

Ανάλυση Διακύμανσης Μονής Κατεύθυνσης

Όπως ακριβώς συνέβη και στο κριτήριο t , τα δεδομένα μας θα πρέπει να έχουν ομαδοποιηθεί χρησιμοποιώντας μια αντίστοιχη μεταβλητή ομαδοποίησης (group variable). Τα δεδομένα που θα χρησιμοποιήσουμε στο πρώτο μέρος αυτής της ενότητας προέρχονται από τον Πίνακα 11.2 στο βιβλίο του Howell. Στη συγκεκριμένη έρευνα είχαμε 5 ομάδες από 10 παρατηρήσεις – συνεπώς 50 συνολικά μετρήσεις. Η μεταβλητή που θα χρησιμοποιήσουμε για την ομαδοποίηση θα έχει κωδικούς από το 1 ως το 5. Ρίξτε μια ματιά στον ακόλουθο πίνακα για να καταλάβετε πώς θα περαστούν τα δεδομένα.

Groups	Scores
1	9
1	8
1	6
...	...
1	7
2	7
2	9
2	6
...	...
...	...
...	...
5	10
5	19
...	...
5	11

Με την κωδικοποίηση αυτή γίνεται σαφής η προέλευση καθεμιάς από τις μετρήσεις. Για να πραγματοποιηθεί η ανάλυση,

- Επιλέγουμε [**Statistics => Compare Means => One-Way ANOVA...**] για να ανοίξει το πλαίσιο διαλόγου.
- Επιλέγουμε και μετακινούμε το "Scores" (ή αλλιώς, την εξαρτημένη μας μεταβλητή) στο **Dependent list:**
- Επιλέγουμε και μετακινούμε το "Groups" στη λίστα **Factor:**
- Πατάμε [**OK**]

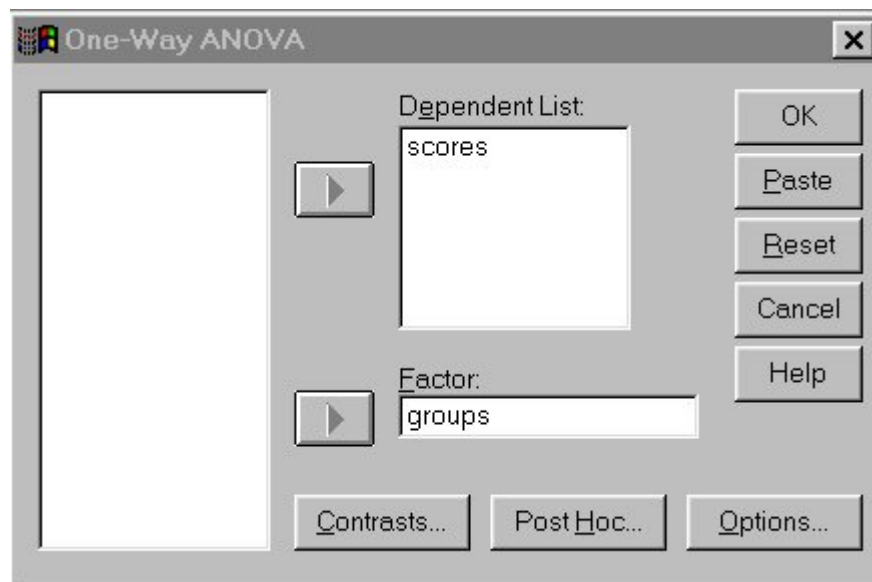
Τα παραπάνω είναι η περιγραφή του ερευνητικού μοντέλου για αυτή την ανάλυση διακύμανσης μονής κατεύθυνσης. Τα αποτελέσματα θα έχουν την ακόλουθη μορφή:

Σχήμα 23. Το αποτέλεσμα από την εφαρμογή του κριτηρίου της ανάλυσης διακύμανσης μόνης κατεύθυνσης όπως εμφανίζεται στο παράθυρο εξόδου

		Sum of Squares	df	Mean Square	F	Sig.
SCORES	Between Groups	351.520	4	87.880	9.085	.000
	Within Groups	435.300	45	9.673		
	Total	786.820	49			

