Θα αποδείξουμε ότι στην ανάλυση της διακύμανσης κατά δύο παράγοντες οι στατιστικές συναρτήσεις F_A και F_B ακολουθούν την F κατανομή.

<u>Πρόταση:</u> Για το μοντέλο $Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$, $\varepsilon_{ij} \sim N(0, \sigma^2)$, με μία παρατήρηση ανά κυψελίδα ισχύει ότι:

1) SS_{res} , SS_A , SS_B είναι ανεξάρτητα.
2) Όταν ισχύει η $H_A : \alpha_i$ =0, $\forall i$ $\square$ $SS_A / \sigma^2 \sim \mathrm{X}^2_{I-1}$.
3) Όταν ισχύει η $H_B : \beta_j$ =0, $\forall j$ $\square$ $SS_B / \sigma^2 \sim$ X^2_{J-1}.
4) $SS_{res} / \sigma^2 \sim \mathrm{X}^2_{(I-1)(J-1)}$ (ανεξάρτητα αν ισχύουν οι H_A , H_B).
5) Η στατιστική συνάρτηση $F_A \sim F_{(I-1);(I-1)(J-1)}$ όταν ισχύει η H_A .
6) Η στατιστική συνάρτηση $F_B \sim F_{(J-1);(I-1)(J-1)}$ όταν ισχύει η H_B .

<u>Απόδειξη:</u>
Τα (5) και (6) είναι συνέπειες των (2), (3), (4) και της ανεξαρτησίας (1).
1) Θα αποδείξουμε ότι SS_A , SS_B και SS_{res} είναι ανεξάρτητα.
$SS_A \,\square\, SS_B$:
Αρκεί να δείξουμε λόγω κανονικότητας ότι οι παραστάσεις της μορφής {$\bar{Y}_{i.}$ - $\bar{Y}_{..}$} και {$\bar{Y}_{.j}$ - $\bar{Y}_{..}$} είναι ασυσχέτιστες.
Υπολογίζω
Cov($\bar{Y}_{i.}$ - $\bar{Y}_{..}$, $\bar{Y}_{.j}$ - $\bar{Y}_{..}$) = Cov($\bar{Y}_{i.}$, $\bar{Y}_{.j}$) - Cov($\bar{Y}_{i.}$ - $\bar{Y}_{..}$) - Cov($\bar{Y}_{..}$, $\bar{Y}_{.j}$) + Cov($\bar{Y}_{..}$, $\bar{Y}_{..}$).
Αλλά
Cov($\bar{Y}_{i.}$, $\bar{Y}_{.j}$) = Cov $\left(\frac{1}{J}\sum_{k=1}^{J} Y_{ik} , \frac{1}{I}\sum_{l=1}^{I} Y_{lj}\right) = \frac{\sigma^2}{IJ}$ (λόγω ανεξαρτησίας).

Cov($\bar{Y}_{i.}$, $\bar{Y}_{..}$)= Cov$\left(\bar{Y}_{i.} , \frac{1}{I}\sum_{k=1}^{I} \bar{Y}_{.k}\right) = \frac{\sigma^2}{IJ}$.

$$\bar{Y}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij} = \frac{1}{IJ}\sum_{i=1}^{I} Y_{i.} = \frac{1}{IJ}\sum_{i=1}^{I} J\bar{Y}_{i.} = \frac{1}{I}\sum_{i=1}^{I} \bar{Y}_{i.}$$

Cov($\bar{Y}_{.j}$, $\bar{Y}_{..}$)=$\sigma^2 / I\ J$ (αποδεικνύεται κατά τον ίδιο τρόπο).
Var($\bar{Y}_{..}$)=$\sigma^2 / I\ J$.

Άρα, Cov($\bar{Y}_{i.}$ - $\bar{Y}_{..}$, $\bar{Y}_{.j}$ - $\bar{Y}_{..}$) = $\frac{\sigma^2}{IJ}$ - $\frac{\sigma^2}{IJ}$ - $\frac{\sigma^2}{IJ}$ + $\frac{\sigma^2}{IJ}$ =0.

