

Bài 3 PHÂN TÍCH PHƯƠNG SAI MỘT NHÂN TỐ

I- NỘI DUNG

Trong chương trước đã trình bày cách so sánh hai trung bình của **hai tổng thể**, mở rộng sang so sánh trung bình của nhiều tổng thể chúng ta có bài toán phân tích phương sai một nhân tố (single factor anova).

Theo dõi ảnh hưởng của **a công thức hay nghiệm thức thí nghiệm (treatement)** đến kết quả thí nghiệm. Công thức có thể chỉ bao gồm **một yếu tố** (Giống, chế độ canh tác, mật độ trồng, loại thuốc trừ sâu bệnh, phương pháp làm đất, chế độ nước ...), cũng có thể bao gồm **nhiều yếu tố** (giống x phân bón, giống x mật độ, mật độ x chế độ nước x phân bón . . .), nhưng **không xét tác động riêng của từng yếu tố mà xét tác động chung của các yếu tố và gọi đó là tác động của một nhân tố**.

Trong tài liệu này nhân tố A được coi là cố định (Fixed)

Việc bố trí thí nghiệm (thiết kế thí nghiệm) để so sánh các trung bình của a công thức được gọi là **bố trí thí nghiệm một nhân tố**, mỗi công thức thí nghiệm là một **mức** của nhân tố. Các mức được coi là định tính và có tên, thường gọi là **nhãn (label)**, để đơn giản gọi **a** mức là $A_1, A_2, \dots, A_a$

Làm thí nghiệm so sánh năng suất của 5 giống ngô thì nhân tố ở đây chỉ gồm một yếu tố có 5 mức là 5 giống ngô, hay còn gọi là 5 công thức. Mỗi giống ngô được thử nghiệm trên một số ô thí nghiệm (hay đơn vị thí nghiệm), mỗi ô được coi là một lần lặp (repetition). Thí dụ nếu mỗi giống lặp lại 3 lần thì phải có $5 \cdot 3 = 15$ ô thí nghiệm.

Thí nghiệm 5 giống ngô và 4 công thức bón phân và chỉ xét tác động chung của tổ hợp Giống x Phân ($G_i \times P_j$) thì có thí nghiệm một nhân tố với $5 \cdot 4 = 20$ công thức thí nghiệm, mỗi công thức được lặp lại 3 lần, như vậy phải có $5 \cdot 4 \cdot 3 = 60$ ô thí nghiệm.

Vì **chỉ quan tâm đến một nhân tố** nên các dữ liệu được sắp thành từng nhóm, mỗi nhóm là các lần lặp của một mức của nhân tố do đó còn gọi việc phân tích số liệu nhằm

tách biệt các phương sai theo hai nguồn biến động **nhân tố** và **sai số** là **bài toán phân tích phương sai một cách sắp xếp (one way anova)**.

Giả sử **công thức A_i** được thực hiện trên **r_i ô thí nghiệm**, các kết quả **x_{ij}** được coi như một mẫu quan sát đối với biến ngẫu nhiên **X_i** và mục đích đặt ra là so sánh các **trung bình m_i** của các biến **X_i** .

Có nhiều kiểu bố trí thí nghiệm để giải quyết bài toán này.

Giả sử nhân tố có **a** mức, mức **i** được lặp lại **r_i** lần, như vậy tổng số có **$n = \sum r_i$** quan sát, hay còn nói là có **n ô thí nghiệm**.

Nếu bố trí **n ô thí nghiệm** hoàn toàn ngẫu nhiên thì kiểu bố trí được gọi là **kiểu bố trí (thiết kế) hoàn toàn ngẫu nhiên** (Completely randomized design).

a - KIỂU BỐ TRÍ HOÀN TOÀN NGẪU NHIÊN **(Completely randomized design CRD)**

Khi tiến hành thí nghiệm kiểu này phải dùng **n** phiếu ghi từ 1 đến **n**, rút thăm ngẫu nhiên **r_1** phiếu để có các ô thí nghiệm đối với công thức 1, rút tiếp **r_2** phiếu để có các ô thí nghiệm đối với công thức 2, . . . , **r_a** ô cuối cùng là của công thức **a**.

Như vậy việc rút thăm ngẫu nhiên được thực hiện trên **toàn bộ các ô thí nghiệm**.

a1- Mô hình toán học

Việc tính toán và kết luận dựa trên một số giả thiết thể hiện ở mô hình sau:

$$x_{ij} = \mu + \alpha_i + e_{ij} \quad (i = 1, \dots, a; j = 1, \dots, r_i) \quad (1)$$

x_{ij} là kết quả của lần lặp thứ **j** của mức **i**, **μ** là trung bình chung, **α_i** là ảnh hưởng của mức **i** của nhân tố, còn **e_{ij}** là sai số ngẫu nhiên. **x_{ij}** có trung bình **$m_i = \mu + \alpha_i$**

Các sai số e_{ij} được giả thiết độc lập, phân phối chuẩn, kỳ vọng 0, phương sai σ^2

Các **α_i** thỏa mãn điều kiện ràng buộc **$\sum \alpha_i = 0$**

a2- Các bước tính

Giả sử có **a** mức, mức **A_i** lặp lại **r_i** lần.

Tổng số ô thí nghiệm (hay số số liệu) **$n = \sum r_i = 24$**

Tổng các số liệu của công thức **i** **$T A_i = \sum_j x_{ij}$** , các trung bình **$\bar{x}_i$**

(xem bảng)

Tổng tất cả các số liệu $ST = \sum_i \sum_j x_{ij}$ trung bình chung $\bar{x} = \frac{ST}{n} ..$

Số điều chỉnh $G = ST^2 / n$

Tính các tổng bình phương:

Tổng bình phương toàn bộ

$$SSTO = \sum_{i=1}^a \sum_{j=1}^{r_i} (x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^{r_i} x_{ij}^2 - G$$

Tổng bình phương do nhân tố:

$$SSA = \sum_{i=1}^a \sum_{j=1}^{r_i} (\bar{x}_{i.} - \bar{x}_{..})^2 = \sum_{i=1}^a \frac{TA_i^2}{r_i} - G$$

Tổng bình phương do sai số:

$$SSE = SSTO - SSA = 260,2148 - 140,6471 = 119,5677$$

Tính các bậc tự do

Bậc tự do của SSTO $dfTO = n - 1$

Bậc tự do của SSA $dfA = a - 1$

Bậc tự do của SSE $dfE = n - a$

Dem các tổng bình phương **SSA** và **SSE** chia cho các bậc tự do tương ứng được các bình phương trung bình **msA**, **msE**.

