Π Α Ν Ε Π Ι Σ Τ Η Μ Ι Ο   Π Ε Ι Ρ Α Ι Ω Σ

Σχολή Χρηματοοικονομικής και Στατιστικής

Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ **ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

ΑΠΟΔΟΤΙΚΟΙ ΣΧΕΔΙΑΣΜΟΙ ΚΡΗΣΑΡΙΣΜΑΤΟΣ ΤΡΙΩΝ ΕΠΙΠΕΔΩΝ ΓΙΑ ΤΗ ΜΕΛΕΤΗ ΠΟΣΟΤΙΚΩΝ ΠΑΡΑΓΟΝΤΩΝ

Τηγανίτας Σωτήριος

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στην *Εφαρμοσμένη Στατιστική*

Πειραιάς
Νοέμβριος 2020

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίσθηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ’ αριθμ. …….. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Ευαγγελάρας Χαράλαμπος (Επιβλέπων)
- Τζαβελάς Γεώργιος
- Πολίτης Κωνσταντίνος

Η έγκριση της Διπλωματική Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνωμών του συγγραφέα.

UNIVERSITY OF PIRAEUS

School of Finance and Statistics

Department of Statistics and Insurance Science

POSTGRADUATE PROGRAM IN
APPLIED STATISTICS

THREE-LEVEL DESIGNS FOR SCREENING QUANTITATIVE FACTORS

By

Tiganitas Sotirios

MSc Dissertation

submitted to the Department of Statistics and Insurance Science of the University of Piraeus in partial fulfilment of the requirements for the degree of Master of Science in Applied Statistics

Piraeus, Greece
November 2020

Στην οικογένεια μου

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τους καθηγητές του μεταπτυχιακού τμήματος «Εφαρμοσμένη Στατιστική» για τις γνώσεις που προσέφεραν αυτά τα δύο χρόνια των σπουδών μου και κυρίως τον κ. Χαράλαμπο Ευαγγελάρα που καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας υπήρξε σημαντικό αρωγός αυτής της προσπάθειας. Ακόμη, θα ήθελα να του εκφράσω τις ειλικρινείς ευχαριστίες μου για τη συνεχή καθοδήγησή που μου προσέφερε, την άριστη συνεργασία μας καθώς και για τις επιπλέον γνώσεις που μου προσέφερε στο κομμάτι αυτό.

Περίληψη

Σε πολλές πειραματικές καταστάσεις στην Στατιστική αυτό που μελετάμε είναι να δούμε αν ένας ή περισσότεροι παράγοντες επιδρούν πάνω σε ένα αντικείμενο που εξετάζουμε. Η πειραματική διαδικασία έχει ουσιώδη ρόλο στην Στατιστική. Σε αυτήν τη διπλωματική θα ασχοληθούμε με κάποιους πειραματικούς σχεδιασμούς που θα μας βοηθήσουν να απαντήσουμε στο ερώτημα “ποιες παραγοντικές επιδράσεις είναι σημαντικές στο πείραμα”.

Ο πιο απλός σχεδιασμός που συναντάμε είναι ο 2^k παραγοντικός σχεδιασμός που αφορά παράγοντες που έχουν από δύο επίπεδα ο καθένας. Ένα πλεονέκτημα τους είναι ότι σε σχέση με παράγοντες που έχουν περισσότερα επίπεδα χρειάζονται λιγότερες εκτελέσεις αλλά όσο αυξάνονται οι παράγοντες τόσο αυξάνονται εκθετικά και οι πειραματικές εκτελέσεις που χρειαζόμαστε για να πραγματοποιηθεί το πείραμα. Γι’ αυτό τον λόγο, χρησιμοποιούμε τους αντίστοιχους κλασματικούς παραγοντικούς σχεδιασμούς, όπου μειώνουμε τις πειραματικές εκτελέσεις, καθώς ασχολούμαστε με μικρότερο αριθμό παραγόντων.

Ένα θέμα που δημιουργείται στους 2^k παραγοντικούς σχεδιασμούς είναι ότι υπάρχει η πιθανότητα να υπάρξει κάποια καμπυλότητα στην επιφάνεια απόκρισης. Κάτι τέτοιο σαφώς δεν μπορεί να ελεγχθεί με δύο επίπεδα στον κάθε παράγοντα . Μια λύση είναι η εισαγωγή κεντρικών σημείων στο σχεδιασμό μας. Μια άλλη λύση είναι η ξεκάθαρη χρήση 3 επίπεδων στους παράγοντες. Σε αυτή την περίπτωση μπορούμε να εξετάσουμε την ύπαρξη κυρτότητας. Το πρόβλημα όμως που δημιουργείται είναι ότι δεν μπορούμε να εξετάσουμε αρκετούς παράγοντες καθώς χρειαζόμαστε αρκετές εκτελέσεις. Έτσι όπως και στους 2^k παραγοντικούς σχεδιασμούς έτσι και στους 3^k παραγοντικούς σχεδιασμούς ισχύουν οι κλασματικοί παραγοντικοί σχεδιασμοί.

Μια οικονομικά λύση, τόσο από άποψη χρόνου όσο και στο θέμα του κόστους είναι οι Definitive Screening Designs (DSD), που είναι και το κυρίως θέμα αυτής της διπλωματικής, καθώς χρειαζόμαστε λιγότερες εκτελέσεις. Οι δύο σχεδιασμοί που προτείνονται είναι αυτοί των Jones & Nachtsheim. και του Xiao et al. Τελικό κομμάτι είναι η μέτρηση της αποδοτικότητας των σχεδιασμών που προτείνουν οι Jones & Nachtsheim.

Abstract

In many experimental situations in Statistics the object of study is to investigate whether one or more factors affect a quantity of interest. This experimental process has a significant role in Statistics. In this Thesis, we will deal with some experimental designs that they help us to answer the question “which factor effects are important in the experiment”.

The simplest factorial design is the 2^k factorial design, which involves factors with two levels each. One of the advantages for such designs is that, in relation to factors that have more levels, fewer replicates are needed. However, as the factors increase, the experimental replicates that we need to carry out grow exponentially.

For this reason, we use fractional factorial designs, which reduce the experimental replicates, as we deal with a smaller number of factors. One issue that arises with 2^k factorial design is that there is a possibility of some curvature in the response surface. This clearly cannot be controlled by two levels in each factor. One solution on this is to add center points into our design. Another solution is the use of factors with 3 levels. In this case we can consider the presence of curvature, but the problem is that we cannot use a lot of factors because we need too many replicates. As in the 2^k factorial designs, in 3^k factorial designs we may have fractional designs.

An economical solution, both in terms of time and cost, is the use of Definitive Screening Designs (DSD), which is the main topic of this dissertation. In such designs, we typically need fewer runs than in the cases mentioned above. The two designs proposed are those of Jones & Nachtsheim and Xiao et al. The final issue is the measurement of the efficiency for the designs proposed by Jones & Nachtsheim.

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Στην σύγχρονη έρευνα, η Στατιστική είναι ένα σπουδαίο εργαλείο για κάθε επιστήμονα και αυτό γιατί τον βοηθάει να μελετήσει ένα φαινόμενο που εξετάζει με σκοπό να βγάλει κάποιο αποτέλεσμα ή συμπέρασμα. Για παράδειγμα, μια φαρμακοβιομηχανία θέλει να παράξει ένα νέο φάρμακο που το θεωρεί κατάλληλο για μια συγκεκριμένη θεραπεία, για να το διαθέσει στην αγορά. Η συμμετοχή της Στατιστικής σε αυτό, είναι στις κλινικές δοκιμές που πραγματοποιούνται, όπου με την βοήθεια των στατιστικών αναλύσεων, προκύπτει η απόφαση αν το φάρμακο είναι καλό για να βγει στην αγορά και αν αποφασιστεί να δοθεί στην αγορά, γίνονται περαιτέρω μελέτες που εξετάζεται η εμπορικότητα του καθώς και η πιθανότητα που υπάρχει να βρεθούν παρενέργειες σε κάποια άτομα. Για να γίνουν όλες αυτές οι μελέτες χρειάζεται να γίνει συλλογή παρατηρήσεων, ώστε να αναλυθούν με τα κατάλληλα εργαλεία, για να εξαχθούν κάποια συμπεράσματα.

Στους περισσότερους επιχειρηματικούς κλάδους τις πιο πολλές φορές, για την εξαγωγή συμπερασμάτων, χρειάζονται να γίνουν κάποια "πειράματα". Η έννοια του πειραματισμού είναι μία διαδικασία που κάνουμε για να ελέγξουμε κατά πόσο ένας ή περισσότεροι παράγοντες επιδρούν σε μια διεργασία που θέλουμε να εξετάσουμε. Οι παράγοντες που μπορεί να επηρεάζουν τη διεργασία που εξετάζουμε μπορεί να είναι πάρα πολλοί. Κάποιοι από αυτούς θα επηρεάζουν πιο πολύ και κάποιοι θα έχουν επίδραση ισχνή. Επίσης, υπάρχει και μία κατηγορία παραγόντων, που ονομάζονται παράγοντες «θορύβου» και είναι παράγοντες που δεν μπορούμε εύκολα να εξαλείψουμε την επίδραση τους, τέτοιοι μπορούν να είναι οι καιρικές συνθήκες ή φθορά των μηχανημάτων κ.λπ. Στόχος μας είναι να εντοπίσουμε και να μελετήσουμε σωστά αυτούς που επηρεάζουν περισσότερο έτσι ώστε να καταφέρουμε να βγάλουμε σωστά συμπεράσματα για μία διαδικασία. Η επίδραση του κάθε παράγοντα εντοπίζεται από τη μελέτη της απόκρισης όταν αυτός λάβει προμελετημένα διάφορες τιμές. Οι τιμές αυτές ονομάζονται επίπεδα και το πλήθος τους συμβολίζονται με s. Οι παράγοντες μπορούν να είναι είτε ποσοτικοί είτε ποιοτικοί. Ποσοτικούς παράγοντες χαρακτηρίζουμε αυτούς που τα επίπεδα επιλέγονται με βάση το εύρος το τιμών που μας ενδιαφέρει όπως η πίεση, η θερμοκρασία κ.ο.κ. Αντίθετα, ποιοτικούς θεωρούμε αυτούς που στα επίπεδα τους εντοπίζουμε κατηγορικές ή ονομαστικές κλίμακες μέτρησης, όπως το φύλο ενός ατόμου ή ομάδα αίματός του κ.ο.κ.

Οι πειραματικοί σχεδιασμοί ή αλλιώς DOE (Design Of Experiments) είναι ένας κλάδος της εφαρμοσμένης στατιστικής που αποτελεί ένα σημαντικό εργαλείο για την ανάπτυξη μιας στρατηγικής, μεγιστοποιώντας τη διαθέσιμη πληροφορία με τη χρήση όσο το δυνατό λιγότερων πόρων (οικονομικών, χρονικών, ανθρώπινων κ.λπ.). Επίσης, είναι ένα εξαιρετικό εργαλείο ανάλυσης, που μπορεί να χρησιμοποιηθεί σε ποικίλες εφαρμογές και καταστάσεις. Οι πειραματικοί σχεδιασμοί επιτρέπουν την ταυτόχρονη μελέτη πολλών παραγόντων, εκτιμώντας την επίδραση τους στην μεταβλητή απόκρισης. Πολλές από τις τωρινές στατιστικές προσεγγίσεις για τον σχεδιασμό πειραμάτων προέρχονται από τον R.A. Fisher στις αρχές του 20ου αιώνα.

Κάθε πείραμα διακρίνεται σε δύο στάδια: Αρχικά στο σχεδιασμό και τη διεξαγωγή του πειράματος, όπου λαμβάνονται και οι πειραματικές αποκρίσεις και στη συνέχεια, στην ανάλυση των δεδομένων που προκύπτουν από το σχεδιασμό. Τόσο ο σχεδιασμός όσο και η ανάλυση των δεδομένων διέπονται από τρεις βασικές αρχές: Την επανάληψη (replicate), την τυχαιοποίηση (randomization) και την ομαδοποίηση (blocking) των δοκιμών (εκτελέσεων) του πειράματος. Οι επαναλήψεις βοηθούν ώστε να εκτιμηθεί το πειραματικό σφάλμα, που είναι βασικό κριτήριο για να αποφασιστεί αν οι παρατηρούμενες διαφορές στα δεδομένα είναι στατιστικά σημαντικές και συνεπώς υπάρχουν σημαντικές επιδράσεις των παραγόντων στην απόκριση. Με την τυχαιοποίηση, εξασφαλίζεται η ανεξαρτησία των παρατηρήσεων, ενώ με την ομαδοποίηση διασφαλίζεται η ομοιομορφία των πειραματικών μονάδων.

Για να πραγματοποιηθεί ένας επιτυχημένος πειραματικός σχεδιασμός θα πρέπει αρχικά να γίνει σαφής διατύπωση και κατανόηση του προβλήματος που θέλουμε να μελετήσουμε. Επίσης, πρέπει να επιλεχθεί, η μεταβλητή απόκρισης που θέλουμε να εξετάσουμε καθώς και οι παράγοντες που πιστεύεται ότι την επηρεάζουν μαζί με τα αντίστοιχα επίπεδά τους. Ύστερα, επιλέγουμε το πειραματικό σχέδιο που θα χρησιμοποιήσουμε, όπου και το εκτελούμε λαμβάνοντας τα αποτελέσματα μας (πειραματικές αποκρίσεις). Στα αποτελέσματα που προκύπτουν κάνουμε στατιστική ανάλυση από όπου θα προκύψουν τα συμπεράσματα μας και οι εισηγήσεις μας.

