

STAT:2010/4200

Lab 6. t-test for 2 independent samples and power

1 Two sample t-test example

A random sample of 130 healthy people is taken and each person's temperature, gender, and heart rate are recorded in "normtemp.dat".

1.1 Recall we did a t-test in lab 5 for a one sample problem (proc univariate)

It is widely believed that the average body temperature for healthy humans is 98.6 F. We think that might not be true, so we decide to do a two-sided significance test at significance level $\alpha = .05$:

$$H_0 : \mu_{\text{temp}} = 98.6$$
$$H_a : \mu_{\text{temp}} \neq 98.6$$

We reformat the variable gender in the data step.

```
proc format ;
value sexfmt 1 = 'M' 2 = 'F' ;
run ;

data normtemp ;
input temp gender heart ;
format gender sexfmt. ;
datalines;
* note: copy and paste data in here ;
;
run ;
```

We performed a t test using **proc univariate**.

(Remember to check the rule of thumb before accepting/interpreting the results.)

```
proc univariate mu0 = 98.6 data = normtemp ;
var temp;
run ;
```

The UNIVARIATE Procedure
Variable: temp

```
Tests for Location: Mu0=98.6
Test          -Statistic-      -----p Value-----
Student's t   t  -5.45482      Pr > |t|  <.0001
Sign          M    -21      Pr >= |M| 0.0002
```

Signed Rank S -1963 Pr >= |S| <.0001

Conclusion: Reject H_0 and have sufficient evidence to support H_1 .

1.2 Two-sample t-test for a two-independent-sample problem (proc ttest)

We might wish to test whether the population mean body temperature is the same in men and in women (at significance level $\alpha = .05$):

$$H_0 : \mu_M - \mu_F = 0$$

$$H_A : \mu_M - \mu_F \neq 0$$

We can view the (fe)males in our study as a random sample from the (fe)male population, then the data contains two independent samples from the two populations of interest.

We can perform a two-sample t test using **proc ttest**. (Check the rule of thumb first.)

```
proc ttest data = normtemp ;
class gender ;
var temp ;
run ;
```

The SAS System

The TTEST Procedure

Variable: temp

gender	N	Mean	Std Dev	Std Err	Minimum	Maximum
M	65	98.1046	0.6988	0.0867	96.3000	99.5000
F	65	98.3938	0.7435	0.0922	96.4000	100.8
Diff (1-2)		-0.2892	0.7215	0.1266		

gender	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
M		98.1046	97.9315 98.2778	0.6988	0.5959 0.8449
F		98.3938	98.2096 98.5781	0.7435	0.6340 0.8990
Diff (1-2)	Pooled	-0.2892	-0.5396 -0.0388	0.7215	0.6429 0.8221
Diff (1-2)	Satterthwaite	-0.2892	-0.5396 -0.0388		

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	128	-2.29	0.0239
Satterthwaite	Unequal	127.51	-2.29	0.0239

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	64	64	1.13	0.6211

Conclusion: The p-value of the 2-sample t-test (the Satterthwaite test) is .0239, smaller than the significance level $\alpha = .05$. Hence the data suggests that the mean body temperature is different in the male and the female population. Specifically, the male has lower body temperatures. We are 95% confident that the mean male body temperature is between .5396 to .0388 degrees lower than the mean female body temperature.

2 Sample size determination and power calculation

2.1 Reconsider the example in lecture 15 page 19 - 24

Suppose we are interested in testing $H_0 : \mu = 120$ vs $H_1 : \mu < 120$. When $\mu = 115$ (hence H_1 is true) and $\sigma = 25$, we would like to determine the sample size needed of a one sample t -test to reach a power of 0.8 at $\alpha = 0.05$.

```
proc power;
onesamplemeans
nullmean = 120
mean     = 115
sides = 1
ntotal = .
stddev = 25
power = 0.8;
run;
```

Computed N Total	
Actual Power	N Total
0.800	156

Remarks: As we can see that there is a small difference between the SAS calculated value (156) and our hand calculated value (155). In fact, SAS calculates the sample size for one sample t -test (a much more realistic situation); however, we calculated by hand the sample size need for one sample z -test (a rare situation). When the sample size is large, the difference is very small (like here). It is relatively easy to calculate power (or determine sample size) for one sample z -test but it is very hard to do so for one sample t -test. Hence we rely on computers.

3 Power calculation

The Trial Urban District Assessment (TUDA) measures educational progress within participating large urban districts. TUDA gives a reading test scored from 0 to 500. A score of 243 is a “Basic” reading level for eight-graders. Suppose scores on the TUDA reading test for eight-graders in your district follow a Normal distribution with standard deviation

$\sigma = 55$. In 2013 the mean score for eight-graders in your district was 255. You plan to give the reading test to a random sample of 25 eighth-graders in your district this year to test whether the mean score μ for all eight-graders in your district is still above the basic level. You will therefore test

$$H_0 : \mu = 243 \quad H_A : \mu > 243$$

If you conduct your test at significance level $\alpha = 0.05$, what is the power against the alternative that $\mu = 255$?

What do your results say about whether you are at risk for either Type I or Type II error?

Solution:

(1) Hand calculation (assuming one sample z -test)

$$\begin{aligned} \text{power} &= P(\text{reject } H_0 | H_1 \text{ true}) \\ &= P\left(\frac{\bar{X} - 243}{55/\sqrt{25}} > 1.645 | H_1 \text{ true}\right) \\ &= P\left(\frac{\bar{X} - 255}{55/\sqrt{25}} > 1.645 - \frac{12}{55/\sqrt{25}}\right) \\ &= P(Z > 0.554) \\ &= 0.290 \end{aligned}$$

(2) Computer calculation (assuming one sample t -test)

```
proc power;
onesamplemeans
test = t
nullmean = 243
mean = 255
sides = 1
ntotal = 25
std = 55
power = .;
run;
```

Computed Power
Power
0.279

$P(\text{make a type I error}) = \alpha = 0.05$

$P(\text{make a type II error}) = 1 - \text{power} = 0.71 = \begin{cases} 0.710 & \text{if doing a } z\text{-test} \\ 0.721 & \text{if doing a } t\text{-test} \end{cases}$

Hence we are very likely to make a type II error.