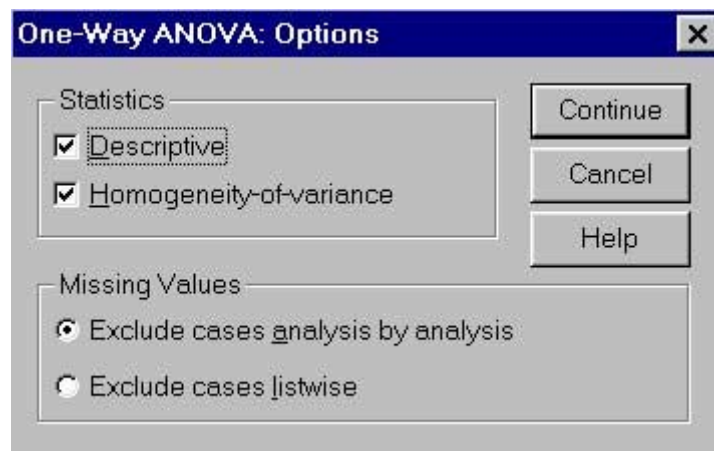
Η ανάλυση που πραγματοποιήθηκε παρέχει σχετικά λίγες πληροφορίες για τα δεδομένα. Αν κοιτάξετε προσεκτικά το πλαίσιο διαλόγου για το στατιστικό αυτό κριτήριο, θα δείτε 3 επιπλέον κουμπιά στο κάτω δεξιά μέρος - **[Contrasts...]**, **[Post Hoc..]**, και **[Options...]**.

Σχήμα 24. Το πλαίσιο διαλόγου για το κριτήριο της ανάλυσης διακύμανσης μόνης κατεύθυνσης



Επιλέγοντας **[Options...]** θα ανοίξει το ακόλουθο πλαίσιο διαλόγου:

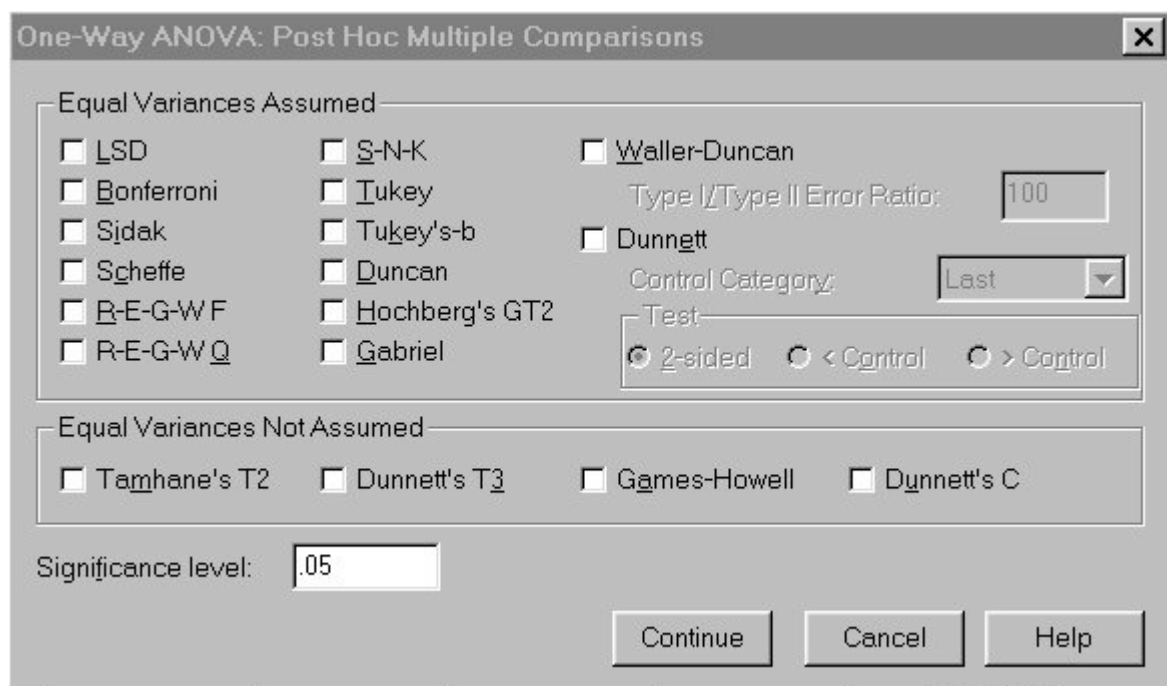
Σχήμα 25. Το πλαίσιο διαλόγου με τις επιλογές (options) για το κριτήριο της ανάλυσης διακύμανσης μονής κατεύθυνσης



Αν επιλέξουμε το **Descriptive**, θα πάρουμε τους περιγραφικούς στατιστικούς δείκτες για καθεμιά από τις συνθήκες της ανεξάρτητης μεταβλητής. Επιλέγοντας το **Homogeneity-of-variance** θα πραγματοποιηθεί ένα τεστ Levene's στα δεδομένα μας.

