

Statistical Inference - Data Science - Quiz 4 - Coursera

Statistical Inference Quiz 4

This is Quiz 4 from the Statistical Inference course in the Data Science Specialization from John's Hopkins.

Questions

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)

Subject Baseline Week 2 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.

(Hint, consider that the observations are paired.)

• **0.087**

```
x <- c(140,138,150,148,135)
y <- c(132,135,151,146,130)
t.test(x,y, paired=TRUE)

##
## Paired t-test
##
## data: x and y
## t = 2.2616, df = 4, p-value = 0.08652
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.7739122 7.5739122
## sample estimates:
## mean of the differences
## 3.4
```

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?

- **1077 to 1123**
-

```
1100+c(-1,1)*10*qt(.975,8)
## [1] 1076.94 1123.06
```

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.

- **0.31**
-

```
#Binomial distribution
pbinom(3,5,.5) - pbinom(2,5,.5)
## [1] 0.3125
```

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?

- **0.03**
-

```
ppois(10,1787/100)
```

```
## [1] 0.03237153
```

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.

- **Less than 0.01**

```
x <- rnorm(9,mean=1,sd=1.5)
y <- rnorm(9,mean=-3,sd=1.8)
t.test(x,y)

##
## Welch Two Sample t-test
##
## data: x and y
## t = 5.9433, df = 15.849, p-value = 2.145e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.584579 5.454068
## sample estimates:
## mean of x mean of y
##  1.576508 -2.442816
```

6. Brain volumes for 9 men yielded a 90% confidence interval of 1,077 cc to 1,123 cc. Would you reject in a two sided 5% hypothesis test of $H_0: \mu = 1,078$?

- **No you wouldn't reject.**

```
2+c(-1,1)*qt(.975,98)*sqrt(2/100)*sqrt(((99*(.5)^2+99*2)/98))  
## [1] 1.576889 2.423111
```

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?

- **0.80**
-

```
power.t.test(n = 100, delta = .01-0, sd = .04, type = "one.sample", alt = "one.sided")$power  
## [1] 0.7989855
```

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?

- **140**
-

```
power.t.test(power = .9, delta = .01-0, sd = .04, type = "one.sample", alt = "one.sided")$n  
## [1] 138.3856
```

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

- **You will get larger power.**
-