Solution for Session 4
Hypothesis Testing

21/11/2023

1 Inference about a proportion

Out of 80 women in a random sample of women in Manchester, 13 were asthmatic; this could be used to calculate a 95% confidence interval for the proportion of women in Manchester with asthma. This confidence interval could be compared to the suggested prevalence of 20% in Northern England. An alternative approach would be to test the hypothesis that the true proportion, π, is 0.20.

1.1 What is the expected proportion of women with asthma under the null hypothesis ?
0.2

1.2 What is the observed proportion of women with asthma ? *13/80 = 0.1625*

1.3 What is the standard error of the expected proportion (remember from last week that the standard error of a proportion p is given by

$$\sqrt{\frac{p(1-p)}{n}}$$

$$\sqrt{\frac{0.2 \times 0.8}{80}} = 0.0447$$

1.4 The appropriate test statistic, T, is given by the formula:

$$\frac{\text{observed proportion - expected proportion}}{\text{standard error of proportion}}$$

Calculate T.

$$T = \frac{0.1625 - 0.2}{0.0447} = -0.839$$

1.5 T should be compared to a t-distribution with how many degrees of freedom ? *79*

1.6 From tables for the appropriate t-distribution, the corresponding p-value is 0.4. Is is reasonable to suppose that these women are a random sample from a population in which the prevalence of asthma is 20% ? *Yes*

2 More inference about a proportion

In the sample heights and weights we have looked at, there were 412 individuals of whom 234 were women. We wish to test that there are equal numbers of men and women in our population.

2.1 What is the null hypothesis proportion of women ? *0.5*

2.2 What is the observed proportion of women ?

$$\frac{234}{412} = 0.568$$

2.3 What is the null hypothesis standard error for the proportion of women ?

$$\sqrt{\frac{0.5 \times 0.5}{412}} = 0.0246$$

2.4 What is an appropriate statistic for testing the null hypothesis ?

$$T = \frac{0.568 - 0.5}{0.0246} = 2.76$$

3 Inference about a mean

Load `htwt.dta` into stata with the commands (each command needs to be entered on a separate line).

```
global datadir http://personalpages.manchester.ac.uk/staff/mark.lunt/stats
use $datadir/2_summarizing_data/data/htwt.dta
```

We wish to test whether the mean height is the same in men and women.

3.1 What is the null hypothesis difference in height between men and women ?
0

3.2 Use the command `ttest nurseht, by(sex)` to test whether the mean height differs between men and women.

3.3 What is the mean height in men ? *173.0cm*

3.4 What is the mean height in women ? *159.8cm*

3.5 What is the mean difference in height between men and women, with its 95% confidence interval ?
-13.2cm, 95z% CI -14.5cm, -11.9cm,

3.6 Which of the three hypothesis tests is the appropriate one in this instance ?
`Ha: diff != 0`*: two sided test that there is a difference*

3.7 What is the p-value from the t-test ? *less than 0.0001*

3.8 What would you conclude ?
Men are very significantly taller than women

4 Two-sample t-test

Compare BMI (based on the measured values, i.e. `bmi`) between men and women in `htwt.dta`, using the command `ttest bmi, by(sex)`.

4.1 Is there a difference in BMI between men and women ? *Yes, but it is small and not statistically significant*

4.2 What is the mean difference in BMI between men and women and its 95% confidence interval.
-0.5kg/m², 95%CI -1.4kg/m², 0.4kg/m²

4.3 Is there a difference in the standard deviation of BMI between men and women ? (This can be tested with the command `sdtest bmi, by(sex)`
Yes: the standard deviation is significantly greater in women than it is in men

4.4 If there is, repeat the t-test you performed above, using the `unequal` option. Are your conclusions any different ?
No: the standard error changes only very slightly, the difference remains non-significant

5 One sample t-test

Load the `bpwide` dataset into stata with the command `sysuse bpwide`. This consists of fiction blood pressure data, taken before and after an intervention. We wish to determine whether the intervention had affected the blood pressure.

