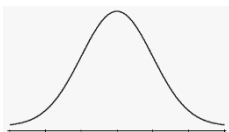Pooled sample: (variances are pooled)
Assume p1 and p2 are the same, we can combine the two samples to create a pooled  sample $\bar{p} = \frac{x_1+x_2}{n_1+n_2}$;

   $\bar{q} = 1 - \bar{p}$
The difference of two sample proportions $\widehat{p_1} - \widehat{p_2}$ will be normally distributed with

 mean = p1 – p2 , sd $= \sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}$. where

x1, x2, n1-x1, n2-x2 are at least 5.



Steps for Testing claim of 2 population proportions:
1) Define populations, record n1, x1, n2, x2, Requirement is x1, n1-x1, x2, n2-x2 are at least 5.
2) Record the claim in symbolic form:
    p1 ( =, < , > or ≠) p2.
     Write H0 (p1 = p2) and Ha (p1 <, > or ≠ p2)  in symbolic form.
     Note: p1 must be on the left, p2 on the right.


3) Identify significant level, type of test (left-tail, right-tail or two-tail test) and the sampling distribution. For proportion, distribution is z-normal.


4) Use **Statdisk/Hypothesis/ proportion Two samples** to find test statistic and p-value.
   - Select Ha: p1 ≠p2 orp1 >p2 or p1<p2.
   - Enter significant level: α
   - Enter x1, n1, x2, n2. Evaluate.
Output:
   - test statistic, z  – describe how many SD the sample  difference $\widehat{p_1} - \widehat{p_2}$ is from the H0 assumption of p1 = p2, (p1 – p2 = 0) in a normal distribution.
   - P-value: probability of getting the sample difference $\widehat{p_1} - \widehat{p_2}$ or worse if p1 = p2 is true.
   - Confidence interval at appropriate C-level lower bound < p1 – p2 < upper bound

5) If P-value ≤ α, reject H0, conclude there is significant difference between p1 and p2.
   If p-value > α, fail to reject H0, conclude there is no significant difference between p1 and p2. The difference $\widehat{p_1} - \widehat{p_2}$ are due to sample variation.

6) Write conclusion about the claim:
    There is (sufficient / not sufficient) evidence to (reject / support) the claim.
   Use "sufficient evidence " if sample is significant.
    Use "reject" if claim is H0, use "support" if claim is Ha.
    Or use Claim conclusion flowchart.


7) Make conclusion using Confidence Interval:
a) Confidence interval is included in an hypothesis test output. OR
b) Use Statdisk/Confidence Interval/Proportion 2 sample to find confidence interval.
Input appropriate C-level:
    two-tail test: C-level = 1 – α.
    Left-tail or right-tail test:  C-level = 1 - 2α

input n1,x1, n2, x2.

Output: lower limit < p1 – p2 < upper limit.

Make conclusion below:

i) If the interval contains zero or both positive and
negative values, conclude p1 and p2 has no
significant difference. Sample is not significant.
(L limit is negative U limit is positive.)

ii) If the interval is all positive, conclude p1 > p2
(L. Limit and U. Limits are both positive.)

iii) If the interval is all negative, conclude p1 < p2
(L. limit and U. Limits are both negative.)

Note: appropriate C-level:

two-tail test: C-level = 1 – α.

Left-tail or right-tail test:  C-level = 1 - 2α

Summary:

| p1 – p2 > 0 implies p1 > p2 |
| p1 – p2 < 0 implies p1 < p2 |
| p1 – p2 = 0 implies p1 = p2 |

Note: Conclusion from hypothesis test is more
accurate than conclusion from Confidence Interval for
Proportion test. Check p-value to confirm if sample
difference is significant.

p-value ≤ α there is significant difference.

p-value > α , there is no significant difference.

Ex1: Given the table below about number of cars with
Rare and Front License plates at two states.

|  | Connecticut | New York |
|---|---|---|
| Cars with rare license plates only | 239 | 9 |
| cars with both front and rare license plates | 1810 | 541 |
| Total | 2049 | 550 |

The sample proportion for Connecticut is 239/2049
(11.7%) The sample proportion for New York = 9/550
(1.6%).  The samples are different, but are they
statistically different?

a) Test with a significant level of 0.05 that the
proportions of car with only rare license plate in
Connecticut and New York are the same.

