Unequal standard deviations

The two-sample t-test tools assume either

$$Y_{ij} \stackrel{ind}{\sim} N(\mu_j, \sigma^2) \qquad \text{or} \qquad Z_{ij} \stackrel{ind}{\sim} N(\mu_j, \sigma^2)$$

depending on whether we were working on the original scale (Y) or log scale (Z), respectively.

But what if we don't believe the variances in the two populations are equal, e.g. in the log transformed miles per gallon data set?

Instead compare

$$Y_{ij} \stackrel{ind}{\sim} N(\mu_j, \sigma_j^2) \qquad \text{or} \qquad Z_{ij} \stackrel{ind}{\sim} N(\mu_j, \sigma_j^2),$$

i.e. the populations have unequal variances. But still test $H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 \neq \mu_2$ or construct a confidence interval for $\mu_2 - \mu_1$.

Welch's SE with Satterthwaite's approximation to df

Estimate of $(\mu_2 - \mu_1)$:

$$\overline{Y}_2 - \overline{Y}_1$$

Standard error:

$$SE_W\left(\overline{Y}_2 - \overline{Y}_1\right) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Degrees of freedom using the Satterthwaite's approximation:

$$df_W = \frac{SE_W\left(\overline{Y}_2 - \overline{Y}_1\right)^4}{\frac{SE\left(\overline{Y}_2\right)^4}{n_2-1} + \frac{SE\left(\overline{Y}_1\right)^4}{n_1-1}}$$

where

$$SE\left(\overline{Y}_2\right) = \frac{s_2}{\sqrt{n_2}} \qquad \text{and} \qquad SE\left(\overline{Y}_1\right) = \frac{s_1}{\sqrt{n_1}}$$

(which is the same formula as in the paired t-test)

Welch's t-test and CI

Welch's t-test has test statistic:

$$t = \frac{\text{(Estimate-Parameter)}}{\text{SE(Estimate)}} = \frac{\overline{Y}_2 - \overline{Y}_1 - (\mu_2 - \mu_1)}{SE_W\left(\overline{Y}_2 - \overline{Y}_1\right)}$$

which has a t distribution with (approximately) df_W degrees of freedom if the null hypothesis is true. Calculate the pvalue

- Two-sided ($H_1 : \mu_2 \neq \mu_1$): $p = 2P(t_{df_W} < -|t|)$
- One-sided ($H_1 : \mu_2 > \mu_1$): $p = P(t_{df_W} < -t)$
- One-sided ($H_1 : \mu_2 < \mu_1$): $p = P(t_{df_W} < t)$

Two-sided $100(1-\alpha)\%$ confidence interval for $\mu_2 - \mu_1$:

$$\overline{Y}_2 - \overline{Y}_1 \pm t_{df_W}(1-\alpha/2)SE_W\left(\overline{Y}_2 - \overline{Y}_1\right)$$

Are the variances equal?

Suppose

$$Y_{ij} \stackrel{ind}{\sim} N(\mu_j, \sigma_j^2)$$

and you want to test $H_0 : \sigma_1 = \sigma_2$ vs $H_1 : \sigma_1 \neq \sigma_2$.

You can use an F-test and its associated pvalue. If the pvalue is small, e.g. less than 0.05, then we reject H_0. If the pvalue is not small, then we fail to reject H_0, but this does not mean the variances are not equal.

(Section 4.5.3) discusses another approach called Levene's test

Welch's test and CI using R

```
var.test(mpg~country,d) # F-test



F test to compare two variances

data:  mpg by country
F = 0.9066, num df = 78, denom df = 248, p-value = 0.6194
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6423 1.3246
sample estimates:
ratio of variances
            0.9066


(t=t.test(mpg~country, d, var.equal=FALSE))



Welch Two Sample t-test

data:  mpg by country
t = 12.95, df = 136.9, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  8.758 11.915
sample estimates:
mean in group Japan    mean in group US
              30.48               20.14
```

SAS code for two-sample t-test

```
DATA mpg;
    INFILE 'mpg.csv' DELIMITER=',' FIRSTOBS=2;
    INPUT mpg country $;

PROC TTEST DATA=mpg;
    CLASS country;
    VAR mpg;
    RUN;
```

SAS output for t-test

```
                              The TTEST Procedure

                                Variable:  mpg

     country           N        Mean     Std Dev     Std Err     Minimum     Maximum

     Japan            79     30.4810      6.1077      0.6872     18.0000     47.0000
     US              249     20.1446      6.4147      0.4065      9.0000     39.0000
     Diff (1-2)              10.3364      6.3426      0.8190

country       Method              Mean      95% CL Mean          Std Dev     95% CL Std Dev

Japan                          30.4810    29.1130  31.8491        6.1077     5.2814   7.2429
US                             20.1446    19.3439  20.9452        6.4147     5.8964   7.0336
Diff (1-2)    Pooled           10.3364     8.7252  11.9477        6.3426     5.8909   6.8699
Diff (1-2)    Satterthwaite    10.3364     8.7576  11.9152

                Method           Variances        df    t Value    Pr > |t|

                Pooled           Equal           326      12.62      <.0001
                Satterthwaite    Unequal      136.87      12.95      <.0001

                             Equality of Variances

                    Method      Num df    Den df    F Value    Pr > F

                    Folded F       248        78       1.10    0.6194
```

```
var.test(log(mpg)~country,d)



F test to compare two variances

data:  log(mpg) by country
F = 0.4617, num df = 78, denom df = 248, p-value = 0.0001055
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3271 0.6745
sample estimates:
ratio of variances
            0.4617


(t = t.test(log(mpg)~country, d, var.equal=FALSE))



Welch Two Sample t-test

data:  log(mpg) by country
t = 14.46, df = 193.3, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.3809 0.5013
sample estimates:
mean in group Japan    mean in group US
              3.396               2.955


exp(t$conf.int)


[1] 1.464 1.651
attr(,"conf.level")
```

SAS code for t-test using logarithms

```
DATA mpg;
  INFILE 'mpg.csv' DELIMITER=',' FIRSTOBS=2;
  INPUT mpg country $;

PROC TTEST DATA=mpg TEST=ratio;
CLASS country;
VAR mpg;
run;
```

SAS output for t-test using logarithms

The TTEST Procedure

Variable: mpg

country	N	Geometric Mean	Coefficient of Variation	Minimum	Maximum
Japan	79	29.8525	0.2111	18.0000	47.0000
US	249	19.2051	0.3147	9.0000	39.0000
Ratio (1/2)		1.5544	0.2928		

country	Method	Geometric Mean	95% CL Mean		Coefficient of Variation	95% CL CV	
Japan		29.8525	28.4887	31.2817	0.2111	0.1820	0.2514
US		19.2051	18.4825	19.9560	0.3147	0.2882	0.3467
Ratio (1/2)	Pooled	1.5544	1.4452	1.6719	0.2928	0.2712	0.3183
Ratio (1/2)	Satterthwaite	1.5544	1.4636	1.6508			

Method	Coefficients of Variation	DF	t Value	Pr > |t|
Pooled	Equal	326	11.91	<.0001
Satterthwaite	Unequal	193.33	14.46	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	248	78	2.17	0.0001

Summary

Two-sample t tools assumptions

- Normality
  - No skewness (take logs?)
  - No heavy tails
- Equal variances
  - Test: F-test or Levene's test
  - Use Welch's two-sample t-test and CI
- Independence (use random effects or avoid)
  - Cluster
  - Serial
  - Spatial