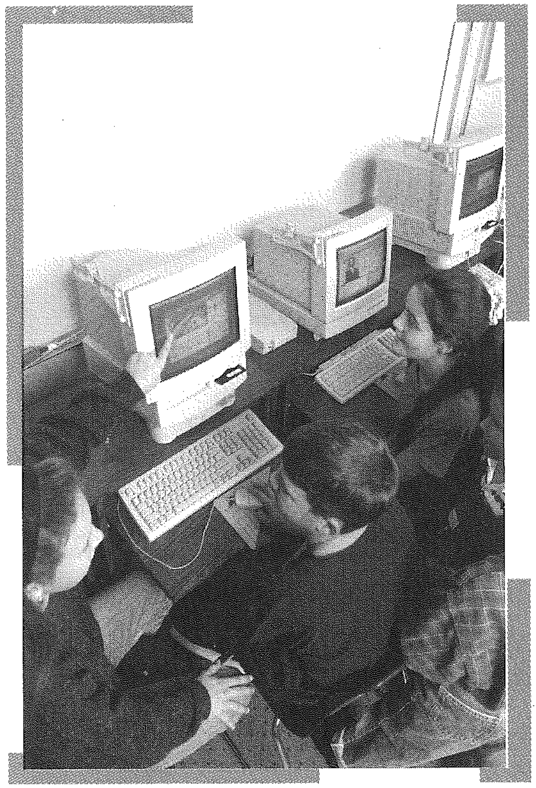
Δημήτρης Λ. Καραγεώργος
ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΑΘΗΝΩΝ
Επιτ. ΣΥΜΒΟΥΛΟΣ ΤΟΥ ΠΑΙΔΑΓΩΓΙΚΟΥ ΙΝΣΤΙΤΟΥΤΟΥ

ΜΕΘΟΔΟΛΟΓΙΑ ΕΡΕΥΝΑΣ
στις επιστήμες της αγωγής

ΜΙΑ ΔΙΔΑΚΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ



Σαββάλας
ΕΚΔΟΣΕΙΣ

$$E_{ij} = \frac{R_i C_j}{v}, \quad \text{όπου } v = 600$$

Οι υπολογισμοί αυτοί έγιναν και γράφονται μέσα στις παρενθέσεις σε κάθε κουτάκι. Για παράδειγμα, στο κουτάκι (1, 1) είναι $f_o = 60$ και $f_e = 33{,}3$. Η f_e υπολογίστηκε ως $E_{11} = \frac{200 \cdot 100}{600} = 33{,}3$. Στο κουτάκι (2, 3) είναι $f_o = 20$ και $f_e = 16{,}7$. Η f_e εδώ υπολογίστηκε ως $E_{23} = \frac{R_2 \cdot C_3}{600} = \frac{100 \cdot 100}{600} = 16{,}7$ κ.λπ.

Βήμα 2ο: Οι βαθμοί ελευθερίας είναι β.ε. $= (r-1)(c-1) = 3 \cdot 3 = 9$, οπότε $x^2_{0,05}(9) = 16{,}919$.

Βήμα 3ο: Προσδιορίζουμε τις αναμενόμενες συχνότητες όταν η H_0 είναι αληθής. Αυτές έχουν υπολογιστεί, όπως είπαμε προηγουμένως, και γράφτηκαν μέσα σε παρένθεση δίπλα στις πραγματικές συχνότητες. Αυτό το κάναμε για να αποφύγουμε την κατασκευή ενός νέου πίνακα με τις $f_{e_{ij}}$, αλλά και γιατί θα υπολογίζονται πιο γρήγορα οι διαφορές $f_{o_{ij}} - f_{e_{ij}}$.

Υπολογίζουμε τώρα το x^2 - στατιστικό από τον τύπο

$$x^2 = \sum \frac{(f_{o_{ij}} - f_{e_{ij}})^2}{f_{e_{ij}}}, \quad i, j = 1, 2, 3, 4$$

Αυτό είναι

$$x^2 = \frac{(60-33{,}3)^2}{33{,}3} + \frac{(40-100)^2}{100} + \frac{(60-33{,}3)^2}{33{,}3} + \frac{(40-33{,}3)^2}{33{,}3} +$$

$$+ \frac{(20-16{,}7)^2}{16{,}7} + \frac{(50-50)^2}{50} + \frac{(20-16{,}7)^2}{16{,}7} + \frac{(10-16{,}7)^2}{16{,}7} +$$

$$+ \frac{(10-16{,}7)^2}{16{,}7} + \frac{(50-50)^2}{50} + \frac{(10-16{,}7)^2}{16{,}7} + \frac{(30-16{,}7)^2}{16{,}7} +$$

$$+ \frac{(10-33{,}3)^2}{33{,}3} + \frac{(160-100)^2}{100} + \frac{(10-33{,}3)^2}{33{,}3} + \frac{(20-33{,}3)^2}{33{,}3} = 174$$

Βήμα 4ο: Η τιμή του στατιστικού $174 > 16{,}919$ πέφτει μέσα στην κρίσιμη περιοχή. Έτσι απορρίπτουμε την H_0 και ο ερευνητής μπορεί να συμπεράνει ότι το χρώμα των ματιών και το χρώμα των μαλλιών έχουν κάποια εξάρτηση μεταξύ τους.