1) Define population 1 = Connecticut, , n1= 2049, x1=
239, population 2 =  New York,  n2 = 550, x2= 9,.

2) Claim: p1 = p2,  H0: p1 = p2 ,  Ha: p1 ≠ p2

3) Significant level = 0.05, Ha use ≠, so two-tail test,
distribution is z.

4) Use Statdisk/Hypothesis/proportion 2 samples,
Select Ha: p1 not = p2.  significance = 0.05,
input n1=2049, x1=239,n2=550,x2=9,. Evaluate.

Test statistic z = 7.11, p-value = 0.0000

5) Since p-value (0.0000) < α(0.05), reject H0

6) Since H0 is the claim and we rejected H0,
use Box1 in flow chart.

Conclude there is sufficient evidence to reject the
claim that the proportion of cars with rare only
License plate are the same with that of New York.

b) The 95% confidence Interval of the difference is
0.083 < p1 – p2 < 0.118

Conclude the difference is between 8.3% to 11.8% at a
confidence level of 95%.

Since the interval contains all positive values, the two
proportions are significantly different and not the
same as stated in the claim.

Note: when x is not given: x = p n
when n is not given: n = successs + failure.

Ex2: An experiment is set up to test the effectiveness
of Aspirin in preventing heart disease. 11,037 adults
were treated with aspirin and another 11,034 adults
were given placebo.  Among the subjects in the
treatment group, 1.26% experienced heart attacks.
Among the subjects given placebos, 2.17%
experienced heart attacks.

a) Use a 0.05 significant level to test the claim that
aspirin is effective in lowering heart attacks.

i) Def population 1 – treatment group,
n1= 11037, x1=0.0126(11037) =139 .
population 2 – placebo group,
n2 = 11034, x2=0.0217(11034) = 239.

ii) Claim p1< p2,   H0: p1 = p2,   Ha: p1< p2

iii) α= 0.05, Ha is "<", so left-tail test, use z-distribution

iv) Use Statdisk/hypothesis/Proportion 2 samples,
select ha:   p1< p2, significance = 0.05,
enter n1, x1, n2, x2, evaluate.

Output: test statistic z = -5.19, p-value = 0.0000
-0.012 < p1 – p2 < -0.006

v) p-value (0.0000) < 0.05 Reject H0, there is
significant difference between p1 and p2.

vi) Since claim is H1, we use support statement,
Since we reject H0, we use "sufficient evidence".
Flow chart will be at box 3.
There is sufficient evidence that to support the
claim that aspirin is effective in lowering heart attacks.

b) The 90% confidence interval is calculated because
the test is a left-tail test. C-level = 1 – 2(0.05) =0.90
-0.012 < p1 – p2 < -0.006

Since confidence interval does not contain zero,
conclude there is significant difference.

Since the whole interval is negative, conclude p1 < p2.

implying aspirin can lower heart attack.
c) Based on the result, would aspirin be recommended for adults to avoid heart attacks?
  Yes, the result conclude that aspirin is effective.
d) Would the result be useful if the treatment group are selected from males only and the placebo group are selected from females only.
The result may not apply to general population because the two groups are not only different on the treatment. The result may be due to other confounding factors between male and female.

Ex3. In a randomly picked year from 1985 to present, there were 1840 Hispanic students at Cabrillo College out of a total of 12328 students. At Lake Tahoe College, there were 321 Hispanic students out of a total of 2441 students.  Test the claim at a significant level of 1% that the precent of Hispanic student is higher in Cabrillo College.

Ans:
1) Define population 1 – Cabrillo College n1 = 12328, x1 = 1840, population 2 – Lake Tahoe College, n2 = 2441, x2 = 321.

