The University of Sydney
STAT2012 Statistical Tests

Semester 2	Computer Practice Week 4	2015

Useful R commands

- The R command to perform an one sample *binomial* test or *sign* test ($p_0 = 0.5$) for the hypotheses: $H_0: \ p = p_0$ against $H_1: \ p \neq p_0$ is

  `binom.test(x,n,p0,alt="two.sided",0.95)`.

  where `x` is the count of success and `n` is the sample size. The alternative can be `greater`, `less` and `two.sided`.

  This command produces the test statistic, the p-value and a 95% CI for p.

- The R command to perform an one sample *Wilcoxon signed-rank* (WSR) test for the hypotheses: $H_0: \ \mu = \mu_0$ against $H_1: \ \mu \neq \mu_0$ is

  `wilcox.test(x,alternative="two.sided",mu=mu0,exact=T,correct=F)` .

  This gives exact p-value. If there are ties or zeros, you may set `exact=F` and normal approximation will be used.

Important points

- You will read data from the course website.
- You will perform sign test and Wilcoxon signed rank test to test for the mean of a population.
- You should state the test statistic and p-value and interpret the test result for each test. This is far more important than merely generating all the test outputs.
- Only some of the R codes are provided.

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 during 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://www.maths.usyd.edu.au/u/UG/IM/STAT2012/r/survey.csv")
attach(survey)
pulse.sf=pulse[smoke==1 & sex==2]
pulse.sf
```

Note that you can use this link to read the data set at home.

1. Test if the mean of the pulse among female students who smoke is 70 or *more* than 70 using the sign test.
   (a) State the null and alternative hypotheses.
   (b) Find the number n of nonzero differences and the number x of positive differences.
   (c) Check the assumption of *symmetric* data distribution using a *boxplot* and comment.
   (d) Perform the test and draw your conclusion about H_0 based on the p-value.
   (e) The p-value of the t-test performed last week is 0.0430. Compare the power and distribution assumption of the two tests.

2. Repeat the test in Question 1 using the Wilcoxon signed rank test.
   (a) Find the signed ranks, W^+, W^- and W. Are there any ties? Should normal approximation be used?
   (b) Perform the test and report the test statistic and p-value. Draw your conclusion about H_0 based on the p-value. Remember that `exact=F` if normal approximation should be used.
   (c) Compare the result with those of the sign test in Question 1 and the t-test performed last week (p-value=0.0430) in terms of the power and the distribution assumption.
   (d) Find the mean and variance of W^+. Hence vertify the p-value of the test.

```
ew.plus = sum(r[d!=0])/2
ew.plus
varw.plus = sum((r[d!=0])^2)/4
varw.plus
z0=(w.plus-ew.plus)/sqrt(varw.plus)
z0
p.value=1-pnorm(z0)
p.value
```