Σημείωση: Τονίζουμε ότι αν σε κάποιο κουτάκι του πίνακα παρουσιαστεί αναμενόμενη συχνότητα μικρότερη του 5, τότε αυτή πρέπει να προστεθεί σε κάποιο άλλο κουτάκι ώστε να ικανοποιείται ο περιορισμός που επιβάλλει να είναι οι αναμενόμενες συχνότητες μεγαλύτερες του 5. Τότε όμως θα ελαττωθούν και οι βαθμοί ελευθερίας, αφού θα ελαττωθεί το πλήθος των γραμμών ή στηλών.

6.4 Εισαγωγή στην ανάλυση της διασποράς (ANOVA)

- *Γενικά*

Στην πράξη πολλές φορές παρουσιάζονται προβλήματα όπου πρέπει να συγκριθούν οι μέσες τιμές για περισσότερα από δύο δείγματα.

Για την αντιμετώπιση των προβλημάτων αυτών αναπτύχθηκε στη Στατιστική μια μέθοδος που ονομάζεται «**ανάλυση της διασποράς** ή **ανάλυση της διακύμανσης**».

Με απλά λόγια, η ανάλυση της διασποράς είναι μια διαδικασία ελέγχου υποθέσεων που τη χρησιμοποιούμε για να εκτιμήσουμε τις διαφορές μεταξύ δύο ή περισσότερων πληθυσμών (θεραπειών). Όπως όλες οι άλλες συμπερασματικές διαδικασίες έτσι και αυτή χρησιμοποιεί ως βάση δεδομένα δείγματος προκειμένου να εξαγάγει συμπεράσματα για τους πληθυσμούς. Συχνά χρησιμοποιούμε t-τεστ για να ελέγξουμε τη διαφορά των μέσων τιμών. Το t-τεστ μας βοηθά μόνο για δύο πληθυσμούς. Για περισσότερους πληθυσμούς χρησιμοποιούμε τη νέα μέθοδο που αναφέραμε, την ανάλυση της διασποράς. Γι' αυτό η ανάλυση της διασποράς παρέχει στους ερευνητές μεγαλύτερη ευελιξία στον σχεδιασμό πειραμάτων και στην ερμηνεία αποτελεσμάτων.

Όπως και το t-τεστ έτσι και η ανάλυση της διασποράς μπορεί να χρησιμοποιηθεί είτε σε διαφορετικά δείγματα (ένα από κάθε πληθυσμό) είτε στο ίδιο δείγμα, εφαρμόζοντας διαφορετικές προσεγγίσεις-πειράματα (θεραπείες, διδακτικές μεθόδους κ.λπ.). Επιπρόσθετα, η ανάλυση της διασποράς μπορεί να χρησιμοποιηθεί στην εκτίμηση των αποτελεσμάτων μιας έρευνας στην οποία εμπλέκονται περισσότερες από μία ανεξάρτητες μεταβλητές. Για παράδειγμα, ένας εκπαιδευτικός είναι δυνατό να θελήσει να συγκρίνει την αποτελεσματικότητα δύο διαφορετικών διδακτικών μεθόδων (πρώτη ανεξάρτητη μεταβλητή) σε τρεις τάξεις διαφορετικών μεγεθών (δεύτερη ανεξάρτητη μεταβλητή). Η εξαρτημένη μεταβλητή εδώ είναι ο βαθμός κάθε μαθητή σ' ένα σταθμισμένο τεστ. Η δομή αυτού του πειράματος, το οποίο περιλαμβάνει τη σύγκριση δειγματικών μέσων για έξι διαφορετικές περιπτώσεις, έχει ως εξής:

		Ανεξάρτητη μεταβλητή 2: Μέγεθος τάξης		
		Μικρή τάξη	Μεσαία τάξη	Μεγάλη τάξη
Ανεξάρτητη μεταβλητή 1: Διδακτικές μέθοδοι	Μέθοδος Α	Δείγμα 1	Δείγμα 2	Δείγμα 3
	Μέθοδος Β	Δείγμα 4	Δείγμα 5	Δείγμα 6

Όπως βλέπουμε, η ανάλυση της διασποράς παρέχει στους ερευνητές μια εξαιρετικά ευέλικτη τεχνική ανάλυσης δεδομένων. Μπορεί να χρησιμοποιηθεί στην εκτίμηση της σημαντικότητας των διαφορών των μέσων τιμών σε μια ποικιλία ερευνητικών καταστάσεων και είναι η διαδικασία που χρησιμοποιείται περισσότερο στον έλεγχο υποθέσεων.

Στα παρακάτω θα κάνουμε απλώς μια εισαγωγή στην πιο απλή μορφή της ανάλυσης της διασποράς για να αντιληφθεί ο αναγνώστης πώς εργαζόμαστε με τη διαδικασία αυτή. Η πλήρης ανάπτυξη της ανάλυσης της διασποράς χρειάζεται χρόνο και εκτεταμένους υπολογισμούς που απωθούν τον αναγνώστη και γι' αυτό θα την αποφύγουμε σ' αυτό το βιβλίο.

Στην ανάλυση της διασποράς με τον όρο **παράγοντα** (factor) εννοούμε μια ανεξάρτητη μεταβλητή. Έτσι, ένας σχεδιασμός με μία μόνο ανεξάρτητη μεταβλητή ονομάζεται **μονοπαραγοντικός σχεδιασμός**. Σχεδιασμοί έρευνας με περισσότερες μεταβλητές (όπως στο προηγούμενο παράδειγμα) ονομάζονται **πολυπαραγοντικοί σχεδιασμοί** ή απλά **παραγοντικοί σχεδιασμοί**.

