Expected Mean Squares and Approximate F Tests

NE Milla, Jr.

2025-11-12

Steps in Deriving Expected Mean Squares

1. Prepare a two way table with the terms of the model as the row labels.
2. Write the subscripts in the model as column headings. Over each subscript, write the number of levels of the factors associated with that subscript and whether the factor is fixed (F) or random (R). Replicates are always treated as random.
3. In each row, write 1 if one of the dead subscripts in the row components matches the subscript in the column.
4. In each row, if any of the live subscripts on the row component match the subscript in the column, write 0 if the column is headed by a fixed factor and 1 if the column is headed by a random factor.
5. In the remaining cells, write the number of levels shown above the column heading.
6. To obtain the EMS for any model component, first cover all columns headed by live subscripts on that component. Then in each row that contains at least the same subscripts as those on the component being considered, take the product of the visible numbers and multiply by the appropriate fixed or random factor. The sum of these quantities is the EMS of the model component being considered.

Three-Factor Factorial Model

$Y_{ijkl} = \mu+\alpha_i+\beta_j+\gamma_k+(\alpha\beta)_{ij}+(\alpha\gamma)_{ik}+(\beta\gamma)_{jk}+(\alpha\beta\gamma)_{ijk}+\epsilon_{ijkl} i = 1, 2, \cdots, a;\ j = 1, 2, \cdots, b;\ k = 1, 2, \cdots, c;\ l = 1, 2, \cdots, r$

Expected Mean Squares (Random Model)

	R	R	R	R
	a	b	c	r
	i	j	k	l
α_i				
β_j				
γ_i				
$(\alpha\beta)_{ij}$				
$(\alpha\gamma)_{ik}$				
$(\beta\gamma)_{jk}$				
$(\alpha\beta\gamma)_{ijk}$				
$\epsilon_{(ijk)l}$				

Model Term	Expected Mean Squares
α_i	$\sigma_\epsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + cr\sigma_{\alpha\beta}^2 + br\sigma_{\alpha\gamma}^2 + bcr\sigma_\alpha^2$
β_j	$\sigma_\epsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + cr\sigma_{\alpha\beta}^2 + ar\sigma_{\beta\gamma}^2 + acr\sigma_\beta^2$
γ_k	$\sigma_\epsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + br\sigma_{\alpha\gamma}^2 + ar\sigma_{\beta\gamma}^2 + abr\sigma_\gamma^2$
$(\alpha\beta)_{ij}$	$\sigma_\epsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + cr\sigma_{\alpha\beta}^2$
$(\alpha\gamma)_{ik}$	$\sigma_\epsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + br\sigma_{\alpha\gamma}^2$
$(\beta\gamma)_{jk}$	$\sigma_\epsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + ar\sigma_{\beta\gamma}^2$
$(\alpha\beta\gamma)_{ijk}$	$\sigma_\epsilon^2 + r\sigma_{\alpha\beta\gamma}^2$
ϵ_{ijkl}	σ_ϵ^2

Approximate F tests

Satterthwaite's approximation method

- Uses linear combinations of expected mean squares, such as

$$MS^{'} = MS_r + \cdots + MS_s$$
$$MS^{"} = MS_u + \cdots + MS_v$$

where:

(1) The mean squares are so chosen that no MS appear simultaneously in $MS^{'}$ and $MS^{"}$ and
(2) $E(MS^{'}) - E(MS^{"})$ is equal to the effect considered in the null hypothesis.

The F statistic is then given by

$$F = \frac{MS^{'}}{MS^{"}}$$

which is distributed approximately as $F_{(p,q)}$, where p and q are the numerator and denominator degrees of freedom, respectively, and are computed as follows:

$$p = \frac{(MS_r + \cdots + MS_s)^2}{\frac{MS_r^2}{df_r} + \cdots + \frac{MS_s^2}{df_s}}$$

and

$$q = \frac{(MS_u + \cdots + MS_v)^2}{\frac{MS_u^2}{df_r} + \cdots + \frac{MS_v^2}{df_s}}$$

We round off to the nearest integer the values of p and q computed using the above formulas in case these are non-integers.

Suppose we want to test $H_0 : \sigma_\alpha^2 = 0$ versus $H_1 : \sigma_\alpha^2 > 0$.

Then an approximate F test is given by:

$$F = \frac{MS_A + MS_{ABC}}{MS_{AB} + MS_{AC}}$$

Why???

What would be the approximate test statistics for testing the hypotheses $H_0 : \sigma_\beta^2 = 0$ and $H_0 : \sigma_\gamma^2 = 0$?

For testing $H_0 : \sigma_\beta^2 = 0$, the approximate F statistic is

$$F = \frac{MS_B + MS_{ABC}}{MS_{AB} + MS_{BC}}$$

While , for testing $H_0 : \sigma_\gamma^2 = 0$, the approximate F statistic is

$$F = \frac{MS_C + MS_{ABC}}{MS_{AC} + MS_{BC}}$$