

**Τ.Ε.Ι ΠΑΤΡΑΣ**

**ΠΑΡΑΡΤΗΜΑ ΑΜΑΛΙΑΔΑΣ**

**Τμήμα Εφαρμογών Πληροφορικής στη Διοίκηση και Οικονομία**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**«ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΣΤΗΝ ΕΡΕΥΝΑ ΥΠΗΡΕΣΙΩΝ  
ΥΓΕΙΑΣ»**

**«Statistical methods in health services research»**

**Όνοματεπώνυμο Σπουδαστών: Νικολάου Η. Γεώργιος**

**Χουντής Δ. Ιωάννης**

**Εποπτεύων Καθηγητής: Ειρήνη Νικολοπούλου**



**Αμαλιάδα – Ιούνιος-2012**

<b>ΠΡΟΛΟΓΟΣ</b> .....	<b>4</b>
<b>ΠΕΡΙΛΗΨΗ</b> .....	<b>5</b>
<b>ABSTRACT</b> .....	<b>6</b>
<b>ΕΙΣΑΓΩΓΗ</b> .....	<b>7</b>
<b>ΚΕΦΑΛΑΙΟ 1<sup>ο</sup></b> .....	<b>8</b>
<b>ΥΓΕΙΑ - ΑΣΘΕΝΕΙΑ</b> .....	<b>8</b>
1.1. ΥΓΕΙΑ .....	8
1.2. ΑΣΘΕΝΕΙΑ .....	9
1.3. ΚΟΥΛΤΟΥΡΑ ΚΑΙ ΑΣΘΕΝΕΙΑ.....	12
<b>ΚΕΦΑΛΑΙΟ 2<sup>ο</sup></b> .....	<b>14</b>
<b>ΥΠΗΡΕΣΙΕΣ ΥΓΕΙΑΣ ΚΑΙ ΕΡΕΥΝΑ</b> .....	<b>14</b>
2.1. ΕΙΣΑΓΩΓΗ .....	14
2.2. ΕΚΤΙΜΗΣΗ ΤΗΣ ΑΝΑΓΚΗΣ .....	14
2.3. ΖΗΤΗΣΗ ΥΠΗΡΕΣΙΩΝ ΥΓΕΙΑΣ .....	17
<b>ΚΕΦΑΛΑΙΟ 3<sup>ο</sup></b> .....	<b>24</b>
<b>ΣΤΑΤΙΣΤΙΚΗ ΣΤΙΣ ΥΠΗΡΕΣΙΕΣ ΥΓΕΙΑΣ</b> .....	<b>24</b>
<b>3.1. ΕΙΣΑΓΩΓΗ</b> .....	<b>24</b>
<b>3.2. ΔΕΙΚΤΕΣ</b> .....	<b>25</b>
<b>3.3. ΕΡΓΑΛΕΙΑ ΤΗΣ ΕΡΕΥΝΑΣ</b> .....	<b>26</b>
<b>3.3.1. ΣΧΕΛΙΑΣΜΟΣ ΤΗΣ ΈΡΕΥΝΑΣ</b> .....	<b>26</b>
<b>3.3.2. ΚΑΤΑΡΤΙΣΗ ΕΡΩΤΗΜΑΤΟΛΟΓΙΩΝ</b> .....	<b>27</b>
<b>3.3.3 Η ΔΟΜΗΜΕΝΗ ΣΥΝΕΝΤΕΥΞΗ</b> .....	<b>28</b>
<b>3.3.4. Η ΕΝΤΟΠΙΣΜΕΝΗ ΣΥΝΕΝΤΕΥΞΗ (FOCUSED INTERVIEW)</b> .....	<b>28</b>
<b>3.3.5. ΕΠΙΛΟΓΗ ΤΟΥ ΔΕΙΓΜΑΤΟΣ</b> .....	<b>28</b>
<b>ΚΕΦΑΛΑΙΟ 4<sup>ο</sup></b> .....	<b>31</b>
<b>ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ</b> .....	<b>31</b>
4.1. ΕΙΣΑΓΩΓΗ .....	31
4.2.2 ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ.....	31
4.2.3 ΚΑΤΑΝΟΜΕΣ $X^2$ , T ΚΑΙ F .....	32
4.2.3.1 ΚΑΤΑΝΟΜΗ $X^2$ .....	33
4.2.3.2 ΚΑΤΑΝΟΜΗ Τ'Η ΚΑΤΑΝΟΜΗ STUDENT (T-DISTRIBUTION 'Η STUDENT DISTRIBUTION) .....	34
4.2.2.3 ΚΑΤΑΝΟΜΗ F (F- DISTRIBUTION).....	35
4.3.1 ΈΛΕΓΧΟΣ $X^2$ ΟΜΟΙΟΓΕΝΕΙΑΣ .....	36
4.3.2 ΔΙΑΧΩΡΙΣΤΙΚΗ ΑΝΑΛΥΣΗ .....	37
4.3.3 ΑΝΑΛΥΣΗ ΕΠΙΒΙΩΣΗΣ.....	39
4.3.4 ΈΛΕΓΧΟΣ $X^2$ ΑΝΕΞΑΡΤΗΣΙΑΣ.....	41
4.3.5 ΈΛΕΓΧΟΣ ΤΑΣΗΣ .....	41
4.3.6 ΣΥΝΤΕΛΕΣΤΗΣ ΓΡΑΜΜΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ ΤΟΥ PEARSON .....	42
4.3.7 ΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ (SIMPLE LINEAR REGRESSION) .....	46
4.3.8 ΠΟΛΛΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ (MULTIPLE LINEAR REGRESSION) .....	52
4.3.9 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ .....	65
4.3.10 ΑΝΑΛΥΣΗ ΤΗΣ ΔΙΑΚΥΜΑΝΣΗΣ (ANOVA) .....	67
4.3.11 ΑΝΑΛΥΣΗ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ.....	70
4.3.12 ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΕΣ ΜΕΤΡΗΣΕΙΣ .....	72
4.4. ΜΕΘΟΔΟΙ ΓΙΑ ΜΗ ΑΝΕΞΑΡΤΗΤΑ ΔΕΙΓΜΑΤΑ .....	73
4.4.1. ΈΛΕΓΧΟΣ MANN-WHITNEY-U .....	73
4.4.2. ΈΛΕΓΧΟΣ KRUSKAL-WALLIS .....	74

4.4.3. ΈΛΕΓΧΟΣ JONCKHEERE .....	74
4.5 ΕΠΙΔΡΑΣΗ PLACEBO – ΕΙΚΟΝΙΚΑ ΦΑΡΜΑΚΑ .....	75
<b>5. ΣΥΖΗΤΗΣΗ – ΣΥΜΠΕΡΑΣΜΑΤΑ.....</b>	<b>77</b>
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ.....</b>	<b>81</b>
<b>ΠΑΡΑΡΤΗΜΑ 1 ΠΙΝΑΚΕΣ. ....</b>	<b>85</b>
ΠΙΝΑΚΑΣ 1 .....	85
ΠΙΝΑΚΑΣ 2 .....	86
ΠΙΝΑΚΑΣ 3 .....	87

•

## **ΠΡΟΛΟΓΟΣ**

Η εργασία αυτή συγγράφηκε στα πλαίσια της εκπόνησης πτυχιακής εργασίας από τους σπουδαστές του Τμήματος Εφαρμογών Πληροφορικής στην Οικονομία και Διοίκηση, ΑΤΕΙ Πατρών (Παράρτημα Αμαλιάδας), Νικολάου Γεώργιο και Χουντή Ιωάννη.

Ευχαριστούμε την Επιβλέπουσα Καθηγήτρια Νικολοπούλου Ειρήνη για την καλή συνεργασία και τη σημαντική βοήθεια που μας παρείχε κατά τη διάρκεια της συγγραφής της εργασίας αυτής.

Νικολάου Γεώργιος

Χουντής Ιωάννης

Αμαλιάδα, 2011

## ΠΕΡΙΛΗΨΗ

Η εφαρμογή των στατιστικών μεθόδων στις υπηρεσίες υγείας, ιδιαίτερα σε μια εποχή όπου η πρόοδος της τεχνολογίας έκανε εφικτή την συλλογή και αποθήκευση μεγάλου όγκου δεδομένων, κρίνεται επιτακτική για την βελτιστοποίηση των παρεχόμενων υπηρεσιών στους ασθενείς. Με τον όρο «υπηρεσίες υγείας» προσδιορίζονται όχι μόνο οι παρεχόμενες φροντίδες υγείας και ιατρικής περίθαλψης αλλά γενικότερα το πλέγμα των δομών με τις οποίες ρυθμίζονται η παραγωγή και διανομή των ιατρικών φροντίδων και καλύπτονται οι ιατρικές ανάγκες, όπως η εκτίμηση της ανάγκης και η ζήτηση. Συνήθως, χρησιμοποιούνται επιδημιολογικές, κοινωνιολογικές, οικονομετρικές, καθώς και καθαρά στατιστικές μέθοδοι και οι οποίες έχουν ως αντικείμενο την κατασκευή δεικτών και υποδειγμάτων. Η διαδικασία διεξαγωγής μιας στατιστικής έρευνας ακολουθεί δύο διακριτά στάδια, το στάδιο σχεδιασμού και το στάδιο υλοποίησης. Οι στατιστικές αυτές μέθοδοι περιλαμβάνουν: τον έλεγχο  $\chi^2$  ομοιογένειας, τον έλεγχο  $\chi^2$  Ανεξαρτησίας, την απλή Γραμμική Παλινδρόμηση, την πολλαπλή Γραμμική Παλινδρόμηση, τη Λογιστική Παλινδρόμηση, την Ανάλυση της Διακύμανσης, την Ανάλυση Συνδιακύμανσης, τις επαναλαμβανόμενες μετρήσεις, τον έλεγχο MANN-WHITNEY-U, τον έλεγχο KRUSKAL-WALLIS και τον έλεγχο Jonckheere. Όταν οι μέθοδοι χρησιμοποιούνται για την αξιολόγηση κλινικών μελετών είναι απαραίτητο να λαμβάνεται υπόψη η επίδραση Placebo. Η ορθή εφαρμογή της στατιστικής έχει αναμφισβήτητη και καταλυτική συνεισφορά στην πρόοδο και στην προαγωγή της δημοσίας υγείας.

## **ABSTRACT**

The application of statistic methods in health services, particularly in an era that technology simplifies the massive collection and storage of data, is imperative in order to optimize the services provided to patients. The term "health services" is used not only to define the health and medical care provided, but in general the structures which regulate and cover the production and distribution of medical care and necessities, including assessing the need and demand. Usually epidemiological, sociological, econometric and purely statistical methods are used, which intend to build indicators and models. The process of conducting a statistical survey follows two distinct stages, design phase and implementation phase. These statistic methods include:  $\chi^2$  homogeneity control,  $\chi^2$  independence, the simple linear regression, multiple linear regression, the logistic regression, analysis of variance, analysis of covariance, repeated measurements, MANN-WHITNEY-U control, KRUSKAL-WALLIS control and When the methods are used to evaluate clinical studies Placebo effect must be taken under consideration. The application of statistics has been of undisputed and catalytic contribution to the advancement of public health.

## ΕΙΣΑΓΩΓΗ

Στατιστική είναι η επιστήμη της συλλογής, οργάνωσης, και ερμηνεία δεδομένων. Ασχολείται με όλες τις πτυχές αυτής της διαδικασίας, συμπεριλαμβανομένων του σχεδιασμού της συλλογής δεδομένων, του σχεδιασμού των ερευνών και των πειραμάτων. Στην επιστήμη της Στατιστικής συνήθως το ενδιαφέρον επικεντρώνεται στην εκτίμηση ενός άγνωστου μεγέθους, το οποίο καλείται παράμετρος και το οποίο συνήθως συνοψίζει κατά κάποιον τρόπο τις τιμές της υπό μελέτης μεταβλητής στον πληθυσμό.

Η Βιοστατιστική είναι η εφαρμογή των στατιστικών στοιχείων σε ένα ευρύ φάσμα θεμάτων στη βιολογία και την ιατρική. Η επιστήμη της βιοστατιστικής περιλαμβάνει το σχεδιασμό πειραμάτων, ιδιαίτερα στον τομέα της ιατρικής και της γεωργίας, τη συλλογή, περίληψη, και ανάλυση των στοιχείων από τα πειράματα αυτά. Τέλος είναι υπεύθυνη για την ερμηνεία όλως των παραπάνω και την εξαγωγή αποτελεσμάτων και συμπερασμάτων.

Εφαρμογές της βιοστατιστικής:

1. Δημόσια υγεία, συμπεριλαμβανομένης της επιδημιολογίας, της υγείας έρευνας για τις υπηρεσίες, τη διατροφή, την υγεία και το περιβάλλον
2. Σχεδιασμός και ανάλυση των κλινικών δοκιμών στον τομέα της ιατρικής γενετικής του πληθυσμού, και των στατιστικών γενετική προκειμένου να συνδέσουν τις διακυμάνσεις γονότυπο με μια παραλλαγή στο φαινότυπο.
3. Οικολογία, οικολογική πρόβλεψη
4. Στατιστικές μέθοδοι αρχίζουν να εντάσσονται στην ιατρική πληροφορική, τη δημόσια υγεία πληροφορικής, τη βιοπληροφορική και την υπολογιστική βιολογία

Από τα παραπάνω γίνεται αντιληπτή η μεγάλη σπουδαιότητα της εφαρμογής των στατικών μεθόδων στις υπηρεσίες υγείας, ιδιαίτερα σε μια εποχή όπου η τεράστια πρόοδος της τεχνολογίας έκανε εφικτή την τεράστια συλλογή και αποθήκευση δεδομένων, η επεξεργασία των οποίων κρίνεται επιτακτική για την βελτιστοποίηση των παρεχόμενων υπηρεσιών στους ασθενείς.

## ΚΕΦΑΛΑΙΟ 1<sup>ο</sup>

### ΥΓΕΙΑ - ΑΣΘΕΝΕΙΑ



**Εικόνα 1. Υγεία**

**Πηγή :** <http://cmhr.anu.edu.au/>

#### **1.1.Υγεία**

Η Υγεία ορίζεται από την WHO ως η κατάσταση πλήρους φυσικής ,πνευματικής και κοινωνικής ευημερίας και δεν περιορίζεται μόνο στην απουσία νόσου ή νοσηρότητας (WHO, 1998). Έτσι αναδεικνύεται η πολυδιάστατη φύση της υγείας, η οποία πρέπει να προσεγγίζεται όχι μόνο με ιατρικούς όρους, αποκλείοντας έτσι ψυχοκοινωνικούς, καθοριστικούς παράγοντες αλλά και με ευρύτερους ανθρωπολογικούς – κοινωνιολογικούς, πολιτισμικούς όρους κι όχι μόνο με γνωσιακούς. Θέτοντας νέα πολιτισμικά πλαίσια για την θεώρηση όποιας κατάστασης, μας βοηθά να αντιμετωπίσουμε προβλήματα που πριν δεν είχαμε τη δυνατότητα, ψυχικά ή άλλα.

Το πώς ορίζεται βέβαια η υγεία και η ασθένεια ως κοινωνική αναπαράσταση διαφέρει ανάλογα με τις πολιτιστικές διαφορές, τις ιδέες και πεποιθήσεις, την εκπαίδευση και την παιδεία αλλά και τις θρησκευτικές αντιλήψεις (για τη ζωή, το θάνατο, την αμαρτία, την τιμωρία, το μίasma), καθώς και τις αντιλήψεις για το σώμα (για τα όρια του, την καθαριότητα, την εικόνα του, το φύλο του) που υπάρχουν σε κάθε κοινωνία. Είναι γνωστό ακόμα ότι ο τεράστιος αριθμός παραγόντων που επηρεάζουν την κατάσταση της υγείας ενός πληθυσμού βρίσκονται σε άμεση συσχέτιση με την



κοινωνική αναπαράσταση που έχει ο πληθυσμός για την υγεία του. Παράγοντες γενετικοί, περιβαλλοντικοί (τρόπος διαβίωσης, κατοικία, εργασία, συνθήκες εργασίας, ρύπανση περιβάλλοντος) παράγοντες συνθηκών ζωής (διατροφή, κάπνισμα, άσκηση, χρήση εθιστικών ουσιών, συμπεριφορά), παράγοντες σχετιζόμενοι με το σύστημα και τις υπηρεσίες υγείας και παράγοντες εκπαίδευσης υγείας επηρεάζουν την υγεία των ανθρώπων αλλά και την κοινωνική αναπαράσταση του δίπολου υγείας-ασθένειας (Αγραφιώτης Δ. 2003).

## **1.2.Ασθένεια**

Η ασθένεια ορίζεται ως το αντικειμενικά προσδιοριζόμενο φαινόμενο και όχι με την αντίληψη που μπορεί να έχει κάθε άτομο και να αποδίδει τον ορισμό ασθένεια ή έλλειψη υγείας και μπορεί να διέπεται από πολιτισμικούς παράγοντες (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

Ο άνθρωπος δεν αντιμετωπίζεται πια σαν μια «βιολογική μηχανή». Χρειάζεται πάρα πολύ μεγάλη διαφώτιση ώστε όλοι όσοι εργάζονται στο χώρο της Υγείας να αρχίσουν να σκέφτονται με αυτό τον τρόπο. Γιατί - τουλάχιστον στον δικό μας πολιτιστικό χώρο - θεωρείται υγιής αυτός που δεν εμφανίζει συμπτώματα ασθένειας. Οι γιατροί μας λένε ότι είμαστε γεροί όταν διάφορες εξετάσεις βρίσκονται στα φυσιολογικά όρια. Δεν συμπεριλαμβάνουν ακόμα στις «μετρήσεις» τους, εάν είμαστε εκνευρισμένοι, μελαγχολικοί, αν νιώθουμε μοναξιά, εάν καπνίζουμε πολύ ή πίνουμε, εάν δεν τρώμε φρέσκιες τροφές, εάν δεν κινούμαστε αρκετά ή αν δεν βρίσκουμε στη ζωή μας νόημα (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995) .

Το καινούργιο «μήνυμα» λοιπόν, είναι ότι, μαζί με τις γνωστές εξετάσεις - που πρέπει να γίνονται , πρέπει να λαμβάνονται υπ' όψη και οι γνωστικές και συναισθηματικές διεργασίες, ο τρόπος της ζωής μας και οι συνθήκες που ζούμε.

Ολοκληρωμένη διάγνωση θα έχουμε μόνο ακολουθώντας αυτόν τον καινούργιο τρόπο ερμηνείας της ασθένειας. Για να γίνει αυτό πρέπει και εμείς -«υγιείς» και μη- να αρχίζουμε να ελέγχουμε την συμπεριφορά μας σε σχέση με την υγεία μας (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

Η ασθένεια και η υγεία δεν είναι στατικές καταστάσεις, αλλά δύο πλευρές (θεωρήσεις) ενός ενιαίου, μεταβαλλόμενου συστήματος, οι οποίες συμπληρώνουν η μία την άλλη. Η ασθένεια δεν είναι κάτι το Κακό, το απορριπτέο, αλλά ένα σήμα κινδύνου, ότι κάτι δεν πάει καλά στη γενική ισορροπία του οργανισμού και μια προσπάθεια του

οργανισμού να επαναφέρει την αρμονία του συστήματος. Ο άνθρωπος δεν είναι μόνο ένα «σωματικό σύστημα» αλλά μία ολότητα, με σωματικές, ψυχοπνευματικές και κοινωνικές διαστάσεις, που έχουν συνεχή αλληλεπίδραση μεταξύ τους. Προβλήματα λοιπόν στο συναισθηματικό χώρο, αντίξοες συνθήκες ζωής, κοινωνικές συνθήκες που δεν αφήνουν περιθώρια εξέλιξης αποδυναμώνουν τις αντιστάσεις του οργανισμού και ευνοούν τις «σωματικές» ασθένειες. Γι' αυτό το λόγο, επειδή οι σωματικές ασθένειες αντικαθρεφτίζουν προβλήματα της ψυχικής και κοινωνικής ζωής, δεν αρκεί να καταπολεμούμε μόνο τα σωματικά συμπτώματα. Πρέπει να ερευνήσουμε από την μια μεριά και να βρούμε τις αιτίες που οδηγούν στην ασθένεια και στη γενική δυσφορία και από την άλλη να καλλιεργήσουμε συνειδητά τις δυνάμεις «άμυνας» του εαυτού μας και να τις ισχυροποιήσουμε (σε ακραίες περιπτώσεις και με ατομική Ψυχοθεραπεία) (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

Ένα από τα πιο σημαντικά στοιχεία της καινούργιας αυτής ερμηνείας της ασθένειας είναι η αυτοϋπευθυνότητα. Ο «ασθενής» παίρνει πια το θέμα της ατομικής του υγείας στα χέρια του, είναι υπεύθυνος και για την πρόληψη και για τη θεραπεία. Δεν είναι πια ένα παθητικό ον που «το ερευνούν» του «μετρούν διάφορες μεταβλητές στις λειτουργίες του οργανισμού» του «καθορίζουν μια φαρμακευτική αγωγή» κι αυτό πρέπει να πειθαρχήσει και θα γίνει καλά.

Η ζωή εμφανίζεται λοιπόν από την οπτική γωνία της ολικής Ιατρικής σαν πορεία μιας συνεχούς εναλλασσόμενης εμπειρίας. Ασθένειες και κρίσεις είναι μέρη αυτής της εμπειρίας, που μας προτρέπουν να ασχοληθούμε σοβαρά με τη ζωή μας: να ερευνήσουμε δηλ. τις αιτιώδεις σχέσεις μεταξύ της ασθένειας, των μέχρι τώρα συνηθειών μας και των πεποιθήσεών μας και να αναθεωρήσουμε τη ζωή μας, εάν θέλουμε να μείνουμε (ή να γίνουμε) υγιείς (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

Υπάρχουν ήδη αρκετές παρατηρήσεις σε ασθενείς, που παίρνουν μέρος σε προγράμματα υγείας, τα οποία βασίζονται στον καινούργιο ορισμό της ασθένειας:

- Ο ασθενής αρχίζει να συνειδητοποιεί ότι η ασθένεια δεν είναι - όπως την έβλεπε πριν - ένας εξωτερικός εχθρός. Κατανοεί δηλαδή πως αυτή έχει σχέση με τον τρόπο που ζει και ότι δεν χρειάζεται να την «πολεμήσει» αλλά να ανακαλύψει τί σημαίνει, τί συμβολίζει .
- Αυτός ο καινούργιος τρόπος «σκέψης» τον οδηγεί να «επικοινωνήσει» με την ασθένεια: προσπαθεί να βρει ποιες συμπεριφορές, ποιες σκέψεις, ποια συναισθήματα προκάλεσαν την ασθένεια, υποβοηθούν δηλαδή την μείωση της άμυνας του οργανισμού. Αυτή η φάση διαρκεί αρκετά, αρχίζει συνήθως με ερωτηματολόγια και έπεται ένα

μεγάλο χρονικό διάστημα εντατικής ενδοσκόπησης (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

- Αυτή η ενασχόληση με τον εαυτό οδηγεί σιγά, σιγά σε μια καινούργια, πιο ολοκληρωμένη αυτογνωσία. Η καινούργια αυτή εικόνα που έχουν αυτά τα άτομα για τον εαυτό τους και για τον περιβάλλον τους (έμπυχο ή μη) σημαίνει ταυτόχρονα και μια ριζική αλλαγή στον τρόπο που αισθάνονται, σκέπτονται και ενεργούν. Αυτός ο καινούργιος «προσανατολισμός στη ζωή» φέρνει συνήθως μια ψυχική ηρεμία η οποία έχει θετικότατο αντίκτυπο είτε στην πορεία της ασθένειας είτε στην εξαφάνιση της ασθένειας.
- Πολύ συχνά, κυρίως σε περιπτώσεις που «ανακαλύπτουν» οι ασθενείς ότι η κύρια αιτία της μείωσης αντοχής του οργανισμού τους είναι το άγχος, μαθαίνουν ειδικά προγράμματα χαλάρωσης (relaxation). Συχνά παρατηρείται ακόμη μία θετική επίπτωση: διότι ο ασθενής δοθείσης της ευκαιρίας, χαλαρώνει και ηρεμεί, επικοινωνεί με τον εαυτό του, κάνει προγράμματα για το μέλλον, βρίσκει τρόπους να εκφράσει επιθυμίες του, ανακαλύπτει καινούργιους δρόμους, που θα τον οδηγήσουν να χαρεί τη ζωή καλύτερα (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

Για να βοηθήσουν οι ειδικοί (Ψυχολόγοι, Ιατροί) καλύτερα τους ασθενείς να συνειδητοποιήσουν τις αιτιώδεις σχέσεις ασθένειας - τρόπου ζωής, πρέπει να ερευνήσουμε ακόμη πολύ. Η σχετική βιβλιογραφία αυξάνει παρόλα αυτά πολλά καίρια ερωτήματα μένουν ακόμα αναπάντητα. Πολλές φορές βοηθάει την ατομική περίπτωση η ερώτηση: «Τι μου συνέβη πριν εμφανισθεί η ασθένεια; Συνέβη κάτι σημαντικό στη ζωή μου, ένα εξωτερικό συμβάν ή είχα μια εσωτερική σύγκρουση;» Αλλά θα ήταν μεγάλο κέρδος να γνωρίζαμε περισσότερα πάνω σ' αυτό το θέμα, κυρίως για να προλάβουμε σωματοποιήσεις μετά από διάφορες κρίσεις της ζωής.

Γνωρίζουμε ήδη ότι πολλές ασθένειες εκδηλώνονται 6 ως 18 μήνες μετά από μία σημαντική αλλαγή στη ζωή, η οποία σήμαινε μια καινούργια προσαρμογή του οργανισμού. Οι ερευνητές δημιούργησαν μία «Σκάλα Κοινωνικής Αναπροσαρμογής», με τη βοήθεια της οποίας μπορεί κανείς να «μετρήσει» αλλαγές στη ζωή, το πόσο σημαντικές είναι για τους περισσότερους ανθρώπους και πόση «ψυχική ενέργεια» χρειάζεται για να προσαρμοστεί κανείς στις αλλαγές των συνθηκών ζωής (lifechangeunits). Όσο πιο σημαντικό είναι ένα γεγονός τόσο περισσότερο καιρό και «ψυχική ενέργεια» χρειάζεται ο άνθρωπος για να «προσαρμοσθεί» αλλά ταυτόχρονα αυξάνει και η πιθανότητα να αρρωστήσει (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

Οι έρευνες σ' αυτό το χώρο συνεχίζονται με ταχύ ρυθμό και φαίνεται πως θα

μας δώσουν πολύτιμες πληροφορίες στο επίπεδο της διάγνωσης δηλαδή για την πορεία που μπορεί να πάρει στον κάθε ένα μας ένα «τραυματικό γεγονός».

Γίνονται επίσης πολλές έρευνες οι οποίες προσπαθούν να ρίξουν φως στις άμεσες σχέσεις «τραυματικό γεγονός - ασθένεια». Γιατί μερικοί άνθρωποι αντιδρούν π.χ. με γαστρίτιδα, ενώ άλλοι π.χ. με δερματικά νοσήματα σε μια ορισμένη δυσκολία στη ζωή; Ακόμη δεν έχει δοθεί μια καλά τεκμηριωμένη απάντηση, αν και ήδη έχουν προταθεί μερικά ερμηνευτικά μοντέλα. Η σωματική ασθένεια είναι μια αντίδραση που φανερώνει την «εσωτερική στάση» που έχει ο καθένας μας για κάθε πρόβλημα. Ορίζουν την «εσωτερική στάση» σαν τον τρόπο που βιώνει κανείς τα γεγονότα, τι συμβαίνει αντικειμενικά και πώς βρίσκει λύσεις. Η αντίδραση του σώματος είναι λοιπόν «λογική απάντηση σε προβλήματα». Παράδειγμα: Διάρροιες. Οι ασθενείς έχουν την επιθυμία να απαλλαγούν από κάτι ή κάποιον («Ηθελα επί τέλους να τελειώνει αυτό το δράμα»). Ερμηνεύεται λοιπόν το σύμπτωμα, η φυσιολογική λειτουργία (κένωση) σαν μια πιθανότητα να απαλλαγεί το σώμα από ουσίες που δεν τις χρειάζεται. Η διάρροια έχει λοιπόν άμεση σχέση με την επιθυμία να απαλλαγεί κάποιους πολύ γρήγορα από κάτι ή κάποιον που τον ενοχλεί (Νέττλετον Σ. 2002, Ποταμιάνος Γρ. 1995).