Ένα παράδειγμα μονοπαραγοντικού σχεδιασμού φαίνεται στο επόμενο σχήμα, όπου έχουμε πάρει τρία διαφορετικά δείγματα για να εκτιμήσουμε τις διαφορές των μέσων τιμών τριών πληθυσμών (θεραπειών, μεθόδων κ.λπ.) με άγνωστες μέσες τιμές.

Πληθυσμός 1	**Πληθυσμός 2**	**Πληθυσμός 3**
$\mu_1 = ;$	$\mu_2 = ;$	$\mu_3 = ;$
$\downarrow$	$\downarrow$	$\downarrow$
Δείγμα 1	**Δείγμα 2**	**Δείγμα 3**
0	2	8
4	8	12
8	14	16
$\bar{x} = 4$	$\bar{x} = 8$	$\bar{x} = 12$

Παρατηρούμε ότι τα τρία δείγματα έχουν διαφορετικά στοιχεία και διαφορετικούς μέσους. Η ανάλυση της διασποράς έχει σκοπό να βοηθήσει τον ερευνητή για να αποφασίσει μεταξύ των ακόλουθων δύο εκδοχών:

1. Στην πραγματικότητα δεν υπάρχουν διαφορές μεταξύ των πληθυσμών (των μέσων τιμών τους) και οι παρατηρούμενες διαφορές μεταξύ των δειγμάτων είναι εντελώς τυχαίες (οφείλονται δηλαδή σε δειγματικό σφάλμα).
2. Οι διαφορές μεταξύ των δειγματικών μέσων απεικονίζουν τις πραγματικές διαφορές μεταξύ των πληθυσμών (των θεραπειών, των μεθόδων κ.λπ.). Δηλαδή οι πληθυσμοί έχουν πράγματι διαφορετικούς μέσους και αυτό ακριβώς παρουσιάζεται και με τα δειγματικά δεδομένα.

Αυτές οι δύο εκδοχές είναι φυσικά αντίστοιχες των υποθέσεων H_0 και H_1. Πριν προχωρήσουμε, τονίζουμε ότι στην ανάλυση της διασποράς εργαζόμαστε με δύο προϋποθέσεις:

1. Οι πληθυσμοί από τους οποίους επιλέγονται τα τυχαία δείγματα **ακολουθούν την κανονική κατανομή**.
2. Οι κανονικοί αυτοί πληθυσμοί **έχουν όλοι την ίδια διακύμανση** σ^2.

- *Στατιστικές υποθέσεις στην ανάλυση της διασποράς*

Ας υποθέσουμε ότι ένας ερευνητής θέλει να μελετήσει την επίδραση της θερμοκρασίας του δωματίου στη μάθηση. Για τον σκοπό αυτό επιλέγει τρεις θερμοκρασίες 15 °C, 25 °C και 35 °C και τρία διαφορετικά δείγματα (τρεις διαφορετικές ομάδες ατόμων), ένα για κάθε θερμοκρασία.

Με στατιστικούς όρους, θέλουμε να αποφασίσουμε μεταξύ δύο υποθέσεων, της μηδενικής (H_0) υπόθεσης, που λέει ότι η θερμοκρασία δεν έχει καμιά επίδραση στη μάθηση, και της εναλλακτικής (H_1) υπόθεσης, που δηλώνει ότι η θερμοκρασία έχει επίδραση στη μάθηση. Δηλαδή:

$H_0: \mu_1 = \mu_2 = \mu_3$

H_1: Τουλάχιστον μία μέση τιμή είναι διαφορετική από τις άλλες.

Σημειώνουμε ότι η H_1 δεν έχει καθοριστεί συγκεκριμένα, γιατί υπάρχουν πολλές διαφορετικές εναλλακτικές δυνατότητες. Για παράδειγμα, μπορούμε να έχουμε:

$H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_1$ (τρεις διαφορετικές μέσες τιμές)

ή $H_1: \mu_1 = \mu_3$ και $\mu_2 \neq \mu_1$

ή $H_1: \mu_1 = \mu_2$ και $\mu_3 \neq \mu_1$ κ.λπ.

Συνήθως ο ερευνητής καθορίζει (με βάση κάποια θεωρία) μια συγκεκριμένη πρόβλεψη, που αφορά την επίδραση της δοκιμασίας (θεραπείας).

Στα παρακάτω, λόγω και της γενικότητας με την οποία αντιμετωπίζουμε το θέμα, θα δηλώνουμε μια γενική εναλλακτική υπόθεση και όχι συγκεκριμένη.

Το στατιστικό τεστ για την ανάλυση της διασποράς είναι όμοιο με το t-τεστ. Για το t-τεστ υπολογίζουμε τον λόγο

$$t = \frac{\text{Διαφορές μεταξύ δειγματικών μέσων}}{\text{Αναμενόμενες διαφορές (σφάλμα)}}$$

Για την ανάλυση της διασποράς το στατιστικό τεστ ονομάζεται F-λόγος και υπολογίζεται ως εξής:

$$F = \frac{\text{Διακύμανση μεταξύ δειγματικών μέσων}}{\text{Αναμενόμενη διακύμανση (σφάλμα)}}$$

Σημειώνουμε ότι το στατιστικό F βασίζεται στη διακύμανση και όχι στη διαφορά δειγματικών μέσων. Αυτό γίνεται γιατί όταν έχουμε δύο δειγματικούς μέσους, υπολογίζουμε αμέσως τη διαφορά, ενώ όταν έχουμε πολλούς δειγματικούς μέσους, είναι αρκετά πολύπλοκο να εκφράσουμε αυτή τη διαφορά.

