

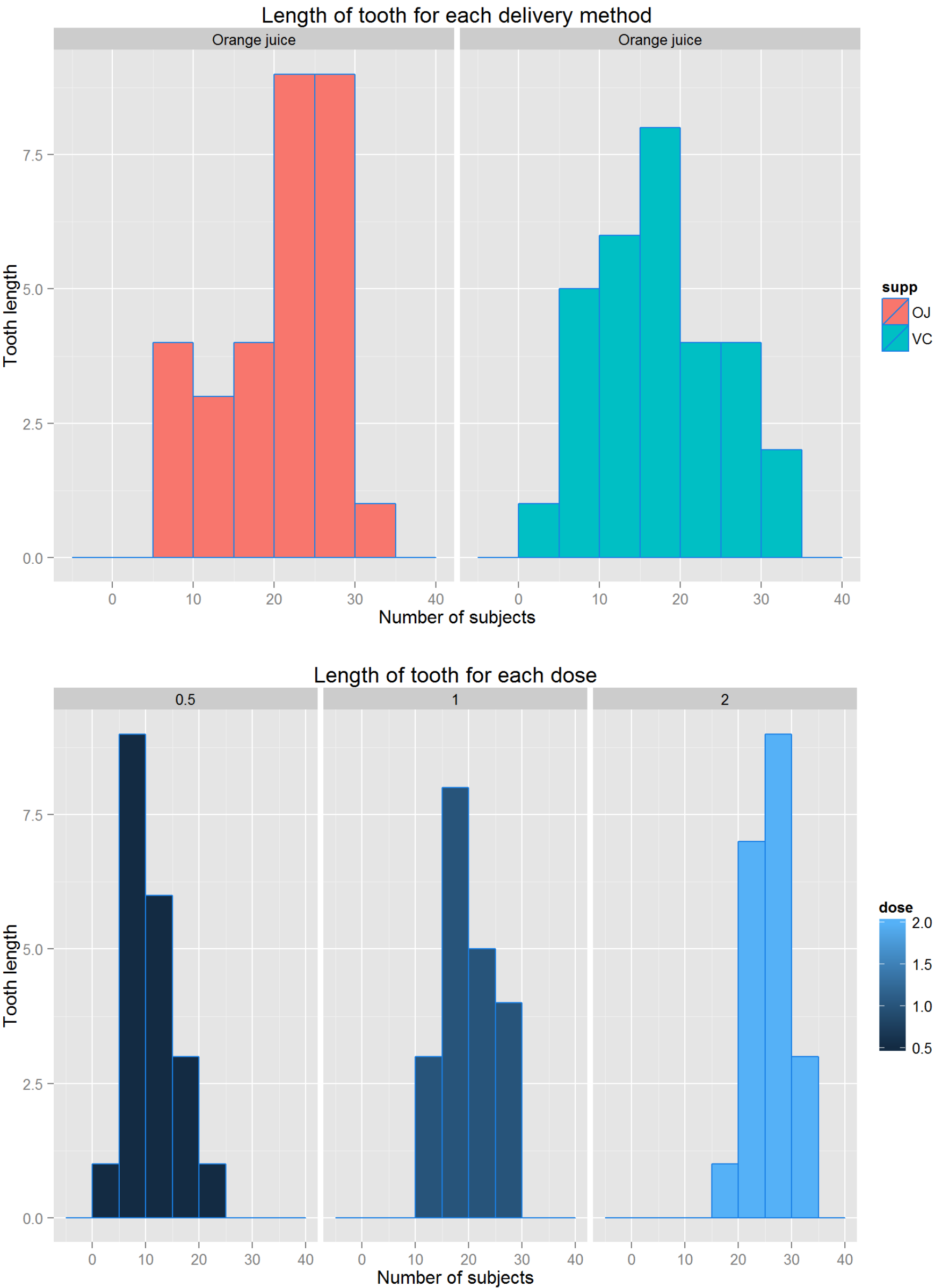
Report for PA of Statistical Inference

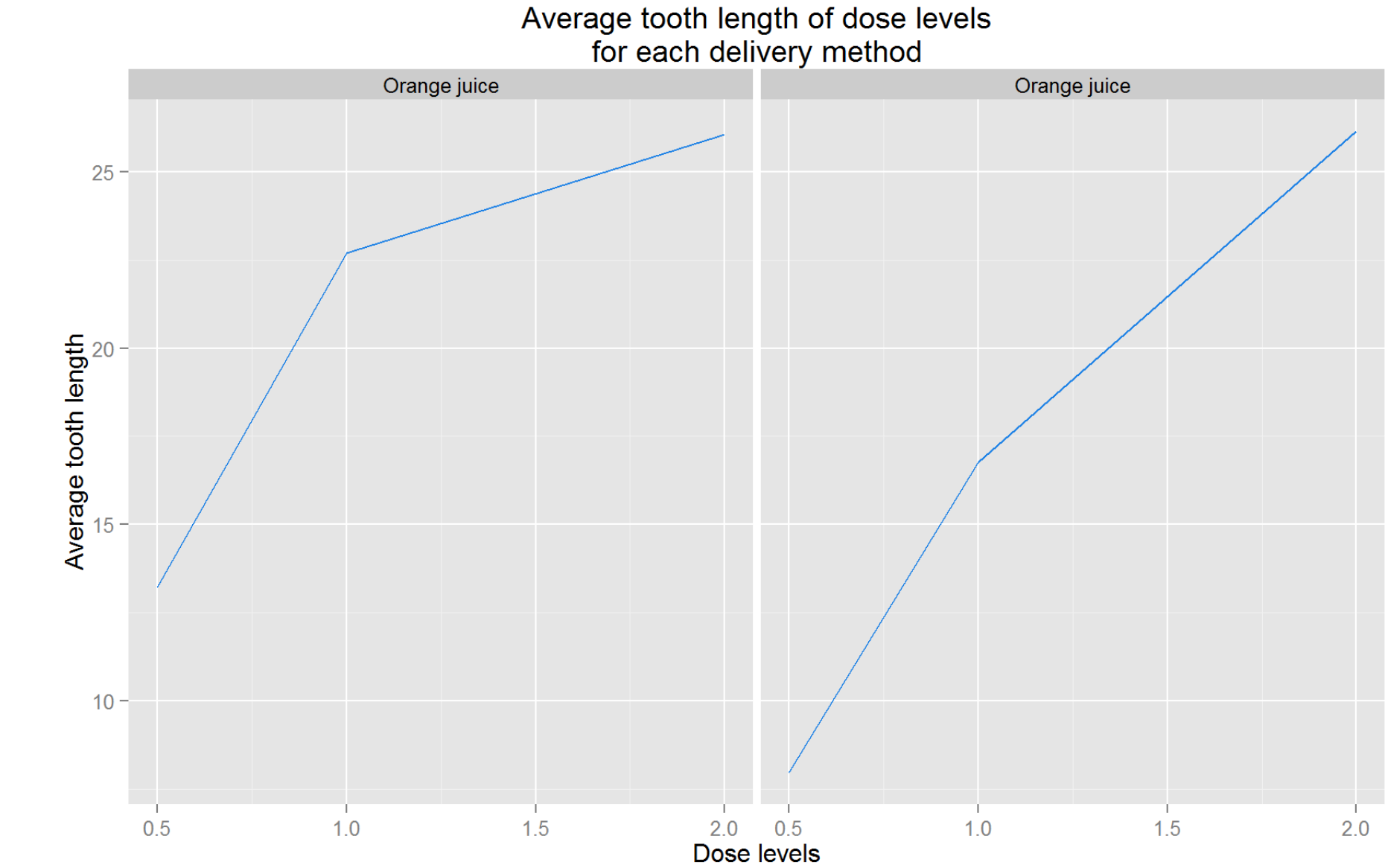
Thuan Bui

Saturday, December 20, 2014

Part 2: Analyze the ToothGrowth data in the R datasets package

1. Basic exploratory data analyses





Basic summary of the data

	len	supp	dose
1	Min. : 4.20	OJ:30	Min. :0.500
2	1st Qu.:13.07	VC:30	1st Qu.:0.500
3	Median :19.25		Median :1.000
4	Mean :18.81		Mean :1.167
5	3rd Qu.:25.27		3rd Qu.:2.000
6	Max. :33.90		Max. :2.000

Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose

The first step, I perform T-statistic for:

1. OJ vs VC delivery method of 0.5 dose
2. OJ vs VC delivery method of 1 dose
3. OJ vs VC delivery method of 2 dose
4. 2mg vs 1 dose of OJ delivery method
5. 1mg vs 0.5mg dose of OJ delivery method
6. 2mg vs 0.5mg dose of OJ delivery method
7. 2mg vs 1mg dose of VC delivery method
8. 1mg vs 0.5mg dose of VC delivery method
9. 2mg vs 0.5mg dose of VC delivery method

Table of T-statistic values

OJ vs VC delivery method

	OJ.vs.VC
0.5	3.17
1	4.03
2	-0.05

As you can see in above table, the T-statistic value of OJ compared with VS almost (except the 2mg dose) greater than 2.26 (the T-statistic of 95% confidence interval), and it greater than 0. It mean the diffence is big enough to say that the OJ give different result from VC and give longer odontoblasts than VC delivery method. With 2mg dose, OJ and VC delivery method almost give no difference

Comparison of 0.5, 1 and 2 dose

	One.mg.vs.Half.mg	Two.mg.vs.One.mg	Two.mg.vs.Half.mg
OJ	5.05	2.25	7.82
VC	7.46	5.47	10.39

As you can see in above table:

- The T-statistic value of 2mg dose compared with 1mg dose and 0.5mg dose greater than 2.26 (the T-statistic of 95% confidence interval, except Two mg vs One mg of OJ delivery method), and it greater than 0 (for both OJ and VC delivery method).
- The T-statistic value of 1mg dose compared with 0.5mg dose greater than 2.26 (the T-statistic of 95% confidence interval), and it greater than 0 (for both OJ and VC delivery method)

It mean the diffence is big enough to say that the 2mg dose give different result and give longer odontoblasts than 1mg. Respectively, 1mg dose give different result than and give longer odontoblasts than 0.5mg dose

Conclusion

With above test, we can see that the delivery method and dose effect to the length of odontoblasts (teeth) in each of 10 guinea pigs.

The orange juice(OJ) give longer odontoblasts than ascorbic acid (VC) delivery method. When increase dose (from 0.5mg to 1mg and 2 mg) the length of odontoblasts is also longer respectively.

Assumptions needed for conclusion

1. Bivariate independent variable (OJ vs Vc delivery methods, dose levels).
2. Continuous dependent variable.
3. Each observation of the dependent variable is independent of the other observations of the dependent variable (its probability distribution isn't affected by their values).
4. Dependent variable has a normal distribution, with the same variance.

Appendix: Code

```
# Group data
byOJSupp <- df %>% filter(supp == "OJ")
byOJSuppHalf <- byOJSupp %>% filter(dose == 0.5)
byOJSuppOne <- byOJSupp %>% filter(dose == 1)
byOJSuppTwo <- byOJSupp %>% filter(dose == 2)

byVCSupp <- df %>% filter(supp == "VC")
byVCSuppHalf <- byVCSupp %>% filter(dose == 0.5)
byVCSuppOne <- byVCSupp %>% filter(dose == 1)
byVCSuppTwo <- byVCSupp %>% filter(dose == 2)

# Test for OJ vs VC
testHalf <- t.test(byOJSuppHalf$len, byVCSuppHalf$len)
testOne <- t.test(byOJSuppOne$len, byVCSuppOne$len)
testTwo <- t.test(byOJSuppTwo$len, byVCSuppTwo$len)

# Test 2mg vs 1mg, 1mg vs 0.5mg, 2mg vs 0.5mg
testOJOneHalf <- t.test(byOJSuppOne$len, byOJSuppHalf$len)
testOJTwoHalf <- t.test(byOJSuppTwo$len, byOJSuppHalf$len)
testOJTwoOne <- t.test(byOJSuppTwo$len, byOJSuppOne$len)

testVCOneHalf <- t.test(byVCSuppOne$len, byVCSuppHalf$len)
testVCTwoHalf <- t.test(byVCSuppTwo$len, byVCSuppHalf$len)
testVCTwoOne <- t.test(byVCSuppTwo$len, byVCSuppOne$len)
```