Μπορεί να υποστηρίξει κανείς πως αυτή η καινούργια θεώρηση της ασθένειας ανοίγει νέους ορίζοντες και για την πρόληψη και για τη θεραπεία ασθενειών.

Ας συμφιλιωθούμε λοιπόν με την ασθένεια, ας την δούμε σαν μια προειδοποίηση, ότι κάτι δεν πάει καλά στη ζωή μας, κι ας αρχίσουμε να ερευνούμε σχολαστικά.

Αυτή η ενδοσκόπηση θα μας δώσει πολλά καινούργια στοιχεία για μια αλλαγή στον τρόπο που ζούμε. Μια τέτοια αλλαγή δεν σημαίνει μόνο καλύτερη Υγεία αλλά και αύξηση της ποιότητας της ζωής μας (Νέττλετον Σ. 2002,)

### **1.3.Κουλτούρα και ασθένεια**

Η κουλτούρα είναι για πολλούς μια ασαφής έννοια η οποία επιδέχεται πάμπολλους ορισμούς λόγω ακριβώς της αφηρημένης σύλληψής της και ανάλογα με τον τομέα τον οποίο μπορεί να προσδιορίζει ή να χρησιμοποιεί ως κριτήριο. Ένας από τους ευρύτερα αποδεχόμενους ορισμούς για την κουλτούρα (L.Kasapi, 2003) είναι του Good enough (1971) όπου ορίζει την κουλτούρα ως ένα σύνολο πεποιθήσεων ή επιπέδων (standards), που μοιράζονται από ένα σύνολο ατόμων, που βοηθάνε το άτομο

να αποφασίσει τι είναι , τι μπορεί να γίνει, πώς να αισθάνεται, τι να κάνει και πώς να το επιτύχει.

Αν και σαφώς δεν εξαρτώνται τα πάντα από την κουλτούρα, επηρεάζει σημαντικά την διαμόρφωση της αντίδρασης και της συμπεριφοράς μας απέναντι σε ένα ευρύ φάσμα ερεθισμάτων και καθημερινών καταστάσεων, κάνοντας διαφορετικά άτομα να αντιδρούν πιθανόν διαφορετικά σε κοινά ερεθίσματα και αναλόγως μη αναμενόμενα σε σχέση με το μέσο όρο. Είναι γνωστό ότι οι μετανάστες υποχρησιμοποιούν τις υπηρεσίες υγείας και ειδικά ψυχικής υγείας. Ένας από τους λόγους θεωρείται πως είναι η αντίληψη που μπορεί να έχει κάποιος για την ασθένεια και κατά πόσο συντρέχει ανάγκη-βάση των πεποιθήσεών τους, της κουλτούρας, των βιωμάτων τους-ιατρικής φροντίδας, δηλαδή αναγνώρισης της ασθένειας ως νόσου.

Η Κοινωνική Ανθρωπολογία έχει προσφέρει τα μέγιστα, μέσα από την ανάλυση και εξέταση των πολιτισμών, ακριβώς σ' αυτό τον τομέα της κατανόησης των συμπεριφορών και των χαρακτηριστικών εκείνων του εκάστοτε πολιτισμού που μπορεί να ευθύνονται ή να εμπλέκονται στην έκλυση των διαταραχών αλλά και στην αντίληψη του συνόλου αναφορικά με την ασθένεια και κατ' επέκταση απέναντι στον ασθενή.

## ΚΕΦΑΛΑΙΟ 2<sup>ο</sup>

### ΥΠΗΡΕΣΙΕΣ ΥΓΕΙΑΣ ΚΑΙ ΕΡΕΥΝΑ

#### 2.1. Εισαγωγή

Με τον όρο «υπηρεσίες υγείας», προσδιορίζονται όχι μόνο οι παρεχόμενες φροντίδες υγείας και ιατρικής περίθαλψης αλλά γενικότερα το πλέγμα των δομών με τις οποίες ρυθμίζονται η παραγωγή και διανομή των ιατρικών φροντίδων και καλύπτονται οι ιατρικές ανάγκες (Κυριόπουλος Γ., 1985).

Στις υπηρεσίες υγείας περιλαμβάνονται:

- Η εκτίμηση της ανάγκης και
- Η ζήτηση υπηρεσιών υγείας

#### 2.2. Εκτίμηση της ανάγκης

Ο Bradshaw (Bradshaw J., 1972) διέκρινε 4 είδη αναγκών:

α) την διαπιστωμένη ανάγκη,

Διαπιστωμένη είναι αυτή, που ο επαγγελματίας υγείας ορίζει ως ανάγκη ανάλογα με την κατάσταση που εξετάζει, μπορεί δε να μεταβάλλεται λόγω της εξέλιξης της γνώσης ή κοινωνικών αλλαγών.

β) την αισθανόμενη,

Αισθανόμενη (συνειδητή) ανάγκη είναι αυτή που αντιλαμβάνεται το άτομο ή το στενό του περιβάλλον.

γ) την εκφραζόμενη

Εκφραζόμενη, είναι η αισθανόμενη ανάγκη όταν μετατρέπεται σε ζήτηση.

δ) την συγκριτική ανάγκη.

Συγκριτική είναι η ανάγκη που προκύπτει όταν οι ανάγκες ενός πληθυσμού που χρησιμοποιεί συγκεκριμένες υπηρεσίες υγείας αντιστοιχισθούν σε έναν πληθυσμό με παρόμοια χαρακτηριστικά.

Η διάκριση σε «συνειδητές» (αντιληπτές) και «μη συνειδητές» (μη αντιληπτές) ανάγκες είναι χρήσιμη λόγω του γεγονότος ότι η ζήτηση υγειονομικών υπηρεσιών προκύπτει από τις «συνειδητές» ιατρικές ανάγκες παρ'όλο που η αντιστοιχία των δύο όρων δεν είναι απόλυτη, αφού αφενός άτομα με «συνειδητές» ιατρικές ανάγκες συχνά δεν προσφεύγουν στις υπηρεσίες υγείας και αφετέρου η ζήτηση

δεν σημαίνει πάντα την ύπαρξη αναγκών αλλά μπορεί να υπαγορεύεται από κοινωνικά κίνητρα, διαφήμιση ή νευρωσικές διαταραχές (Τριχόπουλος Δ., 1982).

Οι παράγοντες που προσδιορίζουν τις ιατρικές ανάγκες μπορεί να είναι επιδημιολογικοί, οικολογικοί, κοινωνιολογικοί και οικονομικοί (Κυριόπουλος Γ. και συν., 1999).

Η προσέγγισή τους επομένως επιτυγχάνεται μέσω δεικτών επιπολασμού ή επιπτώσεως της νόσου, μετρήσεις για τις σύγχρονες και μελλοντικές επιπτώσεις της κοινωνικοοικονομικής ανάπτυξης στα οικοσυστήματα, κοινωνιολογικές έρευνες και οικονομετρικά υποδείγματα σχετικά με την προσφορά, ζήτηση, αποτελεσματικότητα και αποδοτικότητα των υπηρεσιών υγείας.

Ανάγκη λοιπόν είναι, οι καταστάσεις που πιθανώς να γίνουν «ζήτηση», οπότε οι ανάγκες υγείας μπορούν να ορισθούν ως η διαφορά μεταξύ παρατηρούμενων και ιδεωδών επιπέδων υγείας.

Σε σχέση με τις υπηρεσίες υγείας το «ισοδύναμο της ανάγκης» μπορεί να εκφραστεί από τους απαιτούμενους πόρους για να προσφερθούν υπηρεσίες που θα ικανοποιούν αυτή την ανάγκη. Επομένως ισχύει η σχέση:

**πηλίκo αναγκών =απαιτούμενοι πόροι/κόστος ικανοποίησης αναγκών**  
(Donabedian A., 1973).

Ο σχεδιασμός των υπηρεσιών υγείας μέσω της προσέγγισης των αναγκών καταφέρνει να ανταποκρίνεται στο ενδιαφέρον του ατόμου ως ασθενή και ως φορολογούμενου (Mc.Ewin R. et al., 1976).

Ο σχεδιασμός μέσω της «μεθόδου εκτίμησης των αναγκών» (Hall T., Mejia A., 1978; Hornby P. et al., 1980) προσδιορίζει το είδος των υπηρεσιών που είναι απαραίτητο για την διατήρηση του καλού επιπέδου υγείας του πληθυσμού μέσω του προσδιορισμού των δεικτών ειδικής κατά αιτίας νοσηρότητας και θνησιμότητας σε περιφερειακό ή εθνικό επίπεδο, επεξεργασίας του αριθμού, της φύσης, της συχνότητας και της ποιότητας των παρεχομένων υπηρεσιών, υπολογισμού του συνολικού χρόνου απασχόλησης του ιατρονοσηλευτικού δυναμικού και του υπολογισμού του ετήσιου αριθμού ωρών εργασίας, θεωρώντας πως το επίπεδο υγείας βασίζεται σε ιατροτεχνικές προϋποθέσεις, κόστος, δυνατότητα παροχής υπηρεσιών και επιθυμία του πληθυσμού να τις αναζητήσει.

Η επικράτηση του ιατροκεντρικού υποδείγματος στην διανομή των υπηρεσιών που υιοθετείται στην συγκεκριμένη μέθοδο, καθρεφτίζει τους στόχους της ιατρικής κοινότητας, που προσβλέπει μέσω της εκτίμησης των ιατρικών αναγκών σε αποκάλυψη

ειδικών υγειονομικών προβλημάτων, σε επισήμανση πληθυσμιακών ομάδων «υψηλού κινδύνου» για ένα ή περισσότερα νοσήματα και ορθολογικό καθορισμό υγειονομικών προτεραιοτήτων.

Όσον αφορά την Ελλάδα, σύμφωνα πάντα με την έκθεση του Υπουργείου Υγείας (Κέντρο Μελετών Υπηρεσιών Υγείας Εργαστηρίου Υγιεινής & Επιδημιολογίας Ιατρικής Σχολής Αθηνών, 2000) κύριες αιτίες νοσηλείας του ελληνικού πληθυσμού για το 1995 ήταν κατά σειρά τα νοσήματα του κυκλοφοριακού συστήματος, τα νοσήματα του πεπτικού συστήματος, τα νεοπλάσματα, τα ατυχήματα και οι δηλητηριάσεις.

Η αύξηση του προσδόκιμου επιβίωσης, ο σύγχρονος τρόπος ζωής και η δημογραφική γήρανση έχουν αυξήσει τον επιπολασμό χρόνιων νοσημάτων όπως η αρτηριακή πίεση, οι εκφυλιστικές αρθροπάθειες και άλλα νοσήματα του μυοσκελετικού συστήματος, ο σακχαρώδης διαβήτης και οι παθήσεις των αισθητηρίων οργάνων.

Σε σχέση με την ψυχική υγεία των Ελλήνων, έρευνα που πραγματοποιήθηκε από το Εθνικό Κέντρο Κοινωνικών Ερευνών (1978-1980) σε δείγμα 4.083 ενηλίκων ατόμων, προέκυψε πως το 24% των γυναικών του δείγματος και το 11,1% των αντρών θα μπορούσαν να θεωρηθούν άτομα με σοβαρές ψυχικές διαταραχές. Οι κάτοικοι δε της Αθήνας φαίνεται να εμφανίζουν τα μεγαλύτερα ποσοστά ψυχοπαθολογίας. Σε άλλη έρευνα που πραγματοποιήθηκε από την ψυχιατρική κλινική του πανεπιστημίου Αθηνών (1984) τα αντίστοιχα ποσοστά ήταν 34,1% και 18,7% και αποδόθηκαν στις κοινωνικοοικονομικές αλλαγές. Σχετικά με την στοματική υγεία των Ελλήνων, το επίπεδο της παραμένει χαμηλό, αν και έχει βελτιωθεί λόγω της βελτίωσης του βιοτικού επιπέδου.

Οι κύριοι παράγοντες διαμόρφωσης των αναγκών του ελληνικού πληθυσμού είναι οι δημογραφικές και κοινωνικοοικονομικές εξελίξεις.

Είναι χαρακτηριστικό δε πως η ανεργία, η αύξηση των διαζυγίων, η υπερπροσφορά καταναλωτικών αγαθών και ο υπερκορεσμός σε υλικές απολαβές ατόμων εύπορων οικογενειών οδηγεί σε ομάδες που κινούνται έξω από τα καθορισμένα κοινωνικά πρότυπα και αντιμετωπίζουν προβλήματα υγείας που συνδέονται με κατάχρηση οινοπνεύματος, καπνού, χρήση ναρκωτικών, αφροδίσια νοσήματα, ανεπιθύμητες εγκυμοσύνες, ατυχήματα που συνδέονται με την βία καθώς και ψυχολογικά προβλήματα σιγουριάς και ασφάλειας λόγω χαλάρωσης των οικογενειακών θεσμών.



### 2.3. Ζήτηση υπηρεσιών υγείας

Η ζήτηση υπηρεσιών υγείας, εκφράζει την επιθυμία για κατανάλωση υπηρεσιών φροντίδας υγείας που μετατρέπεται σε ενεργή αναζήτηση αυτών των φροντίδων.

Η ζήτηση προσδιορίζεται από αντικειμενικά εκτιμώμενες ανάγκες αλλά εξαρτάται από πλήθος παραγόντων, όπως η διαθεσιμότητα των υπηρεσιών, το φύλο, η ηλικία, το εισόδημα, το επάγγελμα, η ασφαλιστική κάλυψη, το μορφωτικό επίπεδο, το επιδημιολογικό πρότυπο, το προσδόκιμο επιβίωσης, ο τόπος κατοικίας και η απόσταση από τις υπηρεσίες υγείας (Κυριόπουλος και συν., 1999).

Γενικά ισχύει η σχέση:

$$Dy=f(I, P, L),$$

όπου Dy η ζήτηση, I το εισόδημα, P οι τιμές και L η κρισιμότητα της ασθένειας (Σουλιώτης Κ., 2000).

Η δυνατότητα που έχει το ιατρικό σώμα να προκαλεί αύξηση των ζητούμενων ως «αναγκαίων» υπηρεσιών υγείας ή οι διοικητικοί παράγοντες, καθιερώνοντας υποχρεωτικές διαδικασίες, ονομάζεται «προκλητή ζήτηση» (Evans R.G., 1974).

Η ζήτηση υπηρεσιών υγείας διακρίνεται σε «φανερή» και «βουβή» (Alderson M., 1976).

Αν και η «φανερή» ζήτηση μπορεί να υπολογισθεί με βάση τους δείκτες χρησιμοποίησης, η «βουβή» μπορεί να εκτιμηθεί αν η φανερή ζήτηση αφαιρεθεί από την «αναμενόμενη ολική ζήτηση» που υπολογίζεται από ένα άλλο πληθυσμό με συγκρίσιμες ανάγκες, σωστή διαφώτιση και απόλυτη επάρκεια στις αντίστοιχες υπηρεσίες υγείας (Τριχόπουλος Δ., 1982).

Ωστόσο οι δείκτες χρησιμοποίησης δε επαρκούν πάντα για τον υπολογισμό της φανεράς ζήτησης. Τέτοιες είναι οι περιπτώσεις κατά τις οποίες η ζήτηση αν και φανερή δεν οδηγεί σε χρησιμοποίηση υγειονομικών υπηρεσιών λόγω ποσοτικής ανεπαρκείας τους.

Ο σχεδιασμός των υπηρεσιών υγείας μέσω της μεθόδου της ζήτησης προσδιορίζει τον αριθμό και το είδος των υπηρεσιών υγείας που θα χρησιμοποιήσει αποτελεσματικά ο πληθυσμός αναφορικά με το κόστος (Hall T., Mejia A. 1978; Hornby P., 1980). Επιτυγχάνεται λοιπόν εκτίμηση της μελλοντικής ζήτησης των υπηρεσιών υγείας συσχετίζοντας την ποσότητα με μεταβλητές όπως το εισόδημα το φύλο και η

ηλικία. Ωστόσο αυτή η μέθοδος κρίνεται λιγότερο αποτελεσματική συγκρινόμενη με την μέθοδο εκτίμησης αναγκών (Mc.Ewin R., Hall J., 1976).

Στην χώρα μας η μεταβολή της σύνθεσης του πληθυσμού δεν μπορεί παρά να επιφέρει μεταβολές στη ζήτηση των υπηρεσιών υγείας. Αναμένεται λοιπόν αύξηση των ζητούμενων υπηρεσιών υγείας, λόγω της μεγαλύτερης νοσηρότητας των ομάδων του πληθυσμού με ηλικία άνω των 65 ετών αλλά και διαφοροποίηση του είδους των ζητούμενων υπηρεσιών όπου παρατηρείται στροφή σε υπηρεσίες που σχετίζονται με νοσήματα φθοράς. Επίσης η μείωση του δείκτη βρεφικής θνησιμότητας που έχει ως αποτέλεσμα την επιβίωση νεογνών με ολοένα και μικρότερο βάρος αυξάνει την ζήτηση για συγκεκριμένες υπηρεσίες υγείας που θα ανταποκρίνονται στις ανάγκες αυτής της ευπαθούς ομάδας του πληθυσμού (Κέντρο Μελετών Υπηρεσιών Υγείας Εργαστηρίου Υγιεινής & Επιδημιολογίας Ιατρικής Σχολής Πανεπιστημίου Αθηνών, 2000).

Οι κοινωνικοοικονομικοί παράγοντες, πέρα από την σοβαρότητα της ασθένειας και την διαισθανόμενη και εκφραζόμενη νοσηρότητα αποτελούν καθοριστικούς παράγοντες της ζήτησης, είναι σημαντικοί αφού θεωρείται πλέον δεδομένο πως άτομα υψηλού πολιτισμικού επιπέδου εκφράζουν κυρίως ζήτηση για προληπτικές υπηρεσίες υγείας, σε αντίθεση με άτομα χαμηλότερου πολιτισμικού επιπέδου που στρέφονται στην νοσοκομειακή και εξωνοσοκομειακή περίθαλψη (Γείτονα Μ., Κυριόπουλος Γ., Καραλής Γ., 1997).

Ο όρος «προσφορά υπηρεσιών υγείας» εκφράζει την παροχή υπηρεσιών που προέρχονται από προμηθευτές υγείας (γιατρούς, νοσοκομεία κ.τ.λ.). Οι παράγοντες που προσδιορίζουν την προσφορά των υπηρεσιών υγείας είναι:

- α) η επάρκεια των ανθρωπίνων πόρων,
- β) η ποσότητα των διαθέσιμων υλικών πόρων,
- γ) το επιδημιολογικό πρότυπο
- δ) η κοινωνικοοικονομική ανάπτυξη

(Κυριόπουλος Γ. και συν., 1999).

Οι υπηρεσίες υγείας διακρίνονται σε ιδρυματικές (νοσοκομειακές), υπηρεσίες εξωτερικών ιατρείων και σε υπηρεσίες πρωτογενούς περίθαλψης (Τριχόπουλος Δ., 1982).

Τα εξωτερικά ιατρεία των νοσοκομείων και οι υπηρεσίες πρωτογενούς περίθαλψης αποτελούν την «ανοικτή περίθαλψη» (εξωνοσοκομειακή) που έχει ως σκοπό την πρόληψη, θεραπεία και αποκατάσταση στα πλαίσια της κοινότητας ώστε να περιορίζεται η καταφυγή στην νοσοκομειακή περίθαλψη (WHO & UNICEF, 1978).

Σύμφωνα με έκθεση του Υπουργείου Υγείας και Πρόνοιας (Μωραΐτης και συν.,1995) στις αγροτικές περιοχές της χώρας όπου ο πληθυσμός καλύπτεται ασφαλιστικά στην πλειονότητά του από τον Ο.Γ.Α. (70%) και το Ι.Κ.Α. (25%)και στο υπόλοιπο ποσοστό από άλλους ασφαλιστικούς φορείς, η εξωνοσοκομειακή περίθαλψη παρέχεται από τα κέντρα υγείας και τα περιφερειακά τους ιατρεία, από τα μικρά νοσοκομεία/κέντρα υγείας, από τα εξωτερικά ιατρεία των νομαρχιακών νοσοκομείων, από ιδιώτες γιατρούς συμβεβλημένους με τους ασφαλιστικούς φορείς καθώς και από ιδιώτες γιατρούς ή ιδιωτικές μονάδες υγείας (εργαστήρια και διαγνωστικά κέντρα).

Στις αστικές περιοχές όπου ο πληθυσμός καλύπτεται κατά 60% από το Ι.Κ.Α.,25% περίπου από τον Ο.Γ.Α. και το υπόλοιπο από άλλους φορείς, η εξωνοσοκομειακή περίθαλψη παρέχεται από υπηρεσίες του Ι.Κ.Α. και ελάχιστα πολυϊατρεία των άλλων οργανισμών, από τα εξωτερικά ιατρεία των νοσοκομείων , από τα κέντρα υγιεινής των νομαρχιών, από τα ιατρεία του ΠΚΠΑ, από τα ελάχιστα δημοτικά ιατρεία, από ιδιώτες γιατρούς συμβεβλημένους με τους ασφαλιστικούς φορείς και από τον ευρύτερο ιδιωτικό φορέα.

Όσον αφορά την νοσοκομειακή περίθαλψη αυτή χορηγείται από τα δημόσια νοσοκομεία, από ιδιωτικά νοσοκομεία και κλινικές και από νοσοκομεία που συνάπτουν συμβάσεις με τους ασφαλιστικούς φορείς για την περίθαλψη των ασφαλισμένων σε αυτούς.

Η παροχή οδοντιατρικής περίθαλψης στηρίζεται κατά μεγάλο ποσοστό στον ιδιωτικό τομέα αφού παρά το γεγονός ότι το σύνολο του ελληνικού πληθυσμού καλύπτεται για ασθένεια από τους διάφορους φορείς κοινωνικής ασφάλισης ,μόνο το 60% διαθέτει ασφάλιση για οδοντιατρική περίθαλψη. Ωστόσο τα τελευταία χρόνια αναπτύσσονται τα κέντρα υγείας που παρέχουν οδοντιατρικές φροντίδες στα πλαίσια του Ε.Σ.Υ. Παράλληλα η οδοντιατρική περίθαλψη προσφέρεται και από οδοντιάτρους των πολυϊατρείων του Ι.Κ.Α. καθώς και οδοντιάτρους που είναι συμβεβλημένοι με τα ασφαλιστικά ταμεία (Κωλέτση- Κούναρη Χ., Κυριόπουλος Γ., 1992).

Η φαρμακευτική περίθαλψη παρέχεται σχεδόν από ιδιωτικά φαρμακεία(Κυριόπουλος Γ., Γείτονα Μ., Σκουρολιάκου Μ., 1996).

Τα συστήματα υγείας βασίζονται σε μεγάλο βαθμό στους ανθρώπινους πόρους για αυτό και θεωρούνται συστήματα εντάσεως εργασίας. Το ιατρικό, νοσηλευτικό και λοιπό προσωπικό αποτελούν την βάση για την ανάπτυξη μεθόδων αποτελεσματικότερης και αποδοτικότερης χρήσης των πόρων σε ένα Εθνικό Σύστημα Υγείας (Simons S., Bonnett –Jones N. , 1989; Abel-Smith B.,1986).

Στην Ελλάδα ο αριθμός των γιατρών ήταν κατά το έτος 2000, 53.287, των οδοντιάτρων 12.858 και των φαρμακοποιών 11.639. Η αναλογία γιατρών στον πληθυσμό κυμαινόταν σε αρκετά υψηλά επίπεδα αφού αντιστοιχούσε 1 γιατρός σε 190 κατοίκους, ενώ για το νοσηλευτικό προσωπικό η αναλογία αυτή ήταν 1 προς 950 (Μαρούδιας Ν., 2001).

Εκτός από τον υπερπληθωρισμό των γιατρών, εξίσου μεγάλο πρόβλημα είναι και η ανισομερής γεωγραφική τους κατανομή. Στην Αττική όπου κατοικεί το 34% του ελληνικού πληθυσμού εξασκεί το λειτούργημα του το 54% των γιατρών, ενώ αντίθετα στην Στερεά Ελλάδα όπου κατοικεί το 8% του πληθυσμού λειτουργεί το 3% του ιατρικού σώματος. Η ανισοκατανομή αυτή ισχύει και για τις ειδικότητες των γιατρών αφού σε Αθήνα και Θεσσαλονίκη εργάζεται το 80% των αναισθησιολόγων, το 90% των νευροχειρουργών, το 75% των γυναικολόγων, και το 70% των καρδιολόγων. Οι αριθμοί που προαναφέρθηκαν αποκτούν ξεχωριστή σημασία αν εξετασθούν υπό το πρίσμα της παραγωγικότητας που ενώ επηρεάζεται από τον αριθμό των γιατρών, τον χρόνο εργασίας τους και την χρησιμοποίηση βοηθητικού προσωπικού, επηρεάζει την προσφορά των ιατρικών υπηρεσιών και τις αγοραίες τιμές τους (Arrow K. et al., 1961).

Έχει διαπιστωθεί ότι αύξηση στην προσφορά των γιατρών σε μια περιοχή συνοδεύεται από αύξηση της τιμής και της ποσότητας των ιατρικών υπηρεσιών (Ανδριώτη Δ., 1998). Αν και η σχέση αυτή θα έπρεπε να σημαίνει μείωση της ζήτησης βάσει της οικονομικής θεωρίας, στην περίπτωση των ιατρικών υπηρεσιών, λόγω της ατελούς γνώσης του ασθενή-καταναλωτή για την διάγνωση και θεραπεία οδηγεί συχνά σε επιπρόσθετες μη αναγκαίες υπηρεσίες (Monksma G., 1970). Σύμφωνα με τον Phelps (Phelps C.E., 1986) το ερώτημα που ανακύπτει στην προσφορά, είναι κατά πόσο νομιμοποιείται κοινωνικά η συμπεριφορά του παραγωγού (γιατρού) να αποφασίζει κατά το «δοκούν» για την χρησιμοποίηση περιορισμένων ατομικών ή κοινωνικών πόρων.

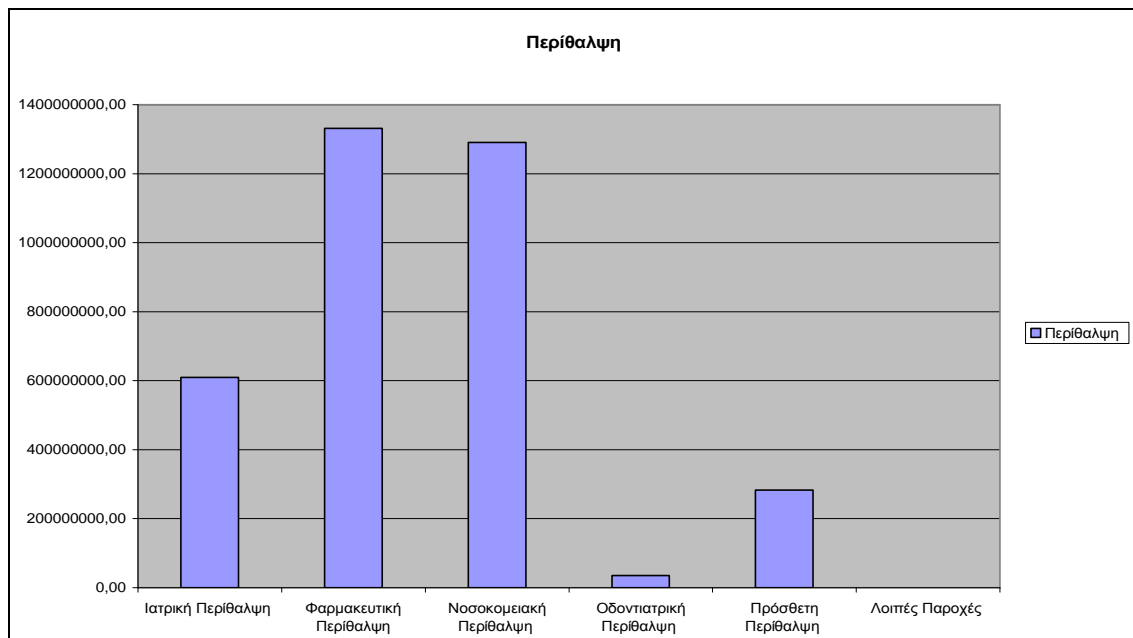
Το Εθνικό Σύστημα Υγείας αντικατοπτρίζει την πρόθεση της πολιτείας να παράγει και να χρηματοδοτεί αγαθά που εν συνεχεία διαθέτει ελεύθερα στον πολίτη χωρίς ή με μικρή συμμετοχή στο κόστος. Με αυτό το σκεπτικό αιτιολογείται ο παρεμβατικός ρόλος του κράτους ή των ταμείων, που αφενός στοχεύουν στην άρση της αβεβαιότητας που δημιουργεί ο μελλοντικός κίνδυνος της ασθένειας (Arrow K.J., 1973) και αφετέρου στην διαμόρφωση προτεραιοτήτων (Williams A., 1988).