Επιλέγοντας [**Post Hoc**] θα ανοίξει το ακόλουθο παράθυρο διαλόγου:

Σχήμα 25. Το πλαίσιο διαλόγου με τα κριτήρια πολλαπλών συγκρίσεων (post-hoc) για το κριτήριο της ανάλυσης διακύμανσης μονής κατεύθυνσης



Μπορούμε εδώ να επιλέξουμε να πραγματοποιηθούν ένα ή περισσότερα κριτήρια πολλαπλών συγκρίσεων (post hoc). Τα αποτελέσματα θα βρίσκονται στο παράθυρο εξόδου. Για παράδειγμα, η επιλογή του κριτηρίου **R-E-G-W F** στα δεδομένα αυτά θα έχει ως αποτέλεσμα τον ακόλουθο πίνακα:

Σχήμα 26. Το αποτέλεσμα από την εφαρμογή του κριτηρίου πολλαπλών συγκρίσεων R-E-G-W F όπως εμφανίζεται στο παράθυρο εξόδου

SCORES

Ryan-Einot-Gabriel-Welsch F

GROUPS	N	Subset for alpha = .05	
		1	2
2.00	10	6.9000	
1.00	10	7.0000	
3.00	10		11.0000
5.00	10		12.0000
4.00	10		13.4000
Sig.		.999	.358

Means for groups in homogeneous subsets are displayed.

Παραγοντική Ανάλυση Διακύμανσης

Προκειμένου να πραγματοποιήσουμε μια παραγοντική ανάλυση διακύμανσης θα πρέπει να ακολουθήσουμε την ίδια λογική με αυτή στην ανάλυση διακύμανσης μόνης κατεύθυνσης. Στον πίνακα που ακολουθεί παρουσιάζονται τα δεδομένα μιας παραγοντικής ανάλυσης διακύμανσης 2×5 . Ο πρώτος παράγοντας (ηλικία – AGE) έχει δύο επίπεδα, και ο δεύτερος (συνθήκη – CONDITION) έχει πέντε επίπεδα. Έτσι, και πάλι καθεμία παρατήρηση θα πρέπει να κωδικοποιηθεί ανάλογα:

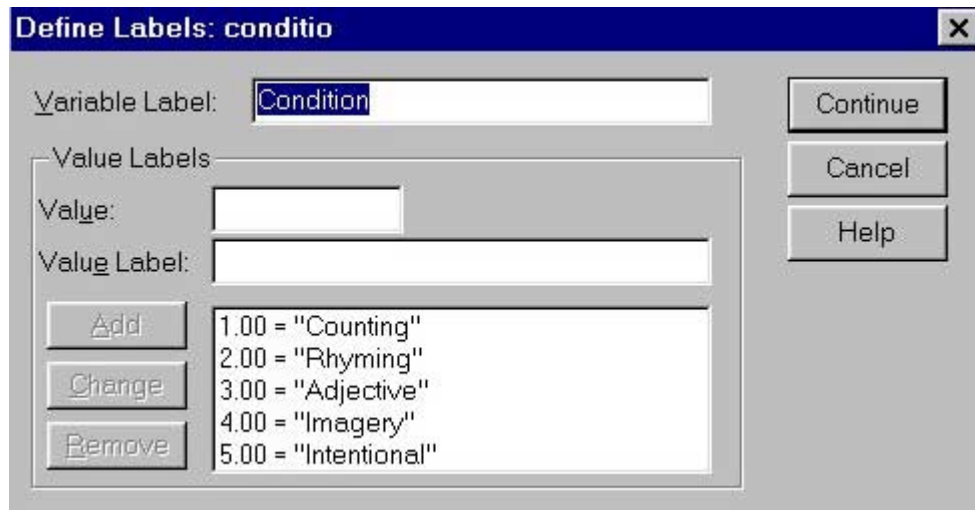
AGE	CONDITION
Old = 1	Counting = 1
Young = 2	Rhyming = 2
	Adjective = 3
	Imagery = 4
	Intentional = 5

