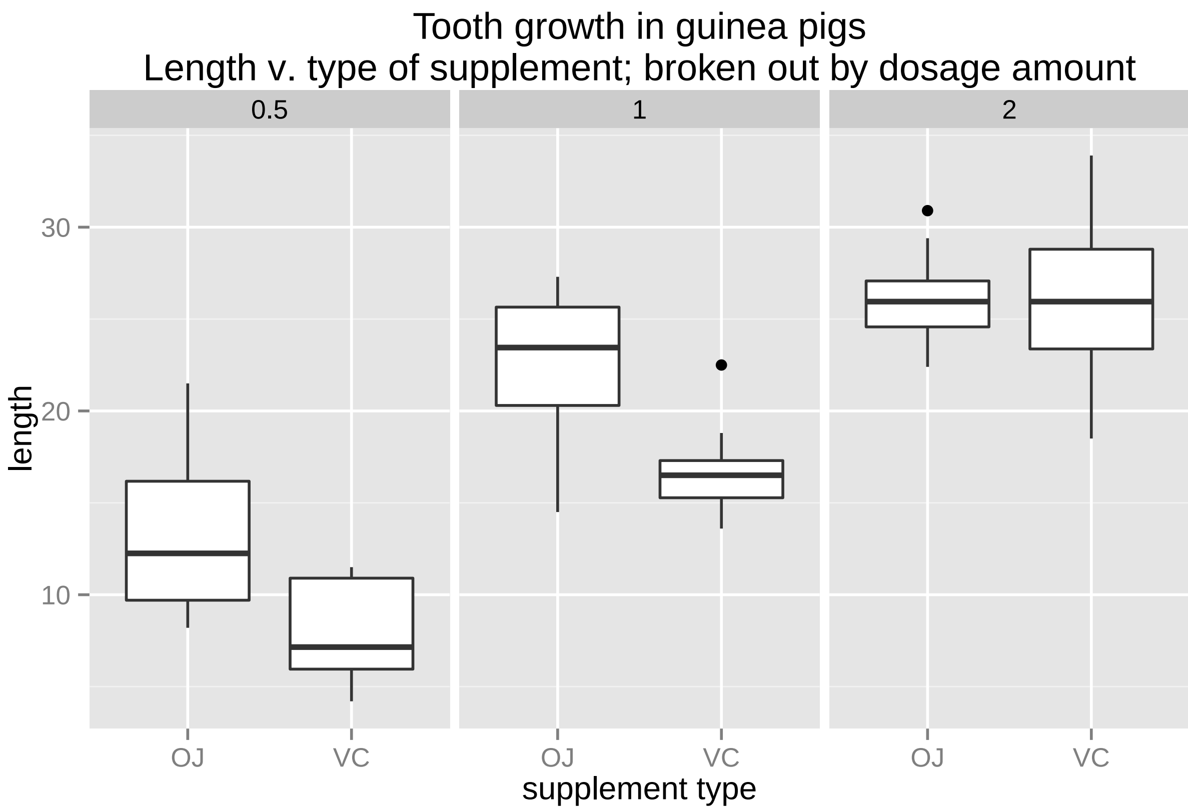
Analysis of the ToothGrowth R Data Set

David Schonberger

Sunday, August 24, 2014

The data set are observations of tooth growth in 10 guinea pigs. Each subject was given vitamin C in two forms–ascorbic acid (VC) and orange juice (OJ)–and in three doses–0.5 mg, 1 mg and 2 mg. It makes sense to look at boxplots for length versus supplement type, broken out by dose amount:

```
library(ggplot2)
data(ToothGrowth)
ggplot(ToothGrowth, aes(supp, len)) + geom_boxplot() + facet_grid(. ~ dose) +
  labs(y = "length") + labs(x = "supplement type") +
  labs(title = "Tooth growth in guinea pigs\nLength v. type of supplement; broken out by dosage amount")
```



Assumptions and Statement of Hypotheses:

1. Data are ordered, i.e. same 10 guinea pigs in same order in each of the six combinations of supplement type and dose.
2. With small samples we use a t-test.
3. Underlying iid Gaussian data, though the t-test is robust and can yield usable results even if the data are not normal.

Thus we do a paired 2-tailed t-test/CI for mean difference in length for the two treatments; with three distinct doses we do three tests/CIs. If we let μ_d be the mean difference in tooth growth between the two supplement types, VC - OJ, (for a given fixed dose), then we hypothesize: $H_0 : \mu_d = 0$ versus $H_a : \mu_d \neq 0$.

Separate and Pair Matched Data:

```
vc_halfmg <- ToothGrowth[1:10,1]
oj_halfmg <- ToothGrowth[31:40,1]
diff_halfmg <- vc_halfmg - oj_halfmg

vc_onemg <- ToothGrowth[11:20,1]
oj_onemg <- ToothGrowth[41:50,1]
diff_onemg <- vc_onemg - oj_onemg

vc_twomg <- ToothGrowth[21:30,1]
oj_twomg <- ToothGrowth[51:60,1]
diff_twomg <- vc_twomg - oj_twomg
```

T-test Calculations and Conclusions:

```
CI_halfmg <- t.test(diff_halfmg)$conf
pval_halfmg <- t.test(diff_halfmg)$p.value
```

With 95% CI bounds of -9.2365 and -1.2635 (which do not contain 0) and a p-value of $0.0155 < 0.05$, we reject H_0 and conclude there is strong evidence (at the $\alpha = 0.05$ level) that there is a difference in tooth growth for guinea pigs receiving ascorbic acid versus orange juice, when the dose is 0.5 mg.

```
CI_onemg <- t.test(diff_onemg)$conf
pval_onemg <- t.test(diff_onemg)$p.value
```

With 95% CI bounds of -9.9081 and -1.9519 (which do not contain 0) and a p-value of $0.0082 < 0.05$, we reject H_0 and conclude there is strong evidence (at the $\alpha = 0.05$ level) that there is a difference in tooth growth for guinea pigs receiving ascorbic acid versus orange juice, when the dose is 1 mg.

```
CI_twomg <- t.test(diff_twomg)$conf
pval_twomg <- t.test(diff_twomg)$p.value
```

With 95% CI bounds of -4.169 and 4.329 (which do 0) and a p-value of $0.967 > 0.05$, we fail to reject H_0 and conclude there is insufficient evidence (at the $\alpha = 0.05$ level) of a difference in tooth growth for guinea pigs receiving ascorbic acid versus orange juice, when the dose is 2 mg.

Additional Observation: Note that for the first two tests, where we reject H_0 in a two-tailed test, we would also reject for a left tailed test where we hypothesize $H_a : \mu_d < 0$. Hence we may conclude that there is *less* tooth growth when the guinea pigs are given ascorbic acid than when given orange juice, for the 0.5 mg and 1 mg doses.