$$F_{tn} = \frac{msA}{msE} \quad \text{Giá trị tới hạn } F_{lt} = F(\alpha, dfA, dfE)$$

Sai số thí nghiệm bình phương là msE, ký hiệu se^2 với **bậc tự do** $dfE = n - a$

Tóm tắt các kết quả vào bảngsau:

Bảng phân tích phương sai

Nguồn biến động	Tổng BP	Bậc tự do	Bình phương trung bình	Ftn	Flt
Giữa các mức	SSA	dfA = a - 1	msA = SSA/dfA	msA/msE	F(α, dfA, dfE)
Sai số ngẫu nhiên	SSE	dfE = n - a	msE = SSE / dfE = se^2		
Toàn bộ	SSTO	dfTO = n - 1			

a3-Kết luận

Dùng bảng phân tích phương sai để kiểm định giả thiết H_0 : “Không có sự khác nhau giữa các trung bình m_i ”, đối thiết H_1 : “Có sự khác nhau giữa các trung bình m_i ”.

Có thể viết lại theo α_i và có giả thiết H_0 : “Các α_i đều bằng 0” với đối thiết H_1 :

” Không phải các α_i đều bằng 0”.

Quy tắc kiểm định:

So F_{tn} với ngưỡng F_{lt}

Nếu $F_{tn} \leq F_{lt}$ chấp nhận giả thiết H_0 : “Không có sự khác nhau giữa các trung bình m_i của các mức của nhân tố”.

Nếu $F_{tn} > F_{lt}$ chấp nhận H_1 : “Có sự khác nhau giữa các trung bình m_i của các mức của nhân tố”.

Sai số của trung bình $\bar{x}_{..}$ $se_{\mu} = \sqrt{\frac{se^2}{n}}$

Sai số của trung bình của các công thức

$$\begin{aligned} se_A &= \sqrt{\frac{se^2}{r_a}} & se_C &= \sqrt{\frac{se^2}{r_c}} \\ se_B &= \sqrt{\frac{se^2}{r_b}} & se_D &= \sqrt{\frac{se^2}{r_d}} \end{aligned}$$

b- KIỂU BỐ TRÍ KHỐI NGẪU NHIÊN ĐẦY ĐỦ

(Randomized complete block design RCBD hay RCB)

Để tiến hành thí nghiệm giả sử có **a** công thức, mỗi công thức lặp lại **r** lần. Tất cả có $n = a \times r$ ô thí nghiệm.

Chọn **r** khối, mỗi khối chia thành **a** ô thí nghiệm. Lấy khối thứ nhất và làm **a** phiếu để bắt thăm xem **a công thức** xếp vào **a ô nào**, sau đó bắt thăm cho khối thứ hai, thứ ba, ..., thứ a. Như vậy **việc chọn ngẫu nhiên được làm riêng cho từng khối**.

Việc chia khối có thể do **không có đủ n ô thí nghiệm đồng đều** nên phải chia thành **r** khối sao cho **a ô** trong mỗi khối tương đối đồng đều. Cũng có khi do thời gian

hạn chế mỗi ngày chỉ làm được **a** thí nghiệm chứ không thể làm tất cả $n = a \times r$ thí nghiệm, như vậy ở đây ngày là khối.

Cũng có khi chia khối **thẳng góc với một hướng biến động** có ảnh hưởng đến kết quả thí nghiệm thí dụ hướng gió, hướng chảy của nước ngầm, hướng nắng, hướng dốc, hướng thay đổi của độ phì của đất . . . nhằm **loại trừ ảnh hưởng của biến động đó** vì mỗi công thức có **mặt một lần ở một mức của biến động**.

Một cái lợi nữa là có thể chọn khối khác nhau về không gian và khác nhau về thời gian (nhưng không được khác nhau quá xa đến mức có sự thay đổi điều kiện thí nghiệm) nên kết luận rút ra có tính khái quát cao hơn là tập trung toàn bộ các thí nghiệm vào một nơi hay cùng một thời gian như thí nghiệm kiểu hoàn toàn ngẫu nhiên.

b1- Mô hình toán học

$$x_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \quad (i=1, a; j=1, r) \quad (2)$$

Khối được coi là yếu tố hạn chế và thường giả thiết là ngẫu nhiên, x_{ij} là kết quả của mức i ở khối j , μ là trung bình chung, α_i là ảnh hưởng của mức i của nhân tố, β_j là ảnh hưởng của khối j

Các sai số e_{ij} được giả thiết **độc lập, phân phối chuẩn, kỳ vọng 0, phương sai σ^2** .

Các tham số thỏa mãn điều kiện: $\sum_i \alpha_i = 0 \quad \sum_j \beta_j = 0$

b2- Các bước tính

Tính các tổng

Nhân tố có a mức bố trí thành r khối.

Tổng số ô thí nghiệm (hay số số liệu) $n = a \cdot r$

Tổng các số liệu của công thức i $TA_i = \sum_j x_{ij}$, các trung bình $\bar{x}_i$.

Tổng các số liệu trong khối j $TK_j = \sum_i x_{ij}$

Tổng tất cả các số liệu $ST = \sum_i \sum_j x_{ij}$

Số điều chỉnh $G = ST^2 / n$

Tính các tổng bình phương:

$$SSTO = \sum_{i=1}^a \sum_{j=1}^r (x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^r x_{ij}^2 - G \quad SSA = \sum_{i=1}^a \sum_{j=1}^r (\bar{x}_{i.} - \bar{x}_{..})^2 = \frac{\sum_{i=1}^a TA_i^2}{r} - G$$

$$SSK = \sum_{i=1}^a \sum_{j=1}^r (\bar{x}_{.j} - \bar{x}_{..})^2 = \frac{\sum_{j=1}^r TK_j^2}{a} - G$$

$$SSE = SSTO - SSA - SSK$$

Tính các bậc tự do:

Bậc tự do của SSTO **dfTO** = n - 1 = a . r - 1

Bậc tự do của SSA **dfA** = a - 1 Bậc tự do của SSK **dfK** = r - 1

Bậc tự do của SSE **dfE** = (a - 1)(r - 1) = dfTO - dfA - dfK

Tính các bình phương trung bình:

$$msK = SSK / dfK \quad msA = SSA / dfA \quad msE = SSE / dfE$$

Chia msK cho msE được F_{tnK} . Tìm giá trị tới hạn $F_{ltK} = F(\alpha, dfK, dfE)$

Chia msA cho msE được F_{tnA} . Tìm giá trị tới hạn $F_{ltA} = F(\alpha, dfA, dfE)$

