Supplementary Material

1. Derivation of the relationship between S_{MS} and F

In the case of the randomised block design a considerable simplification of the CMH MS statistic is possible. In this case $n_{i\bullet j}$ = 1 for all i and j. Consequently $n_{\bullet\bullet j}$ = t, $n_{i\bullet\bullet}$ = b, $n_{\bullet\bullet\bullet}$ = bt and $p_{i\bullet j} = n_{i\bullet\bullet}/n_{\bullet\bullet\bullet}$ = 1/t. Substituting in the definitions given in Section 4 leads to

$$(t-1)S_j^2 = t\sum_h b_{hj}^2 n_{\bullet hj} - (\sum_h b_{hj} n_{\bullet hj})^2,$$

$$\mathrm{cov}(M_j) = S_j^2\{I_t - 1_t 1_t^{\mathrm{T}}/t\}/t \text{ and}$$

$$\mathrm{cov}(M) = \sum_{j=1}^b \mathrm{cov}(M_j) = \{I_t - 1_t 1_t^{\mathrm{T}}/t\}\sum_j S_j^2/t.$$

Now $1_t 1_t^{\mathrm{T}}/t$ is idempotent, as is $I_t - 1_t 1_t^{\mathrm{T}}/t$, and as any idempotent matrix is its own Moore-Penrose inverse

$$\mathrm{cov}^{-}(M) = t\{I_t - 1_t 1_t^{\mathrm{T}}/t\}/(\sum_j S_j^2).$$

In the definition of S_{MS},

$$M - \mathrm{E}[M] = (\sum_j M_{ij}) - (\sum_{h,j} b_{hj} n_{\bullet hj})/t.$$

Put $V = (V_i) = M - \mathrm{E}[M]$ so that $V_i = M_{i\bullet} - \sum_{h,j} b_{hj} n_{\bullet hj}/t$. The V_i are differences between the sum of the treatment scores over all strata and the mean of these. Now

$$1_t^{\mathrm{T}} V = \sum_i (M_{ij} - \mathrm{E}[M_{ij}]) = \sum_{h=1}^c b_{hj} n_{\bullet hj} - n_{\bullet\bullet j}\sum_{h=1}^c b_{hj} n_{\bullet hj}/n_{\bullet\bullet j} = 0.$$

It follows that

$$S_{MS} = V^{\mathrm{T}}\mathrm{cov}^{-}(M)V = t\{\sum_i V_i^2\}/\{\sum_j S_j^2\}.$$

To simplify this first put $x_{ij} = \sum_h b_{hj} N_{ihj}$. Note that $x_{\bullet j} = \sum_h b_{hj} n_{\bullet hj}$. On the jth block there is only one observation of treatment i, so for only one value of h is $n_{ihj} = 1$; otherwise $n_{ihj} = 0$. Thus, for example, $x_{ij}^2 = \sum_h b_{hj}^2 N_{ihj}$. Consider the data set $\{x_{ij}\}$. For the two-way ANOVA of these data write TSS for the treatments sum of squares and ESS for the error sum of squares.

If the data $\{x_{ij}\}$ are analysed as a one-way ANOVA or completely randomised design with blocks as treatments then the error sum of squares is $TSS + ESS$ and is given by

$$TSS + ESS = \sum_{i,j}(x_{ij} - x_{\bullet j}/t)^2 = \sum_{i,j} x_{ij}^2 - \sum_j x_{\bullet j}^2/t = \sum_j\{\sum_i x_{ij}^2 - x_{\bullet j}^2/t\} =$$
$$\sum_j\{\sum_i(\sum_h b_{hj}^2 N_{ihj}) - (\sum_h b_{hj} n_{\bullet hj})^2/t\} = \sum_j\{\sum_h b_{hj}^2 n_{\bullet hj} - (\sum_h b_{hj} n_{\bullet hj})^2/t\} =$$
$$(t-1)\sum_j S_j^2/t.$$

Further, for the two-way ANOVA the treatment sum of squares is (note that $\bar{x}$ means the mean of $\{x_i\}$ while $x_\bullet$ means the sum of the $\{x_i\}$)

$$TSS = \sum_{i,j}(\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet})^2 = b\sum_i\{x_{i\bullet}/b - x_{\bullet\bullet}/(bt)\}^2$$

$$= \frac{1}{b}\sum_i\{\sum_{h,j} b_{h,j} N_{ihj} - \sum_{h,j} b_{h,j} n_{\bullet hj}/t\}^2 = \frac{1}{b}\sum_i V_i^2.$$