Για κάθε συνδυασμό ηλικίας και συνθήκης πήραμε 10 μετρήσεις. Επομένως, 2×5 συνθήκες επί 10 μετρήσεις έχουν ως αποτέλεσμα 100 τιμές, οι οποίες κωδικοποιούνται ως εξής:

AGE	CONDITIO	Scores
1	1	9
1	1	8
1	1	6
1
1	1	7
1	2	7
1	2	9
1	2	6
1
1
1
1	5	10
1	5	19
1
1	5	11
2	1	8
2	1	6
2	1	4
2
2	1	7
2	2	10
2	2	7
2	2	8
2
2
2
2	5	21
2	5	19
2
2	5	21

Εξετάστε τον πίνακα με προσοχή για να κατανοήσετε πώς πραγματοποιήθηκε η κωδικοποίηση των τιμών. Σημείωση: μπορούμε να βελτιώσουμε την ποιότητα του αποτελέσματος που θα πάρουμε στο παράθυρο εξόδου χρησιμοποιώντας το **Value Labels** για τους δύο παράγοντες:

Σχήμα 27. Το πλαίσιο διαλόγου για τον ορισμό των τίτλων των τιμών μιας μεταβλητής



Προκειμένου να υπολογίσουμε το κριτήριο – η απλή μέθοδος,

- Επιλέγουμε [**Statistics => General Linear Model => Simple Factorial...**]
- Επιλέγουμε και μετακινούμε το "Scores" στο πλαίσιο **Dependent**:
- Επιλέγουμε και μετακινούμε το "Age" στο πλαίσιο **Factor(s)**:
- Πατάμε το [**Define Range...**] για να ορίσουμε το εύρος της κωδικοποίησης στον παράγοντα της ηλικίας. Σημειώστε ότι το **1** χρησιμοποιήθηκε για τους ηλικιωμένους (Old) και το **2** για τους νέους (Young). Έτσι η μικρότερη τιμή **Minimum**: είναι <1>, και η μέγιστη **Maximum**: είναι **2**. Πατάμε το [**Continue**].
- Επιλέγουμε και μετακινούμε το "Conditio" στο πλαίσιο **Dependent**:
- Πατάμε στο [**Define Range...**] για να ορίσουμε το εύρος της κωδικοποίησης στον παράγοντα της συνθήκης. Στην περίπτωση αυτή η τιμή **Minimum**: είναι **1** και η **Maximum**: είναι **5**.
- Πατήστε [**OK**] για να υπολογιστεί το αποτέλεσμα της ανάλυσης.

Όπως βλέπετε, η επιλογή **Means and count** είχε ως αποτέλεσμα έναν πολύ καλό συνοπτικό πίνακα με όλες τις μεταβλητές και τις τιμές τους (**Variable Labels** και **Value Labels**) που είχαν οριστεί κατά την εισαγωγή των δεδομένων μας.

Σχήμα 28. Το αποτέλεσμα από την εφαρμογή του κριτηρίου της ανάλυσης διακύμανσης διπλής κατεύθυνσης όπως εμφανίζεται στο παράθυρο εξόδου

Cell Means^b

AGE	Condition	SCORES	
		Mean	N
Old	Counting	7.0000	10
	Rhyming	6.9000	10
	Adjective	11.0000	10
	Imagery	13.4000	10
	Intentional	12.0000	10
	Total	10.0600	50
Young	Counting	6.5000	10
	Rhyming	7.6000	10
	Adjective	14.8000	10
	Imagery	17.6000	10
	Intentional	19.3000	10
	Total	13.1600	50
Total	Counting	6.7500	20
	Rhyming	7.2500	20
	Adjective	12.9000	20
	Imagery	15.5000	20
	Intentional	15.6500	20
	Total	11.6100 ^a	100

a. Grand Mean

b. SCORES by AGE, Condition

ANOVA^a

			Hierarchical Method				
			Sum of Squares	df	Mean Square	F	Sig.
SCORES	Main Effects	(Combined)	1755.190	5	351.038	43.740	.000
		AGE	240.250	1	240.250	29.936	.000
		Condition	1514.940	4	378.735	47.191	.000
	2-Way Interactions	AGE * Condition	190.300	4	47.575	5.928	.000
	Model		1945.490	9	216.166	26.935	.000
	Residual		722.300	90	8.026		
	Total		2667.790	99	26.947		