Οι ερευνητές πραγματοποιούν πειράματα σχεδόν σε όλα τα πεδία της έρευνας. Ένα καλά στρατηγικά σχεδιασμένο και εκτελεσμένο πείραμα ίσως παρέχει καλύτερη πληροφόρηση για την επίδραση που έχει ένας ή περισσότεροι παράγοντες στην μεταβλητή απόκρισης. Αρκετοί πειραματιστές συνήθως χρησιμοποιούν την επιλογή να αλλάζουν τα επίπεδα μόνο ενός παράγοντα σε κάθε δοκιμή. Γενικά, η επιλογή να αλλάζουμε το επίπεδο ενός παράγοντα την φορά σε κάθε δοκιμή σαν προσέγγιση είναι μη αποδοτική. Το πρόβλημα που προκύπτει στην συγκεκριμένη διαδικασία, που στην θεωρία ονομάζεται σχεδιασμός OFAT(One Factor At Time), είναι ότι θα χρειαστεί ένα μεγάλο αριθμό πειραματικών δοκιμών για να εκτελεστεί η διαδικασία και, επίσης, δεν είναι εύκολο να ελεγχθεί η από κοινού επίδραση δύο η περισσότερων παραγόντων στην απόκριση (αλληλεπίδραση). Η τεχνική OFAT λοιπόν, είναι εν τέλει μία λάθος πρακτική, παρότι σαν διαδικασία είναι πιο εύκολη. Οι παραγοντικοί σχεδιασμοί αποτελούν εναλλακτική λύση των OFAT σχεδιασμών καθώς ασχολούνται με πολλούς παράγοντες και έχουν καλύτερες ιδιότητες από ότι στους OFAT. Αυτή είναι μία στρατηγική όπου οι παράγοντες μεταβάλλονται μαζί αντί κάθε φορά από ένας όπως αναφέρθηκε παραπάνω. Ένας παραγοντικός σχεδιασμός είναι μία χρήσιμη τεχνική να ερευνήσουμε την επίδραση που έχει τόσο ο κάθε παράγοντας από μόνος του στην απόκριση (κύρια επίδραση) όσο και τις αλληλεπιδράσεις των παραγόντων στην μεταβλητή απόκρισης. Όταν στο σχεδιασμό εξετάζονται όλοι οι συνδυασμοί των επιπέδων των παραγόντων που συμμετέχουν στο πείραμα, τότε αυτός ο σχεδιασμός ονομάζεται πλήρης παραγοντικός σχεδιασμός. Ο κάθε συνδυασμός των επιπέδων των παραγόντων ονομάζεται θεραπεία.

Οι παραγοντικοί σχεδιασμοί είναι εξαιρετικά χρήσιμοι γενικά στις επιστήμες, διότι δίνουν την δυνατότητα να βρεθεί, αν υπάρχει, ο βέλτιστος συνδυασμός επιπέδων των παραγόντων που βελτιστοποιεί την υπό διερεύνηση διεργασία, μειώνοντας το πειραματικό σφάλμα. Οι παραγοντικοί σχεδιασμοί απλοποιούν την διαδικασία κάνοντας την πιο οικονομική, επιτρέποντας τη μελέτη των παραγόντων σε αρκετά επίπεδα στην ανάλυση μας. Υπάρχουν αρκετά στατιστικά εργαλεία που είναι διαθέσιμα για την ανάλυση, την παρουσίαση και την αξιολόγηση των επιδράσεων των παραγόντων που μελετώνται. Οι μέθοδοι που χρησιμοποιούνται είναι η ανάλυση διακύμανσης η ανάλυση μέσων ή και οι διαδικασίες της πολλαπλής παλινδρόμησης.

Υποθέτουμε ότι υπάρχουν k παράγοντες που θέλουμε να ερευνήσουμε και ο κάθε παράγοντας από αυτούς έχει s επίπεδα. Το συνολικό πλήθος που χρειαζόμαστε για έναν πλήρη παραγοντικό σχεδιασμό είναι ίσο με s^k. Εύκολα καταλαβαίνουμε ότι όσο αυξάνονται τα επίπεδα και οι παράγοντες, οι εκτελέσεις που χρειαζόμαστε για το πλήρες παραγοντικό σχέδιο αυξάνονται ραγδαία. Για παράδειγμα, αν έχουμε επτά παράγοντες με δύο επίπεδα ο καθένας τους, τότε χρειαζόμαστε 128 εκτελέσεις για να γίνει η μελέτη. Γι΄αυτό τον λόγο για εξοικονόμηση πόρων και χρόνου, γίνεται η χρήση των κλασματικών παραγοντικών σχεδιασμών στους οποίους εκτελείται μόνο ένα επιλεγμένο υποσύνολο θεραπειών από ένα πλήρη παραγοντικό σχεδιασμό. Οι κλασματικοί παραγοντικοί σχεδιασμοί χρησιμοποιούνται προκειμένου να πάρουμε πληροφορίες σχετικά με τις κύριες επιδράσεις και τις αλληλεπιδράσεις χαμηλής τάξης, με λιγότερες εκτελέσεις και υποθέτουν ότι οι αλληλεπιδράσεις υψηλής τάξης είναι αμελητέες.

Οι παραγοντικές επιδράσεις θεωρούμε ότι διέπονται από κάποιες βασικές αρχές, την ιεραρχία, την σποραδικότητα και την κληρονομικότητα. Η αρχή της ιεραρχίας αναφέρει ότι οι επιδράσεις χαμηλής τάξης είναι πιθανότερο να είναι σημαντικές από ότι οι επιδράσεις της υψηλής τάξης. Για το λόγο αυτό, τις περισσότερες φορές η μελέτη επικεντρώνεται στις κύριες επιδράσεις και τις αλληλεπιδράσεις δύο παραγόντων. Η αρχή της σποραδικότητας ορίζει ότι ο αριθμός των επιδράσεων που τελικά είναι σημαντικές είναι σχετικά μικρός και, τέλος, η αρχή της κληρονομικότητας τονίζει ότι για να είναι μία αλληλεπίδραση σημαντική θα πρέπει οι παράγοντες που την απαρτίζουν να έχουν σημαντικές κύριες επιδράσεις.

Η κυριότερη χρήση των κλασματικών σχεδιασμών είναι σε πειράματα κρησαρίσματος, πειράματα δηλαδή στα οποία αρχικά μελετώνται πολλοί παράγοντες και πρέπει να αναγνωριστούν όσοι από αυτούς έχουν σημαντικές επιδράσεις με το λιγότερο δυνατό κόστος. Οι πειραματιστές εφαρμόζουν τις τρεις αρχές που διέπουν τις παραγοντικές επιδράσεις σε κλασματικούς παραγοντικούς σχεδιασμούς για να βρουν τους παράγοντες αυτούς που επηρεάζουν σημαντικά την μεταβλητή απόκρισης. Πιλοτικές μελέτες, τα λεγόμενα screening πειράματα, είναι κάποιες από τις τεχνικές που συνήθως χρησιμοποιούνται.

ΚΕΦΑΛΑΙΟ 2

Σχεδιασμοί κρησαρίσματος για παράγοντες με δύο επίπεδα

2.1 Σχεδιασμοί Κρησαρίσματος

Οι σχεδιασμοί κρησαρίσματος είναι ιδιαίτερα χρήσιμοι, όταν χρειάζεται να μελετηθεί ένας μεγάλος αριθμός παραγόντων που πιθανόν να επηρεάζουν την απόκριση ενός πειράματος, αλλά μόνο ένα μικρό υποσύνολο αυτών πιστεύεται ότι είναι σημαντικοί. Σκοπός τους είναι η εύρεση των παραγόντων που έχουν τη σημαντικότερη επίδραση στην εξαρτημένη μεταβλητή του πειράματος στο αρχικό στάδιο την πειραματικής διαδικασίας, ενώ στόχος τους είναι η εξοικονόμηση χρόνου και πόρων. Για το λόγο αυτό, συνήθως επιλέγονται σχεδιασμοί που περιορίζουν το κόστος του πειράματος, σε σχέση με τον αριθμό των πειραματικών εκτελέσεων που πρέπει να πραγματοποιηθούν και που δίνουν τα καλύτερα και πιο εύκολα ερμηνεύσιμα αποτελέσματα. Συνήθως, οι σχεδιασμοί κρησαρίσματος είναι δύο επιπέδων, όπου ο πιο συνήθης συμβολισμός των δύο επιπέδων είναι το +1 και το –1.

Ας θεωρήσουμε ότι έχουμε k παράγοντες προς εξέταση. Σκοπός είναι να πάρουμε τις εκτιμήσεις για τις κύριες επιδράσεις των k παραγόντων που συμμετέχουν στο πείραμα, σύμφωνα με το γραμμικό μοντέλο

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon$$

όπου τα σφάλματα ακολουθούν κανονική κατανομή με μέση τιμή 0 και διασπορά σ^2. Οι εκτιμήσεις των κύριων επιδράσεων στην γενική περίπτωση υπολογίζονται με την μέθοδο των ελαχίστων τετραγώνων και η εύρεση των σημαντικότερων επιδράσεων γίνεται με γνωστούς μεθόδους ανάλυσης.

Οι παράμετροι που έχουν οι σχεδιασμοί κρησαρίσματος είναι τρεις και συμβολίζονται με (n ,k, p), όπου n είναι οι πειραματικές εκτελέσεις (γραμμές του πίνακα σχεδιασμού), k οι παράγοντες (στήλες) που είναι υπό εξέταση και p είναι προβολικότητα του σχεδιασμού. Ένας σχεδιασμός λέμε ότι έχει προβολικότητα p όταν σε κάθε επιλογή p στηλών του σχεδιασμού, όλες οι πειραματικές εκτελέσεις του πλήρους παραγοντικού 2^p σχεδιασμού εμφανίζονται τουλάχιστον μία φορά. Τον όρο της προβολικότητας τον όρισαν οι Box & Tyssedal (1996). Σε ένα σχεδιασμό με προβολικότητα *p*, όλες οι παραγοντικές επιδράσεις αυτών των *p* παραγόντων (μαζί με την μέση τιμή) είναι εκτιμήσιμες με την μέθοδο ελαχίστων τετραγώνων και ο υπολογισμός της διασποράς του σφάλματος να γίνεται από τις πιθανές επαναλαμβανόμενες παρατηρήσεις. Η προβολικότητα του σχεδιασμού αποτελεί χρήσιμο κριτήριο επιλογής σχεδιασμών κρησαρίσματος, διότι παρέχει πληροφορία για την ικανότητα εκτίμησης αλληλεπιδράσεων των παραγόντων που θα βρεθούν σημαντικοί μετά από το αρχικό κρησάρισμα.

Στην κατηγορία των σχεδιασμών κρησαρίσματος ανήκουν οι 2^k πλήρεις παραγοντικοί σχεδιασμοί, οι κλασματικοί παραγοντικοί σχεδιασμοί και οι ορθογώνιοι σχηματισμοί (οι Plackett-Burman σχεδιασμοί, οι σχεδιασμοί Hadamard κ.λπ.). Πρόσφατα, προτάθηκε και μια ακόμη κατηγορία σχεδιασμών, οι Definitive Screening Designs (DSD). Σύμφωνα με τους Myers και Montgomery (1995), η φάση του κρησαρίσματος είναι η πρώτη φάση του πειράματος και η σημαντικότερη, καθώς αν γίνει κάποιο λάθος στο πείραμα αυτό θα έχει αποτέλεσμα να εξάγουμε λανθασμένα συμπεράσματα.

2.2 Ο 2^k πλήρης παραγοντικός σχεδιασμός

Ο 2^k πλήρης παραγοντικός σχεδιασμός είναι ο πιο δημοφιλής σχεδιασμός. Οι συγκεκριμένοι σχεδιασμοί μελετούν *k* παράγοντες που έχουν 2 επίπεδα. Τα επίπεδα αυτά μπορούν να είναι είτε ποσοτικά είτε ποιοτικά. Ένα μεγάλο πλεονέκτημα αυτών των σχεδιασμών είναι ότι μας επιτρέπει για πολλούς παράγοντες να χρησιμοποιήσουμε σχετικά μικρό αριθμό δοκιμών. Επίσης, είναι χρήσιμοι στα πρώιμα στάδια της πειραματικής διαδικασίας, όταν υπάρχουν πολλοί παράγοντες. Μια πλήρης επανάληψη ενός σχεδιασμού με *k* παράγοντες, με δύο επίπεδα ο καθένας, για να εξεταστούν απαιτεί $2 \times 2 \times \ldots \times 2 = 2^k$ παρατηρήσεις και λέγεται 2^k παραγοντικός σχεδιασμός. Επειδή εξετάζουμε μόνο δύο επίπεδα στα πειράματα ο συμβολισμός για αυτά τα επίπεδα που χρησιμοποιούμε είναι απλός. Για την χαμηλή στάθμη του παράγοντα την ορίζουμε με “-1” ή “-” ενώ για την υψηλή στάθμη του “+1” ή “+”.

Εκτός από την εκτίμηση και την αξιολόγηση των παραγοντικών επιδράσεων μας ενδιαφέρει και το στατιστικό μοντέλο που θα μπορεί να περιγράψει τη διαδικασία που μελετάται. Αυτό το πετυχαίνουμε χρησιμοποιώντας ένα γραμμικό μοντέλο παλινδρόμησης. Για παράδειγμα, στην περίπτωση που έχουμε 3 παράγοντες με 2 επίπεδα το μοντέλο θα έχει μορφή

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_2 x_3 + \beta_6 x_1 x_3 + \beta_7 x_1 x_2 x_3 + \varepsilon$$

όπου

- $\boldsymbol{y}$ είναι η μεταβλητή απόκρισης ,
- $\boldsymbol{\beta_i}$ είναι οι παράμετροι του μοντέλου,
- οι όροι $\boldsymbol{x_k}$ είναι οι ανεξάρτητες μεταβλητές που αφορούν τις κύριες επιδράσεις των παραγόντων και παίρνουν τιμές ± 1,
- οι όροι $\boldsymbol{x_i}\, \mathbf{x_j} \dots \mathbf{x_l}$ με $i \neq j \dots \neq l$ δηλώνουν τις αλληλεπιδράσεις των παραγόντων, και
- $\boldsymbol{\varepsilon}$ το τυχαίο σφάλμα που ακολουθεί κανονική κατανομή $N(0,\sigma^2)$

Γενικά όταν μελετάμε *k* παράγοντες στο πλήρες μοντέλο παλινδρόμησης θα έχουμε *k* κύριες επιδράσεις, $\binom{k}{2}$ αλληλεπιδράσεις δύο παραγόντων, κ.ο.κ και μια αλληλεπίδραση *k* παραγόντων. Δηλαδή, σε έναν 2^k πλήρη παραγοντικό σχεδιασμό το πλήρες μοντέλο παλινδρόμησης θα περιέχει συνολικά 2^k-1 επιδράσεις. Η εκτίμηση των επιδράσεων βρίσκονται μέσω του μοντέλου παλινδρόμησης και συγκεκριμένα από τους παραμέτρους που εκτιμήσαμε, καθώς η εκτίμηση της κάθε επίδρασης είναι ίση το διπλάσιο της εκτίμησης της αντίστοιχης παραμέτρου στο μοντέλο.

