Satterthwaite Procedure

In 1946, Satterthwaite proposed a method to estimate the distribution of a linear combination of independent chi-square random variables with a chi-square distribution.

Let $U_1, \ldots, U_k$ be independent $\chi^2$ random variables with respective degree of freedom $\nu_1, \nu_2, \ldots, \nu_k$.

Consider the following linear combination of these random variables:

$$U = a_1 U_1 + a_2 U_2 + \ldots + a_k U_k.$$ 

The Satterthwaite approximation goes as follows:

$$\nu \frac{U}{E[U]}$$

has an approximate $\chi^2$ distribution with the following degrees of freedom

$$\nu = \left( \sum_{i=1}^{k} a_i U_i \right)^2 \frac{\sum_{i=1}^{k} (a_i U_i)^2 / \nu_i}{\sum_{i=1}^{k} (a_i U_i)^2 / \nu_i}.$$ 

**Remark:** We evaluate $\nu$ at the observed values of $U_1, \ldots, U_k$ and the approximation works best when the scalars are positive, i.e. $a_i > 0$.

**Example 1:** Consider $r$ independent normal populations. The $i$th population is $N(\mu_i, \sigma_i^2)$. Let $\bar{Y}_i$ and $S_i^2$ be the sample mean and the sample standard deviation from the $i$th population. The $i$th sample size is $n_i$.

Say we are interested in a linear combination of the means $L = \sum_{i=1}^{r} c_i \mu_i$.

An estimator of $L$ is

$$\hat{L} = \sum_{i=1}^{r} c_i \bar{Y}_i.$$ 

Now let use consider the following statistic:

$$s^2 \{ \hat{L} \} = \sum_{i=1}^{r} c_i^2 S_i^2 / n_i.$$ 

Using the Satterthwaite approximation, we can approximate the sampling distribution of $s^2 \{ \hat{L} \}$. That is

$$\frac{\nu s^2 \{ \hat{L} \}}{E \{s^2 \{ \hat{L} \} \}} = \frac{\nu s^2 \{ \hat{L} \}}{\sigma^2 \{ \hat{L} \}} = \frac{\nu s^2 \{ \hat{L} \}}{\sum_{i=1}^{r} c_i^2 \sigma_i^2 / n_i}.$$
has an approximate $\chi^2$ distribution with the following degrees of freedom

$$
\nu = \frac{(\sum_{i=1}^{r} c_i^2 S_i^2/n_i)^2}{\sum_{i=1}^{r} (c_i^2 S_i^2/n_i)^2/(n_i - 1)}.
$$

To show the latter recall that $U_i = (n_i - 1)S_i^2/\sigma_i^2 \sim \chi^2(n_i - 1)$. Let $a_i = (c_i^2 \sigma_i^2)/(n_i(n_i - 1))$ for $i = 1, \ldots, r$. Therefore,

$$
s^2\{\hat{L}\} = \sum_{i=1}^{r} c_i^2 S_i^2/n_i = \sum_{i=1}^{r} a_i U_i.
$$

This means that we can approximate the distribution of $(\nu s^2\{\hat{L}\})/\sigma^2\{\hat{L}\}$ with a $\chi^2$ distribution with the following degrees of freedom

$$
\nu = \frac{(\sum_{i=1}^{r} a_i U_i)^2}{\sum_{i=1}^{r} (a_i U_i)^2/\nu_i} = \frac{(\sum_{i=1}^{r} c_i^2 S_i^2/n_i)^2}{\sum_{i=1}^{r} (c_i^2 S_i^2/n_i)^2/(n_i - 1)}.
$$

**Example 2 [Welch-Satterthwaite Approximation]:** Consider the context of Example 1. We now consider the studentization of $\hat{L} = \sum_{i=1}^{r} c_i Y_i$, that is

$$
T = \frac{\hat{L} - L}{s\{\hat{L}\}} = \frac{\hat{L} - L}{\sqrt{\sum_{i=1}^{r} c_i^2 S_i^2/n_i}}.
$$

$T$ has an approximate $t$ distribution with the following degrees of freedom:

$$
\nu = \frac{(\sum_{i=1}^{r} c_i^2 S_i^2/n_i)^2}{\sum_{i=1}^{r} (c_i^2 S_i^2/n_i)^2/(n_i - 1)}.
$$