a. SCORES by AGE, Condition

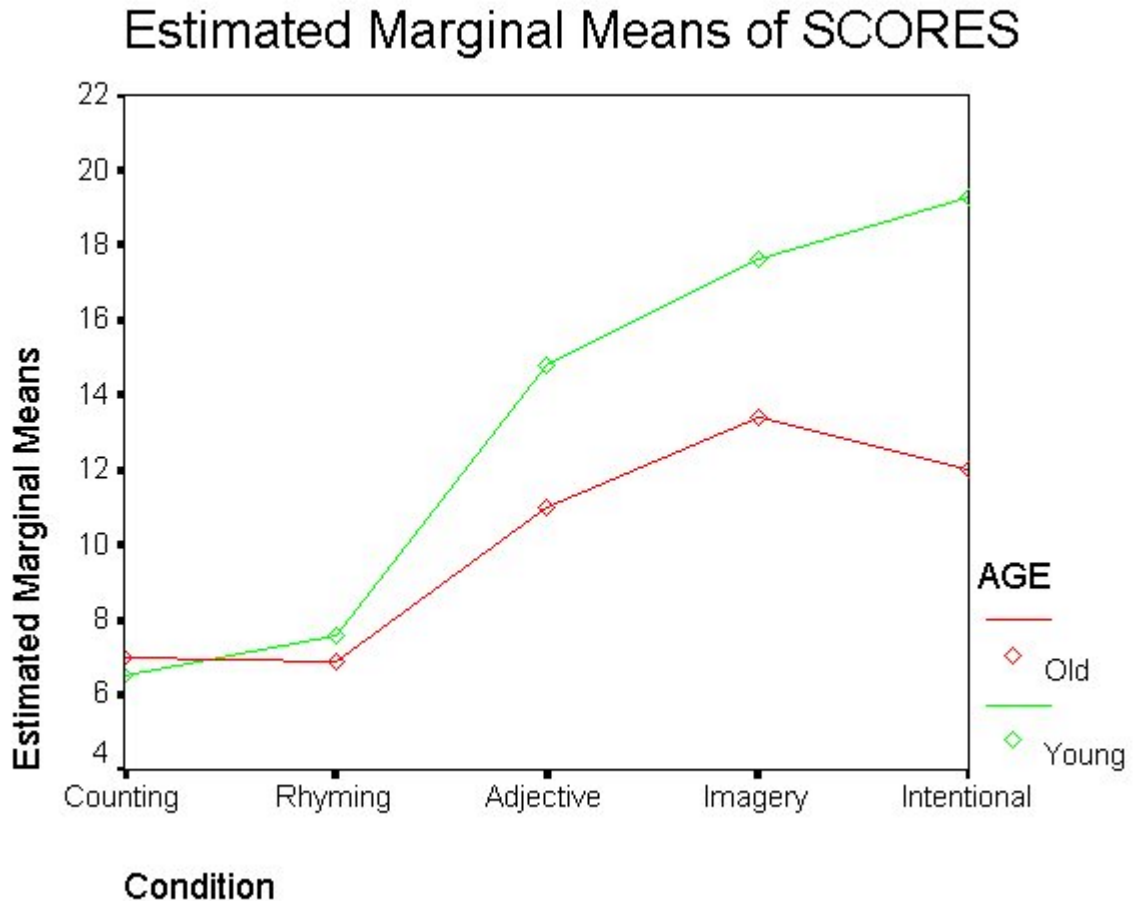
Το αποτέλεσμα περιλαμβάνει επίσης και τον ανακεφαλαιωτικό πίνακα της διακύμανσης όπου οι παράγοντες παρουσιάζονται με τα ονόματα που τους δώσαμε στο **Variable Labels**.