Έτσι λοιπόν, οι ανεπτυγμένες χώρες έχουν οργανώσει την παροχή βασικών υπηρεσιών υγείας στις οποίες η πρόσβαση κατοχυρώνεται είτε μέσω των δικαιωμάτων

του πολίτη (εθνικά συστήματα τύπου Beveridge), είτε μέσω της συμμετοχής σε κοινωνικό επαγγελματικές κατηγορίες (συστήματα τύπου Bismarck) (Roemer M., 1993 ;Wall A., 1996).

Το ελληνικό σύστημα υγείας ανήκει στα συστήματα τύπου Bismarck (βάσει του 1ου άρθρου του νόμου 1397/83, «το κράτος έχει την ευθύνη για την παροχή υγείας στο σύνολο των πολιτών και οι υπηρεσίες υγείας παρέχονται ισότιμα σε κάθε πολίτη ανεξάρτητα από την κοινωνική και επαγγελματική του κατάσταση»), με τους φορείς κοινωνικής ασφάλισης να προσφέρουν παροχές σε είδος (εξωνοσοκομειακή, νοσοκομειακή, και φαρμακευτική περίθαλψη), παροχές σε χρήμα (επιδόματα ασθενείας ή ατυχημάτων) είτε συνδυασμό τους.

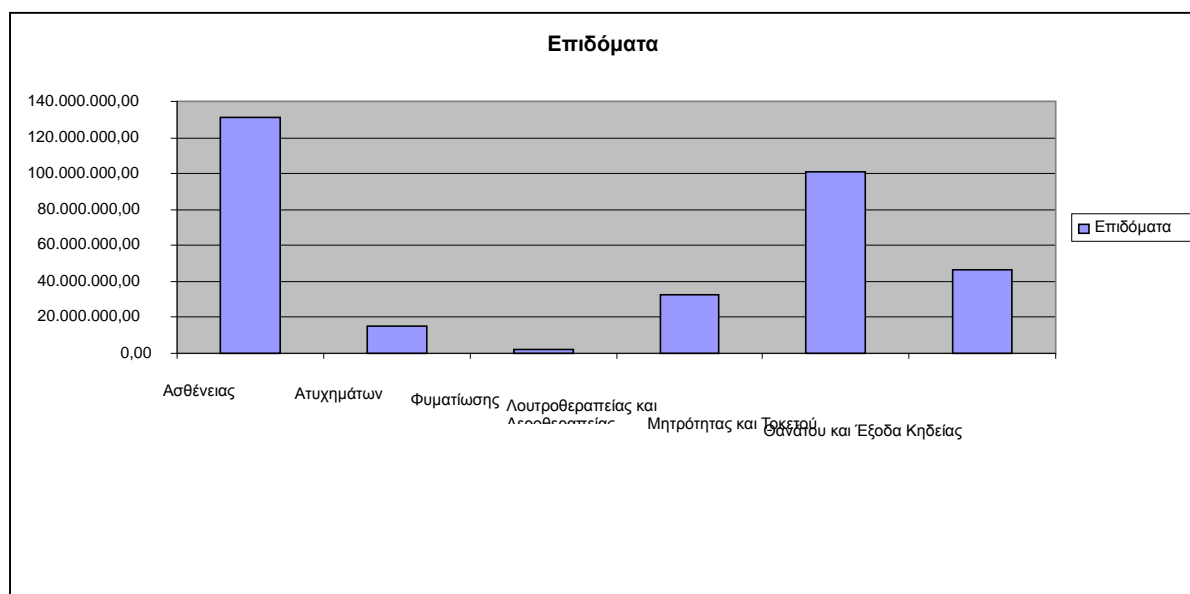
Τα ποσά που διατέθηκαν το 2002 για την υγεία, από τους προϋπολογισμούς του Υπουργείου Υγείας και Πρόνοιας αλλά και των Περιφερειακών Διοικήσεων, ως επιχορήγηση των Νοσηλευτικών Ιδρυμάτων ή ως γενικές δαπάνες του Υπουργείου άγγιξαν τα 2.344.658 χιλιάδες ευρώ. Η δαπάνη για την υγειονομική περίθαλψη των δημοσίων υπαλλήλων, συνταξιούχων, πολιτικών και στρατιωτικών άγγιξε τα 633.144 χιλιάδες ευρώ. Το ποσό που διατέθηκε το έτος 2002 για παροχές ασθένειας από τους φορείς κοινωνικής ασφάλισης που υπάγονται στο Υπουργείο Εργασίας και Κοινωνικών Ασφαλίσεων έφτασε τα 4.008.709.491 ευρώ. Από αυτό το ποσό, τα 3.637.712.693 ευρώ αφορούσε τις παρακάτω παροχές ασθένειας σε είδος:



- Ιατρική Περίθαλψη 609.401.717 ευρώ
- Φαρμακευτική Περίθαλψη 1.331.084.286 ευρώ
- Νοσοκομειακή Περίθαλψη 1.291.188.852 ευρώ

- Οδοντιατρική Περίθαλψη 35.459.505 ευρώ
- Πρόσθετη Περίθαλψη 282.834.311 ευρώ
- Λοιπές Παροχές 87.744.022 ευρώ.

Το υπόλοιπο ποσό των 370.996.798 ευρώ, αφορούσε τις παρακάτω παροχές ασθένειας σε χρήμα:

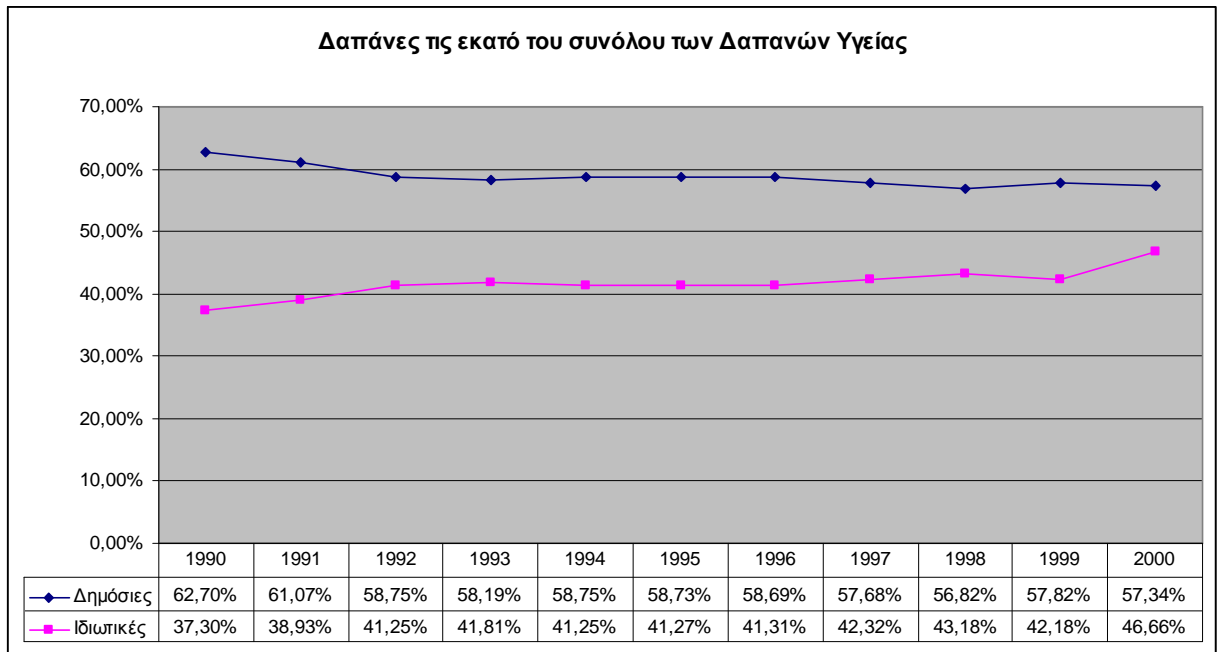


- Επιδόματα Ασθένειας 131.137.338 ευρώ
- Επιδόματα Ατυχημάτων 15.179.222 ευρώ
- Επιδόματα Φυματίωσης 1.831.652 ευρώ
- Επιδόματα Λουτροθεραπείας και Αεροθεραπείας 32.734.914 ευρώ
- Επιδόματα Μητρότητας και Τοκετού 101.058.208 ευρώ
- Βοηθήματα Θανάτου και Έξοδα Κηδείας 46.198.919 ευρώ
- Λοιπές Παροχές 42.856.545 ευρώ (Υπουργείο Εργασίας και Κοινωνικών Ασφαλίσεων, 2002).

Η ανάπτυξη όμως, της ιδιωτικής ασφάλισης και η μερική αδυναμία του δημόσιου τομέα να καλύψει τις ανάγκες υγείας του ελληνικού πληθυσμού, έχει οδηγήσει σε έντονη δραστηριοποίηση του ιδιωτικού τομέα κυρίως σε κέντρα διαγνωστικής ιατρικής και μαιευτηρίων. Στο πίνακα που ακολουθεί καταγράφονται, οι Δημόσιες και Ιδιωτικές Δαπάνες Υγείας ως % επί του Συνόλου των Δαπανών Υγείας την περίοδο 1990-2000 (Πίν. I - Σουλιώτης Κ., 2000).

Όπως παρατηρείτε στο παρακάτω πίνακα που περιγράφει τις Δημόσιες και Ιδιωτικές Δαπάνες τις εκατό του συνόλου των Δαπανών Υγείας την Περίοδο 1990 έως 2000 εμφανίζεται μια σταδιακή αύξηση των δαπανών στον ιδιωτικό τομέα για την υγεία

σε σχέση με τον δημόσιο τομέα.



Ιδιαίτερα από το 1995 και ύστερα η φθίνουσα πορεία του δημόσιου τομέα είναι σταθερή χωρίς κανένα σημάδι ανάκαμψης. Ο λόγος της μετακίνησης αυτής μπορεί να ερμηνευθεί ποικιλοτρόπως, με κυριότερους παράγοντες την ολοένα αυξανόμενη ζήτηση ιατρικής περίθαλψης χωρίς τη δυνατότητα κάλυψης από το δημόσιο τομέα σε συνδυασμό με την «καλύτερη» ποιότητα της περίθαλψης στον ιδιωτικό τομέα.

## ΚΕΦΑΛΑΙΟ 3ο

### ΣΤΑΤΙΣΤΙΚΗ ΣΤΙΣ ΥΠΗΡΕΣΙΕΣ ΥΓΕΙΑΣ

#### 3.1. Εισαγωγή

Στην έρευνα χρησιμοποίησης υπηρεσιών υγείας χρησιμοποιούνται επιδημιολογικές, κοινωνιολογικές, οικονομετρικές, καθώς και καθαρά στατιστικές μέθοδοι. Οι στατιστικές μέθοδοι έχουν ως αντικείμενο την κατασκευή δεικτών και υποδειγμάτων.

Ο Παγκόσμιος Οργανισμός Υγείας ορίζει τους δείκτες ως «μεγέθη που μετρούν μεταβολές» (Global Strategy for Health for All by the Year 2000, 1981), χρησιμοποιούνται δε όταν οι μεταβολές αυτές δεν είναι δυνατόν να μετρηθούν άμεσα.

Η κατασκευή υποδειγμάτων έχει ως σκοπό αφενός την πρόβλεψη και αφετέρου τον εντοπισμό των παραγόντων που επηρεάζουν την χρησιμοποίηση των υπηρεσιών υγείας. Οι εξελίξεις στις σύγχρονες κοινωνίες, τόσο σε κοινωνικοοικονομικό επίπεδο όσο και σε επίπεδο εφαρμογής μεθόδων οργάνωσης και προγραμματισμού στις υπηρεσίες υγείας καθιστούν τα στατιστικά υποδείγματα πολύτιμο εργαλείο σχεδιασμού αλλά και άσκησης πολιτικής.

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, τα κράτη, για να καθορίσουν την πολιτική τους, να προβούν στην θέσπιση θεσμών, να σχεδιάσουν και να οργανώσουν φορείς, δομές, υπηρεσίες και μηχανισμούς, έχουν ανάγκη την πληροφορία.

Αν και οι περισσότερες χώρες έχουν εδώ και αρκετά χρόνια αντιληφθεί την σημασία ερευνών για την υγεία αλλά και την αξία των συμπερασμάτων που προκύπτουν από αυτές σε θέματα όπως η νοσηρότητα, η θνησιμότητα, η ψυχική υγεία ή οι δαπάνες υγείας, δεν έχουν επιχειρήσει εκτενώς την αξιολόγηση των συστημάτων υγείας.

Ο Παγκόσμιος Οργανισμός Υγείας, θεωρώντας αφενός ότι οι κύριοι στόχοι των Συστημάτων Υγείας είναι η βελτίωση υγείας του πληθυσμού, η βελτίωση της ανταποκρισιμότητας και η δίκαιη κατανομή των πόρων και αφετέρου ότι η μέτρηση της «Ανταποκρισιμότητας των Συστημάτων Υγείας» αποτελεί ένα μέσο περιγραφής των αδυναμιών ή των ελλείψεων του Συστήματος Υγείας αλλά και των προσδοκιών του πολίτη από αυτό.



### 3.2. Δείκτες

Η χρησιμοποίηση υπηρεσιών υγείας εκφράζεται ως το ποσοστό ατόμων σε ανάγκη για κάποια συγκεκριμένη υπηρεσία, που τελικά την λαμβάνουν σε κάποια χρονική περίοδο (WHO, 1981).

Ιδιότητες των δεικτών είναι :

- η εγκυρότητα (πρέπει να μετρούν αυτό που καλούνται να μετρήσουν),
- η αμεροληψία (πρέπει να μην επιδέχονται διαφορετική ερμηνεία)
- η ευαισθησία (να αντιλαμβάνονται τις αλλαγές μιας δεδομένης κατάστασης).

Για την πρωτοβάθμια περίθαλψη η χρησιμοποίηση εκφράζεται μέσω δεικτών επισκέψεων σε γενικούς γιατρούς ή άλλους φορείς πρωτοβάθμιας περίθαλψης καθώς και με βάση την συχνότητα συνταγογράφησης, εξαρτώνται δε σε μεγάλο βαθμό από την ηλικία και το φύλο.

Όσον αφορά τους δείκτες που χρησιμοποιούνται για την νοσοκομειακή περίθαλψη εκφράζουν:

- κατοίκους ανά νοσηλευτικό κρεβάτι,
- εισαγωγές ανά 100.000 κατοίκους,
- μέση διάρκεια νοσηλείας
- μέση κάλυψη νοσηλευτικών κρεβατιών.

Συχνότερα χρησιμοποιούνται οι δείκτες:

$$\text{Ποσοστό Κάλυψης} = \frac{\text{Σύνολο Ημερών Νοσηλείας Έτους} * 100}{\text{Αριθμός Κλινών Έτους} * 365}$$

$$\text{Μέση Διάρκεια Νοσηλείας} = \frac{\text{Σύνολο Ημερών Έτους} / \text{Αριθμός Εισαγωγών Έτους}}$$

**Ρυθμός Εισροής Ασθενών (Pκ) = 365 \* Ποσοστό Κάλυψης / 100 \* Μέση Διάρκεια Νοσηλείας**. Εκφράζει τον ρυθμό με τον οποίο χρησιμοποιούνται οι κλίνες σε μια δεδομένη χρονική περίοδο.

$$\text{Διάστημα Εναλλαγής} = (365 / Pκ) - \text{Μέση Διάρκεια Νοσηλείας.}$$

Εκφράζει τον ρυθμό εναλλαγής των ασθενών ή τον μέσο αριθμό ημερών που μένει κενή η κλίνη (Σιτισούρας Α., 2001).

Οι δείκτες που χρησιμοποιούνται για την νοσοκομειακή περίθαλψη έχει παρατηρηθεί ότι εξαρτώνται από την ηλικία και το φύλο. Συγκεκριμένα οι ημέρες

παραμονής των ασθενών στα νοσοκομεία αυξάνουν σε συνάρτηση με την ηλικία (Munozetal., 1989) και είναι περισσότερες για τους άντρες (Anderson JG, 1973).

Οι δείκτες χρησιμοποίησης νοσοκομείων εκτός από το ενδιαφέρον που παρουσιάζουν για την έρευνα των υπηρεσιών υγείας, είναι καθοριστικής σημασίας για την Οικονομία της Υγείας αφού η διαμόρφωση του Νοσοκομειακού Κόστους έχει βρεθεί ότι επηρεάζεται από:

- Τη Μέση Διάρκεια Νοσηλείας
- Το Ποσοστό Κάλυψης των Κλινών
- Τον Αριθμό των Κλινών
- Τον Αριθμό των Ασθενών ανά Κλινική.
- Με δεδομένο το ποσοστό κάλυψης, το μέσο κόστος νοσηλείας ανά περιστατικό μειώνεται όταν μειώνεται η διάρκεια παραμονής των ασθενών (Lave J. Et al ., 1972).

Επιπλέον οι δείκτες χρησιμοποίησης «κλειστής περίθαλψης» αποτελούν ένα εργαλείο αξιολόγησης του εσωτερικού περιβάλλοντος των Νοσοκομείων.

### **3.3. Εργαλεία της έρευνας**

Η διαδικασία διεξαγωγής μιας στατιστικής έρευνας ακολουθεί δύο διακριτά στάδια. Κατά το πρώτο, το στάδιο σχεδιασμού, διατυπώνονται οι σκοποί της έρευνας και προσδιορίζονται τα ζητούμενα βάσει υποθέσεων εργασίας, ακολούθως δε επιλέγεται η μέθοδος πραγματοποίησης της και σχεδιάζεται η βήμα προς βήμα υλοποίησή της. Κατά το δεύτερο, το στάδιο υλοποίησης, συλλέγονται τα απαραίτητα στοιχεία, ακολουθεί η επεξεργασία και ανάλυση των δεδομένων που προκύπτουν και γίνεται η σύνθεσή τους και διατύπωση των σχετικών συμπερασμάτων.

#### **3.3.1. Σχεδιασμός της Έρευνας**

Ο προσδιορισμός και η διατύπωση των ζητούμενων μιας ποσοτικής κοινωνικής έρευνας στηρίζονται σε κάποιες βασικές υποθέσεις εργασίας, των οποίων απαιτείται η περαιτέρω διερεύνηση, ούτως ώστε να διαπιστωθεί κατά πόσον αυτές ισχύουν και να αποτυπωθεί η υφιστάμενη κατάσταση στο συγκεκριμένο πλαίσιο αναφοράς.

### 3.3.2. Κατάρτιση Ερωτηματολογίων

Στις ποσοτικές έρευνες χρησιμοποιείται ευρύτατα η συμπλήρωση ερωτηματολογίων, στα οποία αποτυπώνεται το περιεχόμενο των προσωπικών συνεντεύξεων που λαμβάνονται επί τούτου.

Η συνέντευξη είναι η τεχνική που έχει σκοπό να οργανώσει μία σχέση προφορικής επικοινωνίας ανάμεσα σε δύο πρόσωπα, το συνεντευκτή και τον ερωτώμενο, έτσι ώστε να επιτρέψει στον πρώτο τη συλλογή ορισμένων πληροφοριών απ' τον δεύτερο πάνω σ' ένα συγκεκριμένο αντικείμενο.

Η έρευνα που γίνεται με τη μέθοδο αυτή πρέπει να έχει προετοιμαστεί με κάθε λεπτομέρεια, μια και ο συνεντευκτής οφείλει στο ελάχιστο χρονικό διάστημα, να αποκτήσει πολυάριθμες και τις πιο σημαντικές πληροφορίες.

Αυτή η προετοιμασία οδηγεί στην κατάστρωση ερωτηματολογίου από τον ερευνητή, ο οποίος αναλαμβάνει:

α) Να μετατρέψει τους σκοπούς που επιδιώκει η έρευνα σε επί μέρους ερωτήσεις.

β) Να προσαρμόσει το ερωτηματολόγιο στα πρόσωπα με τα οποία θα γίνει η συνέντευξη.

γ) Να ενημερώσει τους συνεντευκτές γι' αυτά έτσι ώστε να μπορέσουν να εκθέσουν με σαφήνεια τις ερωτήσεις στα πρόσωπα που θα υποβληθούν στη συνέντευξη και να προδιαθέσει το ερωτώμενο πρόσωπο να μεταδώσει αυθόρμητα τις πληροφορίες που περιμένουν από αυτό.

Τα ερωτηματολόγια, τα οποία συμπληρώνονται από τους συμμετέχοντες στην έρευνα, αποτελούνται συνήθως από δύο μέρη: το κοινό μέρος (περιέχει δημογραφικά στοιχεία) και το ειδικό μέρος (περιέχει ερωτήσεις με στοιχεία της έρευνας).

Η συνέντευξη στις Κοινωνικές Επιστήμες - κι' αυτό την ξεχωρίζει από την δημοσιογραφική - είναι μια «συστηματοποιημένη» συνέντευξη - συζήτηση με έναν ορισμένο αριθμό χαρακτηριστικών που καθορίζουν την τεχνική της εφαρμογής της.

Οι μορφές συνέντευξης που χρησιμοποιούνται στις ποσοτικές έρευνες διακρίνονται σε δύο βασικές κατηγορίες, που είναι:

### **3.3.3 Η δομημένη συνέντευξη**

Με τον όρο αυτό εννοούμε τη συνέντευξη εκείνη όπου ο ερωτώμενος προτρέπεται στο να απαντήσει σε μια σειρά ερωτήσεων που ο αριθμός, η σειρά και το περιεχόμενο προκαθορίζεται από το έντυπο της συνέντευξης. Οι απαντήσεις καταγράφονται ή λέξη προς λέξη ή κωδικοποιημένες.

### **3.3.4 Η εντοπισμένη συνέντευξη (FOCUSED INTERVIEW)**

Ο συνεντευκτής θέτει το γενικό πλαίσιο και εντοπίζει τα σημεία ιδιαίτερου ενδιαφέροντος, ούτως ώστε εκεί να εστιασθεί η ανάπτυξη του θέματος (ημι-κατευθυνόμενη συνέντευξη). Η τεχνική αυτή χρησιμοποιείται κυρίως σε πληροφοριοδότες - κλειδιά, για τους οποίους εκ των προτέρων θεωρούμε πως έχουν ιδιαίτερη γνώση επί του διερευνώμενου θέματος.

Μετά την κατάστρωση του κατάλληλου ερωτηματολογίου, γίνεται «πιλοτική έρευνα» (προ-έρευνα) για να προσδιορισθεί η λειτουργικότητα του ερωτηματολογίου και να διαμορφωθεί οριστικά η δομή του. Κατ' αυτή την διαδικασία, σ' ένα βαθμό γίνεται χρήση των τεχνικών της ποιοτικής προσέγγισης. Εν συνεχεία ακολουθεί η επιλογή δείγματος από το σύνολο του πληθυσμού, στο οποίο θα πραγματοποιηθεί η επιτόπια έρευνα με τη διεξαγωγή συνεντεύξεων μέσω ερωτηματολογίων, μια διαδικασία ιδιαίτερα σημαντική για την πορεία και τα αποτελέσματα της έρευνας.

### **3.3.5. Επιλογή του δείγματος**

Το πιο δύσκολο κομμάτι στη διεξαγωγή μιας έρευνας είναι η επιλογή του δείγματος. Προκειμένου να διασφαλιστεί η εγκυρότητα των αποτελεσμάτων, το δείγμα μιας έρευνας πρέπει να είναι αντιπροσωπευτικό του πληθυσμού στον οποίο θα αναχθούν τα τελικά συμπεράσματα. Δυστυχώς κάτι τέτοιο είναι εξαιρετικά δύσκολο. Η επιλογή του δείγματος έχει σχέση με το θέμα της έρευνας και στην συγκεκριμένη περίπτωση αφορά ασθενείς, ιατρούς, νοσηλευτές, άλλους φροντιστές της υγείας, φοιτητές ιατρικής και φοιτητές παραϊατρικών σχολών.

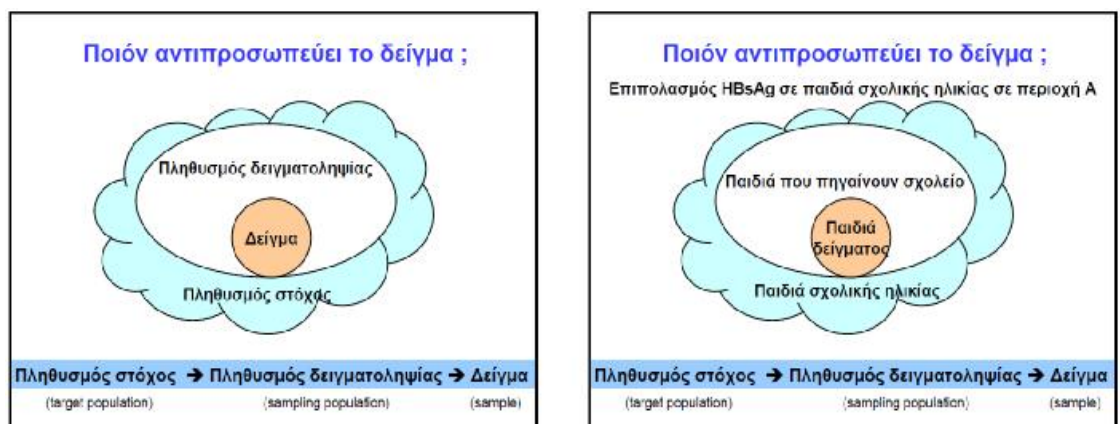
Η δειγματοληψία είναι απαραίτητη μιας και η συλλογή πληροφοριών από ολόκληρο τον πληθυσμό ή πολύ μεγάλο μέρος του αποτελεί μια διαδικασία με μεγάλο

κόστος, μικρή ταχύτητα και ανεπαρκή ακρίβεια.

Η επιλογή του δείγματος θα πρέπει να γίνεται με τέτοιο τρόπο έτσι ώστε τα αποτελέσματα της έρευνας να μπορούν στη συνέχεια να γενικευτούν στον ευρύτερο πληθυσμό.

Σε μελέτες που αφορούν περιγραφική επιδημιολογία ή έρευνες επιπολασμού, στόχος της δειγματοληψίας είναι η αντιπροσωπευτικότητα, ενώ σε έρευνες αναλυτική επιδημιολογίας ή συγκριτικές έρευνες κύριος στόχος είναι η συγκρισιμότητα και δευτερευόντως η αντιπροσωπευτικότητα.

Στην παρακάτω εικόνα (Εικ. 2 ) δίνονται δύο παραδείγματα όσον αφορά την ορθή επιλογή δείγματος σε σχέση με την αντιπροσωπευτικότητα



**Εικόνα 2: Δειγματοληψία**

Για την διασφάλιση της αντιπροσωπευτικότητας είναι αναγκαία η στοχοποίηση των υπό μελέτη χαρακτηριστικών όπως:

- Χαρακτηριστικά ατόμου π.χ. δημογραφικά χαρακτηριστικά (ηλικία, φύλο, κοινωνικο-οικονομικό επίπεδο κλπ)
- Χαρακτηριστικά τόπου π.χ. αστικές/αγροτικές περιοχές
- Χαρακτηριστικά χρόνου π.χ. κατανομή κατά εποχή, ημέρα, ώρα πριν το ξεκίνημα της δειγματοληψίας.

Οι τύποι της δειγματοληψίας που συνήθως αναφέρονται:

- Δειγματοληψία με γνωστή πιθανότητα επιλογής (probability sampling)
- Δειγματοληψία με άγνωστη πιθανότητα επιλογής(non-probability sampling)
- Οι μέθοδοι δειγματοληψίας με γνωστή πιθανότητα επιλογής είναι:
- Απλή τυχαία δειγματοληψία (simple random sampling)
- Συστηματική δειγματοληψία (systematic sampling)
- Δειγματοληψία κατά συστάδες (cluster sampling)

ενώ πρόσθετα στοιχεία δειγματοληπτικού σχεδιασμού(προαιρετικά) αποτελούν:

- Δειγματοληψία με διαστρωμάτωση (stratified sampling)
- Πολυσταδιακή δειγματοληψία (multi stage sampling)

Στη δειγματοληψία με άγνωστη πιθανότητα επιλογής υπάρχει:

- Δειγματοληψία αναλογίας (quota sampling) το δείγμα αντανακλά τη δημογραφική δομή του πληθυσμού,
- Δειγματοληψία ευκολίας (convenience sampling): εύκολη πρόσβαση στον πληθυσμό
- Δειγματοληψία χιονοστιβάδας (snowball sampling) από φίλο σε φίλο, από γείτονα σε γείτονα κλπ.
- Δειγματοληψία σκοπιμότητας (purposive sampling) επιλογή ατόμων που «κρίνεται» ότι πρέπει να μελετηθούν.