5.1 Use the `summarize` command to calculate the mean blood pressure before and after the intervention. Has the blood pressure increased or decreased ?
Decreased: mean is 156 before, 151 after.

5.2 Generate a variable containing the change in blood pressure using the command `gen bp_diff = bp_after - bp_before`

5.3 Use the command `ttest bp_diff = 0` to test whether the change in blood pressure is statistically significant. Is it ?
Yes: $p = 0.0011$

5.4 Give a 95% confidence interval for the change in blood pressure.
-8.1, -2.1

6 Power Calculations

The following questions can all be answered using the `sampsi` command.

6.1 How many subjects would need to be recruited to have 90% power to detect a difference between unexposed and exposed subjects if the prevalence of the condition is 25% in the unexposed and 40% in the exposed, assuming equal numbers of exposed and unexposed subjects ?
432: 216 in each group

6.2 If the exposure was rare, so it was decided to recruit twice as many unexposed subjects as exposed subjects, how many subjects would need to be recruited ?
482: 161 exposed and 321 unexposed

6.3 Suppose it were only possible to recruit 100 subjects in each group. What power would the study then have ? *56%*

6.4 Suppose that we expect a variable to have a mean of 15 and an SD of 5 in group 1, and a mean of 17 and an SD of 6 in group 2. How large would two equal sized groups need to be to have 90% power to detect a difference between the groups ?
161 subjects in each group

6.5 If we wanted 95% power, how large would the groups have to be ?
199 in each group

6.6 Suppose we could only recruit 100 subjects in group 1. How many subjects would we have to recruit from group 2 to have 90% power ?
276

Hint: the last question can only be answered by trying different numbers for the size of group 2 and seeing what power is achieved. Sensible choice of numbers will give a result fairly quickly. The `PageUp` key is your friend.

Stata Log File

```
. global datadir http://personalpages.manchester.ac.uk/staff/mark.lunt/stats

. use $datadir/2_summarizing_data/data/htwt.dta

. ttest nurseht, by(sex)
Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
female	227	159.774	.4247034	6.398803	158.9371	160.6109
male	175	172.9571	.5224808	6.911771	171.9259	173.9884
combined	402	165.5129	.4642267	9.307717	164.6003	166.4256
diff		-13.18313	.6666327		-14.49368	-11.87259

```
    diff = mean(female) - mean(male)                          t = -19.7757
Ho: diff = 0                                 degrees of freedom =      400