Συνήθως, οι πόροι για πειραματισμό είναι περιορισμένοι όποτε ο πειραματιστής αναγκάζεται αρχικά να μειώσει τις επαναλήψεις του πειράματος. Μια μόνο επανάληψη ενός 2^k πλήρους παραγοντικού σχεδιασμού ονομάζεται μη-επαναλαμβανόμενος (unreplicated) πλήρης παραγοντικός σχεδιασμός. Με μόνο μία επανάληψη δεν μπορούμε να εκτιμήσουμε το σφάλμα οπότε δεν μπορούμε να κάνουμε στατιστικούς ελέγχους για τις επιδράσεις των παραγόντων. Μια προσέγγιση στην ανάλυση ενός μη επαναλαμβανόμενου παραγοντικού σχεδιασμού είναι να εκτιμήσουμε το σφάλμα μέσω δημιουργίας ψευδοσφάλματος, συνδυάζοντας τα αθροίσματα τετραγώνων υψηλής τάξης αλληλεπιδράσεων που, με βάση την αρχή της σποραδικότητας και της ιεραρχίας των επιδράσεων δεν είναι πιθανό να είναι σημαντικές. Μία δεύτερη λύση είναι το διάγραμμα κανονική πιθανότητας των επιδράσεων (normal probability plot) ή αλλιώς Daniel's Plot. Το σχεδιάγραμμα αυτό μας βοηθάει στην αξιολόγηση για το ποιες παραγοντικές επιδράσεις είναι σημαντικές στο πείραμα μας. Οι επιδράσεις που δεν είναι σημαντικές τείνουν να πέσουν πάνω σε μία ευθεία, ενώ αυτές που είναι σημαντικές αποκλίνουν από την ευθεία.

2.3 Ο Κλασματικός 2^{k-p} Παραγοντικός σχεδιασμός

Ένα ζήτημα που δημιουργείται στους 2^k παραγοντικούς σχεδιασμούς είναι ότι όσο αυξάνονται οι παράγοντες που μελετάμε, τόσο θα αυξάνονται πιο γρήγορα και οι παρατηρήσεις που απαιτούνται για μια πλήρη επανάληψη του σχεδιασμού. Δηλαδή στην περίπτωση που μελετάμε πέντε παράγοντες θέλουμε 32 εκτελέσεις, όταν έχουμε έξι παράγοντες θέλουμε 64, εφτά παράγοντες 128, με οκτώ παράγοντες θέλουμε 256 παρατηρήσεις κ.ο.κ. Αυτό πρακτικά σημαίνει ότι χρησιμοποιούνται βαθμοί ελευθερίας για την εκτίμηση επιδράσεων υψηλής τάξης που, με βάση τις αρχές της σποραδικότητας και της ιεραρχίας των επιδράσεων δεν είναι πιθανό να είναι σημαντικές. Για παράδειγμα, στην περίπτωση που μελετάμε οκτώ παράγοντες με έναν πλήρη 2^8 παραγοντικό σχεδιασμό, έχουμε 255 βαθμούς ελευθερίας, εκ των οποίων μόνο 8 βαθμοί ελευθερίας χρησιμοποιούνται για την εκτίμηση κύριων επιδράσεων και 28 βαθμοί ελευθερίας για την εκτίμηση αλληλεπιδράσεων δυο παραγόντων. Οι 219 βαθμοί ελευθερίας που απομένουν χρησιμοποιούνται για την εκτίμηση αλληλεπιδράσεων τριών ή και περισσοτέρων παραγόντων, που μπορεί να μην είναι και σημαντικές.

Λαμβάνοντας υπόψη αυτούς τους περιορισμούς, οι 2^{k-p} κλασματικοί παραγοντικοί σχεδιασμοί αποτελούν μια πιο οικονομική και λιγότερο χρονοβόρα επιλογή. Οι σχεδιασμοί αυτοί δημιουργούνται από την ορθή επιλογή ενός υποσυνόλου των εκτελέσεων του πλήρους σχεδιασμού (κλάσμα). Αυτό επιτυγχάνεται με χρήση των οριζουσών σχέσεων, οι οποίες αποτελούν δομικό στοιχείο των κλασματικών σχεδιασμών. Στους σχεδιασμούς αυτούς δεν είναι δυνατή η ανεξάρτητη εκτίμηση όλων των παραγοντικών επιδράσεων. Η διακριτική ικανότητα είναι ένα χρήσιμο κριτήριο για την επιλογή κατάλληλου κλάσματος.

2.3.1 Διακριτική ικανότητα ενός κλασματικού σχεδιασμού

Ένα βασικό κριτήριο για την επιλογή του κατάλληλου κλασματικού παραγοντικού σχεδιασμού είναι η διακριτική ικανότητα (resolution). Το κριτήριο της διακριτικής ικανότητας παρουσιάστηκε για πρώτη φορά από τους Box και Hunter το 1961. Η διακριτική ικανότητα R ενός σχεδιασμού είναι ένας αριθμός που φανερώνει ότι καμία επίδραση i παραγόντων δεν είναι ταυτόσημη με άλλη επίδραση που περιέχει λιγότερους από R-i παράγοντες. Για την διακριτική ικανότητα ενός σχεδιασμού συνήθως χρησιμοποιούμε λατινικούς αριθμούς. Παρακάτω ορίζονται οι σχεδιασμοί διακριτικής ικανότητας III, IV και V οι οποίοι είναι αυτοί που χρησιμοποιούνται ευρύτερα.

- Σχεδιασμοί διακριτικής ικανότητας III. Στους σχεδιασμούς αυτούς, καμία κύρια επίδραση δεν είναι ταυτόσημη με οποιαδήποτε άλλη κύρια επίδραση, αλλά οι κύριες επιδράσεις είναι ταυτόσημες με αλληλεπιδράσεις δυο παραγόντων.
- Σχεδιασμοί διακριτικής ικανότητας IV. Στους σχεδιασμούς αυτούς, καμία κύρια επίδραση δεν είναι ταυτόσημη με οποιαδήποτε άλλη κύρια επίδραση, ή με οποιαδήποτε αλληλεπίδραση δυο παραγόντων, αλλά αλληλεπιδράσεις δυο παραγόντων είναι ταυτόσημες με άλλες αλληλεπιδράσεις δυο παραγόντων.
- Σχεδιασμοί διακριτικής ικανότητας V. Στους σχεδιασμούς αυτούς, καμία κύρια επίδραση ή αλληλεπίδραση δυο παραγόντων δεν είναι ταυτόσημη με οποιαδήποτε άλλη κύρια επίδραση ή αλληλεπίδραση δυο παραγόντων, αλλά αλληλεπιδράσεις δυο παραγόντων είναι ταυτόσημες με αλληλεπιδράσεις τριών παραγόντων.

Γενικά, η διακριτική ικανότητα ενός κλασματικού παραγοντικού σχεδιασμού μπορεί να προσδιοριστεί από τις ορίζουσες σχέσεις του. Η διακριτική ικανότητα ταυτίζεται με το μικρότερο μήκος λέξης που χρησιμοποιείται σε ορίζουσα σχέση.

Ένας σχεδιασμός με μικρή διακριτική ικανότητα, έχει τουλάχιστον μία ορίζουσα σχέση με μικρό μήκος, κάτι το οποίο υποδηλώνει την ύπαρξη ταυτόσημων επιδράσεων (οι επιδράσεις που συγχέονται σε ένα κλασματικό παραγοντικό σχεδιασμό) χαμηλής τάξης. Σύμφωνα με την αρχή της ιεραρχίας, επιδράσεις χαμηλής τάξης είναι πιθανότερο να είναι σημαντικές, άρα δεν ενδείκνυται η επιλογή ενός σχεδιασμού στον οποίο επιδράσεις χαμηλής τάξης συγχέονται μεταξύ τους. Συνήθως είναι επιθυμητό, να χρησιμοποιούνται (για δεδομένο αριθμό παρατηρήσεων) κλασματικοί σχεδιασμοί που έχουν την υψηλότερη δυνατή διακριτική ικανότητα, καθώς η υψηλή διακριτική ικανότητα βάζει λιγότερους περιορισμούς στις υποθέσεις που απαιτούνται, όσον αφορά ποιες αλληλεπιδράσεις είναι αμελητέες, με σκοπό να πάρουμε μια μοναδική ερμηνεία για τα δεδομένα.

Οι σχεδιασμοί κρησαρίσματος τυπικά έχουν διακριτική ικανότητα III. Αυτό που επιτυγχάνεται με την συγκεκριμένη διακριτική ικανότητα είναι ότι μπορούμε να εκτιμήσουμε τις κύριες επιδράσεις των παραγόντων δίχως να χρειάζεται μεγάλος αριθμός εκτελέσεων. Μερικές φορές στους σχεδιασμούς κρησαρίσματος χρησιμοποιούμε σχεδιασμούς με διακριτική ικανότητα IV. Σε αυτούς τους σχεδιασμούς, οι κύριες επιδράσεις συγχέονται στην χειρότερη περίπτωση με αλληλεπιδράσεις τριών παραγόντων. Τέτοιοι σχεδιασμοί είναι προτιμότεροι από όσους έχουν διακριτική ικανότητα ΙΙΙ, όμως απαιτούν περισσότερες εκτελέσεις από ότι οι σχεδιασμοί διακριτικής ικανότητας ΙΙΙ.

2.4 Ορθογώνιοι σχηματισμοί

Οι ορθογώνιοι σχηματισμοί είναι ιδιαίτερα χρήσιμοι σαν σχεδιασμοί κρησαρίσματος για περιπτώσεις όπου εξετάζεται ένας μεγάλος αριθμός παραγόντων, εκ των οποίων ένα μικρό υποσύνολο αναμένεται να έχει σημαντική επίδραση σε μία εξαρτημένη μεταβλητή. Ένας ορθογώνιος σχηματισμός μπορεί να χρησιμοποιηθεί όχι μόνο για τον έλεγχο των κύριων επιδράσεων των παραγόντων άλλα και για την ανίχνευση των αλληλεπιδράσεων μεταξύ ενός υποσυνόλου αυτών των παραγόντων. Αρχικά χρησιμοποιήθηκαν από τον Rao το 1947, ως κλασματικοί παραγοντικοί σχεδιασμοί για τον υπολογισμό των κύριων επιδράσεων και των αλληλεπιδράσεων χαμηλής τάξης.

Οι ορθογώνιοι σχηματισμοί συμβολίζονται με $OA(n,k,s,t)$ όπου n είναι ο αριθμός των παρατηρήσεων, k είναι ο αριθμός παραγόντων, s είναι το πλήθος των επιπέδων των παραγόντων και t είναι η ισχύς του σχεδιασμού. Ένας ορθογώνιος σχηματισμός είναι ένας n x k πίνακας με στοιχεία επιλεγμένα από ένα σύνολο s διακεκριμένων συμβόλων διατεταγμένων έτσι ώστε, για κάθε επιλογή t στηλών του σχεδιασμού, καθένα από τα s^t διανύσματα γραμμών να εμφανίζονται το ίδιο συχνά. Επομένως, είναι εύκολο να παρατηρηθεί ότι το s^t διαιρεί το n. Άρα ισχύς t σημαίνει ότι για οποιαδήποτε επιλογή t στηλών δημιουργείται ένας πλήρης s^t παραγοντικός σχεδιασμός.

Για κάθε επιλογή n, k και s είναι δυνατόν να φτιάξουμε διάφορους ορθογώνιους σχηματισμούς με τις συγκεκριμένες αυτές παραμέτρους. Οι πλήρεις παραγοντικοί σχεδιασμοί, καθώς και τα κλάσματα τους, ανήκουν στην ευρύτερη κατηγορία των Ορθογώνιων Σχηματισμών. Πολλοί από αυτούς όμως είναι ισοδύναμοι. Ισοδύναμοι λέγονται δύο ορθογώνιοι σχηματισμοί αν ο ένας μπορεί να παραχθεί από τον άλλο, με μια σειρά από μεταθέσεις γραμμών ή/και στηλών ή/και συμβόλων σε κάθε στήλη.

2.4.1 Σχεδιασμοί Hadamard

Ιδιαίτερο ενδιαφέρον παρουσιάζει πάντα η εύρεση μη ισοδύναμων ορθογώνιων σχηματισμών με συγκεκριμένες παραμέτρους (n,k,s,t) διότι μπορεί να συμπεριφέρονται καλύτερα σε διάφορες πειραματικές καταστάσεις. Τέτοιους σχεδιασμούς βρίσκουμε στους σχεδιασμούς Hadamard όταν προβάλλονται σε μικρότερους σχεδιασμούς.

Μια μεγάλη κατηγορία ορθογώνιων σχηματισμών $OA(n,n\text{-}1,2,2)$ προκύπτει από τους συγκεκριμένους πίνακες. Συγκεκριμένα ένας πίνακας Hadamard είναι ένας πίνακας n *x* n με στοιχεία ± 1, του οποίου δύο οποιεσδήποτε γραμμές και δύο οποιεσδήποτε στήλες είναι ορθογώνιες. Συνεπώς, για κάθε πίνακα Hadamard H_n ισχύει η σχέση $H_n^T H_n = H_n H_n^T = nI_n$ όπου I_n ο μοναδιαίος πίνακας τάξης n. Αν αφαιρεθεί η πρώτη στήλη με τις μονάδες παίρνουμε έναν κορεσμένο ορθογώνιο σχηματισμό της μορφής $OA(n,n\text{-}1,2,2)$. Κορεσμένος ορθογώνιος σχηματισμός λέγεται όταν ο αριθμός των παραγόντων που μπορεί να μελετήσει είναι ο μεγαλύτερος δυνατός, σε συνάρτηση με τον αριθμό των εκτελέσεων και της ισχύος του.