Προβλήματα αποτελούν η αδυναμία υπολογισμού του τυχαίου σφάλματος, η μεγάλη πιθανότητα συστηματικού σφάλματος ενώ απαιτείται μεγάλη εμπειρία και επανειλημμένη εφαρμογή. (Σχολή Διοίκησης Δημόσιας Υγείας)

Οι κυριότερες μέθοδοι στατιστικών αναλύσεων στον τομέα υγείας είναι τα  $X^2$ , t-test, ανάλυση συνδιακύμανσης, παλινδρόμησης και αναλύονται διεξοδικά στο επόμενο κεφάλαιο.

## ΚΕΦΑΛΑΙΟ 4<sup>ο</sup>

### ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ

#### 4.1. Εισαγωγή

Οι έρευνες στις υπηρεσίες υγείας, επιχειρούν, ανιχνεύοντας τους παράγοντες που προσδιορίζουν την πρόσβαση στις υπηρεσίες υγείας, αφενός την μελέτη της πιθανότητας χρησιμοποίησης των υπηρεσιών υγείας και αφετέρου του μεγέθους της (για την ανάλυση των δεδομένων της έρευνας χρησιμοποιούνται τα στατιστικά πακέτα SAS και SPSS).

Παρακάτω παρουσιάζονται και αναλύονται οι κυριότερες μέθοδοι στατιστικής ανάλυσης στον τομέα της υγείας.

#### 4.2.2 Κανονική κατανομή

Η κανονική κατανομή αναφέρεται σε συνεχείς μεταβλητές αποτελώντας μία συνεχή συνάρτηση πυκνότητας πιθανότητας.

Χρησιμοποιείται ως μία πρώτη προσέγγιση για να περιγραφούν τυχαίες μεταβλητές πραγματικών τιμών, οι οποίες τείνουν να συγκεντρώνονται γύρω από μια μέση τιμή. Η κανονική κατανομή αποτελεί την πιο σημαντική κατανομή της στατιστικής μεθοδολογίας για τους εξής βασικούς λόγους:

- Την κανονική κατανομή ακολουθούν είτε με ακρίβεια είτε με μεγάλη προσέγγιση τα περισσότερα συνεχή φαινόμενα.
- Πολλές ασυνεχείς κατανομές πιθανοτήτων μπορούν να προσεγγιστούν μέσω της κανονικής κατανομής. Για παράδειγμα πολλά πληθυσμιακά χαρακτηριστικά, όπως το ύψος, το βάρος η βαθμολογία σε διαγώνισμα, κ.λπ.
- Η κανονική κατανομή αποτελεί σύμφωνα με το κεντρικό οριακό θεώρημα (το άθροισμα ενός ικανοποιητικά μεγάλου αριθμού ανεξάρτητων και ισόνομων τυχαίων μεταβλητών προσεγγίζεται από την κανονική κατανομή) τη βάση της στατιστικής συμπερασματολογίας ή επαγωγικής στατιστικής.
- Τυχαία σφάλματα που εμφανίζονται σε διάφορες μετρήσεις έχουν κανονική κατανομή. Γι' αυτό το λόγο η Κανονική κατανομή αναφέρεται πολλές φορές

και ως κατανομή σφαλμάτων.

Η γραφική παράσταση της σχετιζόμενης συνάρτησης πυκνότητας πιθανότητας έχει σχήμα "καμπάνας", και είναι γνωστή ως Γκαουσιανή συνάρτηση ή κωδωνοειδής καμπύλη:

$$f(x) = \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(x-\mu)^2}{2s^2}}$$

Όπου:

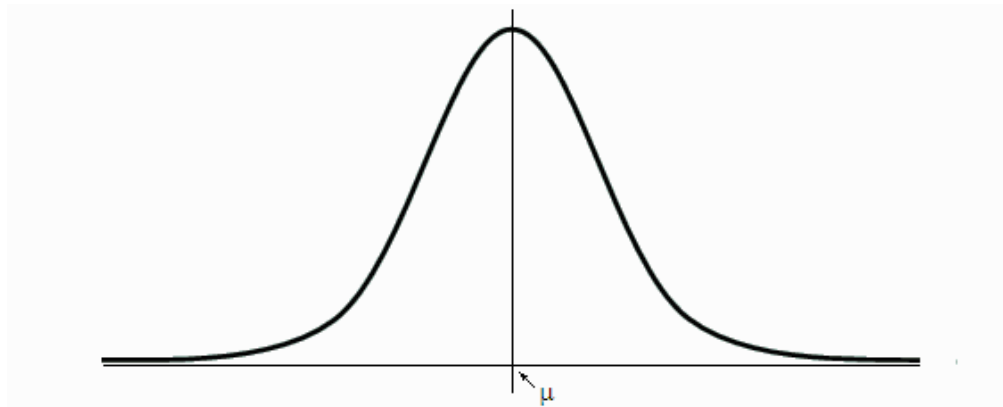
$e$ = η βασική των νεπέρειων λογαρίθμων (2,71828)

$\pi$ =3,1459

$\mu$ = ο μέσος του πληθυσμού

$\sigma$ = η τυπική απόκλιση του πληθυσμού

$x$ = μια τιμή της συνεχούς τυχαίας μεταβλητής στο διάστημα  $-\infty$  έως  $+\infty$



#### 4.2.3 Κατανομές $X^2$ , t και F

Οι κατανομές  $X^2$ , t και F είναι τρεις συνεχείς κατανομές οι οποίες είναι εξίσου σημαντικές και χρήσιμες στην Στατιστική Συμπερασματολογία. Παρατηρείστε στον ορισμό αυτών των κατανομών, ότι πρόκειται για κατανομές συναρτήσεων ανεξάρτητων τυχαίων μεταβλητών τυχαίο δείγμα μεγέθους n από έναν πληθυσμό εννοούμε n ανεξάρτητες και ισόνομες τυχαίες μεταβλητές  $X_1, X_2, \dots, X_n$ .

Την κατανομή  $X^2$  εισήγαγε ο F. R. Helmert το 1876, την κατανομή t το 1908 ο W. L. Gosset και την κατανομή F ο G.W Snedecor (το 1934) ο οποίος της έδωσε το όνομα F προς τιμήν του διακεκριμένου στατιστικού (και γενετιστή) R. A. Fisher, γι' αυτό στη βιβλιογραφία συναντάται και ως κατανομή Fisher, ως κατανομή Snedecor ή ως κατανομή Snedecor-Fisher.

Η κατανομή t είναι γνωστή και ως κατανομή Student, ψευδώνυμο του Gosset με το οποίο δημοσίευε τα άρθρα του.



### 4.2.3.1 Κατανομή $X^2$

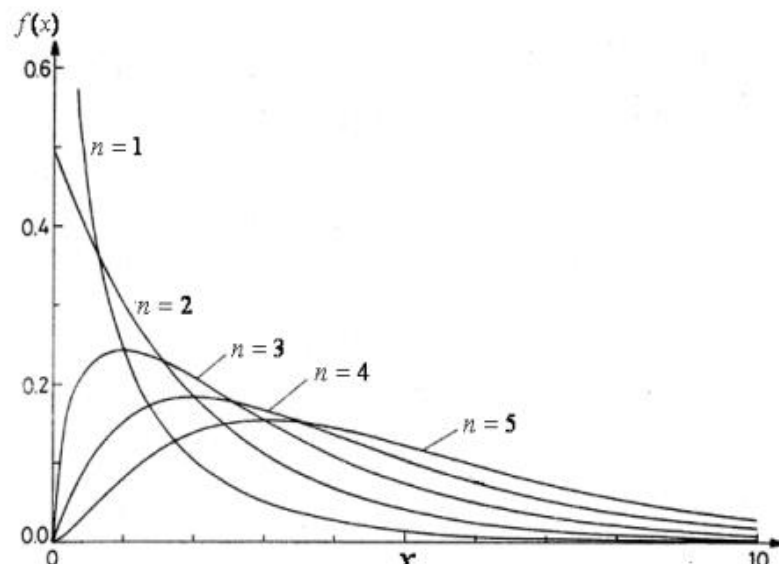
Έχοντας  $n$  ανεξάρτητες τυποποιημένες κανονικές τυχαίες μεταβλητές  $Z_1, Z_2, \dots, Z_n$  και  $Z_i \sim N(0, 1), i = 1, 2, \dots, n$  τότε η κατανομή της τυχαίας μεταβλητής,

$$X = Z_1^2 + Z_2^2 + \dots + Z_n^2$$

ονομάζεται κατανομή χι-τετράγωνο με  $n$  βαθμούς ελευθερίας και συμβολίζεται με  $X_n^2$ . Για κάθε τιμή του  $n$  παίρνουμε και μια άλλη κατανομή χι-τετράγωνο. Η τυχαία μεταβλητή  $X$  που ακολουθεί μια κατανομή  $X_n^2$  δεν παίρνει αρνητικές τιμές.

Στο παρακάτω σχήμα φαίνεται η γραφική παράσταση της συνάρτησης πυκνότητας

της  $X = Z_1^2 + Z_2^2 + \dots + Z_n^2 \sim X_n^2$  για διάφορες τιμές του  $n$ .

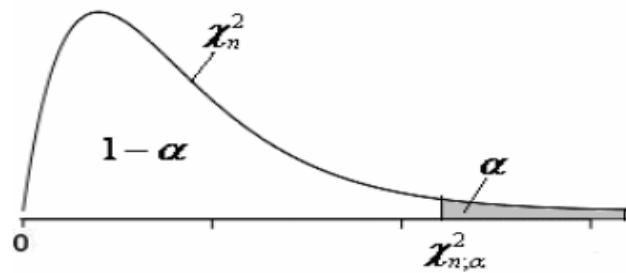


Η μέση τιμή της  $X_n^2$  είναι ίση με  $n$ ,  $E(X)=n$  και η διασπορά της είναι ίση με  $2n$ ,  $V(X) = 2 \cdot n$ .

Στο παραπάνω σχήμα παρατηρούμε ότι όσο το  $n$  αυξάνεται τόσο η γραφική παράσταση της συνάρτησης πυκνότητας της  $X_n^2$  γίνεται πιο συμμετρική. Για  $n > 30$ , προσεγγίζεται πολύ ικανοποιητικά από την κανονική  $N$ .

Μία τυχαία μεταβλητή  $X$  αν ακολουθεί την κατανομή  $X_n^2$  δηλαδή  $X \sim X_n^2$  τότε τα  $\alpha$ -ποσοστιαία σημεία της  $X_n^2$  τα οποία έχει επικρατήσει να συμβολίζονται με  $X_{n,\alpha}^2$  ή με  $X_n^2(\alpha)$  όπως φαίνεται στο παρακάτω σχήμα είναι εκείνη η τιμή  $X$  για την οποία

$$P(X > X_{n;\alpha}^2) = \alpha \text{ η ακόμα } P(X \leq X_{n;\alpha}^2) = 1 - \alpha .$$



Τα  $\alpha$ - ποσοστιαία σημεία της  $X$  τετράγωνο κατανομής δίνονται από πίνακες κρίσιμων τιμών της κατανομής στο παράτημα 1 στον Πίνακα 1.

#### 4.2.3.2 Κατανομή $t$ ή Κατανομή Student (t-distribution ή Student distribution)

Έχοντας μια  $Z$  τυχαία μεταβλητή η οποία ακολουθεί την τυποποιημένη κανονική κατανομή,  $Z \sim N(0,1)$ , και  $S_n$  μια τυχαία μεταβλητή ανεξάρτητη από την  $Z$  η οποία ακολουθεί την κατανομή  $\chi_n^2$  δηλαδή  $S_n \sim \chi_n^2$  τότε η κατανομή της τυχαίας μεταβλητής είναι:

$$T = \frac{z}{\sqrt{\frac{S_n}{n}}}$$

Η κατανομή αυτή ονομάζεται κατανομή  $t$  ή κατανομή Student με  $n$  βαθμούς ελευθερίας και

συμβολίζεται με  $t_n$ . Για κάθε τιμή του  $n$  παίρνουμε και μια άλλη κατανομή  $t$ .

Οι τιμές που παίρνει η κατανομή  $t$  μιας τυχαία μεταβλητή  $T$  που ακολουθεί μια  $t_n$  είναι από το  $-\infty$  έως το  $+\infty$  όπως και η κανονική κατανομή.

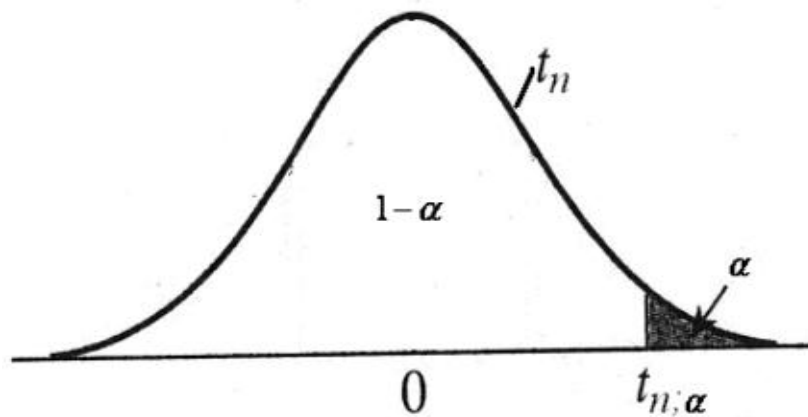
Στο παρακάτω γραφική παράσταση της συχνότητας της  $T \sim t_n$  για  $n=2$  και για  $n=10$  παρατηρείτε εκτός από την κωδωνοειδή συμμετρική μορφή της όπως η  $Z \sim N(0,1)$ , ότι έχει πιο πεπλατυσμένες ουρές. Όσο το  $n$  αυξάνεται η  $t_n$  προσεγγίζεται ικανοποιητικά από την κανονική  $N(0, \frac{n}{n-2})$ . Για  $n > 30$ , προσεγγίζεται πολύ καλά από την τυποποιημένη κανονική  $N(0, 1)$ .

Η μέση τιμή της  $t_n$  είναι ίση με 0 για  $n > 1$  και η διασπορά για  $n > 2$  είναι:

$$V(T) = \frac{n}{n-2}$$

Όπως φαίνεται στο παρακάτω σχήμα, αν μια τυχαία μεταβλητή  $T$  ακολουθεί την

κατανομή  $t_n$  δηλαδή αν  $T \sim t_n$ , τότε το  $t_{n,\alpha}$  είναι εκείνη η τιμή της  $T$  για την οποία ισχύει  $P(T > t_{n,\alpha}) = \alpha$  και  $P(T \leq t_{n,\alpha}) = 1 - \alpha$ .



Οι διάφορες τιμές που παίρνουν τα  $\alpha$  και  $n$  τα  $\alpha$ -ποσοστιαία σημεία της  $t_n$  παρουσιάζονται στο παράτημα 1 στον Πίνακα 2.

### Κατανομή F (F- distribution)

Έχοντας  $S_n$  και  $S_m$  δύο ανεξάρτητες τυχαίες μεταβλητές οι οποίες ακολουθούν τις κατανομές  $\chi_n^2$  και  $\chi_m^2$  αντίστοιχα. Τότε, η κατανομή της τυχαίας μεταβλητής είναι

$$F = \frac{\frac{S_n}{n}}{\frac{S_m}{m}}$$

και ονομάζεται κατανομή F με  $n$  και  $m$  βαθμούς ελευθερίας και συμβολίζεται με  $F_{n,m}$

Η τυχαία μεταβλητή  $X$  που ακολουθεί μια κατανομή  $F_{n,m}$  δεν παίρνει αρνητικές τιμές.

Στο παρακάτω σχήμα φαίνεται η γραφική παράσταση της συνάρτησης πυκνότητας

της  $X = Z_1^2 + Z_2^2 + \dots + Z_n^2 \sim \chi_n^2$  για διάφορες τιμές του  $n$ .

Στην παρακάτω γραφική παράσταση της συνάρτησης πυκνότητας  $F_{n,m}$  παρατηρείται ότι όσο αυξάνονται οι βαθμοί ελευθερίας  $n$  και  $m$  τόσο η θετική ασυμμετρία της γραφικής παράστασης μειώνεται.

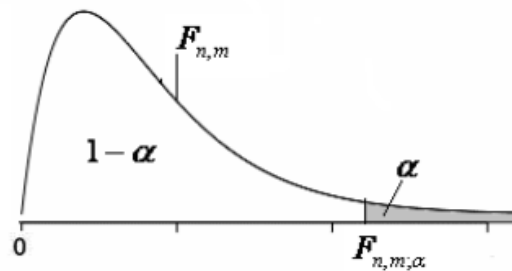
Η μέση τιμή και η διασπορά μιας τυχαίας μεταβλητής  $X$  που ακολουθεί την κατανομή

$F_{n,m} (X \sim F_{n,m})$  είναι για  $m > 2$   $E(X) = \frac{m}{m-2}$  και για  $m > 4$

$$V(X) = \frac{2 \cdot m^2 \cdot (n + m - 2)}{n \cdot (m - 2)^2 \cdot (m - 4)}$$

Η μέση τιμή της  $F_{n,m}$  εξαρτάται μόνο από τους βαθμούς ελευθερίας  $m$ , του παρονομαστή.

Όπως φαίνεται στο παρακάτω σχήμα, αν η τυχαία μεταβλητή  $X$  ακολουθεί την κατανομή  $F_{n,m}$  δηλαδή αν  $X \sim F_{n,m}$ , τότε το  $F_{n,m;a}$  είναι εκείνη η τιμή της  $X$  για την οποία ισχύει  $P(X > F_{n,m;a}) = a$  και  $P(X \leq F_{n,m;a}) = 1 - a$ .



Οι δύο παρακάτω πίνακες δίνουν τα  $\alpha$ -ποσοστιαία σημεία της κατανομής  $F$  με  $n$  και  $m$  βαθμούς ελευθερίας για  $\alpha=0,05$  και  $\alpha=0,01$  αντίστοιχα.

Αν  $X \sim F_{n,m}$ , ισχύει  $P(X > F_{n,m;a}) = a$  και  $F_{n,m;1-a} = \frac{1}{F_{n,m;a}}$

Τα  $\alpha$ -ποσοστιαία σημεία της κατανομής  $F$  με  $n$  και  $m$  βαθμούς ελευθερίας για  $\alpha=0,05$  παρουσιάζονται στο παράτημα 1- Πίνακας 3.

Τα  $\alpha$ -ποσοστιαία σημεία της κατανομής  $F$  με  $n$  και  $m$  βαθμούς ελευθερίας για  $\alpha=0,01$  παρουσιάζονται στο παράτημα 1- Πίνακας 4.

#### 4.3.1 Έλεγχος $\chi^2$ ομοιογένειας

Ο έλεγχος  $\chi^2$  ομοιογένειας χρησιμοποιείται στην περίπτωση  $r$  ανεξαρτήτων δειγμάτων, τα στοιχεία των οποίων είναι ταξινομημένα σε  $c$  κατηγορίες και ελέγχει την υπόθεση:

- $H_0$ : οι πληθυσμοί από τους οποίους προήλθαν τα ανεξάρτητα δείγματα, εκπροσωπούνται σε ίσα ποσοστά στις διάφορες κατηγορίες.
- $H_1$ : τουλάχιστον δύο από τους πληθυσμούς από τους οποίους προήλθαν τα ανεξάρτητα δείγματα, εκπροσωπούνται σε διαφορετικά ποσοστά στις διάφορες κατηγορίες.

Θεωρώντας τον rxc πίνακα :

Κατηγορία							
Δείγμα 1	O <sub>11</sub>	O <sub>12</sub>		O <sub>1j</sub>		O <sub>1c</sub>	n <sub>1</sub>
Δείγμα 2	O <sub>21</sub>	O <sub>22</sub>		O <sub>2j</sub>		O <sub>2c</sub>	n <sub>2</sub>
...							...
Δείγμα r	O <sub>r1</sub>	O <sub>r2</sub>		O <sub>rj</sub>		O <sub>rc</sub>	n <sub>r</sub>
Σύνολο	C <sub>1</sub>	C <sub>2</sub>	..	C <sub>J</sub>		C <sub>g</sub>	N

Με  $n_i = O_{i1} + O_{i2} + \dots + O_{ij} + O_{ic}$ ,  $i=1,2,\dots,r$

$$C_j = O_{1j} + O_{2j} + O_{rj}, j=1,2,\dots,c \text{ και } N = n_1 + n_2 + \dots + n_r,$$

η τιμή της στατιστικής συνάρτησης T την οποία συγκρίνουμε με την τιμή  $\chi^2_{1-\alpha, (r-1)(c-1)}$ , δίνεται από την σχέση  $T = \sum (O_{ij} - E_{ij})^2 / E_{ij}$ , όπου O<sub>ij</sub> οι παρατηρούμενες συχνότητες και E<sub>ij</sub> = n<sub>i</sub>C<sub>j</sub>/N οι αναμενόμενες .

Όταν  $T > \chi^2_{(r-1)(c-1), 1-\alpha}$  η μηδενική υπόθεση απορρίπτεται δηλαδή έχουμε διαφορετικά ποσοστά εκπροσώπησης δύο τουλάχιστον δειγμάτων στις διάφορες κατηγορίες.

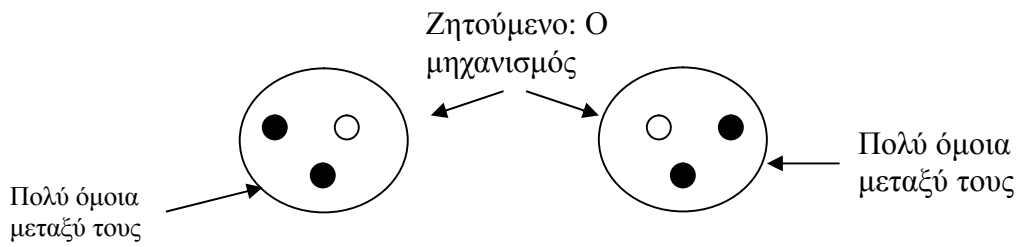
#### 4.3.2 Διαχωριστική Ανάλυση

Το αντικείμενο της Διαχωριστικής Ανάλυσης (Discriminat Analysis) είναι η κατασκευή ενός μοντέλου διαχωρισμού – ομαδοποίησης των υποκειμένων ενός συνόλου δεδομένων σε ένα γνωστό αριθμό από ομάδες.

Η διαχωριστική ανάλυση βρίσκει εφαρμογή κυρίως στην Ιατρική, στην Βιολογία και στην Επιδημιολογία.

Η ομαδοποίηση βασίζεται στις τιμές που λαμβάνουν τα υποκείμενα σε ένα μεγάλο σύνολο χαρακτηριστικών που έχουν μετρήσει για τα συγκεκριμένα υποκείμενα.

Στη διαχωριστική Ανάλυση το ζητούμενο είναι η εκτίμηση του μοντέλου του διαχωρισμού των αντικειμένων στις k γνωστές εκ των προτέρων ομάδες, δεδομένου ενός συνόλου p μεταβλητών  $X_1, X_2, \dots, X_p$  και ενός συνόλου N αντικειμένων.



Για παράδειγμα εάν θέλουμε να ξέρουμε αν κάποιος έχει καρκίνο του πνεύμονα με πρόβλεψη Ναι και Όχι ζητούμενα, οι προγνωστικές μεταβλητές μπορεί να είναι ο αριθμός των τσιγάρων που καπνίζονται ανά μέρα, η συχνότητα η ένταση του βήχα.

Η διαχωριστική ανάλυση δημιουργεί μια ή περισσότερες διαχωριστικές εξισώσεις  $D_i$  τέτοιες ώστε να διαχωρίζουν τις ομάδες όσο το δυνατόν περισσότερο.

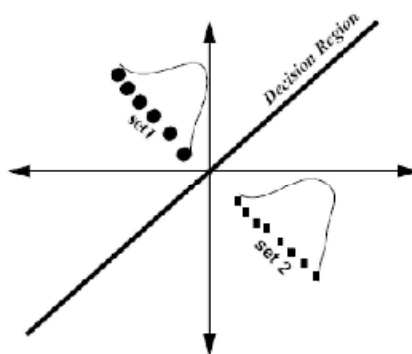
Η διαχωριστική συνάρτηση έχει την μορφή:

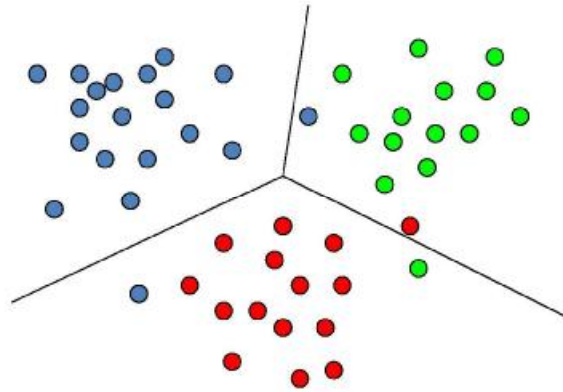
$$D_i = b_0 + \sum_{k=1}^p b_k X_k$$

Η πρώτη διαχωριστική συνάρτηση  $D_1$  διαχωρίζει την πρώτη ομάδα από τις ομάδες 2,3,...N.

Η δεύτερη διαχωριστική συνάρτηση  $D_2$  διαχωρίζει τη δεύτερη ομάδα από τις ομάδες 3 , ... N , και επαναλαμβάνεται N φορές.

Τα βέλτιστα βάρη υπολογίζονται παίρνοντας ένα αντιπροσωπευτικό σύνολο που περιέχει σωστή ταξινόμηση για ένα σύνολο υποκειμένων. Επίσης χρειαζόμαστε και δεδομένα για άτομα για τα οποία γνωρίζουμε με ακρίβεια έχουν καρκίνο του πνεύμονα το οποίο έχει καθοριστεί από εξετάσεις ή από εγχείρηση ή από ακτίνες.





Για την κατάταξη των δειγμάτων σε ομάδες μπορούμε να χρησιμοποιήσουμε τη διαχωριστική συνάρτηση. Η διαδικασία αυτή ονομάζεται ταξινόμηση.

#### 4.3.3 Ανάλυση Επιβίωσης

Η ανάλυση επιβίωσης αναφέρεται στην ανάλυση δεδομένων που αφορούν στο χρόνο που μεσολαβεί μέχρι κάποιο συγκεκριμένο συμβάν. Αρχικά η ανάλυση αναφερόταν στο χρόνο μεταξύ της θεραπείας μέχρι τον θάνατο και για αυτό το λόγο πήρε και το συγκεκριμένο όνομα. Η ανάλυση επιβίωσης όμως μπορεί να εφαρμοστεί σε αρκετές περιπτώσεις, όπως για παράδειγμα στη μηχανολογία, για την ανάλυση του χρόνου μέχρι την εμπλοκή ενός μηχανήματος ή τη υγεία για την ανάλυση του χρόνου επιβίωσης μιας αγωγής ενός ασθενούς.

Ο χρόνος επιβίωσης είναι περιορισμένος στο να είναι πάντα θετικός, και γιατί τα δεδομένα περιέχουν λογοκριμένες παρατηρήσεις. Τα λογοκριμένα δεδομένα είναι αυτά για τα οποία δεν είναι γνωστός ο χρόνος που συμβαίνει το γεγονός. Το μόνο που μπορεί να λεχθεί είναι ότι ο χρόνος επιβίωσής τους είναι μεγαλύτερος από την τιμή που έχει καταγραφεί.

Η κατανομή του χρόνου επιβίωσης στην ανάλυση επιβίωσης περιγράφεται με δυο σημαντικές συναρτήσεις, τη συνάρτηση επιβίωσης και τη συνάρτηση κινδύνου.