Ex4. In a random sample of 100 forests in the United States, 56 were coniferous or contained conifers. In a random sample of 80 forests in Mexico, 40 were coniferous. Is the proportion of conifers in the United States statistically more than the proportion of conifers in Mexico? At a significant level of 1%, construct an appropriate  confidence interval to test the claim and estimate the difference.
Ans:
Define population1 -United States forests, n1 = 100, x1= 56.
Population 2 – Mexico's forests, n2 = 80, x2 = 40
Claim: p1 > p2.  H0: p1 = p2, Ha: p1> p2  (right tail test)
We can do a confidence interval to determine if p1 > p2.
Appropriate C-level = 1 – 2 α  (right-tail test) = 0.98

statdisk/Confidence Interval/proportion 2 samples/
input  Clevel = 0.98  (1 – 2(0.01))
Input n1 = 100, x1 = 56, x2= 80, n2 = 40, evaluate
Output: -0.114 < p1 – p2 < 0.234

Since 0 is in the interval, conclude there is no statistical difference between p1 and p2. p1 = p2, Sample is not significant.  ( use "not sufficient evidence" statement). Ha is claim. (use "support")
Conclusion: There is not sufficient evidence to support the claim that the proportion of conifers in the United States statistically more than the proportion of conifers in Mexico. The difference in percentage can be between -11.4% to 23.4% at 98% confidence.


## Ch 10.1 and 10.4 Hypothesis Test for 2 population means
### Independent and Dependent samples
**Independent samples**: - sample values from one population are not related to or naturally paired or matched with the sample values from the other. Sample size can be same or different.
Summary statistics or data can be given.

**Dependent samples:** - sample values are some how matched, where the matching is based on some inherent relationship. (This can be measurement from the same subject before/after, or each pair consists of matched pairs such as husband/wife, twin of siblings.)
Note: "match paired" does not mean cause/effect.
Sample size must be the same. Paired sample data will be given.

Note: For experiment to compare two populations, dependent samples results are more favorable.

Ex1. sample 1: heights of 14 men.
     sample 2:  heights of 16 women.
     The above samples are independent.

Ex2: sample 1: Heights of husband of a couple.
      sample 2: Heights of the wife of each couple.
      Since each paired value are from the same
       couple. The samples are dependent.

## Use independent samples to compare population means

To compare population mean(μ1 and μ2) from two populations, sample means ($\overline{x_1}$ and $\overline{x_2}$) are collected. If $\overline{x_1}$ and $\overline{x_2}$ are normally distributed, then the difference $\overline{x_1} - \overline{x_2}$ will be also be normally distributed.

The standard deviation of $\overline{x_1} - \overline{x_2}$ will have a value that is square root of the sum of variances.

The most common case is both σ1 and σ2 are <u>unknown and unequal</u>.

Hence variances **are not pooled,** and sample variances are used to standardize the distribution of $\overline{x_1} - \overline{x_2}$ with $\quad t = \dfrac{(\overline{x_1}-\overline{x_2})-(\mu 1-\mu 2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}$

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



Note: by default, when σ1 and σ2 are not given, we assume σ1 and σ2 are unknown and unequal, resulting in not pooling the variances.

## Use paired difference (d) to compare population means.

When samples are dependent, we define a new variable $d$ = x1 – x2 for each paired data.

The distribution of $\bar{d}$ will be normal if X1 and X2 are normal or sample size n1 = n2 = n > 30. The sample mean of d is $\bar{d}$ with standard deviation $s_d$.

The distribution of $\bar{d}$ will have a standard deviation $s_d/\sqrt{n}$. Hence t distribution is used to describe the standard error with $t = \dfrac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$

Note: We are not comparing $\overline{x_1} - \overline{x_2}$ but analyzing a new variable d = (x1 – x2) from each paired sample.

Note: statdisk does not give a new column of d in the output but provide a mean of d and sd of d.

## Inference about means from two populations.

a) Test a claim about μ1 and μ2 where μ1 and μ2 are population mean (of the same type of measurement) from population one and two.

b) Estimate the confidence interval of the difference of μ1 – μ2 or mean differences ($\mu_d$).

Note: The procedure works only if n1, n2 > 30 or X1 , X2 are normal. Assume σ1 and σ2 are unknown and unequal so no pooling of variances is used.

Steps:

1) Determine if the two samples are independent or dependent.

If samples are independent,
   Define population 1, record n1, $\overline{x_1}$ , s1.
   Define population 2, recordn2, $\overline{x_2}$ , s2
   If data are given, enter data to columns in Statdisk.

If samples are dependent,
   Define the *difference* d = (x1 – x2).
   Enter data to 2 columns.

2) Set up Claim, H0 and H1 according to claim statement.

