

# Quiz 3 - Statistical Inference

*Jean-Luc BELLIER*

*16 octobre 2016*

## Question 1

In a a population of interest, a sample of 9 men yielded a sample average brain volume of 1,100cc and standard deviation of 30cc. What is a 95% Students' T confidence interval for a mean brain volume in this new population ?

- [1080,1120]
- [1077,1123]
- [1092,1108]
- [1031,1169]

**Answer :** The confidence interval is given by :

$$CI = \mu \pm \sigma.t_{0.975,n-1}$$

```
mu <- 1100
sigma <- 30
n <- 9
CI <- mu + c(-1,1)*sigma*qt(0.975,df=n-1)/sqrt(n)
CI
```

```
## [1] 1076.94 1123.06
```

*The 95% confidence interval for this t-test is [1077,1123].*

## Question 2

A diet pill is given to 9 subjects over six weeks. The average difference in weight (follow-up - baseline) is -2 pounds. What would the standard deviation of the difference in weight have to be for the upper endpoint of the 95% T confidence interval touch 0 ?

- 2.60
- 2.10
- 1.50
- 0.30

**Answer :** Regarding the previous formula, we must have :

$$CI = \mu + \sigma.t_{0.975,n-1} = 0$$

$$\text{so } -2 + t_{0.975,8} * \frac{\sigma}{\sqrt{(n)}} = 0$$

```
mu <- -2
n <- 9
sigma <- -mu*sqrt(n)/qt(0.975,df=n-1)
sigma
```

```
## [1] 2.601903
```

*The standard deviation would be 2.60.*

### Question 3

In an effort to improve running performance, 5 runners were either given a protein supplement or placebo. Then, after a suitable washout period, they were given the opposite treatment. Their mile times were recorded under both the treatment and placebo, yielding 10 measurements with 2 subjects per period. The researchers intend to use a t-test and interval to investigate the treatment. Should they use a paired or independent group T test and interval ?

- A paired interval
- It's necessary to use both
- You could use either
- Independant groups, since all subjects were seen under both systems

**Answer :** Here we want to know the impact of the substitution of treatment and placebo on two independant groups. *So we should use a paired t-test.*

### Question 4

In a study of emergency room waiting times, investigators consider a new and the standard triage system. To test the systems, administrators selected 20 nights and randomly assigned the new triage system to be used on 10 nights. They calculated the nightly median waiting time (MWT) to see a physician. The average MWT for the new system was 3 hours with a variance of 0.60 while the average MWT for the old system was 5 hours with a variance of 0.68. Consider the 95% confidence interval estimate for the differences of the mean MWT associated with the new system. Assume a constant variance. What is the interval ? Subtract in this order (New System - Old System).

- [1.25,2.75]
- [-2.70,-1.29]
- [-2.75,-1.25]
- [1.29,2.70]

**Answer :** Let us call :

- $S_{new}^2$  and  $S_{old}^2$  the variances for the new and old systems.
- $\mu_{new}$  and  $\mu_{old}$  the average MWT for the new and old systems.
- $n_{new}$  and  $n_{old}$  the number of nights for the new and old systems.

Then we have the pooled variance estimator defined as :

$$S_p^2 = \frac{(n_{new} - 1) * S_{new}^2 + (n_{old} - 1) * S_{old}^2}{n_{new} + n_{old} - 2}$$

And the confidence interval is :

$$CI = \mu_{new} - \mu_{old} \pm t_{0.975, n_{new} + n_{old} - 2} S_p * \sqrt{(1/n_{new} + 1/n_{old})^{1/2}}$$

```
mu_old <- 5
mu_new <- 3
n_old <- 10
n_new <- 10
v_new <- 0.6
v_old <- 0.68
S_p <- sqrt(((n_new-1)*v_new + (n_old-1)*v_old)/(n_new + n_old-2))
CI <- (mu_new - mu_old + c(-1,1)*qt(0.975, n_new+n_old-2)*S_p*sqrt(1/n_new + 1/n_old))
S_p
```

```
## [1] 0.8
```

```
qt(0.975, n_new+n_old-2)
```

```
## [1] 2.100922
```

```
CI
```

```
## [1] -2.751649 -1.248351
```

The confidence interval is then  $[-2.75, -1.25]$ .

## Question 5

Suppose that you create a 95% T confidence interval. You then create a 90% confidence interval using the same data. What can be said about the 90% interval with respect to the 95% interval ?

- The interval will be wider
- The interval will be of same width, but shifted
- The interval will be narrower
- It is impossible to tell

**Answer :** As  $t_{df, 975} > t_{df, 95}$ , *the 90% interval will be narrower than the 95% interval.*