Συμβολίζοντας το χρόνο επιβίωσης με  $T$ , η συνάρτηση επιβίωσης  $S(t)$  ορίζεται ως η πιθανότητα επιβίωσης ενός ατόμου πέραν τη χρονική στιγμή  $t$  και δίνεται από τη σχέση :

$$S(t) = P(T > t) = 1 - F_t$$

Η συνάρτηση επιβίωσης είναι μη αρνητική και μη αύξουσα συνάρτηση του  $t$  με  $S(0) = 1$  και  $S(\infty) = 0$ .

Η γραφική παράσταση της  $S(t)$  συναρτήσεως του  $t$  είναι γνωστή ως καμπύλη επιβίωσης και είναι πολύ σημαντική στην ανάλυση δεδομένων χρόνου επιβίωσης.

Η συνάρτηση κινδύνου,  $h(t)$ , ορίζεται ως η πιθανότητα αποβίωσης (ή πραγμάτωσης του γεγονότος που εξετάζεται) τη χρονική στιγμή  $t$ , δεδομένου ότι το άτομο έχει επιβιώσει μέχρι τη χρονική στιγμή  $t$ . Δηλαδή,

$$h(t) = \lim_{s \rightarrow 0} \frac{P(t \leq T \leq t + s | T \geq t)}{s}$$

Η συνάρτηση κινδύνου δίνει ένα μέτρο του πόσο πιθανό είναι ένα άτομο να αποβιώσει ως συνάρτηση της ηλικίας του ατόμου, για παράδειγμα ο κίνδυνος θανάτου ανάμεσα σε αυτούς που είναι ζωντανοί τη συγκεκριμένη στιγμή.

Στην ανάλυση επιβίωση παίζει μεγάλο ρόλο η εξεύρεση παραγόντων οι οποίοι να σχετίζονται με το χρόνο επιβίωσης. Αυτοί οι παράγοντες θα πρέπει να συμπεριληφθούν στο μοντέλο που θα χρησιμοποιηθεί για τη σχετική ανάλυση των δεδομένων.

Αφού η συνάρτηση κινδύνου είναι μη αρνητική, ο λογάριθμός της μπορεί να εκφραστεί ως γραμμική συνάρτηση επεξηγηματικών μεταβλητών :

$$\ln h(t) = b_0 + b_1 c_1 + b_2 c_2 + \dots + b_p c_p$$

Το μοντέλο αυτό όμως είναι πολύ περιοριστικό αφού υποθέτει ότι η συνάρτηση κινδύνου δεν εξαρτάται από το χρόνο. Υπάρχουν διάφορες μέθοδοι με τις οποίες το μοντέλο θα μπορούσε να υιοθετήσει την εξάρτηση του χρόνου, με την πιο γνωστή να είναι το μοντέλο αναλόγων συναρτήσεων κινδύνου (Cox 1972). Το μοντέλο αυτό δίνεται από

$$\ln h(t) = a(t) + b_1 c_1 + b_2 c_2 + \dots + b_p c_p$$

όπου  $a(t)$  είναι οποιαδήποτε συνάρτηση του χρόνου.

Ο όρος αναλόγων συναρτήσεων κινδύνου είναι λόγω του γεγονότος ότι για οποιαδήποτε άτομα για οποιοδήποτε σημείο του χρόνου, ο λόγος των συναρτήσεων κινδύνου είναι σταθερός. Εξαιτίας του ότι η συνάρτηση κινδύνου  $a(t)$  δεν είναι ανάγκη να οριστεί εξ ολοκλήρου, το μοντέλο αναλόγων συναρτήσεων κινδύνου θεωρείται ως ημιπαραμετρικό.

Ο Cox εισηγήθηκε μια μέθοδο δεσμευμένης πιθανοφάνειας για εκτίμηση των παραμέτρων. Το σημαντικό στοιχείο αυτής της μεθόδου είναι ότι οι εκτιμήσεις εξαρτώνται από τη σειρά με την οποία συμβαίνει το γεγονός και όχι τον ακριβή χρόνο.



#### 4.3.4 Έλεγχος $\chi^2$ Ανεξαρτησίας

Ο έλεγχος  $\chi^2$  ανεξαρτησίας χρησιμοποιείται στην περίπτωση ενός μοναδικού δείγματος, τα στοιχεία του οποίου είναι ταξινομημένα σε  $r$  κατηγορίες ως προς ένα χαρακτηριστικό και σε  $c$  ως προς κάποιο άλλο και ελέγχει την υπόθεση:

- $H_0$ : τα δύο χαρακτηριστικά είναι ανεξάρτητα
- $H_1$ : τα δύο χαρακτηριστικά είναι εξαρτημένα.

Θεωρώντας τον  $r \times c$  πίνακα:

Χαρακτηριστικό A	Σύνολο						
Χαρακτηριστικό B	$O_{11}$	$O_{12}$		$O_{1j}$		$O_{1c}$	$R_1$
	$O_{21}$	$O_{22}$		$O_{2j}$		$O_{2c}$	$R_2$
	$O_{r1}$	$O_{r2}$		$O_{rj}$		$O_{rc}$	$R_r$
Σύνολο	$C_1$	$C_2$	...	$C_j$		$C_c$	$N$

Με  $R_i = O_{i1} + O_{i2} + \dots + O_{ij} + \dots + O_{ic}$ ,  $i=1,2,\dots,r$

$$C_j = O_{1j} + O_{2j} + \dots + O_{rj}, j=1,2,\dots,c \text{ και } N = n_1 + n_2 + \dots + n_r,$$

η τιμή της στατιστικής συνάρτησης  $T$  την οποία συγκρίνουμε με την τιμή  $\chi^2_{1-\alpha, (r-1)(c-1)}$ , δίνεται από τη σχέση  $T = \sum (O_{ij} - E_{ij})^2 / E_{ij}$  όπου  $O_{ij}$  οι παρατηρούμενες συχνότητες και  $E_{ij} = N(R_i/N)(C_j/N)$  οι αναμενόμενες.

Όταν  $T > \chi^2_{(r-1)(c-1), 1-\alpha}$  η μηδενική υπόθεση απορρίπτεται, δηλαδή τα δύο χαρακτηριστικά είναι εξαρτημένα (Rosner, 2000, Pagano-Gauvreau, 2000).

#### 4.3.5 Έλεγχος τάσης

Οι έλεγχοι  $\chi^2$  είναι προτιμότερο να χρησιμοποιούνται για δεδομένα σε ονομαστική κλίμακα. Στην περίπτωση που η μια μεταβλητή είναι σε κλίμακα διάταξης και η άλλη δίτιμη ενδείκνυται έλεγχος τάσης.

	X=1	X=2		X= r	Σύνολο
Y=0	$n_{01}$	$n_{02}$		$n_{0r}$	$n_{0\cdot}$

<b>Y=1</b>	$n_{11}$	$n_{12}$		$n_{1r}$	$n_{1.}$
<b>Σύνολο</b>	$n_{.1}$	$n_{.2}$		$n_{.r}$	$n_{.}$

Θεωρώντας τον πίνακα:

με  $r > \dots > 2 > 1$  για την μεταβλητή  $X$  οι υποθέσεις:

- $H_0$ : δεν υπάρχει τάση
- $H_1$ : υπάρχει τάση, ελέγχονται μέσω της σύγκρισης της τιμής  $b^2/V(b_{xy})$  με την τιμή  $\chi^2_{1(1-\alpha)}$  όπου  $b_{xy} = C_{yx}/C_{xx}$  και  $V(b_{xy}) = (1/n)C_{yy}/C_{xx}$ ,

με  $C_{yx} = \sum y_i x_i - (\sum n_i \cdot y_i)(\sum n_j \cdot X_j)/n$ .

$C_{xx} = \sum n_{.x} \cdot x^2 - (\sum n_i \cdot x_i)^2/n$  και  $C_{yy} = \sum n_i \cdot y_i^2 - (\sum n_i \cdot y_i)^2/n$ .

Όταν  $b^2/V(b_{xy}) > \chi^2_{1(1-\alpha)}$  απορρίπτεται η μηδενική υπόθεση, δηλαδή υπάρχει τάση (Everit, 1980).

#### 4.3.6 Συντελεστής Γραμμικής Συσχέτισης του Pearson

Ο δειγματικός συντελεστής γραμμικής συσχέτισης του Pearson συμβολίζεται με

$r$  και ορίζεται από τον τύπο:  $r = \frac{s_{xy}}{s_x \times s_y}$

Όπου ,

$$s_{xy} = Cov(X, Y) = \frac{\sum_{i=1}^v (x_i - \bar{x}) \cdot (y_i - \bar{y})}{v-1} = \frac{\sum_{i=1}^v x_i y_i - v \cdot \bar{x} \cdot \bar{y}}{v-1}$$

$$s_x = \sqrt{\frac{1}{v-1} \sum_{i=1}^v (x_i - \bar{x})^2} \quad \text{και} \quad s_y = \sqrt{\frac{1}{v-1} \sum_{i=1}^v (y_i - \bar{y})^2}$$

Επομένως

$$r = \frac{s_{xy}}{s_x \cdot s_y} = \frac{\sum_{i=1}^v (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\frac{1}{v-1} \sum_{i=1}^v (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{v-1} \sum_{i=1}^v (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^v x_i y_i - v \cdot \bar{x} \cdot \bar{y}}{\sqrt{\frac{1}{v-1} \sum_{i=1}^v x_i^2 - v \cdot \bar{x}^2} \cdot \sqrt{\frac{1}{v-1} \sum_{i=1}^v y_i^2 - v \cdot \bar{y}^2}}$$

Ο πληθυσμιακός συντελεστής γραμμικής συσχέτισης του Pearson ορίζεται ανάλογα και συμβολίζεται με  $\rho$ .

Ο συντελεστής γραμμικής συσχέτισης  $r$  δίνει ένα μέτρο του μεγέθους της γραμμικής συσχέτισης μεταξύ δύο μεταβλητών. Το  $r$  παίρνει τιμές στο κλειστό διάστημα  $[-1, 1]$ .

Για παράδειγμα για κάθε  $\lambda \in \mathbb{R}$  ισχύει  $\sum ((x_i - \bar{x}) + I(y_i - \bar{y}))^2 \geq 0$

Άρα

$$\begin{aligned} \sum ((x_i - \bar{x})^2 + I^2(y_i - \bar{y})^2 + 2I(x_i - \bar{x})(y_i - \bar{y})) &\geq 0 \text{ ή} \\ \sum ((x_i - \bar{x})^2 + I^2 \sum (y_i - \bar{y})^2 + 2I \sum (x_i - \bar{x})(y_i - \bar{y})) &\geq 0 \text{ ή} \\ I^2 \sum (y_i - \bar{y})^2 + I \cdot 2 \sum (x_i - \bar{x})(y_i - \bar{y}) + \sum (x_i - \bar{x})^2 &\geq 0 \end{aligned}$$

Επειδή η τελευταία ανισότητα ισχύει για κάθε  $\lambda \in \mathbb{R}$ , θα είναι  $\beta^2 - 4\alpha\gamma \leq 0$ . άρα

$$\begin{aligned} 4 \cdot (\sum (x_i - \bar{x})(y_i - \bar{y}))^2 &\leq 4 \cdot \sum (y_i - \bar{y})^2 \cdot \sum (x_i - \bar{x})^2 \Leftrightarrow \\ \left( \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (y_i - \bar{y})^2} \sqrt{\sum (x_i - \bar{x})^2}} \right) &\leq 1 \Leftrightarrow r^2 \Leftrightarrow |r| \leq 1 \Leftrightarrow -1 \leq r \leq 1 \end{aligned}$$

- Αν  $r = \pm 1$  υπάρχει τέλεια γραμμική συσχέτιση.
- Αν  $-0,3 \leq r < 0,3$  δεν υπάρχει γραμμική συσχέτιση. Αυτό, όμως, δεν σημαίνει ότι δεν υπάρχει άλλου είδους συσχέτιση μεταξύ των δύο μεταβλητών.
- Αν  $-0,5 < r \leq -0,3$  ή  $0,3 \leq r < 0,5$  υπάρχει ασθενής γραμμική συσχέτιση.
- Αν  $-0,7 < r \leq -0,5$  ή  $0,5 \leq r < 0,7$  υπάρχει μέση γραμμική συσχέτιση.
- Αν  $-0,8 < r \leq -0,7$  ή  $0,7 \leq r < 0,8$  υπάρχει ισχυρή γραμμική συσχέτιση.
- Αν  $-1 < r \leq -0,8$  ή  $0,8 \leq r < 1$  υπάρχει πολύ ισχυρή γραμμική συσχέτιση.

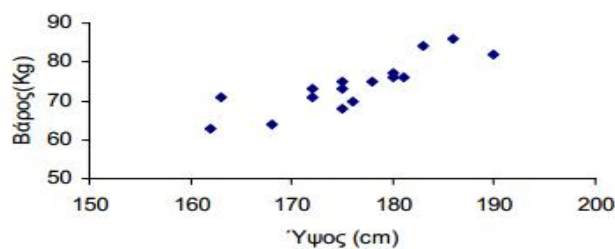
Σημειώνετε ότι οι θετικές τιμές του  $r$  δεν υποδηλώνουν, κατ' ανάγκην μεγαλύτερο βαθμό γραμμικής συσχέτισης από το βαθμό γραμμικής συσχέτισης που υποδηλώνουν αρνητικές τιμές του  $r$ . Ο βαθμός γραμμικής συσχέτισης καθορίζεται από την απόλυτη τιμή του  $r$  και όχι από το πρόσημο του  $r$ . Το πρόσημο του  $r$  καθορίζει το είδος, μόνο, της συσχέτισης (θετική ή αρνητική). Μας πληροφορεί δηλαδή για το αν αύξηση της μιας μεταβλητής αντιστοιχεί σε αύξηση ή σε μείωση της άλλης μεταβλητής.

Για παράδειγμα η τιμή  $r = -0,9$  δείχνει ισχυρότερη γραμμική συσχέτιση από την τιμή  $r = 0,8$  ενώ οι τιμές  $r = -0,6$  και  $r = 0,6$  δείχνουν ίδιο βαθμό γραμμικής συσχέτισης αλλά αντίθετο είδος.

Στον Πίνακα που ακολουθεί είναι καταγεγραμμένες οι παρατηρήσεις του ύψους και του βάρους από 16 εργάτες σε ένα εργοστάσιο.

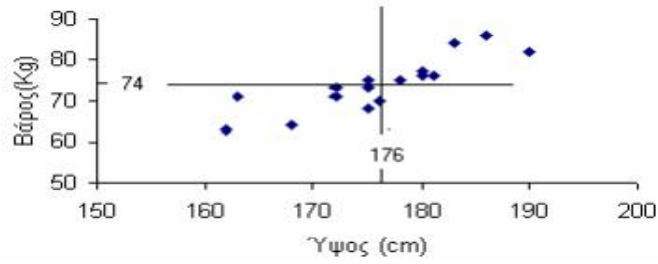
Εργάτες	Ύψος(cm)	Βάρος(Kg)
1	183	84
2	162	63
3	172	71
4	181	76
5	180	77
6	168	64
7	176	70
8	180	76
9	190	82
10	175	68
11	178	75
12	175	73
13	186	86
14	172	73
15	175	75
16	163	71

Από το παρακάτω διάγραμμα διασποράς φαίνεται ότι υπάρχει μια ανάλογη σχέση μεταξύ Ύψους και Βάρους των εργατών.

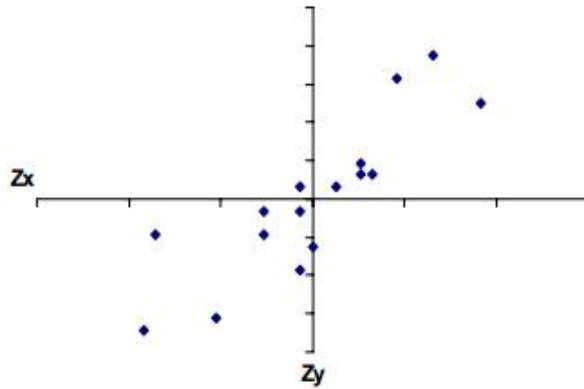


Εξετάζοντας την βασική ιδέα στον ορισμό του συντελεστή Pearson σε δείγμα εργατών όπου το μέσο ύψος τους είναι 176 cm και το μέσο βάρος είναι 74 Kg παρατηρούμε ότι οι εργάτες που έχουν ύψος πάνω από το μέσο ύψος έχουν και βάρος πάνω από το μέσο βάρος και αντίστοιχα οι εργάτες που έχουν μέσο ύψος κάτω από το μέσο ύψος έχουν και βάρος κάτω από το μέσο βάρος.

Παρακάτω στο διάγραμμα έχουμε την συσχέτιση των δύο μεταβλητών ύψους και βάρους.



Επίσης και στο ακόλουθο διάγραμμα:



Παρατηρείται επίσης ότι  $r = \frac{\sum_{i=1}^n z_{x_i} \cdot z_{y_i}}{n-1}$  (όπου  $z_{x_i} = \frac{x_i - \bar{x}}{s_x}$  και  $z_{y_i} = \frac{y_i - \bar{y}}{s_y}$ )

Εργά- τες	$x_i$	$y_i$	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	183	84	7	10	49	100	70
2	162	63	-14	-11	196	121	154
3	172	71	-4	-3	16	9	12
4	181	76	5	2	25	4	10
5	180	77	4	3	16	9	12
6	168	64	-8	-10	64	100	80
7	176	70	0	-4	0	16	0
8	180	76	4	2	16	4	8
9	190	82	14	8	196	64	112
10	175	68	-1	-6	1	36	6

11	178	75	2	1	4	1	2
12	175	73	-1	-1	1	1	1
13	186	86	10	12	100	144	120
14	172	73	-4	-1	16	1	4
15	175	75	-1	1	1	1	-1
16	163	71	-13	-3	169	9	39
	$\sum x_i$ =176	$\sum y_i$ =74			$\sum (x_i - \bar{x})^2$ =54,375	$\sum (y_i - \bar{y})^2$ =38,75	$\sum (x_i - \bar{x})(y_i - \bar{y})$ =39,3125

Ο συντελεστής συσχέτισης Pearson ισούται με

$$r = \frac{\sum_{i=1}^{16} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[ \sum_{i=1}^{16} (x_i - \bar{x})^2 \right] \left[ \sum_{i=1}^{16} (y_i - \bar{y})^2 \right]}} = \frac{39,3125}{\sqrt{54,375 \cdot 38,75}} = 0,856437$$

Από την τιμή του συντελεστή συσχέτισης Pearson προκύπτει ότι υπάρχει πολύ ισχυρή σχέση μεταξύ του Ύψους και του Βάρους σε ένα εργοστάσιο. Δηλαδή όσο εργάτες είναι ψηλότεροι είναι και βαρύτεροι.

#### 4.3.7 Απλή Γραμμική Παλινδρόμηση (Simple Linear Regression)

Αντικείμενο της ανάλυσης με τη χρήση της απλής γραμμικής παλινδρόμησης αποτελεί η σχέση μεταξύ δυο (scaled) μεταβλητών X, Y όπως για παράδειγμα η ηλικία και η πίεση του αίματος ενός ατόμου. (Boutsikas M.V. (2004)

Από έναν, θεωρητικά άπειρο, πληθυσμό λαμβάνεται ένα δείγμα μεγέθους n και για κάθε άτομο του δείγματος καταγράφουμε τις τιμές δύο μεταβλητών X, Y, όπου ως X ορίζεται η ηλικία του κάθε ατόμου και Y η πίεση του αίματος του. Με βάση λοιπόν τα ζεύγη τιμών (X<sub>1</sub>, Y<sub>1</sub>), (X<sub>2</sub>, Y<sub>2</sub>), (X<sub>n</sub>, Y<sub>n</sub>) του δείγματος διερευνείται η σχέση μεταξύ των μεταβλητών X, Y. Στην περίπτωση αυτή θεωρείται ότι η μεταβλητή X καλείται ανεξάρτητη (independent) ή ερμηνευτική μεταβλητή (explanatory variable) και δεν θεωρείται τυχαία, ενώ η μεταβλητή Y, η οποία καλείται εξαρτημένη (dependent) ή μεταβλητή απόκρισης (response variable), θεωρείται τυχαία μεταβλητή.

Ο απλούστερος τύπος μοντέλου που συνδέει τα  $x$  και τα  $y$  λέγεται ευθεία γραμμικής παλινδρόμησης και έχει την μορφή  $y = \alpha + \beta x$ .

Ο συντελεστής  $\alpha$  είναι η τιμή του  $y$  για  $x=0$  και το  $\beta$  είναι η κλίση της ευθείας. Για μια δοσμένη ευθεία τα  $\alpha$  και  $\beta$  είναι σταθερές. Εάν η παρατήρηση  $y$  δεν υπόκειται σε σφάλματα, δηλαδή αν για κάποια τιμή της ανεξάρτητης μεταβλητής  $x$ , μπορούμε να προβλέψουμε ακριβώς την τιμή του  $y$ .

Ένα μοντέλο που δίνει την δυνατότητα στο  $y$ , να μη βρίσκεται ακριβώς στην ευθεία  $y = \alpha + \beta x$  είναι το:  $y = \alpha + \beta x + e$  όπου  $e$  είναι ένα τυχαίο σφάλμα και παριστάνει τη διαφορά της παρατηρούμενης τιμής  $y$ , για δοσμένο  $x$ , από τη θεωρητική τιμή  $\alpha + \beta x$ . Αυτού του είδους τα μοντέλα καλούνται στοχαστικά.

Όσον αφορά τα σφάλματα  $e$  υποθέτουμε ότι είναι τυχαία με μέση τιμή  $E(e) = 0$ . Έτσι, αφού τα  $\alpha$  και τα  $\beta$  είναι άγνωστες σταθερές ισχύει:  $E(y) = \alpha + \beta x$ .

Η ανάλυση απλής παλινδρόμησης είναι μια στατιστική μεθοδολογία η οποία απεικονίζει ουσιαστικά την σχέση μεταξύ 2 μεταβλητών έτσι ώστε η μια μεταβλητή να μπορεί να προβλεφθεί από την άλλη.

Χρησιμοποιείται πλέον ευρέως σαν μεθοδολογία σε τομείς όπως κοινωνικές επιστήμες, βιολογία, οικονομικά αλλά και στην υγεία.

### **Μεθοδολογία**

Στη περίπτωση της απλής γραμμικής παλινδρόμησης έχουμε μια εξαρτημένη μεταβλητή ( $Y$ ) και μια ανεξάρτητη μεταβλητή  $X$  (predictor variable) μέσω της οποίας υπολογίζονται οι τιμές της  $Y$  και η εξίσωση (παλινδρόμησης) η οποία τις συνδέει είναι γραμμική.

$$Y_i = b_0 + b_1 X_i + e_i \text{ όπου}$$

- $Y_i$  είναι η τιμή της ανεξάρτητης μεταβλητής στην  $i$  παρατήρηση της  $X$  μεταβλητής.
- $\beta_0, \beta_1$  παράμετροι, συντελεστές παλινδρόμησης. (Η  $\beta_0$  ονομάζεται και τεταγμένη στην αρχή μηδέν).
- $X_i$  είναι η τιμή της ανεξάρτητης μεταβλητής στην  $i$  παρατήρηση
- $e_i$  είναι το τυχαίο σφάλμα ή κατάλοιπο, δηλαδή η διαφορά μεταξύ της πραγματικής τιμής της  $Y$  και της τιμής πρόβλεψης που προκύπτει από το μοντέλο. Ισχύει  $E(e_i) = 0$  και  $\sigma^2(e_i) = \sigma^2$ . Επίσης τα σφάλματα θεωρούμε ότι είναι ασυσχέτιστα, δηλαδή ισχύει  $\sigma(e_i, e_j) = 0$ , για κάθε  $i \neq j$
- Οι ακριβείς τιμές της  $Y$  διαφέρουν από τις αναμενόμενες κατά ένα

«ανεξήγητο» τυχαίο σφάλμα:  $Y_i = EY_i + e_i = b_0 + b_1 X_i + e_i$ .

- Όλοι οι μέσοι βρίσκονται σε μια ευθεία γραμμή που αποτελεί τη γραμμή παλινδρόμησης του πληθυσμού.
- Για κάθε συγκεκριμένη τιμή της  $X$  αντιστοιχούν πολλές τιμές της  $Y$  (στο δείγμα φυσικά περιλαμβάνεται μία) που κατανέμονται κανονικά.
- Οι τιμές της ανεξάρτητης μεταβλητής  $X$  θεωρούνται γνωστές (όχι τυχαίες), η μοναδική τυχαιότητα στις τιμές της  $Y$  εμφανίζεται από την ύπαρξη του σφάλματος  $e_i$ .
- Τα σφάλματα  $e_i$  κατανέμονται κανονικά με μέσο 0 και διακύμανση  $\sigma^2$ . Τα σφάλματα διαδοχικών παρατηρήσεων είναι ασυσχέτιστα (δεν αλληλοεξαρτώνται).  
Άρα:  $e \sim N(0, \sigma^2)$
- Οι ατομικές παρατηρήσεις της εξαρτημένης μεταβλητής  $Y$  είναι μεταξύ τους ανεξάρτητες.

Το μοντέλο της παλινδρόμησης πρέπει να έχει τα εξής χαρακτηριστικά:

- Η εξαρτημένη μεταβλητή  $Y_i$  αποτελείται από 2 μέρη. Από το σταθερό όρο (ντετερμινιστικό)  $\beta_0 + \beta_1 X_i$  και από τον τυχαίο όρο  $e_i$ . Επομένως η εξαρτημένη μεταβλητή  $Y$  είναι μια τυχαία μεταβλητή.
- Από την στιγμή που ισχύει ότι  $E(e_i) = 0$ , ισχύει ότι  $E(Y_i) = \beta_0 + \beta_1 X_i$ . Η μεταβλητή  $Y_i$  απέχει από την προβλεπόμενη τιμή μέσω της εξίσωσης παλινδρόμησης εξαιτίας του τυχαίου όρου  $e_i$ .
- Ισχύει ότι  $\sigma^2(Y_i) = \sigma^2(\beta_0 + \beta_1 X_i + e_i) = \sigma^2(e_i) = \sigma^2$
- Τα τυχαία σφάλματα  $e_i$  θεωρούμε ότι είναι ασυσχέτιστα.
- Έλεγχος στατιστικής σημαντικότητας ενός υποδείγματος απλής γραμμικής παλινδρόμησης.

Ας υποθέσουμε την ακόλουθη απλή γραμμική παλινδρόμηση.

Η υπόθεση που θέλουμε να ελέγξουμε είναι ότι οι παράμετροι του μοντέλου είναι μηδέν, δηλαδή μηδενικής στατιστικής σημασίας  $H_0 : b_0 = 0$ , η πληθυσμιακή γραμμή διέρχεται από την αρχή των αξόνων.

Έναντι της εναλλακτικής υπόθεσης

$H_1 : b_0 \neq 0$  και  $H_0 : b_1 = 0$ , η κλίση της καλύτερα προσδιοριζόμενης γραμμής είναι



μηδέν, άρα δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών X και Y.

Έναντι της εναλλακτικής υπόθεσης

$$H_0 : b_1 \neq 0$$

Ο έλεγχος γίνεται με το t τεστ. Το τεστ αυτό κατανέμεται σαν μια μεταβλητή t με n-2 βαθμούς ελευθερίας. Σε κάθε περίπτωση η  $H_0$  απορρίπτεται αν ισχύει  $t \leq -t_{n-2, \alpha/2}$  ή  $t \geq t_{n-2, \alpha/2}$ .

n-2,  $\alpha/2$ .

### Πίνακας Ανάλυσης διασποράς (ANOVA)

Model	Sum of Squares	df	Mean Square	F	Sig.(p-value)
<b>Regression</b>	$SSR = \sum (\hat{Y}_i - \bar{Y})^2$	1	$MSR = SSR$	$\frac{MSR}{MSE}$	$1 - F_{F_{1,n-2}} \left( \frac{MSR}{MSE} \right)$
<b>Residuals</b>	$SSE = \sum (Y_i - \hat{Y}_i)^2$	n-2	$MSE = \frac{SSE}{n-2} = S^2$		
<b>Total</b>	$SST = \sum (Y_i - \bar{Y})^2$ $= Y^T \left( I_n - \frac{1}{n} J_n \right) Y$	n-1			

### Πράδειγμα:

Για να γίνει κατανοητή η εφαρμογή της απλής γραμμικής παλινδρόμησης παρουσιάζεται το παρακάτω παράδειγμα:

Στην αρχή θα διερευνήσουμε την σχέση μεταξύ των δύο μεταβλητών X,Y που είναι η Ηλικία και οι πίεση του αίματος.

