ISDS Individual Assignment 2

Chen-Wei Huang

November 7, 2025

Q1.(a) Comparison between OJ and VC (ignoring dose)

Hypotheses. We test whether odontoblast length differs by supplement type:

$$H_0 : \mu_{\text{OJ}} = \mu_{\text{VC}}, \qquad H_1 : \mu_{\text{OJ}} \neq \mu_{\text{VC}}.$$

Test choice. A two-sample independent t-test (two-tailed, $\alpha = 0.05$) was applied. A Levene test indicated no evidence of unequal variances ($p = 0.2752$), so a pooled t-test was used.

Results. The pooled t-test yielded

$$t = 1.9153, \qquad df = 58, \qquad p = 0.06039.$$

The estimated mean difference (OJ $-$ VC) was

$$\widehat{\Delta} = 3.7000,$$

with a 95% confidence interval of

$$[-0.1670, \ 7.5670].$$

Conclusion. At the 5% significance level, we fail to reject H_0. There is no statistically significant evidence that mean odontoblast length differs between guinea pigs receiving orange juice and those receiving ascorbic acid when dosage is ignored, although orange juice shows a higher sample mean.

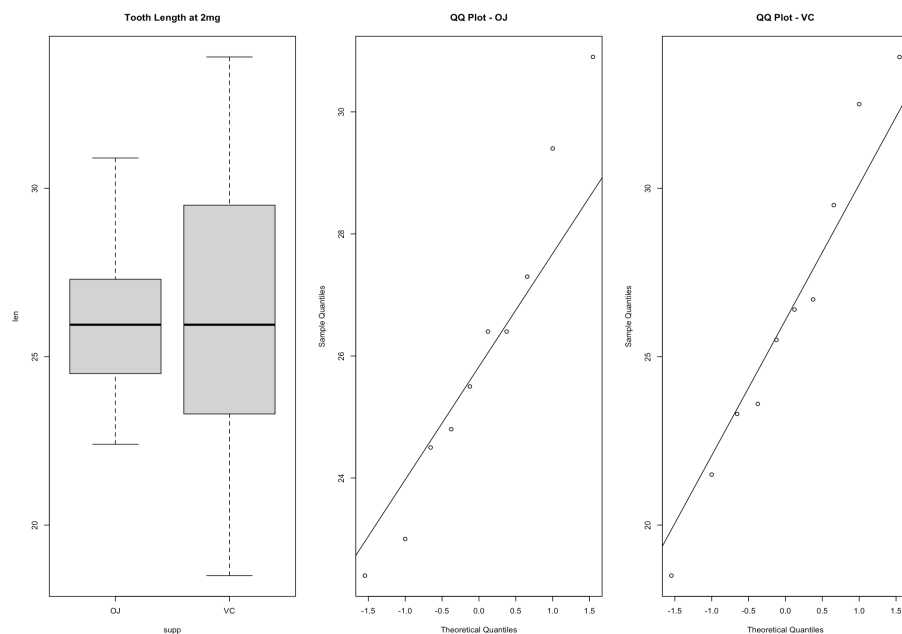## Q1.(b) Assessing assumptions for the 2 mg/day subgroup

Checks required. Before applying a two-sample t-test to the 2 mg/day subgroup, we must verify:

1. Normality within each supplement group;

2. Equality of variances between groups;

3. Independence of observations (assumed from the experimental design).

Normality. Shapiro–Wilk tests were conducted for each group:

$$p_{\mathrm{OJ}} = 0.8148, \qquad p_{\mathrm{VC}} = 0.9194.$$

Since both p-values exceed 0.05, there is no evidence against normality for either group. Q–Q plots further support approximate normality.

Equality of variances. A Levene test yielded

$$p = 0.1303.$$

Thus, the assumption of equal variances is reasonable at the 5% significance level.

Conclusion. The assumptions for a two-sample t-test are satisfied for the 2 mg/day subset. Therefore, applying a two-sample t-test would be appropriate, with the pooled form justified by the variance check.

Q2

Hypotheses. We test whether the mean global surface temperature from 1980–2024 differs from the historical mean of 13.9°C:

$$H_0 : \mu = 13.9 \qquad H_1 : \mu \neq 13.9.$$

Test. A one-sample two-tailed t-test was conducted at the 1% significance level.

Results. The sample statistics were:

$$\bar{x} = 14.48467, \qquad s = 0.28348, \qquad n = 45.$$

The test yielded:

$$t = 13.835, \qquad df = 44, \qquad p < 2.2 \times 10^{-16}.$$

The 1% two-tailed critical value was

$$t_{0.005,44} = 2.692278.$$

The 95% confidence interval for the mean was:

$$[14.39950, \ 14.56983].$$

Conclusion. Since $|t| = 13.835 > 2.692$ and $p < 0.01$, we reject H_0. There is strong statistical evidence that the mean global surface temperature from 1980–2024 differs from the historical average of 13.9°C and, in particular, that recent temperatures are significantly higher.

Q3

(a) **Family-wise error without correction.** Under independence and assuming all 15 null hypotheses are true, the family-wise error rate (FWER) at the per-test level $\alpha = 0.05$ is

$$\mathbb{P}(\text{at least one Type I error}) = 1 - (1 - \alpha)^n = 1 - (1 - 0.05)^{15} \approx 0.5367.$$

Comment. Running 15 tests at the 5% level yields about a 53.7% chance of making at least one false positive if all nulls are true.

(b) **Bonferroni-adjusted per-test level.** The Bonferroni correction sets the per-test level to
$$\alpha_{\text{Bonf}} = \frac{0.05}{15} \approx 0.00333.$$

(c) **FWER with Bonferroni (under independence).** If tests are independent, the resulting FWER is

$$1 - \left(1 - \alpha_{\text{Bonf}}\right)^{15} = 1 - \left(1 - \frac{0.05}{15}\right)^{15} \approx 0.04885.$$

Comment. Bonferroni guarantees FWER ≤ 0.05 by the union bound, and under independence the exact FWER is slightly below 0.05 (here $\approx 4.9\%$), reflecting its conservative nature. However, this strong control of Type I error comes at the cost of increased Type II error, reducing statistical power.