Sai số thí nghiệm se bằng căn bậc hai của msE, bậc tự do dfE = (a-1)(r-1)

Tóm tắt các kết quả vào bảng phân tích phương sai

Phân tích phương sai khối ngẫu nhiên đầy đủ

Nguồn biến động	Tổng BP	BTĐ	Bình phương T _{bình}	F _{tn}	F _{lt}
Khối	SSK	dfK a - 1	msK SSK/dfK	msK/msE	F(α, dfK, dfE)
Nhân tố	SSA	dfA k - 1	msA SSA/dfA	msA/msE	F(α, dfA, dfE)
Sai số	SSE	dfE (a-1)(r-1)	MsE = se ² SSE/dfE		
Toàn bộ	SSTO	dfTO a(r - 1)			

b3- Kết luận

Dùng bảng phân tích phương sai trên để kiểm định giả thiết H_0 : “Không có sự khác nhau giữa các trung bình m_i ”, đối thiết H_1 : “Có sự khác nhau giữa các trung bình m_i ”.

Muốn kết luận phải so F_{tnA} với F_{ltA}

Nếu $F_{\text{tnA}} \leq F_{\text{ltA}}$ chấp nhận H_0 : “Không có sự khác nhau giữa các trung bình của các mức của nhân tố”.

Nếu $F_{\text{tnA}} > F_{\text{ltA}}$ chấp nhận H_1 : “Có sự khác nhau giữa các trung bình của các mức của nhân tố”.

Tính sai số thí nghiệm : $se^2 = msE = SSE / dfE$

bậc tự do $dfE = 9$

Trung bình toàn bộ $\bar{x}_{..} = ST / n$

Hệ số biến động $CV = se * 100 / \bar{x}_{..}$

Sai số của trung bình $\bar{x}_{..}$ $se_{\mu} = \sqrt{\frac{se^2}{n}}$

Sai số của trung bình của các công thức

$$se_A = \sqrt{\frac{se^2}{r}} = se_B = se_C = se_D$$

c- KIỂU BỐ TRÍ Ô VUÔNG LA TINH (Latin square)

Như trên đã thấy khi có một nguồn biến động ảnh hưởng đến khu vực thí nghiệm thì phải chia khối vuông góc với hướng biến động để đảm bảo độ đồng đều của các ô trong một khối.

Trường hợp bên ngoài có 2 hướng biến động trực giao nhau tác động đến khu vực thí nghiệm (thí nghiệm bố trí ở sườn núi vừa chịu ảnh hưởng của hướng gió, vừa chịu ảnh hưởng của độ cao hoặc kết quả thí nghiệm phụ thuộc vào các dụng cụ có chất lượng khác nhau và vào các ngày khác nhau trong tuần . . .), gọi hai tác động này là **hai yếu tố hạn chế**, chúng ta phải bố trí thí nghiệm kiểu ô vuông La tinh.

Gọi yếu tố thứ nhất là hàng, yếu tố thứ hai là cột (đây chỉ là cách nói giản đơn của bố trí thí nghiệm, thí dụ gọi các khối bố trí trên các độ cao khác nhau là hàng, khối bố trí vuông góc với chiều gió là cột và giả thiết hướng gió thổi trực giao với độ dốc. Gọi các dụng cụ là hàng, các ngày trong tuần là cột trong quá trình phân tích ở phòng thí nghiệm.

Trong kiểu bố trí ô vuông La tinh số hàng bằng số cột và bằng số mức a của nhân tố . Mỗi mức được bố trí một lần trên 1 hàng và một lần trên 1 cột. Mức 1 (A_1) được bố trí ở

các ô có ký hiệu A, Mức 2 (A_2) bố trí ở các ô có ký hiệu B, mức 3 (A_3) bố trí ở các ô có ký hiệu C, mức 4 (A_4) bố trí ở các ô ký hiệu D . . .

Căn cứ vào số mức a ta chọn **sơ đồ ô vuông La tinh $a \times a$** có sẵn trong các tài liệu thống kê, sau đó đổi chỗ ngẫu nhiên các hàng, rồi đổi chỗ ngẫu nhiên các cột, để cuối cùng được một sơ đồ ô vuông La tinh cụ thể để bố trí thí nghiệm.

Ô vuông La tinh đơn giản, dễ tính, loại trừ được ảnh hưởng của 2 hướng biến động, nhưng chỉ nên dùng khi đã nắm chắc đó là hai hướng biến động trực giao nhau và có thể chấp nhận mô hình cộng tính (xem phần dưới).

Nhược điểm của ô vuông La tinh là số bậc tự do còn lại cho sai số quá ít do đó thường chỉ dùng ô vuông La tinh tối thiểu là 4×4 và cũng không nên quá to vì phức tạp và khó đảm bảo các điều kiện của mô hình.

c1- Mô hình toán học

$$\text{Mô hình có dạng } X_{ijl} = \mu + \alpha_i + \beta_j + \gamma_l + e_{ijl} \quad (3)$$

$$(\alpha_i \quad i=1, a; \beta_j \quad j=1, a; \gamma_l \quad l=1, a)$$

Hàng, cột và công thức đều được coi là nhân tố cố định

x_{ijl} là kết quả của mức l bố trí ở hàng i, cột j, α_i là ảnh hưởng của hàng i, β_j là ảnh hưởng của cột j, γ_l là ảnh hưởng của công thức l.

Các sai số e_{ijl} được giả thiết **độc lập, phân phối chuẩn, có kỳ vọng 0, phương σ^2**

Các tham số thoả mãn điều kiện:

$$\sum_i \alpha_i = 0 \quad \sum_j \beta_j = 0 \quad \sum_k \gamma_k = 0$$

Gọi a là số mức của nhân tố. Tất cả có $n = a^2$ ô thí nghiệm.

c2- Các bước tính

Tính các tổng:

Tổng số ô thí nghiệm (hay số số liệu) $a = 4$; $n = a \times a$

Tổng các số liệu của hàng i $TH_i = \sum_{j=1}^a x_{ijl}$ (tổng các x_{ijl} trên hàng i)

Tổng các số liệu trong cột j $TC_j = \sum_{l=1}^a x_{ijl}$ (tổng các x_{ijl} trên cột j)

Tổng của các số liệu trong công thức 1

$$TA_l = \sum_{i=1}^a x_{ijl} \quad (\text{tổng các } x_{ijl} \text{ ứng với công thức 1})$$

Tổng tất cả các số liệu $ST = \sum_i \sum_j \sum_l x_{ijl}$ (tổng tất cả các x_{ijl} trong bảng)