Από έναν πληθυσμό λαμβάνουμε ένα δείγμα μεγέθους n και για κάθε άτομο του δείγματος καταγράφουμε τις τιμές δύο μεταβλητών X, Y. Με βάση λοιπόν τα ζεύγη τιμών  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  του δείγματος (π.χ.  $X_i$ : Ηλικία σε έτη i-ατόμου,  $Y_i$ : Πίεση αίματος i-ατόμου) επιθυμούμε να διερευνήσουμε τη σχέση μεταξύ των μεταβλητών X, Y.

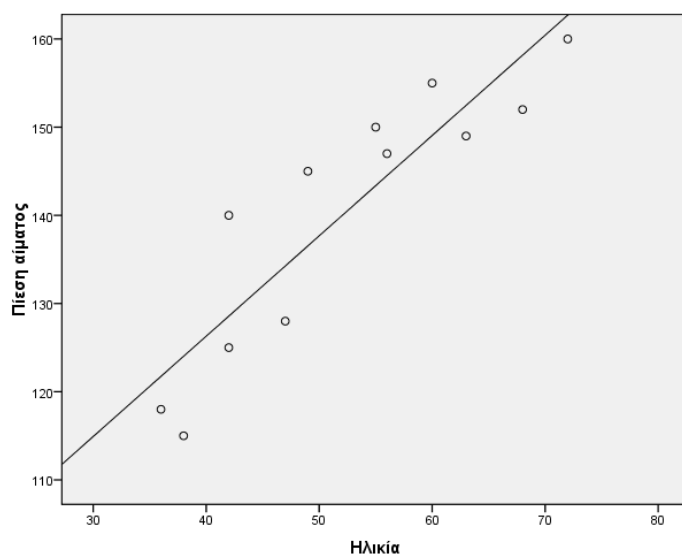
Επιπλέον θεωρούμε ότι η μεταβλητή X η οποία καλείται ανεξάρτητη (independent) ή ερμηνευτική μεταβλητή δεν θεωρείται τυχαία, ενώ η μεταβλητή Y η οποία καλείται εξαρτημένη (dependent) ή μεταβλητή απόκρισης θεωρείται τυχαία μεταβλητή.

Από n = 12 γυναίκες λαμβάνονται οι ακόλουθες τιμές της πίεσης του αίματος και της αντίστοιχης ηλικίας σε έτη:

Ηλικία (X)	36	38	42	42	47	49	55	56	60	63	68	72
Πίεση αίματος (Y)	118	115	125	140	128	145	150	147	155	149	152	160

### Τεχνική:

Εισάγοντας τα δεδομένα στο SPSS δημιουργούμε την γραφική απεικόνιση απλής παλινδρόμησης με την εισαγωγή της γραμμής τάσης εξάγουμε το ακόλουθο διάγραμμα όπου στον άξονα των X τοποθετείται η Ηλικία και στον άξονα Y η πίεση.



### Variables Entered/Removed<sup>b</sup>

Model	Variables Entered	Variables Removed	Method
1	Ηλικία (X) <sup>a</sup>	.	Enter

a. All requested variables entered.

b. Dependent Variable: Πίεση αίματος (Y)

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,896 <sup>a</sup>	,803	,783	7,018

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,896 <sup>a</sup>	,803	,783	7,018

a. Predictors: (Constant), Ηλικία (X)

### ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2008,200	1	2008,200	40,778	,000 <sup>a</sup>
	Residual	492,467	10	49,247		
	Total	2500,667	11			

a. Predictors: (Constant), Ηλικία (X)

b. Dependent Variable: Πίεση αίματος (Y)

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	80,778	9,544		8,464	,000
	Ηλικία (X)	1,138	,178	,896	6,386	,000

a. Dependent Variable: Πίεση αίματος (Y)

### Αποτελέσματα:

Παρατηρείται ότι όσο αυξάνεται η X (Age) τόσο αυξάνεται και η Y (Pressure). Μάλιστα φαίνεται ότι τα σημεία  $(X_i, Y_i)$  βρίσκονται «κοντά» σε μία ευθεία. Οι αποκλίσεις των σημείων από την ευθεία αυτή φαίνονται τυχαίες. Στην γενικότερη περίπτωση, διερευνάται η σχέση μεταξύ των μεταβλητών X, Y. Το απλούστερο μοντέλο που θα μπορούσε να ερμηνεύσει μια τέτοια σχέση θεωρείται, το απλό γραμμικό μοντέλο. Σύμφωνα με το μοντέλο αυτό θεωρούμε ότι τα  $X_i, Y_i$  συνδέονται με τη σχέση:

$$Y_i = b_0 + b_1 X_i + e, i = 1, 2, \dots, n$$

όπου  $b_0$ ,  $b_1$  είναι δυο άγνωστες σταθερές (καλούνται και τεταγμένη ή intercept και κλίση ή slope αντίστοιχα), ενώ οι  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κανονική κατανομή και συνήθως καλούνται «σφάλματα» των μετρήσεων.

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 x_i \Rightarrow \hat{Y} = 80,778 + 1,138 x$$

$$\text{Πίεση} = 80,778 + 1,38 \cdot \text{Ηλικία}$$

Για κάθε αύξηση της ηλικίας κατά 1 χρόνο(μια μονάδα) η αναμενόμενη τιμή της Πίεσης αυξάνεται κατά 1,3664.

$$R^2 = 0,783$$

Το 80% περίπου της μεταβλητότητας της πίεσης εξηγείται από την Ηλικία ή η Ηλικία επηρεάζει κατά 80% την Πίεση.

Αρκετές φορές συμβαίνει οι μεταβλητές  $X$  και  $Y$  να μην έχουν γραμμική σχέση. Σε αυτές τις περιπτώσεις δεν εφαρμόζεται απευθείας το γραμμικό μοντέλο αλλά θα πρέπει τα δεδομένα να μετασχηματιστούν  $X' = f(X)$  και  $Y' = g(Y)$  έτσι ώστε οι  $X'$ ,  $Y'$  να έχουν γραμμική σχέση.

Η στατιστική σημαντικότητα του παράγοντα  $X$  ελέγχεται βάση των κριτηρίων  $F$  και  $t$  τα οποία είναι στατιστικά για οποιοδήποτε επίπεδα ( $\alpha = 5\%$ ,  $\alpha = 1\%$ ) επομένως η μηδενική υπόθεση  $H_0 = b_1 = 0$  απορρίπτεται άρα υπάρχει γραμμική σχέση μεταξύ Ηλικίας και Πίεσης.

#### 4.3.8 Πολλαπλή Γραμμική Παλινδρόμηση (Multiple Linear Regression)

Όπως αναφέρθηκε προηγουμένως στο απλό γραμμικό μοντέλο διερευνάται η εξάρτηση μιας μεταβλητή  $Y$  (απόκρισης ή εξαρτημένη) από μια άλλη μεταβλητή  $X$  (ερμηνευτική ή ανεξάρτητη). Σε αρκετά όμως προβλήματα η μεταβλητή απόκρισης  $Y$  μπορεί να θεωρηθεί ότι επηρεάζεται από περισσότερες από μια ερμηνευτικές μεταβλητές, έστω  $X_1, X_2, \dots, X_{p-1}$ .

Μπορεί λοιπόν και πάλι να χρησιμοποιηθεί ένα γραμμικό μοντέλο για να διερευνηθεί η εξάρτηση της  $Y$  από τις  $X_1, X_2, \dots, X_{p-1}$ . Το μοντέλο αυτό, το οποίο αποτελεί γενίκευση του μοντέλου  $Y = b_0 + b_1 X + \varepsilon$ , θα έχει τη μορφή:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_{p-1} X_{p-1} + \varepsilon$$

για κάποιες παραμέτρους  $b_0, b_1, \dots, b_{p-1}$ . Για να διερευνηθεί η σχέση μεταξύ της  $Y$  και των  $X_1, X_2, \dots, X_{p-1}$  λαμβάνεται δείγμα μεγέθους  $n$  και για κάθε άτομο του δείγματος καταγράφονται οι τιμές των συγκεκριμένων μεταβλητών. Για παράδειγμα για το  $i$ -

άτομο του δείγματος καταγράφονται οι τιμές  $(Y_i, X_{i1}, X_{i2}, \dots, X_{i,p-1})$ ,  $i = 1, 2, \dots, n$ .

Εφαρμόζουμε λοιπόν το μοντέλο:

$$Y_i = b_0 + b_1 X_{i1} + b_2 X_{i2} + \dots + b_{p-1} X_{i,p-1} + e_i, i = 1, 2, \dots, n$$

όπου τα «σφάλματα»  $e_1, e_2, \dots, e_n$  θεωρούνται ανεξάρτητες τ.μ. από την  $N(0,1)$  ενώ οι ερμηνευτικές μεταβλητές  $X_1, X_2, \dots, X_{p-1}$ , όπως και στο απλό γραμμικό μοντέλο, δεν θεωρούνται τυχαίες. Το παραπάνω μοντέλο γράφεται με τη βοήθεια πινάκων στην απλούστερη μορφή:

$$Y = Xb + e$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \mathbf{M} \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_{11} & \mathbf{L} & X_{1,p-1} \\ 1 & X_{21} & \mathbf{L} & X_{2,p-1} \\ \mathbf{M} & & & \\ 1 & X_{n1} & \mathbf{L} & X_{n,p-1} \end{bmatrix}, \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \mathbf{M} \\ b_{p-1} \end{bmatrix}, \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \mathbf{M} \\ e_n \end{bmatrix}$$

Ο πίνακας  $X$  καλείται και πίνακας σχεδιασμού. Στο πολλαπλό μοντέλο θεωρούμε ότι τα σημεία δεν βρίσκονται «κοντά» σε μια ευθεία αλλά «κοντά» σε ένα (υπέρ)επίπεδο στις  $p$  διαστάσεις.

### Μεθοδολογία:

Η εκτίμηση του μοντέλου πολλαπλής γραμμικής παλινδρόμησης γίνεται με τη μέθοδο ελαχίστων τετραγώνων και σύμφωνα με τις ακόλουθες υποθέσεις:

1. Για κάθε παρατήρηση το τυχαίο σφάλμα κατανέμεται κανονικά με μέση τιμή μηδέν και κοινή διακύμανση. Παραβίαση της υπόθεσης σταθερής διακύμανσης δημιουργεί το πρόβλημα της ετεροσκεδαστικότητας.
2. Το σφάλμα αυτό δεν συσχετίζεται με τα σφάλματα των άλλων παρατηρήσεων. Παραβίαση της υπόθεσης για ασυσχέτιστα μεταξύ τους σφάλματα δημιουργεί το πρόβλημα της αυτοσυσχέτισης
3. Κάθε μια από τις ερμηνευτικές μεταβλητές είναι ανεξάρτητη από το τυχαίο σφάλμα.
4. Οι ερμηνευτικές μεταβλητές θεωρούνται σταθερές ποσότητες. Αυτό συνεπάγεται ότι η τυχαιότητα της εξαρτημένης μεταβλητής  $Y$  οφείλεται αποκλειστικά στο τυχαίο σφάλμα.
5. Υπάρχει σωστός αλγεβρικός προσδιορισμός του υποδείγματος σχετικά με τη συναρτησιακή μορφή και τις συμπεριλαμβανόμενες τιμές.
6. Οι ανεξάρτητες μεταβλητές δε συσχετίζονται γραμμικά μεταξύ τους. Παραβίαση αυτής της υπόθεσης δημιουργεί το πρόβλημα της πολυσυγγραμμικότητας.

Έλεγχος στατιστικής σημαντικότητας ενός υποδείγματος πολλαπλής παλινδρόμησης

Ας υποθέσουμε την ακόλουθη πολλαπλή γραμμική παλινδρόμηση.

Η υπόθεση που θέλουμε να ελέγξουμε είναι ότι οι παράμετροι του μοντέλου είναι μηδέν, δηλαδή μηδενικής στατιστικής σημασίας.

$$H_0 : b_0 = b_1 = b_2 = \dots = b_p = 0$$

Έναντι της εναλλακτικής υπόθεσης

$H_1$  : τουλάχιστον μια από τις παραμέτρους είναι στατιστικώς διάφορη από το μηδέν.

Αν θέλουμε να εξετάσουμε την ατομική στατιστική σημαντικότητα μιας μεταβλητής, ο έλεγχος γίνεται όπως στην απλή γραμμική παλινδρόμηση με τη στατιστική t με τη μόνη διαφορά ότι οι βαθμοί ελευθερίας είναι n-p. (Γεώργιος Εμμ. Χάλκος, Οικονομετρία, Θεωρία και Πράξη, Οδηγίες χρήσης σε EVIEWS, MINITAB, SPSS & EXCEL, Β. Γκιούρδας Εκδοτική, Αθήνα, 2006).

### Πίνακας Ανάλυσης διασποράς (ANOVA)

Model	Sum of Squares	df	Mean Square	F-Ratio	Sig.(p-value)
<b>Regression</b>	$SS_{Re g} = \sum (Y_i - \bar{Y})^2$ $= Y^T (P - \frac{1}{n} J_n) Y$	p-1	$MSR = \frac{SSR}{p-1}$	$\frac{MSR}{MSE}$	$1 - F_{F_{p-1, n-p}} \left( \frac{MSR}{MSE} \right)$
<b>Residuals</b>	$SSE = \sum (Y_i - \hat{Y}_i)^2$ $= Y^T (I_n - P) Y$	n-p	$MSE = \frac{SSE}{n-p} = S^2$		
<b>Total</b>	$SST = \sum (Y_i - \bar{Y})^2$ $= Y^T (I_n - \frac{1}{n} J_n) Y$	n-1			

$$(P = X(X^T X)^{-1} X^T, J_n = 1 \cdot 1^T)$$

### Παράδειγμα:

Για να γίνει κατανοητή η εφαρμογή της Πολλαπλής Γραμμικής Παλινδρόμησης παρουσιάζεται το παρακάτω παράδειγμα:

Ενδιαφερόμαστε να προβλέψουμε τον Δείκτη Σωματικού Λίπους από τις παρακάτω σωματομετρικές μετρήσεις.

Πάχος Δέρματος	Περιφέρεια Μηρού	Περιφέρεια Μπράτσου	Δείκτης Σωματικού Λίπους
19,50	43,10	29,10	11,90
24,70	49,80	28,20	22,80
30,70	51,90	37,00	18,70
29,80	54,30	31,10	20,10
19,10	42,20	30,90	12,90
25,60	53,90	23,70	21,70
31,40	58,50	27,60	27,10
27,90	52,10	30,60	25,40
22,10	49,90	23,20	21,30
25,50	53,50	24,80	19,30
31,10	56,60	30,00	25,40
30,40	56,70	28,30	27,20
18,70	46,50	23,00	11,70
19,70	44,20	28,60	17,80
14,60	42,70	21,30	12,80
29,50	54,40	30,10	23,90
27,70	55,30	25,70	22,60
30,20	58,60	24,60	25,40
22,70	48,20	27,10	14,80
25,20	51,00	27,50	21,10

Το εξαρτημένο μέγεθος του Δείκτη Σωματικού Λίπους θα εξεταστεί βάση του συνόλου ανεξάρτητων μεταβλητών.

- Οι μεταβλητές αυτές είναι:
- Πάχος Δέρματος
- Περιφέρεια Μηρού
- Περιφέρειες Μπράτσου

#### **Τεχνική και αποτελέσματα:**

Με την βοήθεια του Στατιστικού πακέτου SPSS έκδοση 17 υπολογίζουμε τα περιγραφικά στατιστικά μέτρα των μεταβλητών που ενδιαφερόμαστε.

**Statistics**

		πάχος δέρματος	περιφέρεια μη- ρού	περιφέρεια μπράτσου	δείκτης σωματι- κού λίπους
N	Valid	20	20	20	20
	Missing	0	0	0	0
	Mean	25,305	51,170	27,620	20,195
	Std. Deviation	5,0233	5,2346	3,6471	5,1062
	Variance	25,233	27,401	13,302	26,073
	Minimum	14,6	42,2	21,3	11,7
	Maximum	31,4	58,6	37,0	27,2



Ο έλεγχος κανονικής κατανομής για τα κατάλοιπα (Residuals) έδειξε ότι η υπόθεση της κανονικής κατανομής δε μπορεί να απορριφθεί, συνεπώς τα κατάλοιπα ακολουθούν κανονική κατανομή σύμφωνα με το τεστ των Kolmogorov – Smirnov και των Shapiro – Wilk με κρίσιμες πιθανότητες  $p = 0.2$  και  $p = 0.670$  αντίστοιχα και επίπεδο σημαντικότητας  $\alpha = 5\%$ .

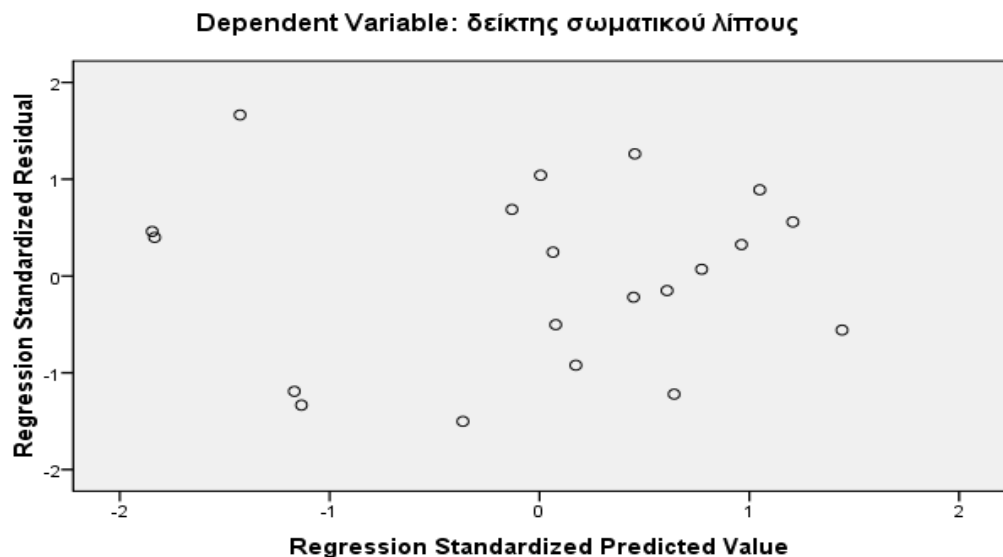
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	,106	20	,200 <sup>*</sup>	,966	20	,670

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

Από τη γραφική παράσταση των καταλοίπων (residuals) ως προς τις εκτιμώμενες τιμές (predicted) δεν προέκυψε ιδιαίτερο πρόβλημα μη σταθερής διακύμανσης.

#### Scatterplot



Η μελέτη των συσχετίσεων γίνεται μέσω του πίνακα συσχετίσεων. Στον πίνακα αυτό απεικονίζεται η τιμή του δείκτη που περιγράφει την συσχέτιση, καθώς και η τιμή ελέγχου σημαντικότητας για κάθε συντελεστή.

Για την αποφυγή δημιουργίας λάθος μοντέλου πριν την εφαρμογή της

γραμμικής παλινδρόμησης να εξετάζονται οι συσχετίσεις των μεταβλητών μεταξύ τους. Η συμμετοχή δύο ανεξάρτητων μεταβλητών αυξάνει τη δημιουργία λάθους μοντέλου.

### Πίνακας συσχετίσεων

#### Correlations

		Πάχος Δέρμα- τος	Περιφέρεια Μηρού	Περιφέρεια Μπράτσου	Δείκτης Σωμα- τικού Λίπους
Πάχος Δέρματος	Pearson Correlation	1	,924**	,458*	,843**
	Sig. (2-tailed)		,000	,042	,000
	N	20	20	20	20
Περιφέρε- ια Μηρού	Pearson Correlation	,924**	1	,085	,878**
	Sig. (2-tailed)	,000		,723	,000
	N	20	20	20	20
Περιφέρε- ια Μπράτ- σου	Pearson Correlation	,458*	,085	1	,142
	Sig. (2-tailed)	,042	,723		,549
	N	20	20	20	20
Δείκτης Σωματι- κού Λίπο- υς	Pearson Correlation	,843**	,878**	,142	1
	Sig. (2-tailed)	,000	,000	,549	
	N	20	20	20	20

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

Από τον πίνακα συσχετίσεων παρατηρούμε ότι από τις ανεξάρτητες μεταβλητές η μεταβλητή Περιφέρεια Μηρού είναι ισχυρότερα συσχετισμένη με το Δείκτη σωματικού λίπους.

Παρατηρείται επίσης ότι οι ανεξάρτητες μεταβλητές Πάχος Δέρματος, Περιφέρεια Μηρού, Περιφέρεια Μπράτσου είναι πολύ ισχυρά συσχετισμένες μεταξύ τους.

Στην συνέχεια με την βοήθεια του Στατιστικού Πακέτου SPSS θα εφαρμόσουμε την τεχνική της γραμμικής παλινδρόμησης εντοπίζοντας τις στατιστικά σημαντικές μεταβλητές για το εξαρτημένο μέγεθος και θα προσδιορίσουμε το μοντέλο που περιγράφει τον τρόπο σύνδεσης των μεταβλητών.

Υπάρχουν διάφοροι τρόποι κατασκευής του μοντέλου της παλινδρόμησης. Η κάθε διαδικασία κατασκευής του μοντέλου ακολουθεί τη δική της λογική και σχεδόν ποτέ δεν καταλήγουν στο ίδιο αποτέλεσμα.

Οι μέθοδοι που χρησιμοποιούνται συχνότερα είναι η Enter, η Forward η Backward και η Stepwise.

Η μέθοδος Enter χρησιμοποιεί στο μοντέλο της παλινδρόμησης όλες τις προτεινόμενες από τον αναλυτή μεταβλητές. Για την χρήση της τεχνικής αυτής ο χρήστης θα πρέπει να αποκλείσει από την ανάλυση τις ανεξάρτητες μεταβλητές που συνδέονται ισχυρά μεταξύ τους.

Η μέθοδος Forward:

- Εντοπίζει μεταξύ των ανεξάρτητων μεταβλητών εκείνη που συνδέεται ισχυρότερα με την εξαρτημένη μεταβλητή
- Κάνει έναν στατιστικό έλεγχο σημαντικότητας της μεταβλητής αυτής και ένα συνολικό έλεγχο του μοντέλου. Σε ένα επόμενο βήμα
- επιλέγει την μεταβλητή εκείνη με την αμέσως υψηλότερη συσχέτιση με την εξαρτημένη και την χαμηλότερη συσχέτιση με την ανεξάρτητη μεταβλητή που συμμετέχει ήδη στο μοντέλο.
- Ακολουθούν οι έλεγχοι για την στατιστική σημαντικότητα των μεταβλητών
- ο έλεγχος για την σημαντικότητα του μοντέλου
- Η διαδικασία αυτή επαναλαμβάνεται μέχρι να μην υπάρχει άλλη στατιστικά σημαντική μεταβλητή ως προς την εξαρτημένη ή όταν οι πληροφορίες των υπόλοιπων ανεξάρτητων μεταβλητών έχουν ήδη περιγραφεί στο μοντέλο μέσω άλλων ανεξάρτητων μεταβλητών.

Η μέθοδος Backward ακολουθεί τον αντίθετο ακριβώς δρόμο από την Forward.

Στο πρώτο βήμα του αλγορίθμου

- Εισάγονται όλες οι μεταβλητές στο μοντέλο και σε κάθε διαδοχικό βήμα.

- Αφαιρείται η μεταβλητή που θεωρείται λιγότερο σημαντική για την εξαρτημένη.
- Η διαδικασία επαναλαμβάνεται τόσες φορές όσες είναι απαραίτητες για να συμμετάσχουν στο μοντέλο, τελικά, μόνο στατιστικά σημαντικές μεταβλητές, ισχυρά συσχετισμένες με την εξαρτημένη μεταβλητή και ασυσχέτιστες μεταξύ τους.

Τέλος, η Stepwise είναι η δημοφιλέστερη μέθοδος και αποτελεί συνδυασμό της Forward, και της Backward.

Η παλινδρόμηση είναι μια πολυπαραγοντική μέθοδος κατά την οποία η προσθήκη ή η αφαίρεση μιας μεταβλητής μπορεί να αυξομειώσει τις τιμές των συντελεστών συσχέτισης. Για το λόγο αυτό οι διαφορετικοί αλγόριθμοι κατασκευής μοντέλων παλινδρόμησης οδηγούν σε διαφορετικά μοντέλα.

Στον παρακάτω πίνακα Variables Entered/Remove παρατηρούμε συνοπτικά την διαδικασία εισαγωγής των μεταβλητών στο μοντέλο.

Model	Variables Entered	Variables Removed	Method
1	περιφέρεια μπράτσου, περιφέρεια μηρού, πάχος δέρματος <sup>a</sup>		. Enter

a. All requested variables entered.

b. Dependent Variable: δείκτης σωματικού λίπους

Στον παρακάτω πίνακα Model Summary απεικονίζονται μερικοί βασικοί δείκτες καλής προσαρμογής του μοντέλου.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,895 <sup>a</sup>	,801	,764	2,4800	2,243

a. Predictors: (Constant), περιφέρεια μπράτσου, περιφέρεια μηρού, πάχος δέρματος

b. Dependent Variable: δείκτης σωματικού λίπους

Ο δείκτης R Square είναι ένδειξη του ποσοστού της διακύμανσης της εξαρτημένης μεταβλητής που επεξηγεί το μοντέλο.

Τιμές του δείκτη κοντά στο ένα είναι ένδειξη ότι οι παράγοντες που συμμετέχουν στην διαδικασία κατασκευής του μοντέλου είναι ικανοποιητικοί για την περιγραφή της κίνησης της εξαρτημένης μεταβλητής.

Τιμές του δείκτη κοντά στο μηδέν είναι ένδειξη ότι οι προτεινόμενες ανεξάρτητες μεταβλητές δεν είναι ικανοποιητικές για την περιγραφή της εξαρτημένης τιμής και επομένως το μοντέλο της παλινδρόμησης δεν θα πρέπει να χρησιμοποιηθεί για πρόβλεψη τιμών.

Στον παρακάτω πίνακα Ανάλυσης Διακύμανσης ANOVA απεικονίζεται ένας συνολικός έλεγχος για την σημαντικότητα του μοντέλου της παλινδρόμησης.

ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	396,985	3	132,328	21,516	,000 <sup>a</sup>
	Residual	98,405	16	6,150		
	Total	495,389	19			

a. Predictors: (Constant), περιφέρεια μπράτσου, περιφέρεια μηρού, πάχος δέρματος

b. Dependent Variable: δείκτης σωματικού λίπους

Ο έλεγχος βασίζεται στην συνάρτηση F και ελέγχει την υπόθεση ότι οι συντελεστές των ανεξάρτητων μεταβλητών που συμμετέχουν στο μοντέλο είναι ταυτόχρονα μηδέν.

Όταν η τιμή του επιπέδου στατιστικής σημαντικότητας (sig.) είναι μικρότερη του 0,05 τότε απορρίπτουμε την αρχική υπόθεση, δηλαδή το μοντέλο είναι στατιστικά σημαντικό. Το σύνολο της διακύμανσης του δείγματος 495,389 είναι άθροισμα της διακύμανσης της παλινδρόμησης 396,985 και της διακύμανσης του λάθους 98,405.

Στον παρακάτω πίνακα Coefficients γίνεται αναλυτική περιγραφή του μοντέλου.