It follows that

$$S_{\mathrm{MS}} = t\{\sum_i V_i^2\}/\{\sum_j S_j^2\} = \frac{bt\,TSS}{t\{TSS+ESS\}/(t-1)} = \frac{b(t-1)F}{(b-1+F)}$$

on using $TMS = TSS/(t-1)$, $EMS = ESS/\{(b-1)(t-1)\}$ and $F = TMS/EMS$, the F test statistic for treatments in the two-way ANOVA.

2. Derivation of the CMH C test statistic with the same scores in every stratum

Suppose that the same treatment and response scores are used on each stratum or block. Then $a_{ij} = a_i$ say, for all i and j, and $b_{hj} = b_h$ say, for all h and j. As $n_{i\bullet j} = 1$ for all i and j, and as a consequence $n_{\bullet\bullet j} = t$ for all j $S_{XXj} = \sum_i a_{ij}^2 n_{i\bullet j} - (\sum_i a_{ij} n_{i\bullet j})^2 / n_{\bullet\bullet j} = \sum_i a_i^2 - (\sum_i a_i)^2/t$, which is independent of j, so we may write $S_{XXj} = S_{XX}$. It is *not* true that S_{YYj} is independent of j.

Next we focus on S_{YYj}. Put $x_{ij} = \sum_h b_{hj} N_{ihj}$. Then $x_{ij}^2 = \sum_h b_{hj}^2 N_{ihj}$. Consider the data set $\{x_{ij}\}$. If these are analysed as a one-way ANOVA or completely randomised design with blocks as treatments then the error sum of squares is

$$\sum_{i,j}(x_{ij} - \bar{x}_{\bullet j})^2 = \sum_{i,j} x_{ij}^2 - t\sum_j \bar{x}_{\bullet j}^2 = \sum_j \{\sum_i x_{ij}^2 - x_{\bullet j}^2/t\} =$$

$$\sum_j \{\sum_i (\sum_h b_{hj}^2 N_{ihj}) - (\sum_h b_{hj} n_{\bullet hj})^2/t\} = \sum_j \{\sum_h b_{hj}^2 n_{\bullet hj} - (\sum_h b_{hj} n_{\bullet hj})^2/t\} = \sum_j S_j^2 .$$

Finally we focus on S_{XYj}. Write $\bar{a} = \sum_i a_i/t$. Then

$$\sum_j C_j = \sum_{i,h,j} a_{ij} b_{hj}\{N_{ihj} - \mathrm{E}[N_{ihj}]\} = \sum_{i,h,j} a_i b_{hj}\{N_{ihj} - n_{\bullet hj}/t\} .$$

Now

$$\sum_{i,h,j}\{a_i - \bar{a}\} b_{hj}\{N_{ihj} - n_{\bullet hj}/t\} = \sum_{i,h,j}\{a_i - \bar{a}\} b_h N_{ihj} - \sum_{i,h,j}\{a_i - \bar{a}\} b_{hj} n_{\bullet hj}/t .$$

Since

$$\sum_{i,h,j}\{a_i - \bar{a}\} b_{hj} n_{\bullet hj}/t = \sum_i \{a_i - \bar{a}\}\sum_h b_h \sum_j n_{\bullet hj}/t = \{\sum_h n_{\bullet h \bullet} b_h\}\sum_i \{a_i - \bar{a}\}/t = 0$$

$$C = \sum_j C_j = \sum_i \{a_i - \bar{a}\}\sum_{h,j} b_h N_{ihj} .$$

Note that $\sum_{h,j} b_h N_{ihj}$ is the sum of the ith treatment scores over responses and blocks, and these are usually easy to calculate directly or are readily available in most packaged analyses. The $\{a_i - \bar{a}\}$ are the centred treatment scores. Thus

$$S_{\mathrm{C}} = \frac{(t-1)C^2}{S_{XX}^2 \sum_j S_{YYj}^2}$$

in which C is the sum of the products of the treatment sums and the centred treatment scores, S_{XX}^2 is the sum of the squares of the centred treatment scores and $\sum_j S_{YYj}^2$ may be read from the output for a one-way ANOVA.