Είναι αυτονόητο ότι, όταν οι δειγματικοί μέσοι διαφέρουν ελάχιστα μεταξύ τους, τότε και το στατιστικό F είναι μικρό, ενώ αν διαφέρουν σημαντικά, το στατιστικό F είναι μεγάλο.

Παρατηρούμε ακόμη ότι το στατιστικό t και ο λόγος F μας δίνουν την ίδια πληροφορία. Σε κάθε περίπτωση ο αριθμητής του λόγου μετρά την πραγματική διαφορά που επιτυγχάνεται από τα δειγματικά δεδομένα και ο παρονομαστής μετρά τη διαφορά που αναμένεται κατά τύχη.

Μια μεγάλη τιμή είτε του στατιστικού t είτε του λόγου F παρέχει ενδείξεις για το ότι η διαφορά των δειγματικών μέσων δεν είναι τυχαία.

- *Η λογική του κριτηρίου της ανάλυσης της διασποράς*

Έγινε φανερό από τα προηγούμενα ότι το πρόβλημα που τίθεται είναι αν οι μέσες τιμές των κανονικών πληθυσμών με κοινή διακύμανση s^2, από τους οποίους προέρχονται τα δείγματα, διαφέρουν σημαντικά ή όχι.

Η απάντηση που δίνει η μέθοδος της ανάλυσης της διασποράς στηρίζεται στη σύγκριση της **μεταβλητότητας μεταξύ των δειγματικών μέσων** και της **μεταβλητότητας των τιμών μέσα σε κάθε δείγμα**.

Γενικά είναι εύκολος ο έλεγχος για τη σύγκριση των μέσων τιμών δύο ανεξάρτητων δειγμάτων που προέρχονται από κανονικές κατανομές με κοινή διακύμανση σ^2. Η χρήση αυτού του κριτηρίου για τη σύγκριση κ μέσων τιμών θα σήμαινε $\binom{\kappa}{2}$ ελέγχους υποθέσεων. Για παράδειγμα, αν κ = 8, θα έπρεπε να γίνουν:

$$\binom{8}{2} = \frac{8!}{2!6!} = \frac{7 \cdot 8}{2} = 28 \text{ έλεγχοι!}$$

Το κυριότερο όμως μειονέκτημα των συγκρίσεων ανά δύο δεν είναι αυτό. Είναι ότι η πιθανότητα να κάνουμε σφάλμα τύπου I αυξάνεται δραματικά με την αύξηση του αριθμού των ελέγχων.

Είναι λοιπόν ανάγκη να αντιμετωπιστεί το πρόβλημα της σύγκρισης των μέσων τιμών συνολικά.

Ας δούμε ένα παράδειγμα για να κατανοήσουμε καλύτερα τη λογική του κριτηρίου της ανάλυσης της διασποράς.

Ας υποθέσουμε ότι ο ερευνητής της προηγούμενης παραγράφου επέλεξε τρεις ομάδες των πέντε ατόμων για να εξετάσει την επίδραση της θερμοκρασίας στη μάθηση και συγκέντρωσε τα ακόλουθα στοιχεία:

Πείραμα 1	**Πείραμα 2**	**Πείραμα 3**
15 °C	25 °C	35 °C
Δείγμα 1	**Δείγμα 2**	**Δείγμα 3**
0	4	1
1	3	2
3	6	2
1	3	0
0	4	0
$\bar{x} = 1$	$\bar{x} = 4$	$\bar{x} = 1$

Όπως παρατηρούμε, οι τιμές που έχουμε δεν είναι ίδιες. Στη γλώσσα της Στατιστικής, οι τιμές είναι μεταβλητές. Δουλειά μας είναι να μετρήσουμε το ποσό της μεταβλητότητας και να εξηγήσουμε από πού προέρχεται.

Το πρώτο βήμα είναι να προσδιοριστεί η μεταβλητότητα για όλα τα δεδομένα μαζί. Έπειτα μπορούμε να προχωρήσουμε στον χωρισμό σε υποσύνολα δεδομένων και να κάνουμε μετρήσεις. Η λέξη **ανάλυση** σημαίνει τον χωρισμό σε μικρότερα τμήματα. Επειδή πρόκειται να αναλύσουμε τη μεταβλητότητα, γι' αυτό η διαδικασία λέγεται **ανάλυση της διασποράς**.

Αυτή η διαδικασία ανάλυσης διαιρεί τη συνολική μεταβλητότητα σε δύο βασικές συνιστώσες.

1. **Μεταβλητότητα μεταξύ των δειγμάτων**. Στα δεδομένα του παραδείγματος παρατηρούμε ότι η μεταβλητότητα στις τιμές οφείλεται στις διαφορές μεταξύ των επιπέδων θερμοκρασίας. Για παράδειγμα, στη θερμοκρασία των 25 °C έχουμε τη μεγαλύτερη μέση τιμή $\bar{x} = 4$, ενώ στη θερμοκρασία των 35 °C έχουμε $\bar{x} = 1$.

Θα υπολογίσουμε τη μεταβλητότητα μεταξύ των δειγμάτων, ώστε να έχουμε ένα συνολικό μέτρο των διαφορών μεταξύ των δειγματικών μέσων.

Υπάρχουν τρεις πιθανές εξηγήσεις γι' αυτή τη μεταβλητότητα.

α) **Η επίδραση της φροντίδας** (θεραπείας, μεθόδου, θερμοκρασίας κ.λπ.). Είναι δυνατό οι διαφορετικές θερμοκρασίες να προκάλεσαν τη διαφορά στους δειγματικούς μέσους.