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	117,085	99,782		1,173	,258		

πάχος δέρματος	4,334	3,016	4,264	1,437	,170	,001	708,843
περιφέρεια μηρού	-2,857	2,582	-2,929	-1,106	,285	,002	564,343
περιφέρεια μπράτσου	-2,186	1,595	-1,561	-1,370	,190	,010	104,606

a. Dependent Variable: δείκτης σωματικού λίπους

- Η στήλη Std.Error περιέχει τις τιμές του τυπικού σφάλματος της εκτίμησης των συντελεστών B.
- Στην στήλη t αναγράφονται οι τιμές της στατιστικής συνάρτησης βάση της οποίας γίνεται ο έλεγχος για την σημαντικότητα ύπαρξης των συντελεστών στην συνάρτηση.
- Στην τελευταία στήλη του πίνακα στην στήλη Sig. δίνεται η τιμή σημαντικότητας βάση της οποίας επιβεβαιώνεται η όχι η διατήρηση της μεταβλητής του μοντέλου.
- Στην στήλη Collinearity statistics δίνονται οι ενδείξεις πολυσυγγραμμικότητας

Collinearity Diagnostics<sup>a</sup>

Dime Model	nsion	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	πάχος δέρματος	περιφέρεια μηρού	περιφέρεια μπράτσου
1	1	3,968	1,000	,00	,00	,00	,00
	2	,021	13,905	,00	,00	,00	,00
	3	,012	18,566	,00	,00	,00	,01
	4	8,648E-6	677,372	1,00	1,00	1,00	,99

a. Dependent Variable: δείκτης σωματικού λίπους

Residuals Statistics<sup>a</sup>

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	11,758	26,786	20,195	4,5710	20
Residual	-3,7263	4,1277	,0000	2,2758	20
Std. Predicted Value	-1,846	1,442	,000	1,000	20
Std. Residual	-1,503	1,664	,000	,918	20

a. Dependent Variable: δείκτης σωματικού λίπους

Από τον πίνακα ANOVA προκύπτει ότι μία τουλάχιστον από τις παραμέτρους είναι στατιστικώς σημαντική γιατί η κρίσιμη πιθανότητα είναι  $p = 0.000 < 0.05$ , για επίπεδο σημαντικότητας  $\alpha = 5\%$ , ενώ από τον πίνακα coefficients προκύπτει ότι καμία από τις παραμέτρους δεν είναι στατιστικώς σημαντική, για επίπεδο σημαντικότητας  $\alpha = 5\%$ , γιατί οι κρίσιμες πιθανότητες για τον σταθερό όρο, το πάχος δέρματος, την περιφέρεια μηρού, και την περιφέρεια μπράτσου είναι αντίστοιχα  $p = 0.258$ ,  $p = 0.17$ ,  $p = 0.285$  και  $p = 0.19$  και είναι όλες μεγαλύτερες από το 0.05. Επομένως για την εκτίμηση του σωματικού λίπους δε χρειάζονται και οι τρεις παράγοντες. Σε αυτήν την περίπτωση, εφαρμόστηκε πάλι πολλαπλή γραμμική παλινδρόμηση με την μέθοδο Forward.

Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	περιφέρεια μηρού		Forward (Criterion: Probability-of-F- to-enter $\leq$ ,050)

a. Dependent Variable: δείκτης σωματικού λίπους

- Η πρώτη μεταβλητή (και μοναδική) που συμμετείχε στο μοντέλο είναι η Περιφέρεια Μηρού.

Η διαδικασία κατασκευής του μοντέλου σταματάει μετά την Εισαγωγή της πρώτης μεταβλητής.

Σε κάθε βήμα εισαγωγής της μεταβλητής στο μοντέλο γίνεται ένας έλεγχος στατιστικής σημαντικότητας για την εισαγωγή της μεταβλητής στην εξίσωση της παλινδρόμησης. Σε κάθε βήμα κατασκευής του μοντέλου ελέγχεται επίσης, η εισαγωγή κάποιας νέας μεταβλητής και η διατήρηση των μεταβλητών που ήδη συμμετέχουν σε αυτό.

Model Summary<sup>b</sup>

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,878 <sup>a</sup>	,771	,758	2,5102	2,500

a. Predictors: (Constant), περιφέρεια μηρού

b. Dependent Variable: δείκτης σωματικού λίπους

Ο έλεγχος της αυτοσυσχέτισης με το τεστ των Durbin – Watson έδειξε ότι δεν υπάρχει πρόβλημα αυτοσυσχέτισης γιατί η τιμή του στατιστικού είναι εντός των επιτρεπτών ορίων,  $dU < DW < 4 - dU$  και δεν υπάρχουν ενδείξεις πολυσυγγραμμικότητας.

ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	381,966	1	381,966	60,617	,000 <sup>a</sup>
	Residual	113,424	18	6,301		
	Total	495,389	19			

a. Predictors: (Constant), περιφέρεια μηρού

b. Dependent Variable: δείκτης σωματικού λίπους

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-23,634	5,657		-4,178	,001		
	περιφέρεια μηρού	,857	,110	,878	7,786	,000	1,000	1,000

a. Dependent Variable: δείκτης σωματικού λίπους

Μαθηματικά το μοντέλο της γραμμικής παλινδρόμησης παρουσιάζεται ως εξής:

$$\text{Δείκτης Σωματικού Λίπους} = -23,634 + (0,857) * \text{Περιφέρεια Μηρού.}$$

Από αυτήν την εξίσωση προκύπτει ότι η Περιφέρεια του Μηρού επηρεάζει το Δείκτη Σωματικού λίπους και για κάθε αύξηση της περιφέρειας του μηρού κατά μια μονάδα, ο δείκτης σωματικού λίπους αυξάνει κατά 0.857.

$$R^2 = 0,771$$

Το 77% περίπου της μεταβλητότητας του Δείκτη Σωματικού Λίπους εξηγείται από την Περιφέρεια του Μηρού ή η Περιφέρεια του Μηρού επηρεάζει κατά 77% τον Δείκτη Σωματικού Λίπους.



Excluded Variables<sup>b</sup>

Model	Beta In	t	Sig.	Partial Correlation	Collinearity Statistics		
					Tolerance	VIF	Minimum Tolerance
1	,219 <sup>a</sup>	,733	,474	,175	,147	6,825	,147
	,069 <sup>a</sup>	,595	,560	,143	,993	1,007	,993

a. Predictors in the Model: (Constant), περιφέρεια μηρού

b. Dependent Variable: δείκτης σωματικού λίπους

Collinearity Diagnostics<sup>a</sup>

Model	Dimensi on	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	περιφέρεια μηρού
1	1	1,995	1,000	,00	,00
	2	,005	20,108	1,00	1,00

a. Dependent Variable: δείκτης σωματικού λίπους

Residuals Statistics<sup>a</sup>

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	12,512	26,559	20,195	4,4837	20
Residual	-4,4949	4,4084	,0000	2,4433	20
Std. Predicted Value	-1,714	1,419	,000	1,000	20
Std. Residual	-1,791	1,756	,000	,973	20

a. Dependent Variable: δείκτης σωματικού λίπους

#### 4.3.9 Λογιστική Παλινδρόμηση

Η μέθοδο της λογιστικής παλινδρόμησης χρησιμεύει στο να εξεταστεί η σχέση μίας δίτιμης ανεξάρτητης τυχαίας μεταβλητής και συνεχών η διακριτών ανεξάρτητων μεταβλητών.

Η λογιστική παλινδρόμηση (Logistic Regression) είναι χρήσιμη σε καταστάσεις στις οποίες επιθυμούμε την πρόβλεψη της ύπαρξης ή της απουσίας ενός χαρακτηριστικού ή ενός συμβάντος. Η πρόβλεψη αυτή βασίζεται στην κατασκευή ενός γραμμικού μοντέλου και συγκεκριμένα στον προσδιορισμό των τιμών που παίρνουν οι

συντελεστές ενός συνόλου ανεξάρτητων μεταβλητών που χρησιμοποιούνται ως μεταβλητές πρόβλεψης. Εκτός από την πρόβλεψη ένα μοντέλο λογιστικής παλινδρόμησης δίνει τη δυνατότητα να εκτιμήσουμε την επίδραση κάθε ανεξάρτητης μεταβλητής στη διαμόρφωση των τιμών της εξαρτημένης μεταβλητής.

Στη λογιστική παλινδρόμηση, σε αντίθεση με την πολλαπλή παλινδρόμηση είναι δυνατό να χρησιμοποιηθούν ως εξαρτημένες μεταβλητές εκτός από αναλογικές αριθμητικές μεταβλητές και κατηγορικές μεταβλητές.

Ουσιαστικά η μέθοδος αυτή γενικεύει τα γραμμικά μοντέλα, έτσι ώστε η εξαρτημένη μεταβλητή να ακολουθεί την εκθετική οικογένεια κατανομών.

Για να γίνει περαιτέρω κατανοητή η χρήση της λογιστικής παλινδρόμησης παραθέτουμε το εξής παράδειγμα (Φωκιανός Κων., 2011):

Έστω μια βιομηχανία η οποία επιθυμεί να εξετάσει αν κάποιος εργάτης πάσχει από κάποια συγκεκριμένη ασθένεια του πνεύμονα. Για τη διενέργεια της μελέτης αυτής συγκεντρώθηκαν οι τιμές για τις ακόλουθες πέντε μεταβλητές :

- φυλή (1=λευκός, 2=άλλο)
- φύλο (1=άρρεν, 2=θήλυ)
- κάπνισμα (1=καπνιστής, 2=μη καπνιστής)
- διάρκεια εργασίας (1= λιγότερο από 10 χρόνια, 2=10-22 χρόνια, 3= περισσότερο από 20 χρόνια) Χάλκος
- ποσοστό σκόνης στον εργασιακό χώρο (1=ψηλό, 2=μέτριο 3=χαμηλό)

Το πρόβλημα για αυτά τα δεδομένα είναι το να εξακριβωθεί κατά πόσο οι επεξηγηματικές μεταβλητές είναι σημαντικές στην εμφάνιση αυτής της ασθένειας. Με άλλα λόγια, ποιες από αυτές τις μεταβλητές μπορούν να χρησιμοποιηθούν για να προβλέψουν κατά πόσο ένας εργάτης πάσχει από ασθένεια του πνεύμονα. Επειδή η ανεξάρτητη μεταβλητή είναι δυαδική, θα χρησιμοποιηθεί η λογιστική παλινδρόμηση για την ανάλυση.

Αντί να χρησιμοποιηθεί ένα γραμμικό μοντέλο για να εξεταστεί η εξάρτηση της πιθανότητας εμφάνισης της ασθένειας του πνεύμονα από τις επεξηγηματικές μεταβλητές, χρησιμοποιείται ο λογιστικός μετασχηματισμός, ο οποίος ορίζεται ως:

$$\log\left(\frac{p}{1-p}\right) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

Στο παράδειγμα,  $p$  είναι η πιθανότητα ένας εργάτης να πάσχει από ασθένεια του πνεύμονα. Στο μοντέλο υπάρχουν  $k$  (στο παράδειγμα 5) επεξηγηματικές μεταβλητές. Οι

συντελεστές παλινδρόμησης εκτιμούνται με τη μέθοδο της μέγιστης πιθανοφάνειας με την υπόθεση ότι η εξαρτημένη μεταβλητή ακολουθεί τη δυναμική κατανομή. Από την εξίσωση, το  $p$  μπορεί να υπολογιστεί από:

$$p = \frac{\exp(b_0 + b_1 X_1 + b_2 X_2 + \mathbf{L} + b_k X_k)}{1 + \exp(b_0 + b_1 X_1 + b_2 X_2 + \mathbf{L} + b_k X_k)}$$

Οι υπολογισμοί πραγματοποιούνται συνήθως με ειδικά λογισμικά σε ηλεκτρονικό υπολογιστή. Από τα πιο πάνω συμπεραίνεται ότι οι μεταβλητές σκόνη, κάπνισμα διάρκεια εργασίας είναι οι πιο σημαντικές για την πρόβλεψη ασθένειας του πνεύμονα, ενώ φαίνεται ότι οι άλλες δύο μεταβλητές δεν είναι τόσο σημαντικές. Στο συμπέρασμα αυτό καταλήγουμε από το τιμή  $p$  τους για τον  $t$  έλεγχο, αλλά και από την συνεισφορά της κάθε μεταβλητής στην απόκλιση όταν αυτή προστεθεί στο μοντέλο. Συνεπώς, εφαρμόζεται ένα νέο μοντέλο λογιστικής παλινδρόμησης με τις τρεις σημαντικές μεταβλητές και το συγκρίνεται με το προηγούμενο.

$$\log\left(\frac{p}{1-p}\right) = -0.1418 - 1.4657 \times \text{ΣΚΟΝΗ} - 0,6778 \times \text{ΚΑΠΝΙΣΜΑ} + 0.3331 \times \text{ΕΡΓΑΣΙΑ}$$

και είναι δυνατόν να υπολογιστεί η εκτιμώμενη τιμή της πιθανότητας κάποιος εργάτης να πάσχει από ασθένεια του πνεύμονα για κάθε συνδυασμό τιμών από τις τρεις επεξηγηματικές μεταβλητές.

#### 4.3.10 Ανάλυση της Διακύμανσης (ANOVA)

Η ανάλυση της διακύμανσης (Analysis Of Variance – ANOVA) είναι μία στατιστική μέθοδος με την οποία η μεταβλητότητα που υπάρχει σ' ένα σύνολο δεδομένων διασπάται στις επιμέρους συνιστώσες της με στόχο την κατανόηση της σημαντικότητας των διαφορετικών πηγών προέλευσής της. Η ανάπτυξη της οφείλεται στον θεμελιωτή της σύγχρονης στατιστικής επιστήμης, άγγλο στατιστικό Sir Ronald Aylmer Fisher (1890-1962). Στην πραγματικότητα η ANOVA περιλαμβάνει μία ομάδα στατιστικών μεθόδων καταλλήλων για την ανάλυση δεδομένων που προκύπτουν από πειραματικούς σχεδιασμούς.

Τα δεδομένα ενός δείγματος ανάλογα με την προέλευσή τους διακρίνονται σε παρατηρήσεις (observational sampling) ή σε πειραματικά (designed sampling). Στην πρώτη κατηγορία ο στατιστικός ερευνητής απλά παρατηρεί τις τιμές που εμφανίζονται χωρίς να έχει δυνατότητα επέμβασης στις αντίστοιχες μεταβλητές. Αντίθετα στη

δεύτερη κατηγορία ο στατιστικός ερευνητής προσπαθεί να ελέγξει τα επίπεδα μιας η περισσότερων ανεξάρτητων (independent) μεταβλητών προκειμένου να προσδιορίσει την επίδραση που έχουν πάνω στην υπό μελέτη μεταβλητή που καλείται εξαρτημένη (dependent) η απόκριση (response). Στόχος κάθε στατιστικού πειράματος είναι ο προσδιορισμός της επίδρασης μιας η περισσότερων ανεξάρτητων μεταβλητών πάνω στην απόκριση. Οι μεταβλητές αυτές αναφέρονται συνήθως σαν παράγοντες (factors) και μπορεί να είναι είτε ποσοτικές είτε ποιοτικές.

Στην απλούστερη μορφή της, η ANOVA μας δίνει τη δυνατότητα να δοκιμάσουμε την υπόθεση ότι οι μέσες τιμές διαφόρων πληθυσμών είναι ίσες. Κάτω από το πλαίσιο αυτό, η ANOVA θεωρείται σαν προέκταση της δοκιμασίας t μόνο που είναι προτιμότερη λόγω της συντομότερης διαδικασίας ανάλυσης και της ακρίβειας διάγνωσης.

Ανάλογα με τον αριθμό των παραγόντων που εξετάζονται διακρίνεται σε ανάλυση διακύμανσης ενός παράγοντα (One factor ANOVA) και ανάλυση διακύμανσης πολλαπλών παραγόντων (Multiple factor ANOVA).

Μελετώντας την στατιστική μελέτη για τον προσδιορισμό της σχέσης μεταξύ φυσικής δραστηριότητας και καθιστικής συμπεριφοράς σε παιδιά προσχολικής ηλικίας. (Στατιστική μελέτη του Επίκουρου Καθηγητή Προπονητικής Τμήμα Επιστήμης Φυσικής Αγωγής και Αθλητισμού του Δημοκρίτειου Πανεπιστήμιο Θράκης Αντώνιο Καμπά)

Το δείγμα της μελέτης αποτέλεσαν 102 παιδιά (49 αγόρια και 53 κορίτσια) ηλικίας 48 έως 79 μηνών (M.O.=60.67, T.A.=7.07).

Από τα αποτελέσματα διαπιστώθηκε στατιστικά σημαντική επίδραση του:

- Δείκτη Μάζας Σώματος ΔΜΣ
- Στον αριθμό των βημάτων/εβδομάδα ( $F=.411, p=.672$ ),
- Στον αριθμό των χλμ./εβδομάδα ( $F=.411, p=.673$ )
- και στις “καθιστικές” ώρες ( $F=.004, p=.996$ ).

Στατιστικά σημαντικές διαφορές παρατηρήθηκαν στο επίπεδο φυσικής δραστηριότητας μεταξύ παιδιών που έβλεπαν τηλεόραση λιγότερο από 1 ώρα ημερησίως αριθμός βημάτων/εβδομάδα: M.O.=61426.75, T.A.=10766.28 και συνολικά χλμ/εβδομάδα: M.O.=24.54, T.A.=4.31) και παιδιών που έβλεπαν τηλεόραση περισσότερο από 5 ώρες ημερησίως (αριθμός βημάτων/εβδομάδα: M.O.=37753.25, T.A.=2975.89 και χλμ. /εβδομάδα: M.O.=15.07, T.A.=1.18).

Επίδραση του Δείκτη Μάζας Σώματος στη Φυσική δραστηριότητα και στη Καθιστική συμπεριφορά.

Από την ανάλυση των δεδομένων δε διαπιστώθηκε στατιστικά σημαντική επίδραση του ΔΜΣ στο συνολικό αριθμό βημάτων/εβδομάδα ( $F_{2,96}=0.411$ ,  $p=0.672$ ) και στη συνολική απόσταση σε χλμ/εβδομάδα ( $F_{2,96}=0.411$ ,  $p=0.0673$ ). Σε απόλυτες τιμές, τα υπέρβαρα και τα παχύσαρκα παιδιά ήταν λιγότερο φυσικά δραστήρια, σε σχέση με τα κανονικά σε βάρος παιδιά (Πίνακας 1).

Επίσης, παρατηρήθηκε στατιστικά σημαντική επίδραση του ΔΜΣ στις ώρες παρακολούθησης τηλεόρασης/ημέρα ( $F_{2,96}=0.004$ ,  $p = 0.996$ ). Οι μέσοι όροι και οι τυπικές αποκλίσεις των μεταβλητών φυσικής δραστηριότητας και οι ώρες/ημέρα της καθιστικής συμπεριφοράς παρουσιάζονται στον παρακάτω πίνακα.

Το δεύτερο υπόδειγμα στον παρακάτω πίνακα περιλαμβάνει των Αριθμός βημάτων ανά εβδομάδα, χλμ και τις ώρες παρακολούθησης τηλεόρασης ανά ημέρα παιδιών προσχολικής ηλικίας σχετικοποιημένα με τον Δείκτη Μάζας Σώματος.

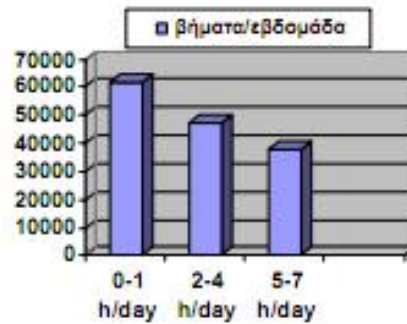
Επίδραση της Καθιστικής συμπεριφοράς στη Φυσική δραστηριότητα.

Φυσική δραστηριότητα	Κανονικά σε Βάρος(N52)		Υπέρβαρα (N=31)		Παχύσαρκα (N=19)	
	M.O.	T.A	M.O.	T.A	M.O.	T.A
Βήματα/εβδομάδα	56.622,17	16.704,37	49.707,80	9.418,99	49.525,33	15.579,53
Χιλιόμετρα/εβδομάδα	22,62	6,69	19,85	3,77	19,78	6,23
Καθιστική συμπεριφορά	2,67	2,42	2,80	2,49	2,67	2,89

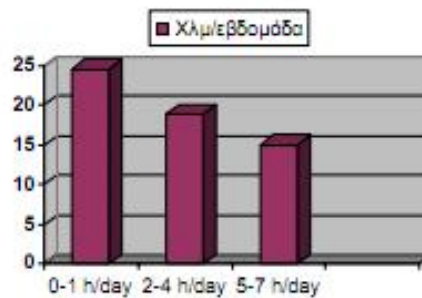
Από την ανάλυση των δεδομένων διαπιστώθηκε στατιστικά σημαντική επίδραση των ωρών παρακολούθησης τηλεόρασης/ημέρα στο συνολικό αριθμό βημάτων/εβδομάδα ( $F_{2,96}=10.239$ ,  $p<0.05$ ) και στα συνολικά χλμ/εβδομάδα ( $F_{2,96}=10.254$ ,  $p<0.05$ ).

Με την μέθοδο των κατά ζευγών συγκρίσεων διαπιστώθηκε στατιστικά σημαντική διαφορά στον αριθμό των βημάτων/εβδομάδα (M.O.=61426.75, T.A.=10766.28) και στα συνολικά χλμ/εβδομάδα (M.O.=24.54, T.A.=4.31) των

παιδιών που παρακολουθούσαν λιγότερο από 1 ώρα τηλεόραση/ημέρα, αναφορικά με τον αριθμό των βημάτων/εβδομάδα (Μ.Ο.=37753.25, Τ.Α.=2975.89) και των χλμ/εβδομάδα (Μ.Ο.=15.07, Τ.Α.=1.18) των παιδιών που παρακολουθούσαν τηλεόραση περισσότερο από 5 ώρες/ημέρα (Διαγράμματα 1 και 2).



Διάγραμμα 1. Διαφορές στις ομάδες μεταξύ συνολικών βημάτων /εβδομάδα και ωρών παρακολούθησης τηλεόρασης .



Διάγραμμα 2. Διαφορές στις ομάδες μεταξύ συνολικών χλμ/εβδομάδα και ωρών παρακολούθησης τηλεόραση.

Τα παιδιά που παρακολουθούσαν τηλεόραση λιγότερο από 1 ώρα/ημέρα βάδισαν κατά μέσο όρο 23.674 περισσότερα βήματα και κάλυψαν εβδομαδιαίως επιπλέον απόσταση 9.5 χιλιομέτρων, από τα παιδιά που παρακολουθούσαν τηλεόραση περισσότερο από 5 ώρες/ημέρα.

#### 4.3.11 Ανάλυση Συνδιακύμανσης

Η ανάλυση συνδιακύμανσης ή συμμεταβλητή ανάλυση συνδυάζει την απλή ανάλυση διακύμανσης (ANOVA) με την παλινδρομική ανάλυση (regression). Υπολογίζει κατά πόσο διαφοροποιούνται οι μέσοι όροι της εξαρτημένης μεταβλητής μεταξύ των ομάδων/επιπέδων ενός ή περισσότερων ανεξάρτητων παραγόντων (όπως στην ANOVA αφού προηγουμένως αφαιρέσει την επίδραση μίας ή περισσότερων

«τρίτων» μεταβλητών ανάλογα με το πώς οι «τρίτες» μεταβλητές είναι σε θέση να προβλέψουν τις τιμές της εξαρτημένης μεταβλητής (όπως στη regression analysis). Ονομάζουμε αυτές τις «τρίτες» μεταβλητές συμμεταβλητές (covariates). (Παυλόπουλος Β., 2008)

Ουσιαστικά, η ανάλυση συνδιακύμανσης διορθώνει τους μέσους όρους της εξαρτημένης μεταβλητής μεταξύ των συγκρινόμενων ομάδων ανάλογα με την επίδραση της συμμεταβλητής στην εξαρτημένη μεταβλητή. Οι διορθωμένοι μέσοι όροι που προκύπτουν, ονομάζονται προσαρμοσμένοι μέσοι όροι (adjusted means). Με τη διαδικασία αυτή προσπαθούμε να μειώσουμε την «εντός των ομάδων» πηγή διασποράς (error variance), αφαιρώντας την επίδραση της συμμεταβλητής.

Έστω τα ότι έχουμε τα ακόλουθα δεδομένα :

- $Y_{jk}$ : η βαθμολογία για τρεις διαφορετικές μεθόδους, A, B, C, και
- $x$ : η ικανότητα μάθησης πριν την διδασκαλία για 7 μαθητές.

Για να εξετάσουμε κατά πόσον υπάρχουν διαφορές μεταξύ μεθόδων δοθέντος της  $x$ , θεωρούμε το πλήρες μοντέλο

$$E(Y_{jk}) = \mu_j + Yx_{jk}, \quad j = 1, 2, 3, k = 1, \dots, 7$$

και το μοντέλο  $E(Y_{jk}) = \mu + Yx_{jk}$

Παρατηρούμε ότι  $j = 1$  αντιστοιχεί στη μέθοδο A,  $j = 2$  στη μέθοδο B και  $j = 3$  στη μέθοδο C.

Έστω,

$$Y_j = \begin{bmatrix} Y_{j1} \\ \mathbf{M} \\ Y_{j7} \end{bmatrix} \quad \text{και} \quad x_j = \begin{bmatrix} x_{j1} \\ \mathbf{M} \\ x_{j7} \end{bmatrix}$$

Το πλήρες μοντέλο δίνεται από

$$E(Y) = Xb$$

με

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}, \quad b = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ Y \end{bmatrix},$$

Και

$$X = \begin{bmatrix} 1 & 0 & 0 & x_1 \\ 0 & 1 & 0 & x_2 \\ 0 & 0 & 1 & x_3 \end{bmatrix}$$

#### 4.3.12 Επαναλαμβανόμενες Μετρήσεις

Επαναλαμβανόμενες μετρήσεις προκύπτουν όταν παρατηρήσεις της ίδιας εξαρτημένης μεταβλητής γίνονται πάνω στον χρόνο, σε κάθε ένα από έναν αριθμό πειραματικών μονάδων (συνήθως υποκείμενα ή ασθενείς), οι οποίοι είναι δυνατόν να ανήκουν σε μία ή περισσότερες θεραπευτικές ή μη ομάδες. Στις επαναλαμβανόμενες μετρήσεις μπορούν να ενταχθούν και μετρήσεις που γίνονται στα υποκείμενα κάτω από διαφορετικές πειραματικές συνθήκες. Το βασικό ενδιαφέρον βρίσκεται στον τρόπο με τον οποίο οι μετρήσεις φαίνονται να αλλάζουν στο χρόνο (ή στις πειραματικές συνθήκες) και πως αυτός επηρεάζεται πιθανώς από την ομάδα στην οποία ανήκει το υποκείμενο ή από άλλες συνοδευτικές μεταβλητές. (Βασδέκης Β.)

Η βασική μορφή των δεδομένων είναι  $n$  υποκείμενα  $\times$   $T$  επαναλαμβανόμενες μετρήσεις για κάθε υποκείμενο  $i$ . Η γενίκευση αυτού του σχήματος βρίσκεται σε δύο κατευθύνσεις. Είτε τα  $n$  υποκείμενα είναι χωρισμένα σε  $m$  ομάδες (παράγοντες μεταξύ των υποκειμένων), είτε οι  $T$  μετρήσεις γίνονται σε περισσότερες από μία μεταβλητές (παράγοντες εντός των υποκειμένων). Συχνά βέβαια συναντάται ο συνδυασμός των δύο αυτών γενικεύσεων.

Οι επαναλαμβανόμενες μετρήσεις σε ένα υποκείμενο είναι συνήθως συσχετισμένες οπότε η υπόθεση της ανεξαρτησίας δεν ισχύει. Επιπλέον, μπορεί η ίδια η συσχέτιση να αποτελεί ερευνητικό ερώτημα εκτός από την τάση πάνω στο χρόνο.

Παραδείγματα επαναλαμβανόμενων μετρήσεων:

1. Ιοί HIV επιτίθενται στα  $T$  λεμφοκύτταρα του αίματος. Όσο πέφτει ο αριθμός τους πέφτει και η ικανότητα άμυνας. Ερώτημα: Ποία είναι η επίδραση του HIV στον αριθμό των  $T$  λεμφοκυττάρων πάνω στον χρόνο;
2. Σε κλινική δοκιμή επιληπτικών μετριέται ο αριθμός των επιληπτικών κρίσεων σε μία σειρά από χρονικές στιγμές. Οι ασθενείς είναι χωρισμένοι σε δύο ομάδες, θεραπείας και placebo. Υπάρχουν μετρήσεις που λαμβάνονται κατά την διάρκεια της περιόδου πριν ξεκινήσει η χορήγηση του φαρμάκου, και μετρήσεις μετά την χορήγηση. Στόχος: Το φάρμακο ελαττώνει τον αριθμό των επιληπτικών κρίσεων; Η ελάττωση αυτή γίνεται με τον ίδιο τρόπο και στις δύο ερευνητικές ομάδες; Πως συγκρίνονται οι αλλαγές αυτές από υποκείμενα με τον ίδιο αριθμό επιληπτικών κρίσεων πριν την έναρξη της θεραπείας;

Κοινά στοιχεία των παραπάνω ερευνών:

- Υπάρχουν επαναλαμβανόμενες παρατηρήσεις σε κάθε υποκείμενο



(πειραματική μονάδα).