For independent samples:
   Claim: μ1 (=, < , > or ≠) μ2,   H0: μ1 = μ2,
                                   Ha: μ1 (<, > or ≠) μ2

For dependent samples:
   Claim: μd (=, < , > or ≠) 0, H0: μd = 0,
                                Ha: μd (<, > or  ≠) 0

3) Identify significant level and type of test. distribution is t and not pooling variances.

4) For independent sample – use
   statdisk/Analysis/Hypothesis/ Mean 2 independent samples. Select alternative Hypothesis and significant level.
    If statistic summaries are given, enter n1, $\overline{x_1}$ , s1 , and n2, $\overline{x_2}$ , s2. If data are given, click "use data tab",  select columns containing sample 1 and 2.
    Use the default "unequal variances, no pooled".
   Evaluate.
   Output: test statistic (t)  and p-value (p)
           confidence interval of μ1 – μ2  at
   appropriate level of 1-α or 1-2α

 For dependent sample –
   statdisk/Analysis/Hypothesis/ Mean Matched pairs.
   Select alternative Hypothesis and significant level.
   Select the column where matched-pair samples are inputted.
   Evaluate.
   Output: test statistic (t ) and p-value (p)
           confidence interval of $\mu_d$

5) If P-value ≤ α, reject H0, conclude there is
   significant difference between µ1 and µ2.
   If p-value > α, fail to reject H0, conclude there is no
   significant difference between µ1 and µ2.
6) Write conclusion about the claim. Use table or
flowchart.
7) If there is significant difference, use the confidence
interval for the difference µ1 - µ2 or $\mu_d$
where C_level is 1 – α (2-tail test) or 1 - 2α (1-tail test)

8 Optional)
Use Confidence Interval With any C-level to test a
claim or estimate mean differences µ1 – µ2 or mean
differences ($\mu_d$).
 For independent samples: Use Statdisk/Analysis /
Confidence intervals/Mean two independent samples/
 For dependent samples: use Statdisk/Analysis/
Confidence intervals/Mean Match pairs
 Enter all inputs or data columns:
 Output
 Confidence interval:   L. limit < µ1 - µ2 < U. limit, or
                        L. limit < $\mu_d$ < U. limit
   p-value:
9) Make conclusion using C-level:
  i)  If the interval contains zero, conclude µ1 and µ2
      have no significant difference. (L limit is negative
      U limit is positive.)
  ii) If the interval is all positive, conclude µ1 > µ2.
     (L. Limit and U. Limits are both positive.)
  iii) If the interval is all negative, conclude µ1 < µ2.
     (L. limit and U. Limits are both negative.)

Note: When H0 is rejected, there is sufficient evidence
to conclude that there is significant difference
between µ1 and µ2.

Note: Conclusion from hypothesis test is exactly the
same as the conclusion from Confidence Interval for
testing of the mean because the standard errors are
the same.

Ex1. A study claims that mean enrollment at two-year
college is lower than at four-year colleges  in the
United States. Two samples are collected, from the 35
two-year colleges surveyed, the mean enrollment was
5,068 with a standard deviation of 4,777. Of the 35
four-year colleges surveyed, the mean enrollment was
5,466 with a standard deviation of 8,191.  Test at sign.
level of 0.05.

The samples are independent.
i) Define population 1 = two-year colleges, n1 = 35, $\bar{x}_1$
= 5068, s1=4777, σ1 = not given,
Define population 2 = four-year colleges, n2=35, ,
$\bar{x}_2 = 5466$, s2 = 8191, σ2 = not given.
ii) Claim: µ1 < µ2; H0 : µ1 = µ2;  Ha= µ1 < µ2
iii) Sign level = 0.05, left-tail test, use t-distribution.
iv) Use Statdisk/Analysis/Mean 2 independent
samples
Select Alt. hypothesis : Pop Mean 1 < Pop Mean 2
Enter 0.05 to significance,
Enter n1 = 35, $\bar{x}_1$ = 5068, s1=4777, σ1 = not given,
n2=35,  $\bar{x}_2 = 5466$, s2 = 8191, σ2 = not given.
Use "Unequal variance, no Pool". Evaluate.
Output  test stat t = -0.248, p-value = 0.4024
90% confidence interval = -3079.7 < µ1 – µ2 < 2283.7
v) since p-value (0.4024) > 0.05, fail to reject H0
vi) Since sample is not significant, claim is Ha.
 There is not suff. evidence to support the claim the
mean enrollment in two-year college is less than those
at four-year college.