β) **Ατομικές διαφορές**. Τα υποκείμενα παίρνουν μέρος στο πείραμα με διαφορετική υποδομή, ικανότητες και διαθέσεις, δηλαδή καθένα άτομο είναι μοναδικό. Οποτεδήποτε συγκρίνουμε διαφορετικά δείγματα, είναι δυνατόν οι διαφορές των ατόμων να είναι η αιτία των διαφορών μεταξύ των δειγμάτων.

γ) **Πειραματικό σφάλμα**. Οποτεδήποτε κάνουμε μια μέτρηση, υπάρχει πιθανότητα λάθους. Αυτό μπορεί να γίνει από απροσεξία, από τη χρήση μη τέλειων οργάνων ή και από απρόβλεπτους παράγοντες που επιφέρουν αλλαγές στο μετρούμενο ποσό. Αυτό το λάθος μπορεί να είναι η αιτία των διαφορών μεταξύ των δειγμάτων.

2. **Μεταβλητότητα μέσα σε κάθε δείγμα**. Εκτός από τη μεταβλητότητα μεταξύ των δειγματικών μέσων, υπάρχει και μεταβλητότητα των τιμών μέσα σε κάθε δείγμα. Υπάρχουν δύο πιθανές εξηγήσεις γι' αυτή τη μεταβλητότητα.

α) **Ατομικές διαφορές**. Οι τιμές επιτυγχάνονται από διαφορετικά άτομα, κάτι που εξηγεί γιατί είναι μεταβλητές.

β) **Πειραματικό λάθος**. Πάντα υπάρχει η πιθανότητα πειραματικού λάθους. Είναι φανερό ότι δεν μπορούμε εδώ να ισχυριστούμε ότι υπάρχουν διαφορές λόγω της επίδρασης των συνθηκών (της θερμοκρασίας), γιατί όλα τα άτομα του δείγματος παίρνουν μέρος κάτω από τις ίδιες συνθήκες.

Μπορούμε τώρα να συγκρίνουμε τις προηγούμενες μεταβλητότητες (μεταξύ δειγμάτων και μέσα στα δείγματα). Αυτό γίνεται με τον υπολογισμό του λόγου F, ο οποίος είναι

$$F = \frac{\text{Διακύμανση μεταξύ δειγμάτων}}{\text{Διακύμανση εντός των δειγμάτων}}$$

Αν αναλύσουμε κάθε μεταβλητότητα στα στοιχεία της, τότε μπορούμε να γράψουμε:

$$F = \frac{\text{Επίδραση της φροντίδας} + \text{Ατομικές διαφορές} + \text{Πειραματικό σφάλμα}}{\text{Ατομικές διαφορές} + \text{Πειραματικό σφάλμα}}$$

Παρατηρούμε ότι οι δύο μεταβλητότητες διαφέρουν μόνο στη μεταβλητότητα που οφείλεται στην επίδραση της φροντίδας (θεραπείας, μεθόδου, συνθηκών πειράματος κ.λπ.). Αυτό είναι σημαντικό γιατί, αν υποθέσουμε ότι η H_0 είναι αληθής, δηλαδή ότι οι συνθήκες του πειράματος δεν έχουν καμία επίδραση στη μεταβλητότητα, που σημαίνει ότι η επίδραση της φροντίδας είναι μηδέν, **τότε ο λόγος F προσεγγίζει τη μονάδα**. Όσο μεγαλύτερη όμως είναι αυτή η επίδραση, τόσο ο λόγος αυτός μεγαλώνει. Άρα, αν μετρήσουμε σε κάθε πείραμα τον λόγο F, ανάλογα με την τιμή του μπορούμε να αποφανθούμε για τον βαθμό επίδρασης των συνθηκών του πειράματος στη μεταβλητότητα των δειγμάτων. Τις τιμές του F τις παίρνουμε από τον αντίστοιχο πίνακα (παράρτημα).

Ο παρονομαστής του λόγου F μετρά μόνο την ανεξέλεγκτη μεταβλητότητα και γι' αυτό ονομάζεται και «όρος σφάλμα».

Συμπερασματικά

1. Η ανάλυση της διασποράς είναι μια στατιστική διαδικασία που εκτιμά τις διαφορές των μέσων τιμών σε δύο ή περισσότερα πειράματα (θεραπείες, μεθόδους, καταστάσεις κ.λπ.).
2. Για να εκτιμήσουμε τις διαφορές των μέσων τιμών, υπολογίζουμε δύο μεταβλητότητες:
   α) τη μεταβλητότητα μεταξύ των μέσων τιμών των δειγμάτων και
   β) τη μεταβλητότητα μέσα σε κάθε δείγμα.
   Η σύγκριση αυτών γίνεται με τον σχηματισμό του λόγου F, που έχει αριθμητή την πρώτη μεταβλητότητα και παρονομαστή τη δεύτερη.
3. Ο παρονομαστής του λόγου F ονομάζεται «σφάλμα» και παρέχει ένα μέτρο για τη διακύμανση που οφείλεται στην «τύχη».
4. Όταν η επίδραση της θεραπείας (φροντίδας, συνθηκών κ.λπ.) δεν επηρεάζει καθόλου την εξέλιξη του φαινομένου, δηλαδή όταν η H_0 είναι αληθής, τότε η αναμενόμενη τιμή του λόγου F είναι 1, γιατί ο αριθμητής και ο παρονομαστής του λόγου εκφράζουν το ίδιο πράγμα, δηλαδή μετρούν την ίδια διακύμανση.
5. Όταν οι διαφορές μεταξύ των μέσων τιμών αυξάνονται, λόγω της επίδρασης των συνθηκών του πειράματος, τότε ο λόγος F αυξάνεται, ενώ όσο η μεταβλητότητα μέσα σε κάθε δείγμα αυξάνεται, τόσο ο λόγος F μικραίνει.