Τα μέχρι σήμερα ακριβή αποτελέσματα για το αριθμό των μη ισοδύναμων πινάκων Hadamard έχουν ως εξής : Για τάξεις n<16 οι πίνακες Hadamard είναι μοναδικοί ως προς την ισοδυναμία. Όταν n=16 και n=20 υπάρχουν πέντε και τρεις μη ισοδύναμοι Hadamard αντίστοιχα. Υπάρχουν ακριβώς 60 μη ισοδύναμοι πίνακες Hadamard τάξης n=24. Για n>32 όμως είναι δύσκολο να βρεθούν. Οι πίνακες Hadamard μελετήθηκαν για πρώτη φορά από τον Sylvester το 1897, ο οποίος στο σύγγραμμα του είχε βρει τέτοιους πίνακες για όλες τις δυνάμεις του 2. Αλλά, οφείλουν το όνομα τους στο Γάλλο μαθηματικό Jacque Hadamard (1893), ο οποίος κατασκεύασε τους τετραγωνικούς πίνακες τάξης 12 και 20 με στοιχεία.

Από πίνακες Hadamard προκύπτουν και οι σχεδιασμοί Plackett-Burman. Οι μικρότεροι από αυτούς είναι 8, 12, 16, 20 και 24 εκτελέσεων και κατασκευάζονται με κυκλική μετάθεση ενός συγκεκριμένου διανύσματος. Τα διανύσματα από τα οποία προκύπτουν οι Plackett και Burman σχεδιασμοί για τις τάξεις n=8, 12, 16, 20 και 24 είναι:

8	+ + + - + - -
12	+ + - + + + - - - + -
16	+ + + + - + - + + - - + - - -
20	+ + - - + + + + - + - + - - - - + + -
24	+ + + + + - + - + + - - + + - - + - + - - - -

Για παράδειγμα, για να κατασκευάσουμε έναν Plackett και Burman σχεδιασμό τάξεως 20, γράφουμε ως πρώτη γραμμή το διάνυσμα που δίνεται για n=20, στη συνέχεια δημιουργούμε τις υπόλοιπες γραμμές με κυκλική μετάθεση των στοιχείων της προηγούμενης γραμμής και προσθέτουμε στο τέλος μια τελευταία γραμμή με όλα τα πρόσημα αρνητικά.

2.5 Σύνοψη των 2^k Παραγοντικών σχεδιασμών

Σύμφωνα με τα όσα αναφέρθηκαν, οι 2^k παραγοντικοί σχεδιασμοί είναι από τους πιο διαδεδομένους πειραματικούς σχεδιασμούς, καθώς κάθε παράγοντας εξετάζεται σε μόνο δύο επίπεδα και συνεπώς μπορούμε να μελετήσουμε ταυτόχρονα πολλούς παράγοντες με σχετικό μικρό αριθμό εκτελέσεων. Όμως ακόμα και για σχετικά μικρό αριθμό παραγόντων, και πάλι χρειαζόμαστε αρκετές εκτελέσεις. Αυτό το αντιμετωπίζουμε με την χρησιμοποίηση των κλασματικών παραγοντικών σχεδιασμών οι οποίοι χρειάζονται λιγότερες εκτελέσεις από ότι ο πλήρης παραγοντικός σχεδιασμός διατηρώντας την επιθυμητή ιδιότητα της ανίχνευσης σημαντικών επιδράσεων, δεδομένης όμως της σύγχυσης κάποιων παραγοντικών επιδράσεων.

Ένα πρόβλημα που ενδέχεται να αναδειχθεί από τη χρήση σχεδιασμών με δύο επίπεδα είναι ότι δεν μπορούμε να ελέγξουμε την ύπαρξη καμπυλότητας στην απόκριση. Ο έλεγχος καμπυλότητας στην απόκριση μπορεί να επιτευχθεί όταν οι παράγοντες μελετώνται σε τρία ή περισσότερα επίπεδα. Για το λόγο αυτό, έχει προταθεί η χρήση σχεδιασμών με τρία επίπεδα, σαν σχεδιασμοί κρησαρίσματος.

ΚΕΦΑΛΑΙΟ 3

Σχεδιασμοί κρησαρίσματος για παράγοντες με τρία επίπεδα

3.1 Η Προσθήκη κεντρικού σημείου στον σχεδιασμό 2^k

Ένα ζήτημα που δημιουργείται στους 2^k παραγοντικούς σχεδιασμούς είναι αν το μοντέλο που εξετάζουμε είναι γραμμικό, δηλαδή αν ισχύει η υπόθεση της γραμμικότητας. Φυσικά, δεν απαιτείται να υπάρχει τέλεια γραμμικότητα, καθώς η διαδικασία των 2^k παραγοντικών σχεδιασμών θα δουλέψει εξίσου καλά ακόμη και όταν η υπόθεση της γραμμικότητας προσεγγίζεται. Στην πραγματικότητα, εάν προσθέσουμε τους όρους της αλληλεπίδρασης σε ένα πρώτης τάξης μοντέλο τότε έχουμε ένα μοντέλο ικανό να υποδείξει κάποια κυρτότητα στην συνάρτηση απόκρισης.

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum\sum_{i<j} \beta_{ij} x_i x_j + \varepsilon$$

Το αποτέλεσμα για την ύπαρξη της κυρτότητας προκύπτει από τους όρους της αλληλεπίδρασης $\beta_{ij}x_ix_j$. Σε κάποιες καταστάσεις, η κυρτότητα της συνάρτηση απόκρισης δεν θα είναι επαρκής σύμφωνα με την προηγούμενη σχέση. Έτσι επιλέγεται να χρησιμοποιείται η παρακάτω συνάρτηση

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum\sum_{i<j} \beta_{ij} x_i x_j + \sum_{j=1}^{k} \beta_{jj} {x_j}^2 + \varepsilon$$

όπου β_{jj} δηλώνει την τετραγωνική επίδραση του x_j παράγοντα και αυτή η εξίσωση εκφράζεται ως δεύτερης τάξης μοντέλο.

Σε ένα πείραμα που οι παράγοντες έχουν δύο επίπεδα, συνήθως προσδοκούμε να ταιριάξουμε το μοντέλο της πρώτης τάξης, αλλά πρέπει να ληφθεί υπόψη η πιθανότητα το μοντέλο των τετραγωνικών επιδράσεων να είναι περισσότερο κατάλληλο. Υπάρχει μία μέθοδος με την εισαγωγή συγκεκριμένων σημείων σε ένα παραγοντικό σχεδιασμό 2^k, όπου είναι ένας καλός τρόπος για να εξετάσουμε την ύπαρξη καμπυλότητας. Η μέθοδος αυτή εφαρμόζεται με την είσοδο κεντρικών σημείων στον σχεδιασμός 2^k. Το σύνολο των σημείων είναι ίσα με n_c και η αντίστοιχη μεταβλητή παίρνει την τιμή $X_i=0$ ($i=1,2,...,k$). Ένας σημαντικός λόγος για την είσοδο συγκεκριμένων εκτελέσεων στο κέντρο του σχεδιασμού είναι να παρέχει ένα μέτρο για την σταθερότητα της διαδικασίας και να ελέγξει την ύπαρξη καμπυλότητας στην επιφάνεια απόκρισης. Όταν προσθέσουμε κεντρικά σημεία, υποθέτουμε ότι οι k παράγοντες είναι ποσοτικοί.

Για εξηγήσουμε την συγκεκριμένη προσέγγιση θεωρούμε έναν σχεδιασμό 2^2 με μία παρατήρηση των παραγοντικών σημείων (- , -) , (+ , -) , (- , +) , (+ , +) και οι n_c παρατηρήσεις στο κεντρικό σημείο (0,0). Ας θεωρήσουμε ως $\overline{y_F}$ την μέση τιμή από τις τέσσερις επαναλήψεις των παραγοντικών σημείων και ως $\overline{y_C}$ μέση τιμή των n_c επαναλήψεων στο κεντρικό σημείο. Αν η διαφορά $\overline{y_F}$ - $\overline{y_C}$ είναι μικρή , τότε το κεντρικά σημεία σχετίζονται ή είναι κοντά στο σχέδιο που περνάει ανάμεσα στα παραγοντικά σημεία και δεν υπάρχει καμπυλότητα. Από την άλλη αν αυτή η διαφορά είναι μεγάλη, τότε η καμπυλότητα είναι παρούσα. Με ένα βαθμό ελευθερίας το άθροισμα των ελαχίστων τετραγώνων της καθαρής τετραγωνικής καμπυλότητας δίνεται από τον τύπο

$$SS_{Pure\ quadratic} = \frac{n_F n_C(\overline{y_F} - \overline{y_C})}{n_F + n_C}$$

όπου n_F είναι το πλήθος των σημείων του παραγοντικού σχεδιασμού. Αυτό το άθροισμα μπορεί να ενσωματωθεί μέσα στον πίνακα ANOVA και μπορεί να συγκριθεί με το μέσο τετραγωνικό σφάλμα ελέγχοντας για την καθαρή τετραγωνική επίδραση. Ειδικότερα, όταν τα σημεία που προστίθενται στο κέντρο του σχεδιασμού 2^k για τον έλεγχο της καμπυλότητας πραγματοποιούνται με τις υποθέσεις

$$H_0: \sum_{j=1}^{k} \beta_{jj} = 0 \text{ έναντι της εναλλακτικής } H_1: \sum_{j=1}^{k} \beta_{jj} \neq 0$$

Επιπλέον, αν στον παραγοντικό σχεδιασμό έχουμε μία επανάληψη, τότε μία χρήση των κεντρικών σημείων είναι να κατασκευάσουν μία εκτίμηση του σφάλματος με n_c -1 βαθμούς ελευθερίας. Ας δούμε την περίπτωση που έχουμε δύο παράγοντες στον σχεδιασμό, υποθέτοντας ότι η καμπυλότητα είναι σημαντική, έτσι το μοντέλο μας θα είναι δεύτερης τάξης και θα είναι το εξής

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} {x_1}^2 + \beta_{22} {x_2}^2 + \varepsilon$$

Ένα πρόβλημα που δημιουργείται είναι ότι δεν μπορούν να εκτιμηθούν οι άγνωστοι παράμετροι (τα β δηλαδή), καθώς στο μοντέλο υπάρχουν έξι παράμετροι να εκτιμήσουμε και το σύνολο των ανεξάρτητων δοκιμών είναι 5, όπου 4 προέρχονται από τον σχεδιασμό 2^2 και μία είναι το κεντρικό σημείο.

Μία απλή και πολύ αποτελεσματική λύση στο πρόβλημα είναι να χρησιμοποιήσουμε τους κεντρικούς σύνθετους σχεδιασμούς (central composite designs-CCD). Οι CCD είναι οι πιο κοινά χρησιμοποιημένοι σχεδιασμοί στις αποκριτικές επιφάνειες. Είναι κατάλληλοι όταν θέλουμε να ταιριάξουμε ένα τετραγωνικό μοντέλο ή ένα μοντέλο δεύτερης τάξης. Ένας σχεδιασμός τέτοιας μορφής αποτελείται από τρία μέρη:

1. Τα n_f που απαρτίζονται από το πλήθος των σημείων που έχουν σαν τιμές $x_i = -1,-1$ *για* $i=1,2,\ldots k$.
2. Τα $2k$ αξονικά σημεία(n_a) που έχουν την μορφή $(0 ,\ldots\ldots, x_i , \ldots\ldots, 0)$ και $x_i = \alpha$ ή $-\alpha$ για $i=1,2,\ldots k$.
3. Το κεντρικά σημεία n_c όπου $x_i =0$ α για $i=1,2,\ldots k$.

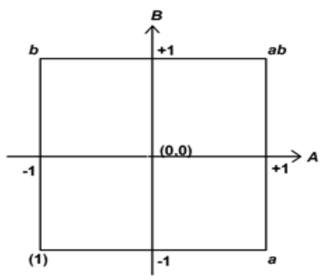
Όσον αφορά τα αξονικά σημεία, έχουν την ιδιότητα στον άξονα της κάθε μεταβλητής να γίνεται επιλογή δύο σημείων σε απόστασή α από το κέντρο του σχεδιασμού. Η αναπαράσταση ενός κεντρικού σύνθετου σχεδιασμού έχει είτε κυβοειδή μορφή (για α=1) είτε σφαιρική. Τα σημεία που προκύπτουν από τον παραγοντικό σχεδιασμό οδηγούν σε ένα σχεδιασμό βέλτιστης διασποράς για ένα μοντέλο πρώτης τάξης και συνεισφέρουν σημαντικά στην εκτίμηση των όρων της αλληλεπίδρασης. Τα κεντρικά σημεία είναι εκείνα που δίνουν στοιχεία σχετικά με την καμπυλότητα της συνάρτησης απόκρισης. Εφόσον, υπάρχει ένδειξη κυρτότητας προσθέτουμε τα αξονικά σημεία για την εκτίμηση των δευτεροβάθμιων όρων, αφού χωρίς αυτά έχουμε πληροφορία μόνο για το άθροισμά τους.

Τα κεντρικά σημεία μπορούν διασκορπίζονται όσο κατά το δυνατόν ομοιόμορφα σε ολόκληρο τον πίνακα σχεδιασμού. Το πλήθος των κεντρικών σημείων που προσθέτονται είναι γενικά κατά προσέγγιση 3 έως 5 επαναλήψεις σε ένα πλήρη ή κλασματικό παραγοντικό σχεδιασμό. Ένα πράγμα που χρειάζεται να τονίσουμε για το πλήθος των κεντρικών σημείων που πρέπει να εισαχθούν είναι ότι με τα κεντρικά σημεία συνηθίζουμε να εκτιμούμε το «καθαρό» σφάλμα για τον έλεγχο καλής προσαρμογής. Ο έλεγχος καλής προσαρμογής μας δείχνει κατά πόσο τα δεδομένα ταιριάζουν καλά με το μοντέλο που προσαρμόσαμε. Στην περίπτωση που έχουμε λιγότερες από 5 ή 6 επαναλήψεις, ο έλεγχος θα έχει πολύ χαμηλή ισχύ.