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

```
. ttest bmi, by(sex)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
  female |     225    25.85984    .3268714    4.903072    25.21571    26.50398
    male |     175    26.35534    .3152409    4.170245    25.73315    26.97753
---------+--------------------------------------------------------------------
combined |     400    26.07662    .2298956    4.597912    25.62467    26.52858
---------+--------------------------------------------------------------------
    diff |           -.4954959    .4633426               -1.406401    .4154089
------------------------------------------------------------------------------
    diff = mean(female) - mean(male)                              t =  -1.0694
Ho: diff = 0                                     degrees of freedom =      398

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.1428         Pr(|T| > |t|) = 0.2855          Pr(T > t) = 0.8572

. sdtest bmi, by(sex)

Variance ratio test
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
  female |     225    25.85984    .3268714    4.903072    25.21571    26.50398
    male |     175    26.35534    .3152409    4.170245    25.73315    26.97753
---------+--------------------------------------------------------------------
combined |     400    26.07662    .2298956    4.597912    25.62467    26.52858
------------------------------------------------------------------------------
    ratio = sd(female) / sd(male)                                 f =   1.3823
Ho: ratio = 1                                    degrees of freedom = 224, 174

    Ha: ratio < 1               Ha: ratio != 1                 Ha: ratio > 1
  Pr(F < f) = 0.9874         2*Pr(F > f) = 0.0253          Pr(F > f) = 0.0126

. ttest bmi, by(sex) unequal

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
  female |     225    25.85984    .3268714    4.903072    25.21571    26.50398
    male |     175    26.35534    .3152409    4.170245    25.73315    26.97753
---------+--------------------------------------------------------------------
combined |     400    26.07662    .2298956    4.597912    25.62467    26.52858
---------+--------------------------------------------------------------------
    diff |           -.4954959    .4541164               -1.388285     .397293
------------------------------------------------------------------------------
    diff = mean(female) - mean(male)                              t =  -1.0911
Ho: diff = 0                     Satterthwaite´s degrees of freedom =  394.793

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.1379         Pr(|T| > |t|) = 0.2759          Pr(T > t) = 0.8621

. sysuse bpwide
(fictional blood-pressure data)

. summarize bp*

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   bp_before |       120      156.45    11.38985        138        185
    bp_after |       120    151.3583    14.17762        125        185

. gen bp_diff =  bp_after- bp_before
```

```
. ttest bp_diff == 0

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 bp_diff |     120   -5.091667    1.525736     16.7136   -8.112776   -2.070557
------------------------------------------------------------------------------
    mean = mean(bp_diff)                                          t =  -3.3372
Ho: mean = 0                                     degrees of freedom =      119

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 0.0006         Pr(|T| > |t|) = 0.0011          Pr(T > t) = 0.9994

. sampsi 0.25 0.4, power(0.9)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.9000
            p1 =   0.2500
            p2 =   0.4000
         n2/n1 =   1.00

Estimated required sample sizes:

            n1 =      216
            n2 =      216

. sampsi 0.25 0.4, power(0.9) r(0.5)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.9000
            p1 =   0.2500
            p2 =   0.4000
         n2/n1 =   0.50

Estimated required sample sizes:

            n1 =      321
            n2 =      161

. sampsi 0.25 0.4, n(100)

Estimated power for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
            p1 =   0.2500
            p2 =   0.4000
sample size n1 =      100
            n2 =      100
         n2/n1 =     1.00

Estimated power:

         power =   0.5618
```

```
. sampsi 15 17, sd1(5) sd2(6) power(.90)
Estimated sample size for two-sample comparison of means
Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
         power =   0.9000
            m1 =       15
            m2 =       17
           sd1 =        5
           sd2 =        6
         n2/n1 =     1.00
Estimated required sample sizes:
            n1 =      161
            n2 =      161

. sampsi 15 17, sd1(5) sd2(6) power(.95)
Estimated sample size for two-sample comparison of means
Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
         power =   0.9500
            m1 =       15
            m2 =       17
           sd1 =        5
           sd2 =        6
         n2/n1 =     1.00
Estimated required sample sizes:
            n1 =      199
            n2 =      199

. sampsi 15 17, sd1(5) sd2(6) n1(100) n2(200)
Estimated power for two-sample comparison of means
Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
            m1 =       15
            m2 =       17
           sd1 =        5
           sd2 =        6
sample size n1 =      100
            n2 =      200
         n2/n1 =     2.00
Estimated power:
         power =   0.8621

. sampsi 15 17, sd1(5) sd2(6) n1(100) n2(280)
Estimated power for two-sample comparison of means
Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
            m1 =       15
            m2 =       17
           sd1 =        5
           sd2 =        6
sample size n1 =      100
            n2 =      280
         n2/n1 =     2.80
Estimated power:
         power =   0.9016
```

```
. sampsi 15 17, sd1(5) sd2(6) n1(100) n2(275)
Estimated power for two-sample comparison of means
Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
            m1 =       15
            m2 =       17
           sd1 =        5
           sd2 =        6
sample size n1 =      100
            n2 =      275
         n2/n1 =     2.75
Estimated power:
         power =   0.8998

. sampsi 15 17, sd1(5) sd2(6) n1(100) n2(277)
Estimated power for two-sample comparison of means
Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
            m1 =       15
            m2 =       17
           sd1 =        5
           sd2 =        6
sample size n1 =      100
            n2 =      277
         n2/n1 =     2.77
Estimated power:
         power =   0.9005

. sampsi 15 17, sd1(5) sd2(6) n1(100) n2(276)
Estimated power for two-sample comparison of means
Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
            m1 =       15
            m2 =       17
           sd1 =        5
           sd2 =        6
sample size n1 =      100
            n2 =      276
         n2/n1 =     2.76
Estimated power:
         power =   0.9002
```