Số điều chỉnh $G = ST^2 / n$

Tính các tổng bình phương:

$$SSTO = \sum_i \sum_j \sum_l x_{ijl}^2 - G \quad SSH = \sum_i TH_i^2 / a - G$$

$$SSC = \sum_j TC_j^2 / a - G \quad SSA = \sum_l TA_l^2 / a - G$$

$$SSE = SSTO - SSA - SSH - SSC$$

Tính các bậc tự do:

Bậc tự do của SSTO $dfTO = n - 1$ Bậc tự do của SSA $dfA = a - 1$

Bậc tự do của SSH $dfH = a - 1$ Bậc tự do của SSC $dfC = a - 1$

Bậc tự do của SSE $dfE = n - 3(a-1) = (a-1)(a-2)$

Tính các bình phương trung bình:

$msH = SSH/dfH$; $msA = SSA/dfA$; $msC = SSC/dfC$; $msE = SSE/dfE = se^2$

Giá trị F thực nghiệm $F_{tnA} = msA / msE$, giá trị tới hạn $F_{ltA} = F(\alpha, dfA, dfE)$

Bảng phân tích phương sai

Nguồn biến động	TổngBP	BTD	Tbình	Ftn	Flt
Hàng	SSH	a-1	msH		
Cột	SSC	a-1	msC		F_{ltA}
Nhân tố	SSA	a-1	msA	F_{tnA}	$F(\alpha, dfA, dfE)$
Sai số	SSE	dfE	msE		
Toàn bộ	SSTO	dfTO			

c3- Kết luận

So F_{tnA} với F_{ltA} ở mức ý nghĩa α với dfA và dfE bậc tự do

Nếu $F_{tnA} \leq F_{ltA}$ kết luận: ” Không có sự khác nhau giữa các trung bình của các mức của nhân tố ”.

Nếu $F_{tnA} > F_{ltA}$ kết luận: “Có sự khác nhau giữa các trung bình của các mức của nhân tố”.

Tính sai số thí nghiệm se:

$$se^2 = msE = SSE / dfE$$

bậc tự do $dfE = 6$

Trung bình toàn bộ $\bar{x}_{..} = ST / n$

Hệ số biến động $CV = se * 100 / \bar{x}$

Sai số của trung bình $\bar{x}$ $se_{\mu} = \sqrt{\frac{se^2}{n}}$

Sai số của trung bình của các công thức

$$se_A = \sqrt{\frac{se^2}{a}} = se_B = se_C = se_D$$

Trên đây là 3 kiểu bố trí thí nghiệm khi khảo sát một nhân tố.

Để khảo sát một nhân tố mà cần phải chia thành khối thì ngoài kiểu bố trí khối đầy đủ còn kiểu khối không đầy đủ (Randomized incomplete Block design) trong đó đáng chú ý là loại không đầy đủ cân đối (Balanced incomplete block design) tiếp theo là lưới ô vuông (Lattice design, rất hay dùng trong các nghiên cứu ban đầu về giống).

Để khảo sát một nhân tố khi có hai yếu tố hạn chế (hai nguồn biến động trực giao) ngoài ô vuông La tinh còn kiểu ô vuông La tinh thiếu (chữ nhật Youden)

Khi có 3 yếu tố hạn chế thì dùng ô vuông La tinh Hy Lạp (Graeco-Latin squares).

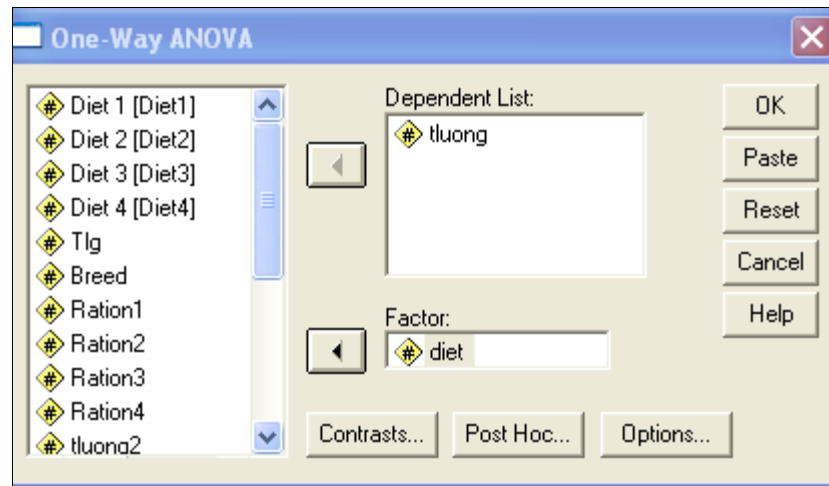
II- XỬ LÝ TRONG SPSS

a- Phân tích phương sai một nhân tố kiểu CRD

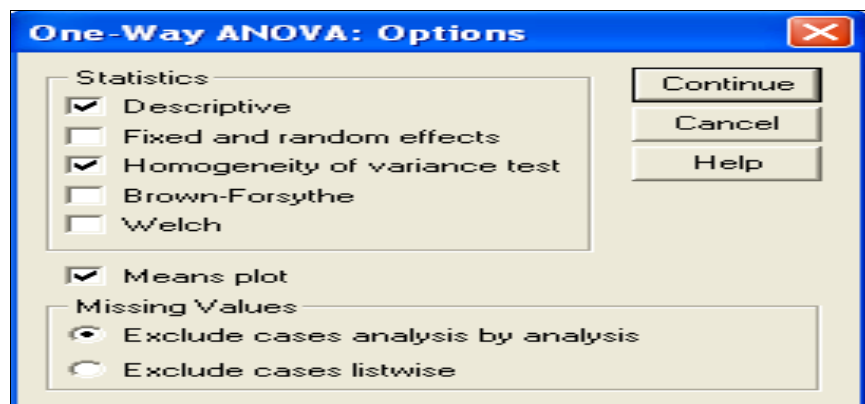
Mở tệp Baitap3 Vào Analyse Compare means One way anova