Θα δείξω ότι $SS_{res} \,\square\, SS_A$:
Cov(Y_{ij} - $\bar{Y}_{i.}$ - $\bar{Y}_{.j}$ + $\bar{Y}_{..}$, $\bar{Y}_{i.}$ - $\bar{Y}_{..}$) = Cov(Y_{ij} , $\bar{Y}_{i.}$) - Cov(Y_{ij} , $\bar{Y}_{..}$) - Cov($\bar{Y}_{i.}$, $\bar{Y}_{i.}$) + Cov($\bar{Y}_{i.}$, $\bar{Y}_{..}$) ($\square$)
αφού {$\bar{Y}_{.j}$ - $\bar{Y}_{..}$} είναι ανεξάρτητο του {$\bar{Y}_{i.}$ - $\bar{Y}_{..}$}.

Αλλά Cov(Y_{ij} , $\bar{Y}_{i.}$) = σ^2 / J }
Cov(Y_{ij} , $\bar{Y}_{..}$) = $\sigma^2 / I\ J$ } $\square$ ($\square$)=0 και συνεπώς
Var($\bar{Y}_{i.}$) = σ^2 / J } $SS_{res} \,\square\, SS_A$.
Cov($\bar{Y}_{i.}$, $\bar{Y}_{..}$) = $\sigma^2 / I\ J$ }

Παρόμοια αποδεικνύω ότι $SS_{res} \,\square\, SS_B$.

2) Θα δείξω ότι $SS_A/\sigma^2 \sim \mathrm{X}^2_{I-1}$ όπου $SS_A = J\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{Y}_{..})^2$, όταν ισχύει η $H_A: \alpha_i$ =0.

Όταν ισχύει η μηδενική υπόθεση, έχω ότι Y_{ij} ~N($\mu + \beta_j, \sigma^2$).

Άρα $\overline{Y}_{i.}$ ~N($\mu, \sigma^2/I$), αφού $\sum_{j=1}^{J}\beta_j = 0$.

Αφού τα {$\overline{Y}_{i.}$} είναι ανεξάρτητα έχω ότι $\dfrac{\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{Y}_{..})^2}{\sigma^2/J} \sim \mathrm{X}^2_{I-1}$

□ $SS_A/\sigma^2 \sim \mathrm{X}^2_{I-1}$.

3) Η απόδειξη του (3) γίνεται ακριβώς όπως του (2).

4) Θα χρησιμοποιήσω την παρακάτω ταυτότητα:

5) $\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \mu - \alpha_i - \beta_j)^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}_{..})^2 + I\,J\,(\overline{Y}_{..} - \mu)^2$

$J\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{Y}_{..} - \alpha_i)^2 + I\sum_{j=1}^{J}(\overline{Y}_{.j} - \overline{Y}_{..} - \beta_j)^2$ (□□).

Έχω ότι $(Y_{ij} - \mu - \alpha_i - \beta_j)^2/\sigma^2 \sim \mathrm{X}^2_1$ □

Άρα $\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \mu - \alpha_i - \beta_j)^2/\sigma^2 \sim \mathrm{X}^2_{IJ}$.

Επειδή $\overline{Y}_{..}$ ~N($\mu, \sigma^2/I\,J$) □ $\dfrac{IJ(\overline{Y}_{..} - \mu)^2}{\sigma^2} \sim \mathrm{X}^2_1$.

Ακόμα έχω ότι $\overline{Y}_{i.}$ ~N($\mu + \alpha_i, \sigma^2/J$)

□ $\overline{Y}_{i.} - \alpha_i$ ~N($\mu, \sigma^2/J$).

Θέτω $Z_i = \overline{Y}_{i.} - \alpha_i$. Έχω ότι

$J\sum_{i=1}^{I}(Z_i - \overline{Z})^2/\sigma^2 \sim \mathrm{X}^2_{I-1}$ }

Αλλά $\overline{Z} = \overline{Y}_{..}$ } □ $J\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{Y}_{..} - \alpha_i)^2/\sigma^2 \sim \mathrm{X}^2_{I-1}$.

Ανάλογα δείχνω ότι $I\sum_{j=1}^{J}(\overline{Y}_{.j} - \overline{Y}_{..} - \beta_j)^2/\sigma^2 \sim \mathrm{X}^2_{J-1}$.

Από την (□□) και λόγω ανεξαρτησίας των όρων του δευτέρου μέλους της ισότητας έχω ότι:

$$SS_{res}/\sigma^2 \sim \mathrm{X}^2_{(I-1)(J-1)}.$$