We will show the latter result by using the approximation of Example 1. First note that $Y_i$ and $S_i^2$ are independent for all $i = 1, \ldots, r$, since the populations are normals. Thus, the numerator and denominator of $T$ are independent. Now divide the numerator and the denominator of $T$ by the constant $\sigma\{\hat{L}\}$ to give

$$
T = \frac{(\hat{L} - L)/\sigma\{\hat{L}\}}{s\{\hat{L}\}/\sigma\{\hat{L}\}} = \frac{(\hat{L} - L)/\sigma\{\hat{L}\}}{\sqrt{\nu (s^2\{\hat{L}\}/\sigma^2\{\hat{L}\})/\nu}}.
$$

Let $Z$ be the numerator and $W/\nu$ be the denominator in the last expression. Clearly, $Z \sim N(0, 1)$ and $W \sim \chi^2(\nu)$ approximately with

$$
\nu = \frac{(\sum_{i=1}^{r} c_i^2 S_i^2/n_i)^2}{\sum_{i=1}^{r} (c_i^2 S_i^2/n_i)^2/(n_i - 1)}.
$$

Furthermore, $Z$ and $W$ are independent, this $T$ has an approximate $t(\nu)$ distribution.
Remark: The above $t$ approximation for $T$ was first proposed by Welch in 1947. He proposed a few different ways to approximate the degrees of freedom. However, the approximation that is mostly used is the above Satterthwaite approximation. So in the context of the $t$ approximation, it is sometimes called the Welch approximation, the Welch-Satterthwaite approximation or simply the Satterthwaite approximation.

Where does the approximation come from? Recall that for a chi-square distribution with $df$ degrees of freedom, then its mean is $df$ and variance is $2 \cdot df$. This means that $E\{U_i\} = \nu_i$ and that $\sigma^2\{U_i\} = 2 \cdot \nu_i$.

The idea behind the Satterthwaite approximation was to find a chi-square distribution $\chi^2(\nu)$ whose first two moments agreed with the first two moments of the true sampling distribution of the statistic $(\nu U) / E\{U\}$, where 

$$U = a_1 U_1 + \ldots + a_k U_k.$$ 

The first moment of the statistic is 

$$E\left\{ \frac{\nu U}{E\{U\}} \right\} = \frac{\nu}{E\{U\}} E\{U\} = \nu.$$ 

This means that we should use a chi-square with $\nu$ degrees of freedom. Now let us consider the second moment. The variance of the statistic is 

$$\sigma^2\left\{ \frac{\nu U}{E\{U\}} \right\} = \frac{\nu^2}{(E\{U\})^2} \sum_{i=1}^{k} a_i^2 \sigma^2\{U_i\}$$ 

$$= \frac{\nu^2}{(E\{U\})^2} \sum_{i=1}^{k} a_i^2 \cdot (2\nu_i)$$ 

$$= \frac{2 \nu^2}{(E\{U\})^2} \sum_{i=1}^{k} (a_i \nu_i)^2 / \nu_i$$ 

$$= \frac{2 \nu^2}{(E\{U\})^2} \sum_{i=1}^{k} (a_i E\{U_i\})^2 / \nu_i$$

The variance of $\chi^2(\nu)$ is $2\nu$. So equating variances, we get 

$$2\nu = \frac{2 \nu^2}{(E\{U\})^2} \sum_{i=1}^{k} (a_i E\{U_i\})^2 / \nu_i.$$
Now using sample moments instead of population moments means replacing $E\{U_i\}$ by $U_i$ and replacing $E\{U\}$ by $U$. We get

$$2\nu = \frac{2\nu^2}{(U)^2} \sum_{i=1}^{k} (a_i U_i)^2 / \nu_i.$$\nonumber

Solving for $\nu$ gives

$$\nu = \frac{U^2}{\sum_{i=1}^{k} (a_i U_i)^2 / \nu_i} = \frac{\left(\sum_{i=1}^{k} a_i U_i\right)^2}{\sum_{i=1}^{k} (a_i U_i)^2 / \nu_i}.$$\nonumber