Chọn Tluong (trọng lượng) vào Dependent list (danh sách biến phụ thuộc),

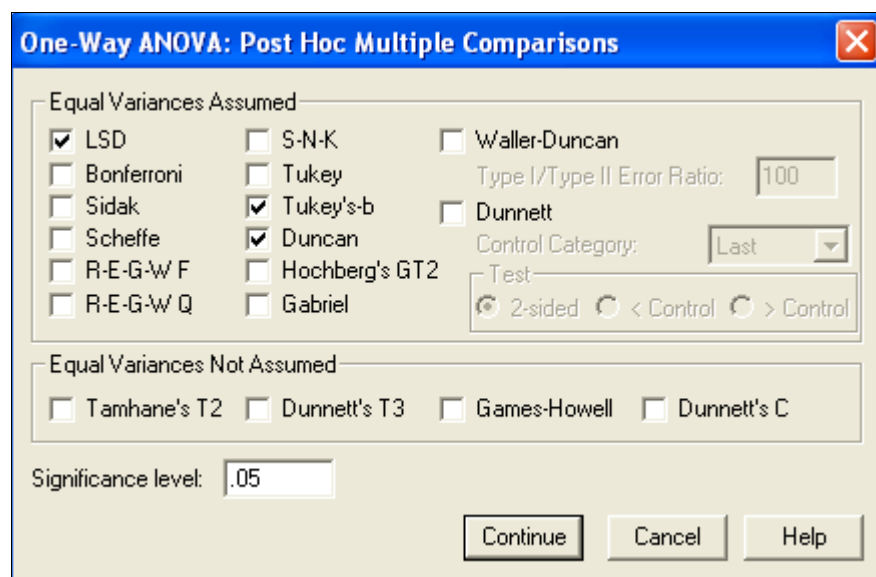
Factor (nhân tố) chọn diet (thức ăn)



Trong Options chọn Descriptive Homogeneity of variance test và Means plot



Trong Post hoc (kiểm định sau phân tích phương sai) chọn LSD, Tukey và Duncan



Kết quả

Descriptives

tluong

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	5	79.00	24.474	10.945	48.61	109.39	38	99
2	5	71.00	31.024	13.874	32.48	109.52	30	112
3	5	81.40	22.876	10.230	53.00	109.80	42	97
4	5	142.80	34.903	15.609	99.46	186.14	85	169
Total	20	93.55	39.523	8.838	75.05	112.05	30	169

Tluong

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.386	3	16	.765

Tluong

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	16466.950	3	5488.983	6.647	.004
Within Groups	13212.000	16	825.750		
Total	29678.950	19			

Multiple Comparisons

Dependent Variable: tluong

	(I) diet	(J) diet	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1	2	8.000	18.174	.666	-30.53	46.53
		3	-2.400	18.174	.897	-40.93	36.13
		4	-63.800*	18.174	.003	-102.33	-25.27
	2	1	-8.000	18.174	.666	-46.53	30.53
		3	-10.400	18.174	.575	-48.93	28.13
		4	-71.800*	18.174	.001	-110.33	-33.27
	3	1	2.400	18.174	.897	-36.13	40.93
		2	10.400	18.174	.575	-28.13	48.93
		4	-61.400*	18.174	.004	-99.93	-22.87
	4	1	63.800*	18.174	.003	25.27	102.33
		2	71.800*	18.174	.001	33.27	110.33
		3	61.400*	18.174	.004	22.87	99.93

*. The mean difference is significant at the .05 level.

Phân chia nhóm theo Tukey và Duncan

Homogeneous Subsets

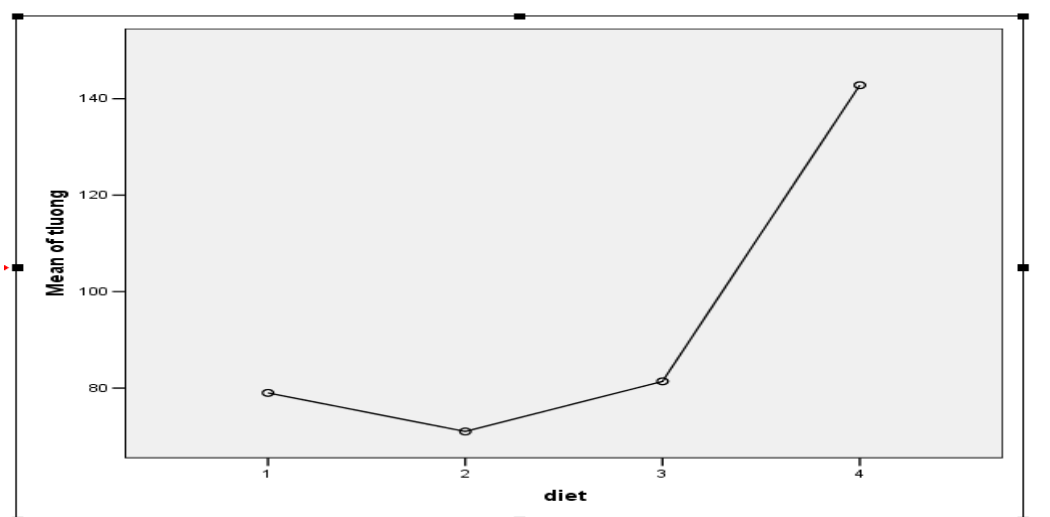
tuong				
diet		N	Subset for alpha = .05	
			1	2
Tukey B ^a	2	5	71.00	142.80
	1	5	79.00	
	3	5	81.40	
	4	5		
Duncan ^a	2	5	71.00	142.80
	1	5	79.00	
	3	5	81.40	
	4	5		
	Sig.			.596

Means for groups in homogeneous subsets are displayed.

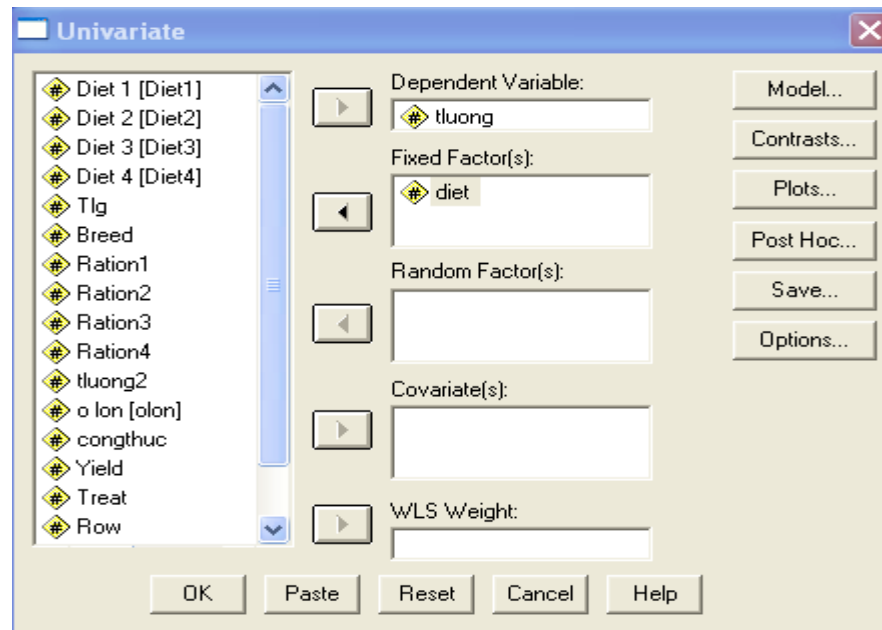
a. Uses Harmonic Mean Sample Size = 5.000.