3.2 Ο 3^k Παραγοντικός σχεδιασμός

Οι παραγοντικοί σχεδιασμοί δύο επιπέδων είτε πλήρεις είτε κλασματικοί που συζητήθηκαν στο προηγούμενο Κεφάλαιο είναι χρήσιμοι στην έρευνα και στην ανάπτυξη των προϊόντων επιχειρήσεων και οργανισμών. Αν και είναι απλοί σχεδιασμοί, όπως είδαμε δεν μπορούν να ελέγξουν πιθανή καμπυλότητα στην απόκριση. Οι πειραματιστές αρκετές φορές ενδιαφέρονται στην μελέτη παραγόντων με περισσότερα από δύο επίπεδα. Σε αυτή την παράγραφο θα περιγράψουμε τους παραγοντικούς σχεδιασμούς τριών επιπέδων, καθώς και με τους αντίστοιχους κλασματικούς, όπου συμβολίζονται 3^k και 3^{k-p} αντίστοιχα . Οι σχεδιασμοί αυτοί μας επιτρέπουν την μελέτη και των τετραγωνικών επιδράσεων, παρότι ίσως χρειαστούν πολλές εκτελέσεις για να πραγματοποιηθούν, δοθέντος της επιλογής των τιμών που θα γίνουν στα k και p. Εδώ σαν συμβολισμούς χρησιμοποιούμε "-1" την χαμηλή στάθμη, "0" την μεσαία στάθμη, "+1" για υψηλή στάθμη. Κάποιοι πειραματιστές χρησιμοποιούν και τους συμβολισμούς "0","1","2" αντιστοιχίζοντας κατά αύξουσα σειρά και τις στάθμες τους παράγοντα.

Στις τρεις διαστάσεις η περιοχή του σχεδιασμού γίνεται ένας κύβος, ενώ από τους τέσσερις παράγοντες και πάνω είναι δύσκολο να παρασταθεί γραφικά.



Σχήμα 3.2.2: Περίπτωση 2 παραγόντων με 3 επίπεδα

3.2.1 Ο 3^2 Παραγοντικός σχεδιασμός

Ο πιο απλός 3^k πλήρης παραγοντικός σχεδιασμός αφορά τη μελέτη 2 παραγόντων και είναι ο 3^2 σχεδιασμός. Ο σχεδιασμός αυτός έχει 9 θεραπείες, που προσφέρουν 8 βαθμούς ελευθερίας για εκτίμηση παραγοντικών επιδράσεων. Οι κύριες επιδράσεις των παραγόντων έχουν από 2 βαθμούς ελευθερίας και η αλληλεπίδραση τους έχει 4 βαθμούς ελευθερίας. Αν γίνουν n

επαναλήψεις του σχεδιασμού, τότε προκύπτουν 3^2(n-1) βαθμοί ελευθερίας για την εκτίμηση του σφάλματος.

Για την ανάλυση με τη μέθοδο της γραμμικής παλινδρόμησης και για ποσοτικούς παράγοντες, κάθε κύρια επίδραση αντιπροσωπεύεται από δύο συνιστώσες με έναν βαθμό ελευθερίας η καθεμία: μια γραμμική και μία τετραγωνική επίδραση. Ομοίως, η αλληλεπίδραση των δύο παραγόντων μπορεί να διασπαστεί σε τέσσερις διαφορετικές συνιστώσες με έναν βαθμό ελευθερίας η καθεμία τους. Αυτές οι συνιστώσες είναι αφορούν την γραμμική x γραμμική επίδραση, την γραμμική x τετραγωνική επίδραση, την τετραγωνική x γραμμική επίδραση και την τετραγωνική x τετραγωνική επίδραση. Επίσης μπορούμε να αντιστοιχίσουμε τις άνωθι επιδράσεις που αναφέρθηκαν με τους αντίστοιχους όρους $\beta_{12} x_1 x_2$,$\beta_{122} x_1 x_2^2$, $\beta_{112} x_1^2 x_2$, και $\beta_{1122} x_1^2 x_2^2$ ενός γραμμικού μοντέλου πολλαπλής παλινδρόμησης.

3.2.2 Ο 3^3 Παραγοντικός σχεδιασμός

Έστω ότι τώρα έχουμε 3 παράγοντες για μελέτη όπου κάθε παράγοντας έχει τρία επίπεδα. Οι 27 θεραπείες που προκύπτουν προσφέρουν 26 βαθμούς ελευθερίας. Κάθε κύρια επίδραση έχει 2 βαθμούς ελευθερίας, σε κάθε αλληλεπίδραση δύο παραγόντων αντιστοιχούν 4 βαθμοί ελευθερίας και η αλληλεπίδραση των τριών παραγόντων έχει 8 βαθμούς ελευθερίας. Εάν το πείραμα μας έχει n επαναλήψεις τότε οι βαθμοί ελευθερίας του σφάλματος είναι $3^3(n\text{-}1)$ και το σύνολο των βαθμών ελευθερίας είναι $3^3n\text{-}1$.

Και σε αυτή την περίπτωση, αν οι παράγοντες είναι ποσοτικοί, οι κύριες επιδράσεις διαχωρίζονται στη γραμμική και την τετραγωνική τους συνιστώσα, με έναν βαθμό ελευθερίας για την καθεμία. Οι αλληλεπιδράσεις δύο παραγόντων επίσης διασπώνται στις γραμμική x γραμμική, γραμμική x τετραγωνική, τετραγωνική x γραμμική και τετραγωνική x τετραγωνική συνιστώσες τους, επίσης με έναν βαθμό ελευθερίας η καθεμία. Τέλος, η αλληλεπίδραση των τριών παραγόντων διασπάται επίσης, με παρόμοια λογική (γραμμική x γραμμική x γραμμική, γραμμική x γραμμική x τετραγωνική κοκ), σε 8 συνιστώσες που επίσης έχουν ένα βαθμό ελευθερίας. Όμως, γενικά η διάσπαση της αλληλεπίδρασης των τριών παραγόντων δεν είναι χρήσιμη.

3.2.3 Ο Γενικός 3^k Παραγοντικός σχεδιασμός

Στις προηγούμενες ενότητες έγιναν αναφορές για τους 3^k παραγοντικούς σχεδιασμούς στις περιπτώσεις που έχουμε στον σχεδιασμό 2 ή 3 παράγοντες. Η διαδικασία αυτή γενικεύεται και στις περιπτώσεις που έχουμε k παράγοντες. Οι συνολικές εκτελέσεις είναι ίσες με 3^k και δίνουν 3^k-1 βαθμούς ελευθερίας. Υπάρχουν k κύριες επιδράσεις με 2 βαθμούς ελευθερίας η καθεμία, $\binom{k}{2}$ αλληλεπιδράσεις δύο παραγόντων, με 4 βαθμούς ελευθερίας η καθεμία, κ.ο.κ., και μία αλληλεπίδραση k παραγόντων με 2^k βαθμούς ελευθερίας. Γενικά μία αλληλεπίδραση z παραγόντων έχει 2^z βαθμούς ελευθερίας. Στην περίπτωση που έχουμε επαναλήψεις τότε προκύπτουν $3^k(n\text{-}1)$ βαθμούς ελευθερίας για το σφάλμα και $3^kn\text{-}1$ βαθμούς ελευθερίας συνολικά.

Σαν ένα γενικό κανόνα μοντελοποίησης αναφέρουμε ότι οι κύριες επιδράσεις και οι αλληλεπιδράσεις διασπώνται σε συνιστώσες με έναν βαθμό ελευθερίας η καθεμιά. Κάθε αλληλεπίδραση h παραγόντων έχει 2^h συνιστώσες. Οι συνιστώσες των αλληλεπιδράσεων υψηλής τάξης είναι δύσκολο και ερμηνευτούν και συνήθως δεν έχουν κάποια φυσική ερμηνεία.

Σημειώνεται ότι το μέγεθος των σχεδιασμού αυξάνεται ραγδαία, όσο αυξάνεται το k. Για παράδειγμα, για k=3 έχουμε 27 θεραπείες για κάθε πλήρη επανάληψη, για k=4 έχουμε 81 θεραπείες κ.ο.κ. Συνήθως λοιπόν στην πράξη γίνεται μία μόνο επανάληψη του 3^k πλήρους παραγοντικού σχεδιασμού και οι υψηλότερης τάξης αλληλεπιδράσεις χρησιμοποιούνται για τη δημιουργία ψευδοσφάλματος ώστε να έχουμε μια εκτίμηση του σφάλματος για τους ελέγχους. Για παράδειγμα, αν οι αλληλεπιδράσεις τριών παραγόντων και άνω θεωρηθούν ασήμαντες, τότε από μία μόνο επανάληψη του 3^3 πλήρους παραγοντικού σχεδιασμού «προκύπτουν» 8 βαθμοί ελευθερίας για το σφάλμα, ενώ για από μία μόνο επανάληψη του 3^4 πλήρους παραγοντικού σχεδιασμού «προκύπτουν» 48 βαθμοί ελευθερίας για το σφάλμα. Έτσι και υπό την αρχή της ιεραρχίας των επιδράσεων, για μεγάλο αριθμό παραγόντων οι πλήρεις σχεδιασμοί δεν είναι τόσο χρήσιμοι.

3.3 Ο Κλασματικός 3^{k-p} παραγοντικός σχεδιασμός

Όπως στις περιπτώσεις που μελετάμε παράγοντες με δύο επίπεδα υπάρχουν κλάσματα του πλήρους σχεδιασμού προς χρήση, έτσι και τώρα, στην περίπτωση που έχουμε παράγοντες με τρία επίπεδα, μπορούμε να κατασκευάσουμε κλασματικούς παραγοντικούς σχεδιασμούς, που συμβολίζονται με το σύμβολο 3^{k-p}. Αυτό γίνεται, όπως και στην περίπτωση των δύο επιπέδων, για την οικονομία του πειράματος τόσο σε πόρους όσο και σε χρόνο. Η διαδικασία για την κατασκευή ενός 3^{k-p} κλασματικού παραγοντικού σχεδιασμού είναι παρόμοια με αυτή που κατασκευάζονται οι κλασματικοί σχεδιασμοί δύο επιπέδων. Η δομή των ταυτόσημων επιδράσεων στην περίπτωση αυτή, είναι βέβαια πολυπλοκότερη.

ΚΕΦΑΛΑΙΟ 4

Definitive screening Designs

4.1 Definitive screening Designs

Οι Jones και Nachtsheim πρότειναν τη χρήση των Definitive screening Designs (DSD) για χρήση σε πειράματα κρησαρίσματος, σαν οικονομικότερες και συνάμα αποδοτικές επιλογές σχεδιασμού από τους 3^k παραγοντικούς σχεδιασμούς. Οι DSD μελετούν τους παράγοντες σε τρία επίπεδα, χρησιμοποιώντας συνολικά $2k+1$ εκτελέσεις και συνεπώς απαιτούν μικρότερο αριθμό εκτελέσεων από τους παραδοσιακούς σχεδιασμούς τριών επιπέδων. Είναι ιδανικοί σε περιπτώσεις μελέτης ποσοτικών παραγόντων, ενώ δεν είναι τόσο χρήσιμοι όταν έχουμε κατηγορικούς παράγοντες. Στην εργασία τους, δίνονται σχεδιασμοί με τουλάχιστον 6 παράγοντες και προτείνεται, στις περιπτώσεις που έχουμε λιγότερους, να αφαιρούνται μία ή δύο στήλες χρησιμοποιώντας από τον DSD για 6 παράγοντες. Οι DSD μας επιτρέπουν να κάνουμε έλεγχο και για ύπαρξη κάποιας καμπυλότητας

Για την κατασκευή των DSD χρησιμοποιείται μια διαδικασία που καλείται fold-over διαδικασία. Σαν fold over διαδικασία, και συγκεκριμένα για την κατασκευή DSD, εννοούμε την δημιουργία ενός πίνακα που έχει την παρακάτω δομή,

$$M=\begin{pmatrix} A \\ -A \\ \mathbf{0} \end{pmatrix}$$

όπου A είναι ένας πίνακας διαστάσεων $k\ x\ k$ που στην διαγώνιο έχει μηδενικά και στις υπόλοιπες θέσεις έχει είτε την τιμή -1 ή την τιμή 1. Ο πίνακας $-A$ προκύπτει από τον A με αλλαγή των πρόσημων. Τέλος το $\mathbf{0}$ είναι το μηδενικό διάνυσμα που περιέχει μόνο μηδενικά που είναι διάστασης $n_c\ x\ k$, όπου n_c είναι το πλήθος τον κεντρικών σημείων που θέλουμε να προσθέσουμε. Ιδανικά, ο A πρέπει να είναι ορθογώνιος πίνακας, δηλαδή όταν υπολογιστεί ο A^TA θα πρέπει να προκύψει ένας διαγώνιος πίνακας. Η τεχνική foldover και η ορθογωνιότητα του A έχει ως συνέπεια τις εξής ιδιότητες σύγχυσης των παραγοντικών επιδράσεων:

1. Οι κύριες επιδράσεις είναι ανεξάρτητες και μεταξύ τους και από αλληλεπιδράσεις δύο παραγόντων .
2. Οι κύριες επιδράσεις είναι ανεξάρτητες από τις τετραγωνικές επιδράσεις.