Γενικεύοντας τώρα τον τύπο που δίνει την κοινή διακύμανση δύο πληθυσμών A και B, δηλαδή τον τύπο $s^2 = \dfrac{(\nu_A - 1)s_A^2 + (\nu_B - 1)s_B^2}{\nu_A + \nu_B - 2}$ για κ πληθυσμούς παίρνουμε:

$$S^2 = \frac{(\nu_1 - 1)s_1^2 + (\nu_2 - 1)s_2^2 + \ldots + (\nu_\kappa - 1)s_\kappa^2}{\nu_1 + \nu_2 + \ldots + \nu_\kappa - \kappa} \qquad (1)$$

που είναι ο εκτιμητής του σ^2, δηλαδή η μεταβλητότητα μεταξύ των δειγμάτων.

Ο εκτιμητής (1) γράφεται:

$$S^2 = \frac{\sum_{i=1}^{\nu_1}(x_{1i} - \bar{x}_1)^2 + \sum_{i=1}^{\nu_2}(x_{2i} - \bar{x}_2)^2 + \ldots + \sum_{i=1}^{\nu_\kappa}(x_{\kappa i} - \bar{x}_\kappa)^2}{\nu_1 + \nu_2 + \ldots + \nu_\kappa - \kappa} = \frac{SSE}{N - \kappa} \qquad (2)$$

όπου x_{ij} είναι η j - παρατήρηση του i - δείγματος, $\bar{x}_i$ είναι η μέση τιμή του i δείγματος, **SSE είναι το άθροισμα των τετραγώνων των σφαλμάτων**, το οποίο μετρά τη συνολική μεταβλητότητα **μέσα** στα δείγματα, και $N = \nu_1 + \nu_2 + \ldots + \nu_\kappa$.

Ένα μέτρο της μεταβλητότητας **μεταξύ** των δειγμάτων δίνεται από τη σχέση

$$\frac{1}{\kappa - 1} \cdot \sum_{i=1}^{\kappa} \nu_i(\bar{x}_i - \bar{x})^2 = \frac{SSA}{\kappa - 1} \qquad (3)$$

όπου $\bar{x}$ είναι ο γενικός μέσος όρος όλων των δειγμάτων μαζί και **SSA είναι το άθροισμα των τετραγώνων των αποκλίσεων των μέσων τιμών των δειγμάτων του παράγοντα από τον γενικό μέσο όρο**.

Με τις ποσότητες (2) και (3) έχουμε δύο εκτιμητές της κοινής διακύμανσης σ^2 των πληθυσμών. Αυτές οι ποσότητες είναι ίσες ($F = 1$) όταν η H_0 είναι αληθής.

Ο λόγος F, τον οποίο λέμε και στατιστικό F, με τις ποσότητες (2) και (3) παίρνει τη μορφή

$$F = \frac{\dfrac{SSA}{\kappa - 1}}{\dfrac{SSE}{N - \kappa}} = \frac{MSA}{MSE} \qquad (4)$$

όπου το $MSA = \dfrac{SSA}{\kappa - 1}$ λέγεται **μέσο τετράγωνο των δειγμάτων** και το $MSE = \dfrac{SSE}{N - \kappa}$ λέγεται **μέσο τετράγωνο των σφαλμάτων**.

Αν με **SST** συμβολίσουμε τη **συνολική** μεταβλητότητα, δηλαδή αν

$$SST = \sum_{i=1}^{\kappa}\sum_{j=1}^{\nu_j}(x_{ij} - \bar{x})^2$$

τότε αποδεικνύεται ότι

$$\sum_{i=1}^{\kappa}\sum_{j=1}^{\nu_j}(x_{ij} - \bar{x})^2 \quad = \quad \sum_{i=1}^{\kappa}\nu_i(\bar{x}_i - \bar{x})^2 \quad + \quad \sum_{i=1}^{\kappa}\sum_{j=1}^{\nu_j}(x_{ij} - \bar{x}_i)^2 \qquad (5)$$

Ολικό άθροισμα τετραγώνων SST	Άθροισμα τετραγώνων των δειγμάτων SSA	Άθροισμα τετραγώνων των σφαλμάτων SSE
β.ε. $= N - 1$	β.ε. $= \kappa - 1$	β.ε. $= N - \kappa$

Σημείωση: Για την απόδειξη της σχέσης (5) αντικαθιστούμε το $x_{ij} - \bar{x}$ με $(x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x})$ και κάνουμε πράξεις.

Τα αθροίσματα των τετραγώνων, οι βαθμοί ελευθερίας, τα μέσα τετράγωνα και το στατιστικό F συγκεντρώνονται σ' έναν πίνακα που λέγεται **πίνακας ανάλυσης διασποράς** ή **ANOVA πίνακας**, όπως ο ακόλουθος.