Vẽ các trung bình

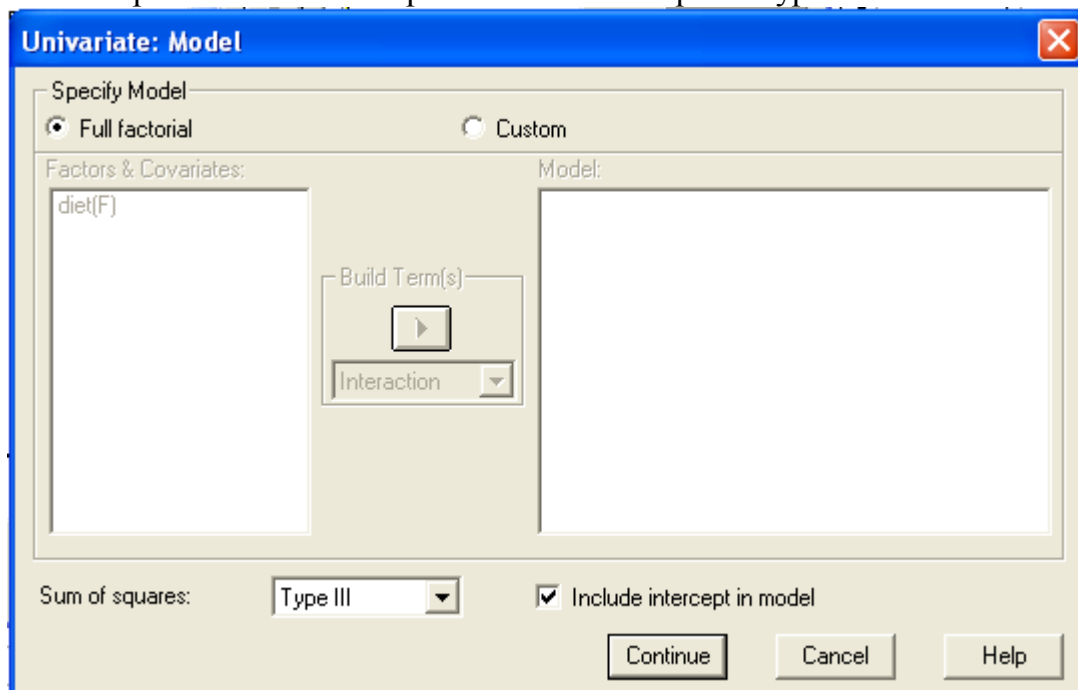
Means Plots



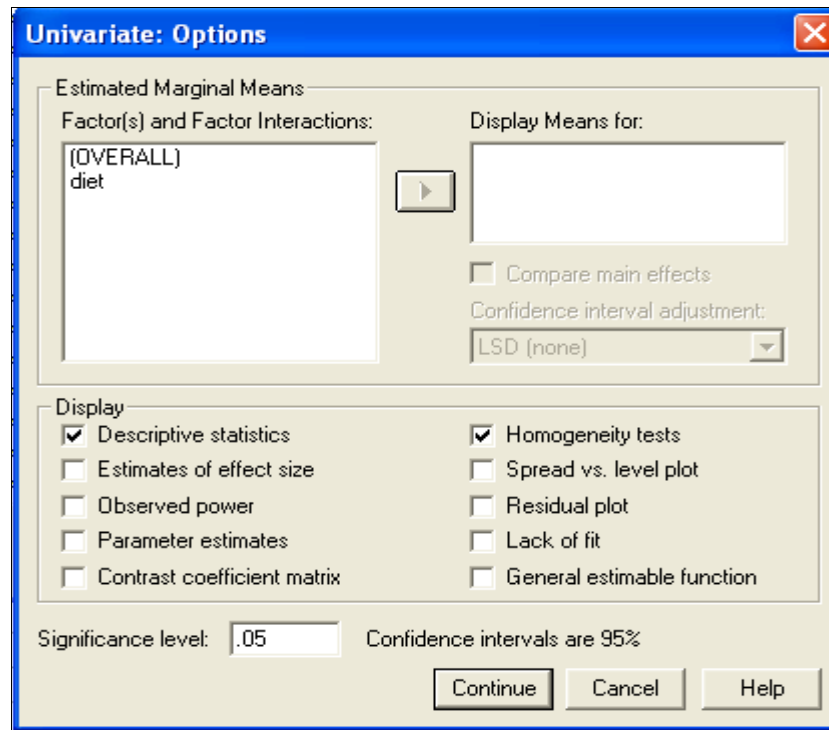
Có thể phân tích phương sai bằng General linear model
Vào General linear model Univariate sau đó làm giống như one way anova
Chọn Dependent variable Tuong, fixed factor diet



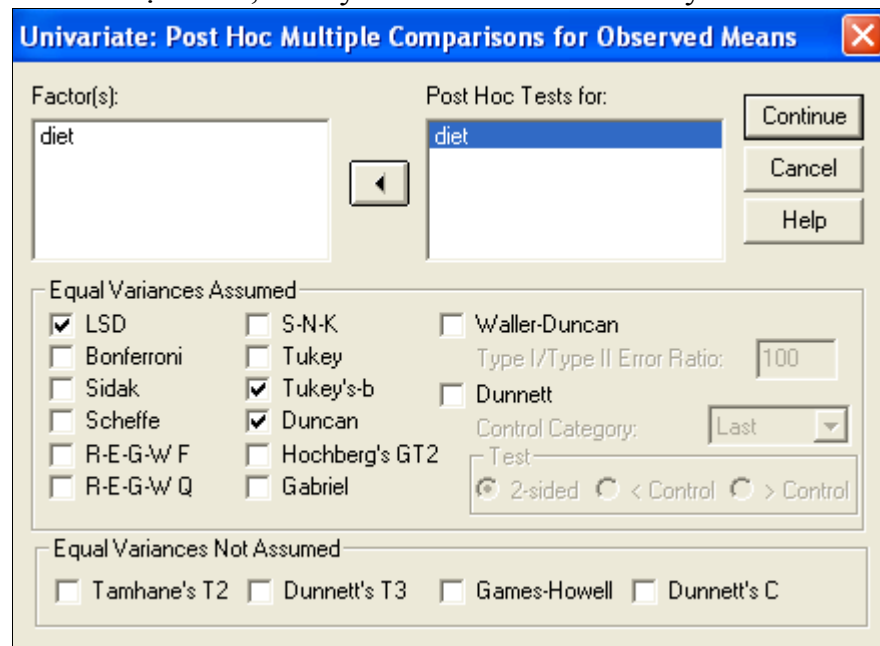
Trong Model chọn Full factorial, nếu chọn Custom thì phải đưa diet sang khung model, nháy ô include intercept in model hoặc không nháy đều được. SPSS chọn sẵn cách phân tích theo kiểu phân tích Sum of squares type III.