Ας δώσουμε ένα παράδειγμα για καλύτερη κατανόηση του σχεδιασμού για την περίπτωση που k=6. Ο πίνακας A που χρησιμοποιείται προκύπτει από τη μέθοδο των Jones & Nachtsheim (2011) για τον συγκεκριμένο αριθμό παραγόντων. Ο πίνακας 4.1.1 παρουσιάζει τον A και ο πίνακας 4.1.2 τον A^TA, όπου και φαίνεται ότι είναι ένας διαγώνιος πίνακας.

$$A=\begin{bmatrix} 0 & +1 & -1 & -1 & -1 & -1 \\ +1 & 0 & -1 & +1 & +1 & -1 \\ -1 & -1 & 0 & +1 & -1 & -1 \\ -1 & +1 & +1 & 0 & +1 & -1 \\ +1 & -1 & +1 & -1 & 0 & -1 \\ +1 & +1 & +1 & +1 & -1 & 0 \end{bmatrix}$$

Πίνακας 4.1.1

$$A^TA=\begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

Πίνακας 4.1.2

Αυτό σημαίνει ότι και ο πίνακας M είναι ορθογώνιος και η χρήση του ως DSD θα μας εφοδιάσει με τις επιθυμητές ιδιότητες σύγχυσης των επιδράσεων που αναφέρθηκαν παραπάνω.

Στον Πίνακα 4.1.3 παρουσιάζεται η γενική δομή των Definitive Screening Designs για k παράγοντες. Για τους παράγοντες χρησιμοποιείται ο συμβολισμός $X_{i,j}$, j=1,2,…, k και οι τιμές που μπορούν να πάρουν είναι -1, 1, 0. Για k παράγοντες έχουμε $2k$+1 εκτελέσεις στο πείραμα. Ο Πίνακας 4.1.3 είναι λίγο διαφορετικά σχεδιασμένος σε σχέση με ότι περιγράψαμε πριν αλλά δεν έχει καμία διαφορά με όσα περιεγράφηκαν.

Foldover pair	*Run (i)*	*Επίπεδα παραγόντων*				
		$X_{i,1}$	$X_{i,2}$	$X_{i,3}$	⋯	$X_{i,k}$
1	1	0	±1	±1	⋯	±1
	2	0	∓1	∓1	⋯	∓1
2	3	±1	0	±1	⋯	±1
	4	∓1	0	∓1	⋯	∓1
3	5	±1	±1	0	⋯	±1
	6	∓1	∓1	0	⋯	∓1
⋮	⋮	⋮	⋮	⋮	⋱	⋮
m	2k-1	±1	±1	±1	⋯	0
	2k	∓1	∓1	∓1	⋯	0
Κεντρικό σημείο	2k+1	0	0	0	0	0

Πίνακας 4.1.3

4.2 Κατασκευή σχεδιασμών

Οι Jones και Nachtsheim (2011) κατασκεύασαν DSD μέσω της χρήσης ενός αλγορίθμου, μεγιστοποιεί την ορίζουσα του πίνακα πληροφορίας του μοντέλου των κύριων επιδράσεων, με ταυτόχρονη εφαρμογή της foldover δομής. Στις περιπτώσεις όπου το πλήθος των παραγόντων k είναι περιττό, όπως και στην περίπτωση που το k=12, δεν προέκυψαν ορθογώνιοι σχεδιασμοί. Σε αυτές τις περιπτώσεις παρουσιάστηκαν οι βέλτιστοι σχεδιασμοί που προέκυψαν. Στους παρακάτω πίνακες 4.1.1 – 4.1.9 παρουσιάζονται οι DSD σχεδιασμοί των Jones και Nachtsheim με 4 έως 12 παράγοντες.

	k=4				*k*=5					*k*=6						*k*=7						
	Παράγοντες				Παράγοντες					Παράγοντες						Παράγοντες						
i	1	2	3	4	1	2	3	4	5	1	2	3	4	5	6	1	2	3	4	5	6	7
1	0	+	-	-	0	+	+	-	-	0	+	-	-	-	-	0	+	-	+	-	+	-
2	0	-	+	+	0	-	-	+	+	0	-	+	+	+	+	0	-	+	-	+	-	+
3	-	0	-	+	+	0	-	-	+	+	0	-	+	+	-	-	0	+	-	+	+	-
4	+	0	+	-	-	0	+	+	-	-	0	+	-	-	+	+	0	-	+	-	-	+
5	-	-	0	-	+	-	0	+	-	-	-	0	+	-	-	+	-	0	+	+	+	+
6	+	+	0	+	-	+	0	-	+	+	+	0	-	+	+	-	+	0	-	-	-	-
7	-	+	+	0	+	-	+	0	+	-	+	+	0	+	-	+	-	-	0	+	-	-
8	+	-	-	0	-	+	-	0	-	+	-	-	0	-	+	-	+	+	0	-	+	+
9	0	0	0	0	+	+	+	+	0	+	-	+	-	0	-	-	-	+	+	0	-	-
10					-	-	-	-	0	-	+	-	+	0	+	+	+	-	-	0	+	+
11					0	0	0	0	0	+	+	+	+	-	0	-	+	-	+	+	0	+
12										-	-	-	-	+	0	+	-	+	-	-	0	-
13										0	0	0	0	0	0	+	+	+	+	+	-	0
14																-	-	-	-	-	+	0
15																0	0	0	0	0	0	0
	Πίνακας 4.2.1				Πίνακας 4.2.2					Πίνακας 4.2.3						Πίνακας 4.2.4						

	k=8								*k*=9									*k*=10									
	Παράγοντες								Παράγοντες									Παράγοντες									
i	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	10
1	0	-	+	+	-	+	+	+	0	+	+	+	+	+	+	+	+	0	+	+	-	+	+	+	+	-	+
2	0	+	-	-	+	-	-	-	0	-	-	-	-	-	-	-	-	0	-	-	+	-	-	-	-	+	-
3	-	0	-	+	+	+	+	-	+	0	+	-	+	-	-	+	-	+	0	-	+	+	-	+	+	-	-
4	+	0	+	-	-	-	-	+	-	0	-	+	-	+	+	-	+	-	0	+	-	-	+	-	-	+	+
5	-	-	0	+	+	-	-	+	-	+	0	-	+	-	+	-	-	-	+	0	-	-	-	+	-	-	-
6	+	+	0	-	-	+	+	-	+	-	0	+	-	+	-	+	+	+	-	0	+	+	+	-	+	+	+
7	+	-	+	0	+	+	-	-	-	-	+	0	+	-	-	-	+	-	+	+	0	+	-	-	+	+	-
8	-	+	-	0	-	-	+	+	+	+	-	0	-	+	+	+	-	+	-	-	0	-	+	+	-	-	+
9	-	-	+	-	0	-	+	-	+	-	+	-	0	+	+	-	-	-	-	-	-	0	+	+	+	+	-
10	+	+	-	+	0	+	-	+	-	+	-	+	0	-	-	+	+	+	+	+	+	0	-	-	-	-	+
11	+	-	-	-	+	0	+	+	-	-	-	-	+	0	+	+	+	-	+	-	+	+	0	+	-	+	+
12	-	+	+	+	-	0	-	-	+	+	+	+	-	0	-	-	-	+	-	+	-	-	0	-	+	-	-
13	-	+	+	-	+	+	0	+	+	+	-	-	+	+	0	-	+	+	+	-	-	-	-	0	+	+	+
14	+	-	-	+	-	-	0	-	-	-	+	+	-	-	0	+	-	-	-	+	+	+	+	0	-	-	-
15	+	+	+	+	+	-	+	0	-	-	-	+	+	+	-	0	-	+	+	+	+	-	+	+	0	+	-
16	-	-	-	-	-	+	-	0	+	+	+	-	-	-	+	0	+	-	-	-	-	+	-	-	0	-	+
17	0	0	0	0	0	0	0	0	-	+	+	-	-	+	-	+	0	+	+	-	-	+	+	-	-	0	-
18									+	-	-	+	+	-	+	-	0	-	-	+	+	-	-	+	+	0	+
19									0	0	0	0	0	0	0	0	0	+	-	+	-	+	-	+	-	+	0
20																		-	+	-	+	-	+	-	+	-	0
21																		0	0	0	0	0	0	0	0	0	0
	Πίνακας 4.2.5								Πίνακας 4.2.6									Πίνακας 4.2.7									

	k=11											k=12											
	Παράγοντες											Παράγοντες											
i	1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9	10	11	12
1	0	-	+	-	-	-	-	-	+	-	+	0	-	-	+	-	+	-	+	+	+	-	+
2	0	+	-	+	+	+	+	+	-	+	-	0	+	+	-	+	-	+	-	-	-	+	-
3	-	0	-	-	+	-	-	-	-	+	+	-	0	+	+	+	+	+	+	+	-	-	-
4	+	0	+	+	-	+	+	+	+	-	-	+	0	-	-	-	-	-	-	-	+	+	+
5	-	-	0	+	+	+	+	-	-	-	+	+	+	0	-	+	+	-	+	+	-	+	+
6	+	+	0	-	-	-	-	+	+	+	-	-	-	0	+	-	-	+	-	-	+	-	-
7	-	-	-	0	-	+	+	-	+	+	-	+	-	-	0	+	-	+	-	+	-	-	+
8	+	+	+	0	+	-	-	+	-	-	+	-	+	+	0	-	+	-	+	-	+	+	-
9	+	-	-	+	0	+	-	+	+	+	+	+	+	+	+	0	-	+	+	+	+	+	+
10	-	+	+	-	0	-	+	-	-	-	-	-	-	-	-	0	+	-	-	-	-	-	-
11	-	-	+	+	-	0	-	+	-	+	-	+	-	+	-	+	0	+	+	-	+	-	+
12	+	+	-	-	+	0	+	-	+	-	+	-	+	-	+	-	0	-	-	+	-	+	-
13	-	-	-	+	+	-	0	+	+	-	-	+	+	+	+	-	+	0	-	-	-	-	+
14	+	+	+	-	-	+	0	-	-	+	+	-	-	-	-	+	-	0	+	+	+	+	-
15	-	+	+	+	-	-	+	0	+	+	+	-	-	+	+	+	-	-	0	-	-	+	+
16	+	-	-	-	+	+	-	0	-	-	-	+	+	-	-	-	+	+	0	+	+	-	-
17	-	+	-	-	-	+	-	+	0	-	+	+	-	+	+	+	+	-	-	0	+	+	-
18	+	-	+	+	+	-	+	-	0	+	-	-	+	-	-	-	-	+	+	0	-	-	+
19	+	-	-	-	-	-	+	+	-	0	+	+	+	-	+	+	-	-	+	-	0	-	-
20	-	+	+	+	+	+	-	-	+	0	-	-	-	+	-	-	+	+	-	+	0	+	+
21	+	+	-	+	-	-	-	-	-	-	0	-	+	-	+	+	+	+	-	-	+	0	+
22	-	-	+	-	+	+	+	+	+	+	0	+	-	+	-	-	-	-	+	+	-	0	-
23	0	0	0	0	0	0	0	0	0	0	0	+	-	-	+	-	+	+	+	-	-	+	0
24												-	+	+	-	+	-	-	-	+	+	-	0
25												0	0	0	0	0	0	0	0	0	0	0	0
	Πίνακας 4.2.8											Πίνακας 4.2.9											

4.3 Ιδιότητες Σχεδιασμών

Οι ερευνητές χρησιμοποιούν συχνά τους σχεδιασμούς κρησαρίσματος (screening designs) όταν θέλουν να ταυτοποιήσουν ποιοι παράγοντες επηρεάζουν σημαντικά την μεταβλητή απόκρισης που εξετάζουν σε μια διεργασία. Οι ορθογώνιοι Definitive screening designs προσφέρουν περισσότερα πλεονεκτήματα από ότι οι καθιερωμένοι σχεδιασμοί screening. Ένα σημαντικό τους πλεονέκτημα είναι το γεγονός ότι αποφεύγεται η σημαντική σύγχυση των επιδράσεων ενώ παράλληλα γίνεται εφικτή η ανάδειξη των παραγόντων αυτών που έχουν μη γραμμική επίδραση στην μεταβλητή απόκρισης. Ένα επίσης πλεονέκτημα που έχουν οι DSD είναι η ανεξαρτησία των κύριων επιδράσεων με τις αλληλεπιδράσεις δύο παραγόντων αλληλεπιδράσεις, κάτι που συνήθως δεν ισχύει στους καθιερωμένους σχεδιασμούς κρησαρίσματος (κάτι τέτοιο συναντάται σε σχεδιασμούς που έχουν διακριτική ικανότητα V ή υψηλότερη). Επιπλέον, στους DSD καμία από τις αλληλεπιδράσεις δύο παραγόντων δεν συγχέεται πλήρως με άλλες αλληλεπιδράσεις δύο παραγόντων και τέλος, στους DSD πάντα μπορεί να εκτιμηθούν και όλες οι τετραγωνικές επιδράσεις ταυτόχρονα με τις κύριες επιδράσεις. Αυτό συμβαίνει διότι οι τετραγωνικές επιδράσεις είναι ορθογώνιες με τις κύριες επιδράσεις, πράγμα που δεν γίνεται να συμβεί στους απλούς σχεδιασμούς κρησαρίσματος με δύο επίπεδα. Αναφορικά με τις τετραγωνικές επιδράσεις, είναι χρήσιμο να αναφερθεί ότι στους DSD δεν συγχέονται πλήρως με τις αλληλεπιδράσεις δύο παραγόντων.

Η χρησιμότητα των DSD είναι πολύ σημαντική στον πειραματικό σχεδιασμό, καθώς αποτρέπει στο μοντέλο να υπάρξει ασάφεια, δίνοντας την δυνατότητα να αναδείξει ποιοι παράγοντες είναι σημαντικοί πιο γρήγορα και πιο αποδοτικά. Επιπλέον, αναδεικνύει ποιοι από αυτούς έχουν μη γραμμική επίδραση, ενώ ταυτόχρονα απομακρύνεται η σύγχυση στους όρους της δεύτερης τάξης. Έτσι όχι μόνο μπορούμε να εντοπίσουμε μη γραμμικότητα στην απόκριση, που στους απλούς σχεδιασμούς κρησαρίσματος με δύο επίπεδα κάτι τέτοιο μπορεί να συμβεί με τη χρήση των κεντρικών σημείων, αλλά και να βρούμε ποιοι παράγοντες είναι υπεύθυνοι για την μη γραμμικότητα.