Όπως αναφέρθηκε προηγουμένως, η ανάλυση αυτή είναι η απλούστερη μορφή παραγοντικής ανάλυσης διακύμανσης. Αν χρησιμοποιήσουμε [**Statistics => General Linear Model => GLM - General Factorial...**], τότε έχουμε περισσότερες επιλογές. Ο ορισμός των εξαρτημένων και

ανεξάρτητων μεταβλητών (**Dependent** και **Independent**) γίνεται με τον ίδιο τρόπο που περιγράψαμε στην απλή παραγοντική ανάλυση (**Simple Factorial**). Επιπλέον, οι επιλογές περιλαμβάνουν:

- Επιλέγοντας [**Model...**], μπορούμε να ζητήσουμε το μοντέλο **Custom**. Όπως θα διαπιστώσετε, το μοντέλο **Fully Factorial** είναι προεπιλεγμένο, αλλά με την επιλογή **Custom** μπορούμε να προσδιορίσουμε με ακρίβεια τις επιδράσεις που μας ενδιαφέρουν.
- Η επιλογή **Contrasts** μας επιτρέπει να ελέγξουμε τις διαφορές μεταξύ των επιπέδων ενός παράγοντα.
- Διάφορες γραφικές παραστάσεις μπορούμε να ζητήσουμε μέσω της επιλογής [**Plots...**]. Για παράδειγμα, προκειμένου να κατασκευάσουμε το γράφημα για το συνδυασμό των μέσων όρων των δύο παραγόντων ("conditio*age") μπορούμε να επιλέξουμε τον παράγοντα "Conditio" στον οριζόντιο άξονα (**Horizontal Axis:**) και τον παράγοντα "Age" στο **Separate Lines:**

Σχήμα 29. Η γραφική παράσταση των μέσων όρων των δύο ανεξάρτητων μεταβλητών του παραδείγματός μας όπως εμφανίζεται στο παράθυρο εξόδου



- Τα τυπικά κριτήρια πολλαπλών συγκρίσεων (post-hoc) για καθέναν από τους παράγοντες υπολογίζονται κάνοντας τις κατάλληλες επιλογές στο **[Post Hoc...]**. Αυτό που πρέπει να κάνουμε είναι να επιλέξουμε τους παράγοντες που θα αναλυθούν και τα κατάλληλα κριτήρια post-hoc.
- Το παράθυρο διαλόγου **[Options...]** προσφέρει ορισμένα διαγνωστικά και περιγραφικά χαρακτηριστικά. Από εκεί υπολογίζουμε, ανάμεσα στα άλλα, τους περιγραφικούς στατιστικούς δείκτες, τις εκτιμήσεις του μεγέθους της επίδρασης, και τα τεστ για την ομοιογένεια της διακύμανσης. Ένα παράδειγμα ανάλυσης κατά την οποία ζητήθηκαν ορισμένοι από τους δείκτες αυτούς είναι το ακόλουθο:

Σχήμα 30. Το αποτέλεσμα από την εφαρμογή του κριτηρίου της ανάλυσης διακύμανσης διπλής κατεύθυνσης με τον υπολογισμό επιπλέον δεικτών όπως εμφανίζεται στο παράθυρο εξόδου

Tests of Between-Subjects Effects

Dependent Variable: SCORES

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared	Noncent. Parameter	Observed Power ^a
Corrected Model	1945.490 ^b	9	216.166	26.935	.000	.729	242.412	1.000
Intercept	13479.21	1	13479.21	1679.536	.000	.949	1679.536	1.000
AGE	240.250	1	240.250	29.936	.000	.250	29.936	1.000
CONDITIO	1514.940	4	378.735	47.191	.000	.677	188.765	1.000
AGE * CONDITIO	190.300	4	47.575	5.928	.000	.209	23.712	.980
Error	722.300	90	8.026					
Total	16147.00	100						
Corrected Total	2667.790	99						

a. Computed using alpha = .05

b. R Squared = .729 (Adjusted R Squared = .702)

Η χρήση της διαδικασίας **GLM - General Factorial** προσφέρει πολύ περισσότερες εναλλακτικές από το **Simple Factorial**. Ανάλογα με τον ερευνητικό σχεδιασμό που έχετε ακολουθήσει και τις υποθέσεις σας, η πρώτη από τις δύο διαδικασίες μπορεί να φανεί πολύ πιο χρήσιμη στην ερμηνεία των δεδομένων σας.

Η παραγοντική ανάλυση διακύμανσης με περισσότερες ανεξάρτητες μεταβλητές γίνεται με τον ίδιο ακριβώς τρόπο που γίνεται και η ανάλυση διακύμανσης διπλής κατεύθυνσης. Ωστόσο, θα πρέπει να είμαστε πολύ προσεκτικοί με την κωδικοποίηση των παραγόντων.