Trong options có thể chọn Descriptives Statistics và Homogeneity tests để kiểm tra mô hình



Trong Post hoc chọn LSD, Tukey và Duncan như one way anova

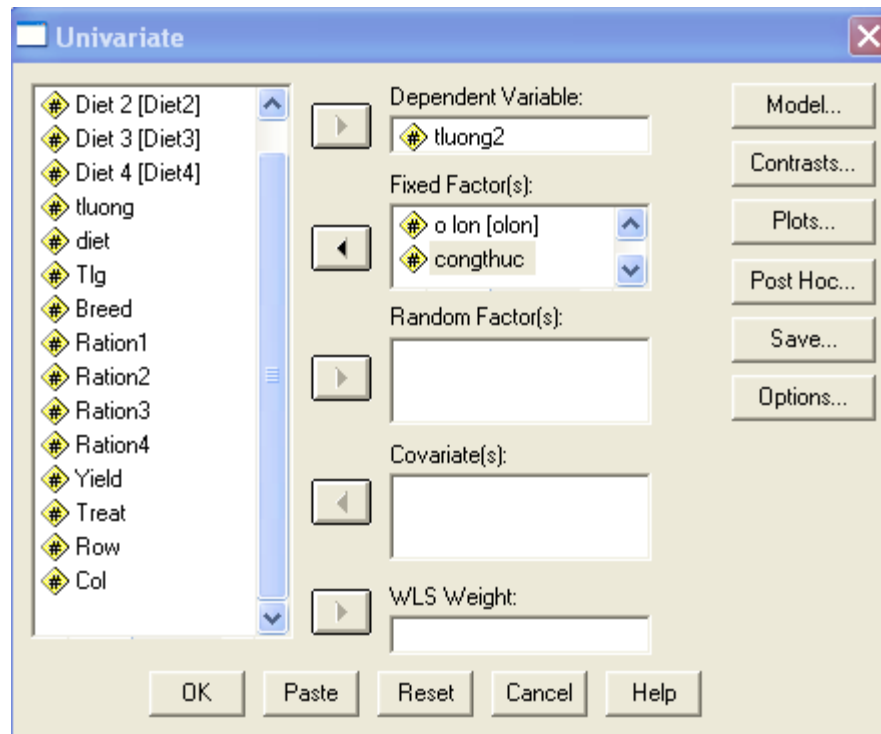


Kết quả giống như one way anova

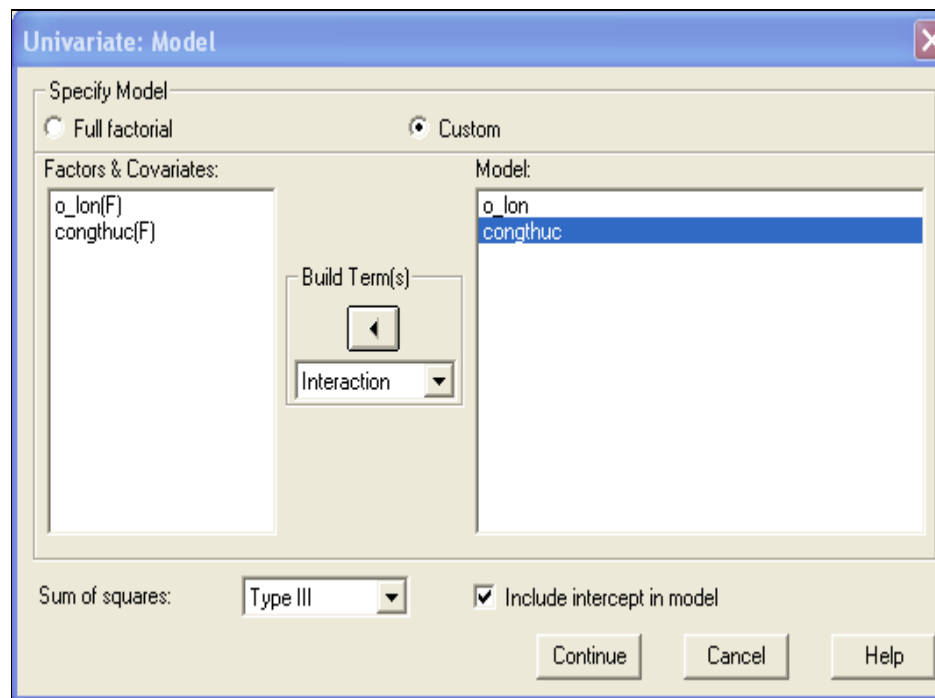
b- Phân tích phương sai một nhân tố kiểu khối ngẫu nhiên (RCBD)

Vào General linear model univariate

Chọn biến `tluong2` vào Dependent variable, chọn khối và công thức vào Fixed factor



Trong Model chọn Custom rồi đưa khai và congthuc sang khung model, nháy ô include intercept in model hoặc không nháy đều được. SPSS chọn sẵn cách phân tích theo kiểu phân tích Sum of squares type III. Trong post hoc chọn Post hoc tests for Congthuc sau đó chọn LSD, Tukey, Duncan.