Ex2. Use confidence interval at 95% confidence to
estimate the difference in mean life of Duracell and
Eveready battery. Sample results are as follow:
Duracell: n1 = 8, $\bar{x}_1$=41 hr, s1= 18 hr
Eveready: n2 = 10, $\bar{x}_2$= 45 hr, s2 = 20 hr
i) Samples are independent.
Define population 1 – Duracell
Define population 2 – Eveready
ii) Use Statdisk/Analysis/Confidence Intervals/
Mean 2 independent samples
Enter C-level = 0.95,
Enter sample 1 : n1 = 8, $\bar{x}_1$=41, s1= 18 hr
Enter sample 2: n2 = 10, $\bar{x}_2$= 45 hr, s2 = 20 hr
Use 'Unequal variances, no Pool". Evaluate
Output: p-value = 0.6618,  95% confidence interval
-23.05 <  µ1-µ2 < 15.05 hr
iii) Since zero contains in the interval, there is no suff.
evidence conclude there is significant difference.
Conclude that the mean life for the two batteries has
no significant difference.

Ex3: A study of seat belt use involved children who
were hospitalized after car accidents. For 123 children
who were wearing seat belts, the number of days in
ICU has a mean of 0.83 days and a standard deviation
of 1.77 days. For the sample of 290 children who were
not wearing seat belts, the number of days in ICU has
a mean of 1.39 days and a standard deviation of 3.06

days. Test the claim at α = 0.05 by hypothesis and Confidence interval that seat belt use for children is effective in lowering the degree of injuries.

i) The two samples are independent. Values are summarized.
Define population 1: children who wear seat belts
n1 = 123, $\overline{x_1} = 0.83, s1 = 1.77$
Define population 2: children who do not wear seat belts. $n2 = 290, \overline{x_2} = 1.39, s2 = 3.06$
ii) Claim: μ1 < μ2, H0: μ1 = μ2, Ha: μ1 < μ2
iii) Use α = 0.05, left-tail test, use t distribution.
iv) Use Statdisk/Analysis/Hypothesis Testing/Mean 2 independent samples/
Select alternative hypothesis Mean 1 < Mean 2
Enter significance = 0.05,
Enter $n1, \overline{x_1}, s1;\ \ n2, \overline{x_2}, s2$
Unequal variances, no pool.
Output: t = -2.33 p-value = 0.0102.
90% confidence interval is -0.9563< μ1 – μ2 < -0.1637
v) P-value (0.0102) < 0.05 Reject H0. There is significant difference between mean ICU times.
vi) There is sufficient evidence to support the claim that seat belt use for children is effective in lowering the degree of injuries.
 Confidence interval is -0.96 < μ1 - μ2 < -0.16 days
 Since interval contains negative number, conclude there is sufficient evidence to show that μ1 < μ2, so conclude degree of injuries for children wearing seat belts are lower.
Conclusion from hypothesis test and confidence interval are the same.


Ex4: Listed below are course evaluation scores for courses taught by female professors and male professors.
a) Use α = 0.05 to test the claim that there is a difference in mean evaluation scores of course taught by female professors and male professors.

| Female | 4.3 | 4.3 | 4.4 | 4.0 | 3.4 | 4.7 | 2.9 | 4.0 | 4.3 | 3.4 | 3.4 | 3.3 | | |
| Male | 4.5 | 3.7 | 4.2 | 3.9 | 3.1 | 4.0 | 3.8 | 3.4 | 4.5 | 3.8 | 4.3 | 4.4 | 4.1 | 4.2 | 4.0 |

i) The samples are independent because pairs of value are not matched or related.
Define population 1 – female professor scores
        population 2 – male professor scores
        Enter data to Statdisk.
ii) Claim: μ1 ≠ μ2, H0: μ1 = μ2, Ha: μ1 ≠ μ2.
iii) Use α = 0.05, Two-tail test, use t distribution.
iv) Use Statdisk/Hypothesis Testing/Mean 2 independent samples / Click "use data" tab

Select "Population mean 1 not = pop mean 2" as alternative hypothesis, select columns for data, sample 1 is female professors' scores, sample 2 is male professors' scores.
Use Unequal variances, No Pooled. Evaluate.
Output: t = -0.66, p-value(p) = 0.5172, df=19.06
Confidence interval (95%) -0.53 < μ1 - μ2 < 0.27.
v) p-value (0.5172 ) > 0.05, fail to reject H0. Conclude no significant difference.
vi) Since the claim is in H1: Conclusion about the claim: There is not sufficient evidence to support the claim that there is a difference in evaluation scores between course taught by female and male professor. Conclude plausibly no difference in scores.

vii) The confidence interval -0.53 < μ1 - μ2 < 0.27, contains zero, conclude there could be no significant difference between the two means. (Not sufficient evidence to support the claim of difference.