Για να δούμε όσα περιεγράφηκαν προηγουμένως πιο αναλυτικά, ας υποθέσουμε ότι χρησιμοποιούμε έναν ορθογώνιο DSD και το «πραγματικό» μοντέλο είναι το πλήρες τετραγωνικό μοντέλο και γράφεται στην παρακάτω μορφή

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon,$$

όπου X_1 είναι ένας πίνακας $(2k+1)\ x\ k$ και περιέχει τους όρους για τις γραμμικές κύριες επιδράσεις και το X_2 είναι ένας πίνακας με διαστάσεις $(2k + 1) \times (1 + k(k + 1)/2)$ και περιέχει όλους τους υπόλοιπους όρους (τετραγωνικούς και αλληλεπιδράσεις). Αν ο αναλυτής εργαστεί μόνο με τις κύριες επιδράσεις τότε θα χρησιμοποιήσει το μοντέλο $y = X_1\beta_1 + \varepsilon^*$, όπου το β_1 εκτιμάται από την μέθοδο των ελαχίστων τετραγώνων και η αναμενόμενη τιμή του εκτιμάται από $E(\widehat{\beta_1}) = \beta_1 + A\beta_2$, όπου ο A υπολογίζεται από τον τύπο $A = (X_1^T X_1)^{-1} X_1^T X_2$. Στους DSD προκύπτει ότι A=0 και συνεπώς $E(\widehat{\beta_1}) = \beta_1$.

4.4 Οι DSD σχεδιασμοί των Xiao et al.

Ένα ζήτημα που υπάρχει στους DSD που πρότειναν οι Jones & Nachtsheim είναι ότι δεν ισχύει η ορθογωνιότητα των στηλών για όλους τους σχεδιασμούς που δόθηκαν. Οι Xiao at el (2012) προτείνουν τη χρήση των Conference πινάκων για την κατασκευή DSD. Οι σχεδιασμοί που προκύπτουν έχουν το ίδιο μέγεθός και όλες τις επιθυμητές ιδιότητες των αντίστοιχων των Jones & Nachtsheim, όμως επιπλέον, είναι πάντα ορθογώνιοι. Ειδικότερα, αυτοί οι σχεδιασμοί είναι κορεσμένοι σχεδιασμοί για την εκτίμηση του σταθερού όρου, όλων των κύριων επιδράσεων, καθώς και όλων τετραγωνικών επιδράσεων. Επιπλέον εξασφαλίζουν ότι όλες οι κύριες επιδράσεις είναι ορθογώνιες με τις τετραγωνικές επιδράσεις, καθώς και με τις αλληλεπιδράσεις δύο παραγόντων.

Οι Conference πίνακες συμβολίζονται με *C* και έχουν την ιδιότητα να περιέχουν μόνο μηδενικά στη διαγώνιο και εκτός αυτής στοιχεία που είναι ή -1 ή +1. Οι πίνακες αυτοί είναι τετραγωνικοί διάστασης *n*, όσοι θα είναι και οι παράγοντες που τελικά θα υπάρχουν στον DSD σχεδιασμό. Σε αυτούς τους πίνακες ισχύει πάντα ότι $C^T C = (n\text{-}1)\ I_n$ όπου I_n ο μοναδιαίος πίνακας. Συνεπώς, οι στήλες των Conference πινάκων είναι ορθογώνιες. Για να δημιουργήσουμε DSD σχεδιασμούς χρησιμοποιούμε την τεχνική foldover, δημιουργώντας έναν νέο πίνακα της μορφής

$$\begin{pmatrix} C \\ -C \\ 0 \end{pmatrix}$$

Για παράδειγμα και για καλύτερη κατανόηση, παρακάτω φαίνεται η κατασκευή ενός DSD με 21 γραμμές και 10 στήλες, χρησιμοποιώντας ένα 10 x 10 Conference πίνακα *C*. Η προηγούμενη σχέση θα ισούται $C^T C = 9\ I_{10}$ και ο σχεδιασμός θα είναι ως εξής:

Μέρος	***Επανάληψη***	***X1***	***X2***	***X3***	***X4***	***X5***	***X6***	***X7***	***X8***	***X9***	***X10***
C	***1***	0	1	1	1	1	1	1	1	1	1
	2	1	0	-1	-1	-1	-1	1	1	1	1
	3	1	-1	0	-1	1	1	-1	-1	1	1
	4	1	-1	-1	0	1	1	1	1	-1	-1
	5	1	-1	1	1	0	-1	-1	1	-1	1
	6	1	-1	1	1	-1	0	1	-1	1	-1
	7	1	1	-1	1	-1	1	0	-1	-1	1
	8	1	1	-1	1	1	-1	-1	0	1	-1
	9	1	1	1	-1	-1	1	-1	1	0	-1
	10	1	1	1	-1	1	-1	1	-1	-1	0
-C	***-10***	-1	-1	-1	1	-1	1	-1	1	1	0
	-9	-1	-1	-1	1	1	-1	1	-1	0	1
	-8	-1	-1	1	-1	-1	1	1	0	-1	1
	-7	-1	-1	1	-1	1	-1	0	1	1	-1
	-6	-1	1	-1	-1	1	0	-1	1	-1	1
	-5	-1	1	-1	-1	0	1	1	-1	1	-1
	-4	-1	1	1	0	-1	-1	-1	-1	1	1
	-3	-1	1	0	1	-1	-1	1	1	-1	-1
	-2	-1	0	1	1	1	1	-1	-1	-1	-1
	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
0	***11***	0	0	0	0	0	0	0	0	0	0

Πίνακας 4.4.1 Παράδειγμα κατασκευής DSD Σχεδιασμού με την χρήση Conference Πίνακα

ΚΕΦΑΛΑΙΟ 5

Αποδοτικότητα για τους DSD σχεδιασμούς των Jones & Nachtsheim

5.1 Αποδοτικότητα σχεδιασμών

Για τον υπολογισμό της αποδοτικότητας των σχεδιασμών DSD των Jones & Nachtsheim χρησιμοποιούμε ένα μέτρο που μας δείχνει την αποδοτικότητα της εκτίμησης των παραμέτρων ενός δοθέντος μοντέλου. Το μέτρο αυτό είναι η D-αποδοτικότητα ή D-efficiency. Ο τύπος υπολογισμού της αποδοτικότητας είναι ο

$$D_{efficiency} = |W^T W|^{1/p},$$

όπου W είναι ο κανονικοποιημένος πίνακας του γραμμικού μοντέλου που αξιολογείται. Αν λοιπόν ο πίνακας μοντέλου συμβολιστεί με X, ο πίνακας W προκύπτει διαιρώντας το κάθε στοιχείο της κάθε στήλης του X με το μέτρο της. Το μέτρο κάθε στήλης είναι η ρίζα του αθροίσματος κάθε στοιχείου στα τετράγωνο, δηλαδή αν κάθε στοιχείο της στήλης συμβολίζεται με c_j τότε $\|c\| = \sqrt{\sum_{j=1}^{n} c_j^{2}}$.

Το διάστημα τιμών που μπορεί να πάρει η D-efficiency είναι από 0 έως 1 και μπορεί να εκφραστεί και σε μορφή ποσοστού. Προφανώς, όσο πιο κοντά στο 1 προκύψει τιμή αποδοτικότητας, τόσο καλύτερα.

5.2 Αποδοτικότητα των DSD σχεδιασμών για το τετραγωνικό μοντέλο

Σε ό,τι ακολουθεί θα αξιολογήσουμε τους σχεδιασμούς των Jones και Nachtsheim βρίσκοντας την αποδοτικότητα των σχεδιασμών που περιέχουν από k=4 έως 12 παράγοντες και παρουσιάστηκαν στο τέταρτο Κεφάλαιο. Το μοντέλο το οποίο θα αξιολογηθεί υπολογίζοντας την αποδοτικότητα των σχεδιασμών είναι το τετραγωνικό μοντέλο

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \beta_{jj} x_j^{2} + \varepsilon$$

Από αυτό το μοντέλο θα προκύπτει ο πίνακας μοντέλου που θα είναι διαστάσεων $(2k+1)x(2k+1)$. Παρακάτω θα παρουσιαστεί αναλυτικά όλη η διαδικασία για τον υπολογισμό της αποδοτικότητας για τις περιπτώσεις που έχουμε 4 και 5 παράγοντες.

5.2.1 Αποδοτικότητα του σχεδιασμού DSD με 4 παράγοντες

Ο πίνακας σχεδιασμού του DSD με 4 παράγοντες φαίνεται στον Πίνακα 4.2.1. Το τετραγωνικό μοντέλο που προκύπτει σε αυτήν την περίπτωση έχει 9 παραμέτρους και είναι το

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} {X_1}^2 + \beta_{22} {X_2}^2 + \beta_{33} {X_3}^2 + \beta_{44} {X_4}^2 + \varepsilon$$

Ο αντίστοιχος πίνακας μοντέλου X είναι ένας πίνακα διάστασης 9 x 9 και δείχνει τις τιμές που παίρνει κάθε τυχαία μεταβλητή που είναι -1,0,1 και όπου X_0 θεωρούμε μια μεταβλητή που παίρνει τιμή 1 και είναι η αντίστοιχη για την παράμετρο β_0. Ο πίνακας X ακολουθεί.

$$X=\begin{array}{ccccccccc}
X_0 & X_1 & X_2 & X_3 & X_4 & X_1^2 & X_2^2 & X_3^2 & X_4^2 \\
+1 & 0 & +1 & -1 & -1 & 0 & +1 & +1 & +1 \\
+1 & 0 & -1 & +1 & +1 & 0 & +1 & +1 & +1 \\
+1 & -1 & 0 & -1 & +1 & +1 & 0 & +1 & +1 \\
+1 & +1 & 0 & +1 & -1 & +1 & 0 & +1 & +1 \\
+1 & -1 & -1 & 0 & -1 & +1 & +1 & 0 & +1 \\
+1 & +1 & +1 & 0 & +1 & +1 & +1 & 0 & +1 \\
+1 & -1 & +1 & +1 & 0 & +1 & +1 & +1 & 0 \\
+1 & +1 & -1 & -1 & 0 & +1 & +1 & +1 & 0 \\
+1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Για τον υπολογισμό του W, βρίσκουμε το μέτρο της πρώτη στήλης που είναι ίσο με 3, ενώ για τις υπόλοιπες στήλες είναι ίσο με $\sqrt{6}$.

$$W=\begin{array}{ccccccccc}
1/3 & 0 & 1/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\
1/3 & 0 & -1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\
1/3 & -1/\sqrt{6} & 0 & -1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} \\
1/3 & 1/\sqrt{6} & 0 & 1/\sqrt{6} & -1/\sqrt{6} & 1/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} \\
1/3 & -1/\sqrt{6} & -1/\sqrt{6} & 0 & -1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} & 0 & 1/\sqrt{6} \\
1/3 & 1/\sqrt{6} & 1/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} & 0 & 1/\sqrt{6} \\
1/3 & -1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} & 0 \\
1/3 & 1/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} & 0 \\
1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Από τον W προκύπτει ο παρακάτω πίνακας πληροφορίας W^TW.

$$W^TW=\begin{array}{ccccccccc}
1 & 0 & 0 & 0 & 0 & 0.816 & 0.816 & 0.816 & 0.816 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0.816 & 0 & 0 & 0 & 0 & 1 & 0.667 & 0.667 & 0.667 \\
0.816 & 0 & 0 & 0 & 0 & 0.667 & 1 & 0.667 & 0.667 \\
0.816 & 0 & 0 & 0 & 0 & 0.667 & 0.667 & 1 & 0.667 \\
0.816 & 0 & 0 & 0 & 0 & 0.667 & 0.667 & 0.667 & 1
\end{array}$$

Η ορίζουσα του W^TW είναι ίση με 0.0124 και, αν υψωθεί στην 1/9, προκύπτει η *D-efficiency*=0.6143.

5.2.2 Αποδοτικότητα του σχεδιασμού DSD με 5 παράγοντες

Ο DSD σχεδιασμός με 5 παράγοντες φαίνεται στον Πίνακα 4.2.2. Το τετραγωνικό μοντέλο που προκύπτει σε αυτήν την περίπτωση έχει 11 παραμέτρους και είναι το

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \ \beta_{11} {X_1}^2 + \ \beta_{22} {X_2}^2 + \beta_{33} {X_3}^2 + \ \beta_{44} {X_4}^2 + \ \beta_{55} {X_5}^2 + \varepsilon$$

Ο αντίστοιχος πίνακα μοντέλου X είναι ένας ο πίνακας διάστασης 11 x 11 που ακολουθεί.

$$X = \begin{array}{ccccccccccc}
X_0 & X_1 & X_2 & X_3 & X_4 & X_5 & X_1^2 & X_2^2 & X_3^2 & X_4^2 & X_5^2 \\
+1 & 0 & +1 & +1 & -1 & -1 & 0 & +1 & +1 & +1 & +1 \\
+1 & 0 & -1 & -1 & +1 & +1 & 0 & +1 & +1 & +1 & +1 \\
+1 & +1 & 0 & -1 & -1 & +1 & +1 & 0 & +1 & +1 & +1 \\
+1 & -1 & 0 & +1 & +1 & -1 & +1 & 0 & +1 & +1 & +1 \\
+1 & +1 & -1 & 0 & +1 & -1 & +1 & +1 & 0 & +1 & +1 \\
+1 & -1 & +1 & 0 & -1 & +1 & +1 & +1 & 0 & +1 & +1 \\
+1 & +1 & -1 & +1 & 0 & +1 & +1 & +1 & +1 & 0 & +1 \\
+1 & -1 & +1 & -1 & 0 & -1 & +1 & +1 & +1 & 0 & +1 \\
+1 & +1 & +1 & +1 & +1 & 0 & 1 & +1 & +1 & +1 & 0 \\
+1 & -1 & -1 & -1 & -1 & 0 & +1 & +1 & +1 & +1 & 0 \\
+1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Για τον υπολογισμό του W, βρίσκουμε το μέτρο της πρώτη στήλης που είναι ίσο με $\sqrt{11}$, ενώ για τις υπόλοιπες στήλες είναι ίσο με $\sqrt{8}$.

$$W = \begin{array}{ccccccccccc}
\frac{1}{\sqrt{11}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & 0 & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & \frac{1}{\sqrt{8}} & 0 & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & \frac{-1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & \frac{1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & \frac{1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & 0 & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} \\
\frac{1}{\sqrt{11}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 \\
\frac{1}{\sqrt{11}} & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & 0 \\
\frac{1}{\sqrt{11}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Ο πίνακας πληροφορίας $W^T W$ που προκύπτει είναι ο

$$W^T W = \begin{array}{ccccccccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0.853 & 0.853 & 0.853 & 0.853 & 0.853 \\
0 & 1 & -0.25 & 0.25 & 0.25 & 0.25 & 0 & 0 & 0 & 0 & 0 \\
0 & -0.25 & 1 & 0.25 & -0.25 & -0.25 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.25 & 0.25 & 1 & 0.25 & -0.25 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.25 & -0.25 & 0.25 & 1 & -0.25 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.25 & -0.25 & -0.25 & -0.25 & 1 & 0 & 0 & 0 & 0 & 0 \\
0.853 & 0 & 0 & 0 & 0 & 0 & 1 & 0.750 & 0.750 & 0.750 & 0.750 \\
0.853 & 0 & 0 & 0 & 0 & 0 & 0.750 & 1 & 0.750 & 0.750 & 0.750 \\
0.853 & 0 & 0 & 0 & 0 & 0 & 0.750 & 0.750 & 1 & 0.750 & 0.750 \\
0.853 & 0 & 0 & 0 & 0 & 0 & 0.750 & 0.750 & 0.750 & 1 & 0.750 \\
0.853 & 0 & 0 & 0 & 0 & 0 & 0.750 & 0.750 & 0.750 & 0.750 & 1
\end{array}$$

με ορίζουσα ίση με 0.0006 και *D-efficiency*=0.51447.

