

Coursera Statistical Inference Quiz 4

AChoi

Saturday, May 02, 2015

Question 1

A pharmaceutical company is interested in testing a potential blood pressure lowering medication. Their first examination considers only subjects that received the medication at baseline then two weeks later. The data are as follows (SBP in mmHg)

```
(data <- data.frame(base=c(140,138,150,148,135),
                    later=c(132,135,151,146,130)))
##   base later
## 1  140   132
## 2  138   135
## 3  150   151
## 4  148   146
## 5  135   130
```

Consider testing the hypothesis that there was a mean reduction in blood pressure? Give the P-value for the associated two sided T test.

```
t.test(data$base, data$later, paired=T, var.equal=T, alt="two.sided")$p.value
## [1] 0.08652278
```

Question 2

A sample of 9 men yielded a sample average brain volume of 1,100cc and a standard deviation of 30cc. What is the complete set of values of μ_0 that a test of $H_0: \mu = \mu_0$ would fail to reject the null hypothesis in a two sided 5% Students t-test?

```
n <- 9
mu <- 1100
sd <- 30
t <- qt(1-0.05/2, n-1)
mu + c(-1, 1) * t * sd / sqrt(n)
## [1] 1076.94 1123.06
```

Question 3

Researchers conducted a blind taste test of Coke versus Pepsi. Each of four people was asked which of two blinded drinks given in random order that they preferred. The data was such that 3 of the 4 people chose Coke. Assuming that this sample is representative, report a P-value for a test of the hypothesis that Coke is preferred to Pepsi using a one sided exact test.

```
# Probability of more than 2 out of 4 people choosing coke assuming that people
# are indifferent between coke and pepsi.
pbinom(2, size=4, prob=0.5, lower.tail=F)
## [1] 0.3125
```

Question 4

Infection rates at a hospital above 1 infection per 100 person days at risk are believed to be too high and are used as a benchmark. A hospital that had previously been above the benchmark recently had 10 infections over the last 1,787 person days at risk. About what is the one sided P-value for the relevant test of whether the hospital is *below* the standard?

```
lambda <- 1 / 100 * 1787
# Probability of 10 or less people infected given expected lambda people
infected.
ppois(10, lambda)
## [1] 0.03237153
```

Question 5

Suppose that 18 obese subjects were randomized, 9 each, to a new diet pill and a placebo. Subjects' body mass indices (BMIs) were measured at a baseline and again after having received the treatment or placebo for four weeks. The average difference from follow-up to the baseline (followup - baseline) was ???3 kg/m² for the treated group and 1 kg/m² for the placebo group. The corresponding standard deviations of the differences was 1.5 kg/m² for the treatment group and 1.8 kg/m² for the placebo group. Does the change in BMI appear to differ between the treated and placebo groups? Assuming normality of the underlying data and a common population variance, give a pvalue for a two sided t test.

```
# Treated group
trt.n <- 9
trt.mean <- -3
trt.sd <- 1.5

# Placebo group
plc.n <- 9
plc.mean <- 1
plc.sd <- 1.8

# Calculate t-value
(df <- trt.n + plc.n - 2)
## [1] 16
sp <- ((trt.n - 1) * trt.sd^2 + (plc.n - 1) * plc.sd^2) / df
```

Question 7

Researchers would like to conduct a study of 100 healthy adults to detect a four year mean brain volume loss of .01 mm³. Assume that the standard deviation of four year volume loss in this population is .04 mm³. About what would be the power of the study for a 5% one sided test versus a null hypothesis of no volume loss?

```
power.t.test(n=100, delta=.01, sd=.04, alt="one.sided", type="one.sample",
             sig.level=.05)$power
## [1] 0.7989855
```

Question 8

Researchers would like to conduct a study of n healthy adults to detect a four year mean brain volume loss of .01 mm³. Assume that the standard deviation of four year volume loss in this population is .04 mm³. About what would be the value of n needed for 90% power of type one error rate of 5% one sided test versus a null hypothesis of no volume loss?

```
power.t.test(delta=.01, sd=.04, alt="one.sided", type="one.sample",
sig.level=.05,
              power=.9)$n
## [1] 138.3856
```

Question 9

As you increase the type one error rate, ??, what happens to power?

```
s=seq(.01, .1, by=.01)
cbind(s,
      sapply(s, function (d) {
        power.t.test(n=100, delta=.01, sd=.04, alt="one.sided",
type="one.sample", sig.level=d)$power
      })
)
##           s
## [1,] 0.01 0.5554571
## [2,] 0.02 0.6625986
## [3,] 0.03 0.7246828
## [4,] 0.04 0.7672743
## [5,] 0.05 0.7989855
## [6,] 0.06 0.8237900
## [7,] 0.07 0.8438488
## [8,] 0.08 0.8604652
## [9,] 0.09 0.8744841
## [10,] 0.10 0.8864829
```