Ex5: Assume Freshmen year for college student is from September to April of the following year. Use the sample data given below with a 0.05 significance level to test the claim that there is no difference(ie the same) in mean weight change from September to April for students in their Freshman year. Samples are SRS and weights are normally distributed.

| student | A | B | C | D | E |
| --- | --- | --- | --- | --- | --- |
| September | 67 | 53 | 64 | 71 | 70 |
| April | 66 | 52 | 68 | 69 | 71 |

i) The samples are dependent because pairs of value are from the same student.
 define d = September weight – April weight
  Enter data to Statdisk.
ii) Claim: μd = 0, H0: μd = 0, H1: μd ≠0 (claim is H0)
iii) α= 0.05, two-tail test, use t -distribution.
iv) Use Statdisk/Analysis/Hypothesis Testings/Mean match pairs.
Select alt. hypothesis: mean of difference not = 0
Enter significance = 0.05,
Select sample 1 for September weights,
Select Sample 2 April's weights. Evaluate
Output t = -0.19, p-value(p) = 0.8605
95% confidence interval is -3.16 < $\mu_d$ < 2.76 kg
v) P-value(0.8605) > 0.05, fail to reject H0.
vi) There is not sufficient evidence to reject the claim that there is no mean weight change from September to April for students in their freshmen year.
Conclude the claim is plausibly true.
vii) Since Confidence interval is  -3.16 kg < $\mu_d$ < 2.76 kg. Interval contains zero. We are 95% confidence that the mean change in weight from September to April

can be between an increase of 3.16 kg to a decrease of 2.76 kg.

Since the difference contains zero, conclude there is no mean difference at September and April's weights.

Note: Conclusion from Hypothesis test and Confidence Interval are the same.

Ex 6: A study was conducted to investigate the effectiveness of hypnotism in reducing pain. Randomly selected subjects are given hypnotic treatment and their pain level before and after measured and recorded below.  Higher level corresponds to greater level of pain.
a) Conduct a hypothesis test to test the claim that hypnotism treatment can reduce pain. Use $\alpha = 0.05$.
b) Construct an appropriate confidence interval for the mean difference of pain before and after hypnotism treatment.

| Patients | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| before | 6.6 | 6.5 | 9 | 10.3 | 11.3 | 8.1 | 6.3 | 11.6 |
| After | 6.8 | 2.4 | 7.4 | 8.5 | 8.1 | 6.1 | 3.4 | 2 |

i) The samples are dependent because each pair of values are from the same patient.
define d = before – after pain level.
 ( d > 0 if hypnotic treatment reduce pain)
Enter before and after data to Statdisk.
ii) Claim: $\mu_d > 0$,  H0: $\mu_d = 0$,  H1: $\mu_d > 0$
iii) $\alpha = 0.05$, two-tail test, use t -distribution
iv)  Use Statdisk/Analysis/Hypothesis Testing/Mean match pair.
Select alt. hypothesis : Mean of differences > 0
Enter significance 0.05.
Select before data for sample 1,
Select after data for sample 2. Evaluate
$\bar{d} = 3.125$, sd= 2.91
 Output t = 3.04 and p-value = 0.0095
 90% confidence interval is 1.1748 <  $\mu_d$ <  5.0752
v) P-value (0.0095) < 0.05 reject H0.
vi) There is sufficient evidence to support the claim that hypnotism treatment can reduce pain level.
vii ) 90% Confidence interval is 1.17 <  $\mu_d$ < 5.08
Since the interval does not contain zero, there is a significant difference. The interval contains only positive values. So pain level before > pain level after. Conclude hypnotism treatment can reduce pain.

Note: Conclusion from Hypothesis test and Confidence interval are the same.