5.2.3 Αποδοτικότητα των σχεδιασμών DSD με 6 έως 12 παράγοντες

Με παρόμοιο τρόπο υπολογίζεται η D-efficiency των DSD σχεδιασμών για 6 ως 12 παράγοντες, υπό το τετραγωνικό μοντέλο. Στις περιπτώσεις με 6 έως και 8 παράγοντες θα δίνεται ο πίνακας W^TW που προκύπτει και η τιμής της αποδοτικότητας. Στις περιπτώσεις με 9, 10, 11 και 12 παράγοντες θα δοθεί μόνο η τιμή της αποδοτικότητας, λόγω μεγέθους των πινάκων που προκύπτουν. Ο τρόπος εργασίας είναι όμοιος με αυτόν που περιεγράφηκε στις ενότητες 5.2.1 και 5.2.2.

DSD με 6 παράγοντες

	1	0	0	0	0	0	0.877	0.877	0.877	0.877	0.877	0.877	0.877
	0	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0	0	0	0	0	0	0
	0	0	0	0	1	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	0	0	0	0	0	0	0
$W^TW=$	0	0	0	0	0	0	1	0.8	0.8	0.8	0.8	0.8	0.8
	0.877	0	0	0	0	0	0.8	1	0.8	0.8	0.8	0.8	0.8
	0.877	0	0	0	0	0	0.8	0.8	1	0.8	0.8	0.8	0.8
	0.877	0	0	0	0	0	0.8	0.8	0.8	1	0.8	0.8	0.8
	0.877	0	0	0	0	0	0.8	0.8	0.8	0.8	1	0.8	0.8
	0.877	0	0	0	0	0	0.8	0.8	0.8	0.8	0.8	1	0.8
	0.877	0	0	0	0	0	0.8	0.8	0.8	0.8	0.8	0.8	1

Η ορίζουσα σε αυτήν την περίπτωση είναι ίση με 9.4×10^{-6} την όποια την υψώνουμε στην 1/13, άρα προκύπτει η *D-efficiency*=0.5003.

DSD με 7 παράγοντες

	1	0	0	0	0	0	0	0	0.894	0.894	0.894	0.894	0.894	0.894	0.894
	0	1	-0.167	-0.167	0.167	0.167	-0.167	-0.167	0	0	0	0	0	0	0
	0	-0.167	1	-0.167	0.167	-0.167	0.167	0.167	0	0	0	0	0	0	0
	0	-0.167	-0.167	1	-0.167	0.167	-0.167	-0.167	0	0	0	0	0	0	0
	0	0.167	0.167	-0.167	1	0.167	-0.167	0.167	0	0	0	0	0	0	0
	0	0.167	-0.167	0.167	0.167	1	-0.167	0.167	0	0	0	0	0	0	0
$W^TW=$	0	-0.167	0.167	-0.167	-0.167	-0.167	1	0.167	0	0	0	0	0	0	0
	0	0.167	0.167	-0.167	0.167	0.167	0.167	1	0	0	0	0	0	0	0
	0.894	0	0	0	0	0	0	0	1	0.833	0.833	0.833	0.833	0.833	0.833
	0.894	0	0	0	0	0	0	0	0.833	1	0.833	0.833	0.833	0.833	0.833
	0.894	0	0	0	0	0	0	0	0.833	0.833	1	0.833	0.833	0.833	0.833
	0.894	0	0	0	0	0	0	0	0.833	0.833	0.833	1	0.833	0.833	0.833
	0.894	0	0	0	0	0	0	0	0.833	0.833	0.833	0.833	1	0.833	0.833
	0.894	0	0	0	0	0	0	0	0.833	0.833	0.833	0.833	0.833	1	0.833
	0.894	0	0	0	0	0	0	0	0.833	0.833	0.833	0.833	0.833	0.833	1

Η ορίζουσα σε αυτήν την περίπτωση είναι ίση με 4.83×10^{-6} άρα προκύπτει η *D-efficiency*=0.4422.

DSD με 8 παράγοντες

$$W^TW=\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.907 & 0.907 & 0.907 & 0.907 & 0.907 & 0.907 & 0.907 & 0.907 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.857 & 1 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.857 & 0.857 & 1 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.857 & 0.857 & 0.857 & 1 & 0.857 & 0.857 & 0.857 & 0.857 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.857 & 0.857 & 0.857 & 0.857 & 1 & 0.857 & 0.857 & 0.857 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 1 & 0.857 & 0.857 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 1 & 0.857 \\
0.907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 0.857 & 1
\end{matrix}$$

Η ορίζουσα σε αυτήν την περίπτωση είναι ίση με 5.1×10^{-7} την όποια αν την υψώσουμε στην 1/17, προκύπτει η *D-efficiency*=0.4264.

DSD με 9 παράγοντες

Η ορίζουσα του πίνακα W^TW σε αυτήν την περίπτωση είναι ίση με $1,1\times10^{-8}$ και η *D-efficiency* προκύπτει ίση με 0.3862.

DSD με 10 παράγοντες

Η ορίζουσα του πίνακα W^TW σε αυτήν την περίπτωση είναι ίση με 1.09×10^{-9} την όποια την υψώνουμε στην 1/21, άρα προκύπτει η *D-efficiency*=0.3742.

DSD με 11 παράγοντες

Η ορίζουσα του πίνακα W^TW σε αυτήν την περίπτωση είναι ίση με 2.46×10^{-11} την όποια την υψώνουμε στην 1/23, άρα προκύπτει η *D-efficiency*=0.3457.

DSD με 12 παράγοντες

Η ορίζουσα σε αυτήν την περίπτωση είναι ίση με 9×10^{-13} την όποια την υψώνουμε στην 1/25, άρα προκύπτει η D-efficiency=0.3324.

5.3 Συγκεντρωτικά αποτελέσματα για την αποδοτικότητα των σχεδιασμών

Σε αυτήν την ενότητα θα συνοψίσουμε τα αποτελέσματα που προέκυψαν για την αποδοτικότητα των σχεδιασμών που προτείνουν οι Jones & Nachtsheim έχοντας σαν μοντέλο το τετραγωνικό. Το πεδίο τιμών που μπορεί να πάρει η αποδοτικότητα είναι [0,1] αλλά εκφράζεται σαν ποσοστό όποτε το εναλλακτικό πεδίο τιμών είναι [0,100%]

Αριθμός Παραγόντων	Αριθμός Παραμέτρων	D-efficiency
4	9	61.43%
5	11	51.44%
6	13	50.03%
7	15	44.22%
8	17	42.64%
9	19	38.62%
10	21	37.42%
11	23	34.57%
12	25	33.24%

Πίνακας 5.3.1

Ένα πρώτο συμπέρασμα που εύκολα παρατηρούμε είναι ότι όσο αυξάνονται οι παράγοντες, συνεπώς και οι παράμετροι, τόσο μειώνεται και η αποδοτικότητα του σχεδιασμού. Ένα άλλο συμπέρασμα ότι με βάση το τετραγωνικό μοντέλο χωρίς να υπάρχουν μέσα οι αλληλεπιδράσεις, οι τιμές κυμαίνονται από το 30% έως και λίγο παραπάνω από το 60%. Όταν έχουμε 11 ή 12 παράγοντες η αποδοτικότητα είναι χαμηλή, ενώ μέτρια αποδοτικότητα είναι από 7 έως 10 παράγοντες και τέλος όταν έχουμε 4 έως 6 παράγοντες έχουμε μια ικανοποιητική αποδοτικότητα καθώς ξεπερνούν το 50% ,αλλά υπάρχει μια σημαντική διαφορά περίπου 10% ανάμεσα στην χρήση 4 παραγόντων με αντίστοιχη χρήση 5 ή 6 παραγόντων.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Διπλωματικές-Διδακτορικά

- Χ. Ευαγγελάρας, Ορθογώνιοι σχηματισμοί και προβολικές ιδιότητες, Διδακτορική διατριβή, 2005, Εθνικό Μετσόβιο Πολυτεχνείο
- Σέργη Ανθή Μαρία, “Σχεδιασμοί Κρησαρίσματος και Εφαρμογές”, Διπλωματική Εργασία, 2007, Πανεπιστήμιο Αιγαίου, Τμήμα Στατιστικής και Χρηματοοικονομικών Μαθηματικών.
- Πέτρος-Παύλος Δ. Υψηλάντης, “Μεθοδολογία Αποκριτικών Επιφανειών-Σχεδιασμοί Υποσυνόλου”, Διπλωματική Εργασία, 2012
- Ευθυμίου Ιφιγένεια, «Παραγοντικοί σχεδιασμοί και οι ιδιότητες τους», Διπλωματική Εργασία, 2010, Πανεπιστήμιο Αιγαίου, Τμήμα Στατιστικής και Αναλογιστικών-Χρηματοοικονομικών Μαθηματικών.
- Άγγελος Ιωάννης Χ. Γαλής, «Μελέτη αξιολόγηση και επιλογή σχεδιασμών κρησαρίσματος δυο επιπέδων σε παραγοντικά πειράματα», Διπλωματική Εργασία, 2014, Πανεπιστήμιο Πειραιά, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

Άρθρα

- Bradley Jones & Christopher J. Nachtsheim, “A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects”, Journal of Quality Technology, 43 (2012), 1-15
- Lili Xiao, Dennis K. J. Lin, Fengshan Bai , “Construting Defenitive Screening Desings Using Conference Matrices” , 44 (2012), 2-8
- S. D. Georgiou, S. Stylianou, and M. Aggarwal. Efficient three-level screening desings using weighing matrices. Statistics,48 (2013), 815-833
- Nam-Ky Nguyen and Stella Stylianou, Constructing Definitive Screening Designs Using Cyclic Generators, Journal of Statistical Theory and Practice, 7 (2013), 713–724
- J.J. Sylvester, Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to

Newton's rule, ornamental tile-work, and the theory of numbers, Philosophical Magazine, 34 (1867), 461–475

- C. R. Rao, Factorial experiments devirable from combinatorial arrangements of arrays, Journal of the Royal Statistics Society, 9 (1947), 128-139
- G. E. P. Box and J. S. Hunter, The 2^{k-p} fractional factorial designs Technometrics, Volume 3 (1961), 311-351
- G. Box and J. Tyssedal, Projective properties of certail orthogonal arrays, Biometrika, 83 (1996), 950-955
- H. Kharaghani and B. Tayfch-Rezaie. A Hadamard matrix of order 428, Journal of Combinatorial Designs, 13 (2005) .435-440
- Hadamard, J., "Résolution d'une question relative aux déterminants", Bulletin des Sciences Mathématiques. 17 (1893), 240–246

Βιβλία

- Douglas C. Montgomery , “Design and Analysis of Experiments”, Wiley, 2012
- C. F. Jeff Wu, Michael S. Hamada, «Planning, Analysis, and Optimization Second Edition», Wiley, 2011
- Thomas P. Ryan, «Modern Experimental Design», Wiley, 2007
- Myers, R.H. and Montgomery, D.C., Response Surface Methodology: Process and Product Optimization Using Designed Experiments, John Wiley and Sons, New York, 1995
- Tapan Kr. Barman & Prasanta Sahoo, ANN modelling of fractal dimension in machining, Mechatronics and Manufacturing Engineering,Woodhead Publishing, 2012
- Akm Samsur Rahman, Effects of nanofibers on properties of geopolymer composites, Nanotechnology in Eco-efficient Construction (Second Edition),Woodhead Publishing Series, 2019
- Dr .Subhash C. Mandal M. Pharm & Dr.Anup Kumar Das M. Pharm. & Vivekananda Mandal, Innovative Extraction Process Design and Optimization Using Design of Experimental Approach., Essentials of Botanical Extraction, Academic Press, 2015

Ιστοσελίδες

- Engireering Statistics Handbook, Process Improvement, Section 5, in itl.nist.gov
- Dr. Neil Polhemus, Definitive Screening Designs, in statgraphics.com/blog, (2018)
- Martyn Shuttleworth, Factorial Design, in explorable.com
- JMP Community, Overview of Definitive Screening Design, (2020) in jmp.com/support/help