ANOVA πίνακας για τη σύγκριση των μέσων τιμών κ δειγμάτων

Μεταβλητότητα	**Άθροισμα τετραγώνων (SS)**	**β.ε.**	**Μέση μεταβολή**	**F**
Μεταξύ δειγμάτων (Παράγοντας)	$SSA = \sum_{i=1}^{\kappa}\nu_i(\bar{x}_i - \bar{x})^2$	$\kappa - 1$	$MSA = \dfrac{SSA}{\kappa - 1}$	$F = \dfrac{MSA}{MSE}$
Μέσα στα δείγματα (Σφάλμα)	$SSE = \sum_{i=1}^{\kappa}\sum_{j=1}^{\nu_j}(x_{ij} - \bar{x}_i)^2$	$N - \kappa$	$MSE = \dfrac{SSE}{N - \kappa}$	
Συνολική	$SST = \sum_{i=1}^{\kappa}\sum_{j=1}^{\nu_j}(x_{ij} - \bar{x})^2$	$N - 1$		

Είναι φανερό τώρα ότι ο έλεγχος της σημαντικότητας του παράγοντα γίνεται με το F-τεστ, όπου $F = \dfrac{MSA}{MSE}$.

Αν $F > F_{\kappa-1, N-\kappa, \alpha}$, $N - \kappa$ οι β.ε. του παρονομαστή, τότε απορρίπτουμε την H_0, διαφορετικά δεχόμαστε την H_0.

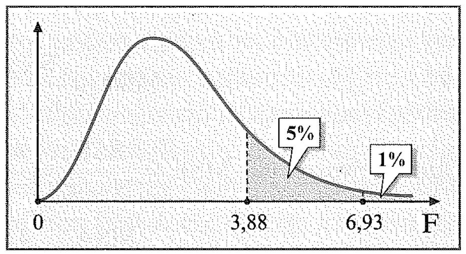
Για το παράδειγμα της επίδρασης της θερμοκρασίας στη μάθηση μπορούμε να συμπληρώσουμε τον προηγούμενο πίνακα, για να διευκολυνθούμε στους υπολογισμούς, ως εξής:

Πείραμα 1	**Πείραμα 2**	**Πείραμα 3**	
15 °C	25 °C	35 °C	
Δείγμα 1	**Δείγμα 2**	**Δείγμα 3**	
0	4	1	
1	3	2	
3	6	2	$N = 15$
1	3	0	$\kappa = 3$
0	4	0	
$\sum x_{1i} = 5$	$\sum x_{2i} = 20$	$\sum x_{3i} = 5$	
$\bar{x}_1 = 1$	$\bar{x}_2 = 4$	$\bar{x}_3 = 1$	
$\nu_1 = 5$	$\nu_2 = 5$	$\nu_3 = 5$	

οπότε εύκολα βρίσκουμε SSA = 30, SSE = 16, SST = 46, $MSA = \dfrac{SSA}{\kappa - 1} = \dfrac{30}{2} = 15$

και $MSE = \dfrac{SSE}{15 - 3} = \dfrac{16}{12} \simeq 1{,}33$, οπότε το στατιστικό $F = \dfrac{MSA}{MSE} = \dfrac{15}{1{,}33} \simeq 11{,}28$.

Αν πάρουμε ως επίπεδο $\alpha = 0{,}05$, τότε η κριτική τιμή του F είναι $F_{\kappa-1, N-\kappa, \alpha} = F_{2, 12, 0{,}05} = 3{,}8853$. Επειδή $F = 11{,}28 > F_{2, 12, 0{,}05} = 3{,}8853$, απορρίπτουμε την H_0, δηλαδή δεχόμαστε ότι υπάρχει επίδραση της θερμοκρασίας του δωματίου στη μάθηση.



Αν θεωρούσαμε το $\alpha = 0{,}01$, τότε $F_{2, 12, 0{,}01} = 6{,}9266$, οπότε και πάλι θα απορρίπταμε την H_0.

Σημειώνουμε ότι στο παράδειγμά μας αυτό έχουμε μόνο έναν παράγοντα (μονοπαραγοντική ανάλυση), τη θερμοκρασία, με τρία επίπεδα (15 °C, 25 °C, 35 °C). Τα επίπεδα αυτά είναι οι **τιμές που παίρνει ο παράγοντας** και λέγονται **στάθμες του παράγοντα**.

Παράδειγμα

Ένα Λύκειο έχει στην Α΄ τάξη τέσσερα τμήματα. Με τυχαία δειγματοληψία επιλέγουμε από κάθε τμήμα μια ομάδα έξι μαθητών. Σε όλες τις ομάδες διδάξαμε την ίδια ενότητα από τα Μαθηματικά (συστήματα α΄ βαθμού) με διαφορετική μέθοδο σε κάθε ομάδα. Μετά την ολοκλήρωση της διδασκαλίας δώσαμε ένα κοινό τεστ σε όλες τις ομάδες, προκειμένου να δούμε την κατανοησιμότητα της ύλης που διδάχθηκε. Τα αποτελέσματα (με άριστα το 50) παρουσιάζονται στον ακόλουθο πίνακα.

Ομάδα 1	Ομάδα 2	Ομάδα 3	Ομάδα 4
37	49	33	41
42	38	34	48
45	40	40	40
49	39	38	42
50	50	47	38
45	41	36	41
$\sum = 268$	$\sum = 257$	$\sum = 228$	$\sum = 250$

Ο ερευνητής θέλει να ελέγξει αν οι μέσοι όροι της βαθμολογίας των τεσσάρων πληθυσμών από τους οποίους επιλέχτηκαν τα δείγματα είναι ίσοι.

