The University of Sydney
STAT2012 Statistical Tests

Semester 2	Computer Practice Week 3	2015

Useful R commands

- The R command to perform an one sample t-test for the hypotheses: $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$ is

  `t.test(x,mu=mu0,alternative=".")`.

  The alternative can be `greater`, `less` and `two.sided`.

  This command produces the test statistic, the p-value and a 95% CI for μ (default is $\alpha = 0.05$). If $(1-\alpha)$% CI is needed, one should set `alpha=0.xx` and add "`conf=1-alpha`" to `t.test`, that is

  ```
  alpha=0.xx
  t.test(x,mu=mu0,alternative=".",conf=1-alpha)
  ```

- The R command for k such that $\Pr(t_{n-1} < k) = p$ is `qt(p,n-1)`.
- The R commands for calculating sample mean and sample standard deviation for `x` are

  ```
  smean=mean(x)
  ssd=sd(x)
  ```

Important points

- You will read data from the course website.
- You will perform suitable tests on the data.
- You will calculate rejection region and compare the sample mean to draw your decision of the test.

Practice Problems

Open the data set `survey` containing measurements of the following variables from 95 students:

sex	1=male; 2=female
age	Year
height:	Inches
credit:	Number of credit cards in possession
pulse:	Number of heartbeats in one minute
pulse.ex:	Number of heartbeats in one minute after regular exercise over a period
exercise:	Number of hours of exercise last week
smoke:	1=yes; 2=no
hand:	1=left-handed; 2=right-handed; 3=ambidextrous

Read the data `survey`. Set `pulse.sf` to contain the `pulse` among female students who smoke.

```
survey=read.csv("http://rome/u/UG/IM/STAT2012/r/survey.csv")
attach(survey)
pulse.sf=pulse[smoke==1 & sex==2]
pulse.sf
```

1. Test if the mean of the pulse among female students who smoke (`pulse.sf`) is 70 or more than 70.

   (a) State the null and alternative hypotheses.

   (b) Draw a *qq-plot* of the data and comment the normality assumption of the data.

   (c) Find the sample mean named as `smean` and sample standard deviation named as `ssd` of the pulse among female students who smoke (`pulse.sf`).

   (d) Perform a suitable test. Report the test statistic and p-value. Draw your conclusion about the null hypothesis regarding the mean of the pulse among female students who smoke based on the p-value.

   (e) State the rejection region for the sample mean when the level of significance is $\alpha = 0.05$. Draw your conclusion about the null hypothesis based on the rejection region.

   ```
   n=length(pulse.sf)
   n
   cv=qt(0.95,n-1)
   cv
   mu0=70
   rr.bound=mu0+cv*ssd/sqrt(n)
   rr.bound
   ```

   (f) Using the test result of (d), state the one-sided 95% confidence interval for the true mean. Draw your conclusion about the null hypothesis based on the confidence interval.

   (g) Do the three ways of testing the hypotheses, in (d), (e) and (f), give consistent result?