Univariate: Post Hoc Multiple Comparisons for Observed Means

Factor(s):
 a_lon
 congthuc

Post Hoc Tests for:
 congthuc

Continue
 Cancel
 Help

Equal Variances Assumed

☒ LSD ☐ S-N-K ☐ Waller-Duncan
☐ Bonferroni ☒ Tukey Type I/Type II Error Ratio: 100
☐ Sidak ☐ Tukey's-b ☐ Dunnett
☐ Scheffe ☒ Duncan Control Category: Last
☐ R-E-G-W F ☐ Hochberg's GT2 Test:
☒ 2-sided ☐ < Control ☐ > Control
☐ R-E-G-W Q ☐ Gabriel

Equal Variances Not Assumed

☐ Tamhane's T2 ☐ Dunnett's T3 ☐ Games-Howell ☐ Dunnett's C

Kết quả:

Dependent Variable: tluong2 Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	8.249(a)	7	1.178	22.198	.000
Intercept	711.625	1	711.625	13405.799	.000
congthuc	1.846	3	.615	11.589	.001
olon	6.403	4	1.601	30.155	.000
Error	.637	12	.053		
Total	720.510	20			
Corrected Total	8.886	19			

a R Squared = .928 (Adjusted R Squared = .886)

Homogeneous Subsets					
tluong2					
	Cong thuc	N	Subset		
			1	2	3
Tukey B ^{a,b}	4	5	5.660		
	2	5	5.720		
	3	5	6.060	6.060	
	1	5		6.420	
Duncan ^{a,b}	4	5	5.660		
	2	5	5.720		
	3	5		6.060	
	1	5			6.420
	Sig.		.688	1.000	1.000

Means for groups in homogeneous subsets are displayed.
 Based on Type III Sum of Squares
 The error term is Mean Square(Error) = .053.
 a. Uses Harmonic Mean Sample Size = 5.000.
 b. Alpha = .05.

Nếu sau khi vào General model Univariate khai báo Congthuc vào Fixed factor còn khoi vào Random factor thì có kết quả tương tự.

Tests of Between-Subjects Effects

Dependent Variable: tuong2

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	711.625	1	711.625	444.557	.000
	Error	6.403	4	1.601(a)		
congthuc	Hypothesis	1.846	3	.615	11.589	.001
	Error	.637	12	.053(b)		
o_lon	Hypothesis	6.403	4	1.601	30.155	.000
	Error	.637	12	.053(b)		

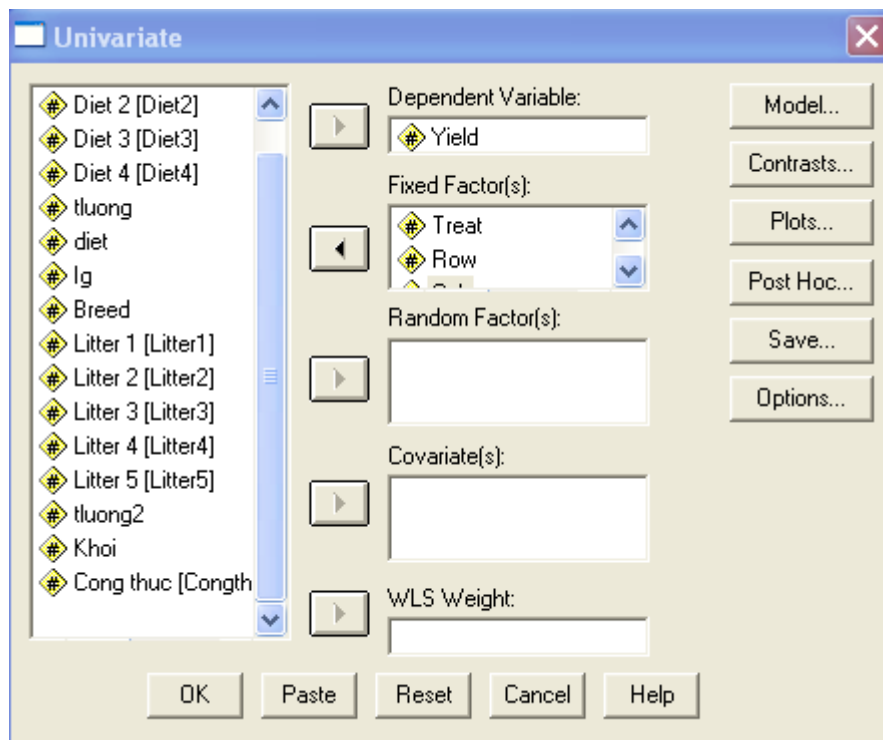
a MS(o_lon)

b MS(Error)

c- Phân tích phương sai một nhân tố kiểu ô vuông La tinh (LS)

Analyse General linear model Univariate

Đưa Yield vào Dependent variable, đưa trset, row, column vào Fixed factor



Trong Model chọn Custom rồi đưa Row, Column, treat sang model

Trong Post hoc chọn Post hoc test for Treat.

Univariate: Model

Specify Model

☐ Full factorial ☒ Custom

Factors & Covariates:

Row(F)
Col(F)
Treat(F)

Build Term(s)

Interaction

Model:

Row
Col
Treat

Sum of squares: Type III

☐ Include intercept in model

Continue Cancel Help

Univariate: Post Hoc Multiple Comparisons for Observed Means

Factor(s):

Row
Col
Treat

Post Hoc Tests for:

Treat

Continue
Cancel
Help

Equal Variances Assumed

☒ LSD ☐ S-N-K ☐ Waller-Duncan
☐ Bonferroni ☐ Tukey
☐ Sidak ☒ Tukey's-b
☐ Scheffe ☒ Duncan
☐ R-E-G-W F ☐ Hochberg's GT2
☐ R-E-G-W Q ☐ Gabriel

Type I/Type II Error Ratio: 100

Dunnett

Control Category: Last

Test

☒ 2-sided ☐ < Control ☐ > Control

Equal Variances Not Assumed

☐ Tamhane's T2 ☐ Dunnett's T3 ☐ Games-Howell ☐ Dunnett's C

Dependent Variable: Yield **Tests of Between-Subjects Effects**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	113372.675(a)	13	8720.975	517.936	.000
Row	2326.386	4	581.597	34.541	.000
Col	901.374	4	225.344	13.383	.000
Treat	1284.510	4	321.128	19.072	.000
Error	202.055	12	16.838		
Total	113574.730	25			

a R Squared = .998 (Adjusted R Squared = .996)

Homogeneous Subsets

Yield					
	Treat	N	Subset		
			1	2	3
Tukey B ^{a,b}	1	5	52.920		
	2	5		65.680	
	3	5		67.760	67.760
	4	5		69.000	69.000
	5	5			74.580
Duncan ^{a,b}	1	5	52.920		
	2	5		65.680	
	3	5		67.760	
	4	5		69.000	69.000
	5	5			74.580
Sig.			1.000	.247	.053

Means for groups in homogeneous subsets are displayed.
 Based on Type III Sum of Squares
 The error term is Mean Square(Error) = 16.838.
 a. Uses Harmonic Mean Sample Size = 5.000.
 b. Alpha = .05.