Απάντηση

Εδώ έχουμε $\nu_1 = \nu_2 = \nu_3 = \nu_4 = 6$, $\kappa = 4$, $N = 24$, $\bar{x}_1 = 44{,}67$, $\bar{x}_2 = 42{,}83$, $\bar{x}_3 = 38{,}00$ και $\bar{x}_4 = 41{,}67$. Κάνοντας τους απαραίτητους υπολογισμούς βρίσκουμε:

$$\bar{x} = \frac{268 + 257 + 228 + 250}{24} = 41{,}79$$

$$SSA = \sum_{i=1}^{\kappa} \nu_i (\bar{x}_i - \bar{x})^2 =$$

$$= 6(44{,}67 - 41{,}79)^2 + 6(42{,}83 - 41{,}79)^2 + 6(38 - 41{,}79)^2 + 6(41{,}67 - 41{,}79)^2 =$$

$$= 6 \cdot 8{,}29 + 6 \cdot 1{,}08 + 6 \cdot 14{,}36 + 6 \cdot 0{,}01 \simeq 142{,}44$$

$$MSA = \frac{SSA}{\kappa - 1} = \frac{142{,}44}{3} \simeq 47{,}48$$

Ομοίως παίρνουμε $MSE = \dfrac{SSE}{N - \kappa} = \dfrac{439{,}5}{20} = 21{,}98$.

Άρα το στατιστικό $F = \dfrac{MSA}{MSE} = \dfrac{47{,}48}{21{,}98} = 2{,}16$.

Επιλέγουμε ως επίπεδο σημαντικότητας $\alpha = 0{,}05$ και δηλώνουμε τις υποθέσεις μας:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_1: Ένας τουλάχιστον μέσος μ_i είναι διαφορετικός από τους άλλους.

Από τον πίνακα F (παράρτημα) βρίσκουμε την κριτική τιμή του F για β.ε. $\kappa - 1 = 3$, $N - \kappa = 20$ και $\alpha = 0{,}05$ να είναι

$$F_{3,20,0{,}05} = 3{,}0984$$

Επειδή είναι $F = 2{,}16 < F_{3,20,0{,}05} = 3{,}0984$, αποδεχόμαστε την H_0, δηλαδή οι μέσες τιμές της βαθμολογίας των τεσσάρων τμημάτων δεν διαφέρουν μεταξύ τους ή οι τέσσερις διαφορετικές διδακτικές μέθοδοι έχουν σχεδόν το ίδιο αποτέλεσμα.

Πιστεύουμε ότι δώσαμε τα απαιτούμενα στοιχεία για να κατανοήσει ο αναγνώστης τι είναι η ανάλυση διασποράς (ANOVA) και πώς εργαζόμαστε. Όπως είπαμε και στην αρχή του κεφαλαίου, δεν θα προχωρήσουμε σε βαθύτερη μελέτη, γιατί οι υπολογισμοί είναι τεράστιοι.

6.5 Συσχετιστική έρευνα

Στη συσχετιστική έρευνα (correlational research) οι ερευνητές εξετάζουν πιθανές σχέσεις μεταξύ μεταβλητών χωρίς καμιά προσπάθεια να τις επηρεάσουν. Αν και οι συσχετιστικές μελέτες δεν είναι δυνατό να προσδιορίσουν τις αιτίες αυτών των σχέσεων μεταξύ των μεταβλητών, μπορούν να τις υποδείξουν. Αυτές οι υποδείξεις συχνά παρέχουν την αφορμή για περαιτέρω πειραματικές μελέτες. Στα παρακάτω θα δούμε τη φύση της συσχετιστικής έρευνας, θα δώσουμε μερικά παραδείγματα συσχετιστικών ερευνών και θα περιγράψουμε μερικά προβλήματα που εμπλέκονται όταν επιχειρούμε τέτοιες έρευνες.

- *Η φύση της συσχετιστικής έρευνας*

Η συσχετιστική έρευνα στην πιο απλή της μορφή ερευνά τη δυνατότητα των σχέσεων μεταξύ μόνο δύο μεταβλητών, αν και έρευνες με περισσότερες από δύο μεταβλητές είναι συνηθισμένο φαινόμενο. Σε αντίθεση με την πειραματική έρευνα, εδώ ο ερευνητής δεν επεμβαίνει καθόλου πάνω στις μεταβλητές, τις οποίες μετρά με τα κατάλληλα μέσα και με τη βοήθεια στατιστικών τεχνικών βρίσκει κάποιους δείκτες, οι οποίοι τον βοηθούν να αντιληφθεί τον βαθμό της σχέσης που υπάρχει μεταξύ των μεταβλητών.

Η συσχετιστική έρευνα, μερικές φορές, αναφέρεται και σαν ένας τύπος περιγραφικής έρευνας επειδή περιγράφει μια υφιστάμενη σχέση μεταξύ μεταβλητών. Ο τρόπος όμως που περιγράφει αυτή τη σχέση είναι τελείως διαφορετικός από τις περιγραφές που βρίσκονται σε άλλες μελέτες. Η συσχετιστική μελέτη περιγράφει τον βαθμό κατά τον οποίον δύο ή περισσότερες ποσοτικές μεταβλητές συσχετίζονται μεταξύ τους και αυτό φυσικά γίνεται με τη βοήθεια του συντελεστή συσχέτισης. Γι' αυτό συνιστούμε στον αναγνώστη πριν προχωρήσει στη μελέτη της συσχετιστικής έρευνας να επαναλάβει την § 2.6, όπου έχει αναπτυχθεί η εύρεση και ο ρόλος του συντελεστή συσχέτισης.