- Οι μετρήσεις διαφορετικών υποκειμένων μπορούν να θεωρηθούν ανεξάρτητες.
- Μετρήσεις εντός των υποκειμένων τείνουν να είναι συσχετισμένες.
- Το αντικείμενο της έρευνας είναι να περιγραφεί η εξάρτηση της απόκρισης από επεξηγηματικές μεταβλητές.
- Η επιλογή του στατιστικού μοντέλου εξαρτάται από την φύση της απόκρισης.

Κατόπιν μπορεί κάποιος να αναλύσει τα δεδομένα χρησιμοποιώντας μεθόδους οι οποίες κυμαίνονται από πολύ απλές έως πολύ σύνθετες. Στην ιατρική έρευνα μία ιδιαίτερα απλή μέθοδος είναι η σύγκριση των θεραπευτικών ομάδων σε κάθε χρονική στιγμή χρησιμοποιώντας δοκιμασίες t, ANOVA για μη ανεξάρτητα δείγματα (Mann-Whitney για δύο ομάδες ή Kruskal-Wallis για πολλές ομάδες).

#### 4.4. Μέθοδοι για μη ανεξάρτητα δείγματα

Παρακάτω αναλύονται οι κυριότερες μέθοδοι που χρησιμοποιούν για την ανάλυση μη ανεξαρτήτων δειγμάτων, όπως οι μέθοδοι KRUSKAL-WALLIS και MANN-WHITNEY-U. Επιπλέον γίνεται αναφορά κι στον έλεγχο Jonckheere.

##### 4.4.1. Έλεγχος MANN-WHITNEY-U

Ως έλεγχος Mann-Whitney θεωρείται το μη παραμετρικό ανάλογο του t-test για δύο ανεξάρτητα δείγματα ελέγχει δε τις υποθέσεις:

- $H_0$ : οι  $n_1$   $x_i$  παρατηρήσεις του πληθυσμού X κατανέμονται ακριβώς όπως και οι  $n_2$   $y_j$  παρατηρήσεις του πληθυσμού Y.
- $H_1$ : οι  $n_1$   $x_i$  παρατηρήσεις του πληθυσμού X κατανέμονται διαφορετικά από τις  $n_2$   $y_j$  παρατηρήσεις του πληθυσμού Y.

Σε σχέση με τις πρακτικές εφαρμογές, η διαφορά μεταξύ των συναρτήσεων κατανομής δύο πληθυσμών, αναφέρεται στην διαφορά της θέσης των δύο κατανομών οπότε οι υποθέσεις που ελέγχονται είναι:

$$H_0 : E(X) = E(Y)$$

$$H_1 : E(X) \neq E(Y)$$

Η ελεγχουσυνάρτηση που χρησιμοποιείται στην περίπτωση που υπάρχει ταύτιση τιμών είναι η :

$$T = \frac{\sum R(x_i) - \frac{n_1(N+1)}{2}}{\sqrt{\frac{n_1 n_2}{N(N-1)} \sum R_i^2 - \frac{n_1 n_2 (N+1)^2}{4(N-1)}}$$

με  $R(x_i)$ : βαθμός των  $x_i$  παρατηρήσεων στο συνενωμένο δείγμα  $N = n_1 + n_2$  και  $\Sigma^2$ : άθροισμα τετραγώνων των  $N$  μέσων βαθμών που χρησιμοποιούνται στο συνενωμένο δείγμα. Ο έλεγχος πραγματοποιείται με σύγκριση της  $T$  με τα ποσοστιαία σημεία της τυποποιημένης κανονικής κατανομής (Ξεκαλάκη, 2001).

#### 4.4.2. Έλεγχος KRUSKAL-WALLIS

Ο έλεγχος KRUSKAL-WALLIS είναι η μη παραμετρική ανάλυση διακύμανσης και ελέγχει τις υποθέσεις:

- $H_0$ : οι συναρτήσεις κατανομής  $k$  πληθυσμών είναι ίσες
- $H_1$ : δύο τουλάχιστον από τους  $k$  πληθυσμούς έχουν διαφορετικές μέσες τιμές.

Η συνάρτηση ελέγχου που χρησιμοποιείται είναι η:

$$T = \frac{12}{N(N+1) \cdot \left[ \frac{\sum R_1^2}{n_1} + L + \frac{\sum R_k^2}{n_k} \right]} - 3(N+1)$$

με  $R_i$  = τις τάξεις μεγέθους των  $k$  ανεξάρτητων πληθυσμών σε συνενωμένο δείγμα

$$N = n_1 + \dots + n_k$$

Η απόφαση λαμβάνεται με σύγκριση της τιμής της συνάρτησης ελέγχου  $T$  με την τιμή  $\chi^2_{(k-1), 1-\alpha}$  (Mason-Lind, 1996).

#### 4.4.3. Έλεγχος Jonckheere

Ο έλεγχος Jonckheere χρησιμοποιείται όταν  $k$  ανεξάρτητες ομάδες τις οποίες θέλουμε να ελέγξουμε ως προς την ισότητα των διαμέσων τους είναι εκφρασμένες σε κλίμακα διάταξης. Αντιστοιχίζοντας τον αριθμό 1 στην ομάδα με την μικρότερη υποτιθέμενη διάμεσο και το  $k$  ( $k > k-1 > \dots > 1$ ) στην ομάδα με την μεγαλύτερη διάμεσο οι υποθέσεις που διατυπώνονται είναι οι:

- $H_0$ : διάμεσος<sub>1</sub>= διάμεσος<sub>2</sub>=...=διάμεσος<sub>κ</sub>
- $H_1$ : διάμεσος<sub>1</sub><διάμεσος<sub>2</sub><...<διάμεσος<sub>κ</sub>

Η συνάρτηση ελέγχου που χρησιμοποιείται είναι η  $J = \sum \sum U_{ij}$ , με  $U_{ij} = \sum X_{hi,j}$  και  $\sum X_{hi,j}$  να εκφράζει τον αριθμό των φορών που η μονάδα  $X_{hi}$  είναι μικρότερη από μια μονάδα στο δείγμα  $j$ , για  $i < j$ .

Για μεγάλα δείγματα χρησιμοποιείται η συνάρτηση ελέγχου  $J = J_{\mu_j}/\sigma_j$ , η οποία ακολουθεί κανονική κατανομή με μέσο 0 και τυπική απόκλιση 1 (Siegel-Castellan, 1988.)

#### 4.5 Επίδραση Placebo – Εικονικά Φάρμακα

Το placebo είναι μια ψευδής ή εικονική ιατρική παρέμβαση, η οποία χρησιμοποιείται ως θεραπεία ελέγχου, κατά τη μελέτη μιας πραγματικής νέας θεραπείας. Μερικές φορές οι ασθενείς που λαμβάνουν μια θεραπεία placebo εμφανίζουν βελτίωση είτε πραγματική είτε αντιληπτή.

Κοινές θεραπείες placebo συμπεριλαμβάνουν μη ενεργά φάρμακα μπορεί να έχει τη μορφή ενός «χαπιού» ζάχαρης, αλλά ένα εικονικό φάρμακο μπορεί επίσης να είναι μια ένεση, ένα υγρό, ψευδείς εγχειρίσεις ή οποιοδήποτε άλλο είδος θεραπείας που δεν επηρεάζει άμεσα την ασθένεια αντιμετωπίζεται. Σε συνήθεις διαδικασίες placebo δίνεται στον ασθενή ένα χάπι, μη ενεργό, και του επισημαίνεται ότι η υγεία του θα βελτιωθεί. Μια τέτοια διαδικασία μπορεί να κάνει τον ασθενή να «πιστέψει» ότι το νέο φάρμακο έχει αλλάξει την κατάσταση της υγείας του.

Η χρήση του placebo φανερώνει την επίδραση που μπορεί να έχει η ψευδαίσθηση της καλύτερευσης της υγείας του ασθενούς στην πραγματική του υγεία. Επομένως μπορεί να χρησιμοποιηθεί ως ένα συγκριτικό μέτρο διόρθωσης των στατιστικών μελετών, συμπεριλαμβάνοντας στη μελέτη την ψυχοσωματική επιρροή που μπορεί να έχει η όλη διαδικασία στον ασθενή.

Στατιστικά η ανάλυση γίνεται με εφαρμογή κριτηρίων για μη ανεξάρτητα δείγματα ή με τεχνικές με επαναλαμβανόμενες μετρήσεις όταν πρόκειται για τον ίδιο τον ασθενή.

Μια από τις πρώτες έρευνες οι οποίες ανέδειξαν την αξία της χρήσης του placebo και από την οποία εξάχθηκαν πολύ σημαντικά αποτελέσματα αποτελεί αυτή του Jellinek το 1946. Σκοπός της έρευνα ήταν η διερεύνηση της αποτελεσματικότητας του κάθε συστατικού φαρμάκων έναντι των πονοκεφάλων. Τα φάρμακα αυτά

αποτελούνταν από τρία συστατικά: a, b και c. Το συστατικό b ήταν σε έλλειψη και εξεταζόταν η πιθανότητα να εξαλειφθεί από τη συνταγή. Η μελέτη του Jellinek περιελάμβανε 199 ασθενείς οι οποίοι εμφάνιζαν συχνούς πονοκεφάλους. Οι ασθενείς χωρίστηκαν σε τέσσερις ομάδες στους οποίους δόθηκαν τα ακόλουθα τέσσερα φάρμακα:

- Φάρμακο **A**: περιείχε **a**, **b** και **c**.
- Φάρμακο **B**: περιείχε **a** και **c**.
- Φάρμακο **C**: περιείχε **a** και **b**.
- Φάρμακο **D**: περιείχε μόνο γαλακτικό(placebo)

Τα χάπια είχαν όλα το ίδιο σχήμα, μέγεθος, χρώμα και γεύση. Κάθε δύο εβδομάδες οι ομάδες έπαιρναν διαφορετικό φάρμακο έτσι ώστε μέσα σε οκτώ εβδομάδες ο κάθε ασθενής να έχει δοκιμάσει και τέσσερα φάρμακα.

Από τους 199 ασθενείς οι 120 αντέδρασαν στο placebo ενώ οι 79, ποσοστό ιδιαίτερα μεγάλο το οποίο αποδεικνύει την ισχυρή επίδραση μερικές φορές των placebo φαρμάκων.

Όσον αφορά την επίδραση του συστατικού b, αποδείχθηκε ότι ήταν σημαντικό και δεν μπορούσε να παραληφθεί. Αξίζει να αναφερθεί ο τρόπος που χειρίστηκε ο Jellinek τα αποτελέσματα για να καταλήξει στο συμπέρασμα αυτό. Δεδομένης της επίδρασης placebo, απέκλεισε τα 120 άτομα και ως πληθυσμό έρευνας θεώρησε τα 79 άτομα που δεν είχαν το placebo effect. Σε αυτά τα άτομα τα φάρμακα A, B και C είχαν ποσοστά επιτυχίας 88%, 67% και 77% αντίστοιχα, συνηγορώντας υπέρ της ουσίας B.

Μερικές φορές η επίδραση προκαλεί δυσάρεστα ή χειρότερα συμπτώματα. Αυτά μπορεί να περιλαμβάνουν πονοκεφάλους, νευρικότητα, ναυτία, ή δυσκοιλιότητα, για να αναφέρουμε μερικά. Οι δυσάρεστες συνέπειες που συμβαίνουν μετά από τη λήψη του εικονικού φαρμάκου ή ανενεργό θεραπεία αποκαλείται επίδραση placebo.

## 5. ΣΥΖΗΤΗΣΗ – ΣΥΜΠΕΡΑΣΜΑΤΑ

Στις ημέρες μας υπάρχει πλέον καθολική αποδοχή της προσφοράς της στατιστικής στην παροχή υπηρεσιών υγείας. Η προσφορά αυτή δεν περιορίζεται στα στενά όρια της αξιολόγησης δεδομένων από ιατροβιολογικές έρευνας, αλλά επεκτείνεται και σε πιο καθημερινές εκφράσεις αυτής της παροχής υπηρεσιών. Πιο συγκεκριμένα, η στατιστική μπορεί να εφαρμοστεί σε επιδημιολογικές, κοινωνιολογικές, διοικητικές και οικονομικές παραμέτρους.

Για παράδειγμα, επωφελής οικονομικά μπορεί να αποδειχθεί, ιδιαίτερα σε μια τόσο δύσκολη περίοδο για την οικονομία της Ελλάδας, η σύγκριση μακροοικονομικών δεικτών σχετικών με τις δαπάνες υγείας ανάμεσα στην Ελλάδα και άλλες χώρες. Μέχρι τώρα η σύγκριση αυτή δεν είναι αξιόπιστη, λόγω μεθοδολογικών προβλημάτων που υπεισέρχονται στις μετρήσεις. Οι διεθνείς οργανισμοί δε χρησιμοποιούν κατ' ανάγκη τα ίδια πρότυπα καταγραφής στοιχείων. Χαρακτηριστικό είναι το παράδειγμα της μελέτης του Κοντοζαμάνη Β. (2000) που αφορά τα μεθοδολογικά προβλήματα στον υπολογισμό της φαρμακευτικής δαπάνης στην Ελλάδα, σύμφωνα με την οποία στα πρότυπα καταγραφής του Οργανισμού Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ – Organization for Economic Co-operation and Development - OECD), η Συνολική Φαρμακευτική Δαπάνη δεν περιλαμβάνει τα φάρμακα που καταναλώνονται στα νοσοκομεία, τα οποία εντάσσονται στη δαπάνη ενδονοσοκομειακής περίθαλψης. Η Ελλάδα, όπως προκύπτει από τα στοιχεία της Βάσης Δεδομένων Υγείας του ΟΟΣΑ 2000, έχει συμπεριλάβει στα στοιχεία που έχει υποβάλλει ως Συνολική Φαρμακευτική Δαπάνη στον ΟΟΣΑ και τη φαρμακευτική κατανάλωση εντός των νοσοκομείων, με αποτέλεσμα να υπερεκτιμάται η συνολική δαπάνη, όταν αυτή συγκρίνεται με τις χώρες του ΟΟΣΑ. Η Δημόσια και Συνολική Φαρμακευτική Δαπάνη στην Ελλάδα, ως ποσοστό επί του Ακαθάριστου Εθνικού Προϊόντος, σύμφωνα με τα πρότυπα καταγραφής του ΟΟΣΑ και μετά τις απαραίτητες διορθώσεις, που δεν ήταν όλες εφικτές, όπως π.χ. η αφαίρεση από τη φαρμακευτική δαπάνη της κατανάλωσης σε ιδιωτικές κλινικές και ιδιωτικά νοσοκομεία, είναι πολύ κοντά στο μέσο όρο των χωρών-μελών του ΟΟΣΑ.

Η αναξιοπιστία των στοιχείων οφείλεται στο γεγονός ότι η επεξεργασία τους παρουσιάζει μεθοδολογικά σφάλματα και η παροχή αυτών προς διεθνείς οργανισμούς, όπως ο ΟΟΣΑ, δε γίνεται με βάση τα πρότυπα καταγραφής που ακολουθεί ο κάθε οργανισμός. Επομένως, στην προσπάθεια να συγκριθούν μακροοικονομικοί δείκτες σχετικοί με τις δαπάνες υγείας ανάμεσα στην Ελλάδα και άλλες χώρες, λόγω των

αντικρουόμενων και αναξιόπιστων στοιχείων η φαρμακευτική δαπάνη δεν μπορεί να αποτελέσει εργαλείο σύγκρισης. Γίνεται εύκολα κατανοητό ότι μια σωστή στατιστική μελέτη θα μπορούσε να προσφέρει πολύτιμα στοιχεία στην Ελλάδα, τα οποία θα είχαν ως αποτέλεσμα την καλύτερη αντιμετώπιση της αλόγιστης φαρμακευτικής δαπάνης που παρατηρείται.

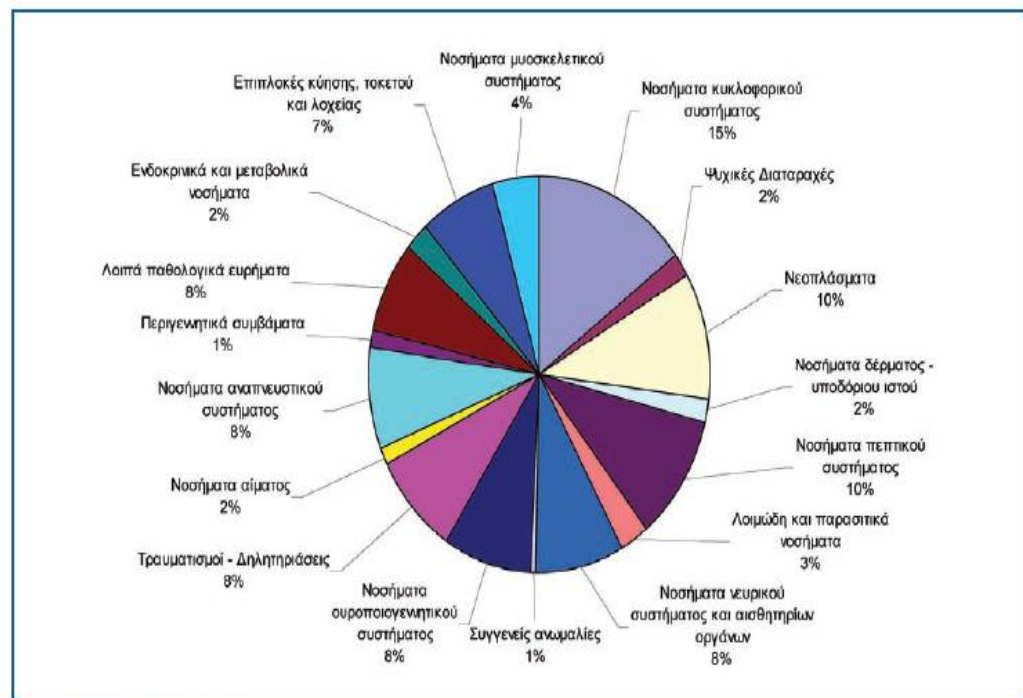
Μια ακόμα μελέτη στο οποίο φαίνεται η τεράστια συνεισφορά που μπορεί να έχει η στατιστική στις υπηρεσίες υγείας είναι η ακόλουθη: Όπως προκύπτει και από τα στατιστικά δεδομένα, έχει αυξηθεί σημαντικά ο αριθμός των προσερχόμενων στα Κέντρα Υγείας και τα Περιφερειακά Ιατρεία τόσο για πρωτοβάθμια περίθαλψη όσο και για πρόληψη, Η παρατηρούμενη αύξηση θεωρείται ότι παρουσιάζεται όχι μόνο λόγω νοσηρότητας αλλά κυρίως για τους παρακάτω λόγους όπως προκύπτει μετά από ερωτήσεις σε ασθενείς (Στοιχεία από το Γενικό Νοσοκομείο Κατερίνης):

- Η εμπιστοσύνη των πολιτών στις υποδομές των Κέντρων Υγείας αυξήθηκε.
- Ο αριθμός επισκέψεων στις υποδομές της Πρωτοβάθμιας Φροντίδας Υγείας βρίσκονται σε συνεχή άνοδο.
- Όσο επενδύουμε στις υποδομές Πρωτοβάθμιας Φροντίδας Υγείας (κτιριακές, ιατροτεχνολογικές, στελέχωση), τόσο μεγαλύτερη είναι η ανταπόκριση των πολιτών.
- Οι παρεχόμενες υπηρεσίες έχουν βελτιωθεί και
- Οι πολίτες έχουν ευαισθητοποιηθεί σημαντικά, λόγω της πληροφόρησης, σε θέματα κυρίως πρόληψης.

Τα στατιστικά αυτά δεδομένα προσέλευσης των ασθενών στα Κέντρα Υγείας και τα Περιφερειακά Ιατρεία μπορούν να χρησιμοποιηθούν από τους αρμόδιους φορείς για να ενισχυθεί ο εξοπλισμός τους και η στελέχωση τους με το κατάλληλο προσωπικό για να ανταποκριθεί στις νέες αυξημένες απαιτήσεις.

Η επάνδρωση των νοσοκομείων και ο εξοπλισμός τους δεν εξαρτάται μόνο από την ανάλυση του αριθμού των προσερχόμενων ασθενών, αλλά και από την αιτιολογία που αυτοί προσέρχονται. Συγκρίνοντας πρόσφατα στατιστικά στοιχεία εξελθόντων ασθενών με τα αντίστοιχα του 1997 (Τούντας Γ., 2007), παρατηρείται μια μεταβολή στην κατάταξη των νοσολογικών κατηγοριών, ως προς τον αριθμό των εισαγωγών. Τα νεοπλάσματα, από 4η αιτία το 1997 ανεβαίνουν στη 2η θέση και τα νοσήματα του νευρικού συστήματος και των αισθητηρίων οργάνων από την 9η στην 4η θέση. Επίσης, από τη συγκριτική μελέτη των στοιχείων του 1997 και του 2003, διαπιστώνεται ότι οι

τραυματισμοί και οι δηλητηριάσεις, που οφείλονται συνήθως σε ατυχήματα, το 1997 ήταν η 3η συχνότερη αιτία εισαγωγής σε νοσοκομείο, ενώ το 2003 ήταν μόλις η 6η. Τέλος, πρέπει να σημειωθεί ότι, τόσο συνολικά, όσο και ανά κατηγορία νοσημάτων, ο αριθμός των εξελθόντων αυξήθηκε στη διάρκεια της χρονικής περιόδου που εξετάζεται. Ο συνολικός αριθμός εισαγωγών στα ελληνικά νοσοκομεία ήταν 1.552.940 το 1997 και έφτασε το 1.940.512 το 2003, σημειώνοντας αύξηση κατά 20% .



Πηγή: ΕΣΥΕ, Ίδιοι υπολογισμοί (2004)

Σχήμα: Ποσοστιαία κατανομή των εξελθόντων ασθενών κατά κατηγορία πάθησης στην Ελλάδα (2003).

Η αξιοποίηση τέτοιων στατιστικών αποτελεσμάτων βοηθούν τους αρμόδιους φορείς να προσαρμόζουν τις ανάγκες των ιατρικών μονάδων ανάλογα με τη νοσηρότητα που παρουσιάζει ο εκάστοτε ο πληθυσμός, για παράδειγμα αυξημένα καρδιακά επεισόδια στη Χ περιοχή της Ελλάδος ή αύξηση των νεοπλασμάτων στην Ελλάδα, και τη στρατηγική τους όσον αφορά την προληπτική ιατρική.

Τα αξιόπιστα στατιστικά στοιχεία βοηθούν στην ορθή απεικόνιση της υφιστάμενης κατάστασης, συμβάλλουν στον εποικοδομητικό διάλογο, βοηθούν στη λήψη τεκμηριωμένων αποφάσεων και προάγουν την επιστημονική έρευνα. Για τους παραπάνω λόγους, κρίνεται απαραίτητη η αναβάθμιση της Στατιστικής Υπηρεσίας του υπουργείου Υγείας και θεωρείται σκόπιμη η συνεργασία της υπηρεσίας αυτής με τους εμπλεκόμενους φορείς. Μέσω της συνεργασίας αυτής θα επιτευχθεί η συστηματική συλλογή, καταγραφή και ανάλυση των στοιχείων των υπηρεσιών υγείας στη χώρα μας

και παράλληλα, θα παρέχονται ακριβή στοιχεία σε διεθνείς οργανισμούς. Ήδη από το 1998 λειτουργεί στην Ιατρική Σχολή του Εθνικού Καποδιστριακού Πανεπιστημίου Αθηνών το Κέντρο Μελετών Υπηρεσιών Υγείας του Εργαστηρίου Υγιεινής , Επιδημιολογίας και Ιατρικής Στατιστικής το οποί έχει ως σκοπό να συμβάλλει στην ανάπτυξη του τομέα των Υπηρεσιών Υγείας και της Προαγωγής - Αγωγής Υγείας, μέσω της εκπαίδευσης προπτυχιακών και μεταπτυχιακών φοιτητών και την εκπόνηση στατιστικών μελετών και προώθηση της προληπτικής Ιατρικής.

Η ορθή εφαρμογή της στατιστικής έχει αναμφισβήτητη και καταλυτική συνεισφορά στην πρόοδο και στην προαγωγή της δημοσίας υγείας. Τα προβλήματα δημιουργούνται όταν δεν τηρούνται οι απαιτούμενοι κανόνες δεοντολογίας από τους εμπλεκόμενους φορείς στις διάφορες φάσεις της έρευνας, από την αρχική, της σχεδίασής της, έως την τελική, της δημοσιοποίησης των αποτελεσμάτων της, στο ευρύ κοινό και την αξιοποίηση τους. Η ανάπτυξη της στατιστικής σκέψης στον σύγχρονο πολίτη του 21ου αιώνα, αντιμέτωπο με τόσες αλλαγές, αβεβαιότητες και κινδύνους, είναι περισσότερο από ποτέ απαραίτητη, όπως με μεγάλη διορατικότητα είχε προφητεύσει ο Samuel Wilks.



## **ΒΙΒΛΙΟΓΡΑΦΙΑ**

**Arrow K., Chenery B., Solow R., (1961).** Capital-Labor Substitution and Economic Efficiency, *Rev Econ Stat* XLII, 225-250.

**Alderson M (1976)** An introduction to epidemiology

**Bernard A. Rosner** Fundamentals Of Biostatistics

**Bradshaw J., (1972).** “A Taxonomy of Social Need” , in G.Mc Lachlan ( ed. ) , Problems and Progress in Medical Care : Essays on Current Research. Seventh Series, Oxford University Press for Nuffield Provincial Hospital Trusts

**Donabedian A., (1973).** Aspects of Medical Care Administration, Harvard University Press , Cambridge Ma.

**Ewin R. , Hall J., (1976).** Planning Health Services , Demand or Need, *Med J*,

**Evans, R.G. (1974),** Models, Markets and Medical Care

**Hornby P., Ray D., Shipp P., Hall T., (1980).** Guidelines for Health Manpower Planning, WHO, Geneva.

**Kasapi L., 2003** “Culture and Psychology. Special Feature: The role of culture in Psychology”, *PsyPag: Quarterly*

**Kirmayer, Groleau, Guzder, Blake, Jarvis,** Cultural consultation: a model of mental health service for multicultural societies. *Can J Psychiatry* 2003 Apr;48(3):145-53

**Lave L. , Silverman L.P. , (1972).** Hospital Cost Estimation Controlling for Case Mix, *Applied Economics*, No 3.

**Monsma G. (1970).** Marginal Revenue and Demand for Physician's Services in

**Klarkman H. (eds.),** Empirical Studies in Health Economics, Johns Hopkins University Press ,Baltimor.

**Mason, Robert D. - Lind, Douglas A.** Statistical techniques in business and economics

**M.Pagano's K.Gauvreau's** Principles of Biostatistics

**Phelps C.E., (1986).** Induced Demand -Can we Ever Know it's Extent, *Journal of Health Economics*, 5, 355-365.

Placebo effect : Άρθρο στην American Cancer Society

Placeboeffect: <http://www.cancer.org/Treatment/TreatmentsandSideEffects/TreatmentTypes/placebo-effect>

**Rubel Arthur (1977),** «The epidemiology of a folk illness: Susto in Hispanic America», στο *Culture, Disease, and Healing: Studies in Medical Anthropology*, Macmillan, New York, σ. 119-128.

- Simons S., Bonnett-Jones N., (1989).** Human Resource Development The Management, Planning and Training of Health Personnel, Evaluation & Planning Center For Health Care
- Williams A. ,(1988).** Health Economics : The End of Clinical Freedom, British Medical Journal, 297, 1183-1186.
- WHO (1998)** Health 21: an introduction to the health for all policy framework for the WHO European region. World Health Organization, Regional Office for Europe, Series No. 5, Copenhagen
- Wolf E. (1982),** Europe and the People Without History, University of California Press, Berkeley.
- Αγραφιώτης Δ. (2000),** Επιστήμη, Τεχνολογία, Κοινωνία, Ελληνικά Γράμματα, Αθήνα.
- Ανδριώτη Δ., Γεωργούση Ε., Κυριόπουλος Γ.,(1996).** Ο Γενικός Γιατρός στη Ελλάδα και τις Χώρες της Ευρωπαϊκής Ένωσης: Συγκλίσεις και Διαφοροποιήσεις, στο **Κυριόπουλος Γ., Φιλαλήθης Τ., (επιμ.),** Η πρωτοβάθμια Φροντίδα Υγείας στην Ελλάδα, Θεμέλιο, Αθήνα.
- Αγραφιώτης Δ. (2003),** Υγεία, Αρρώστια, Κοινωνία. Τόποι και τρόποι σύμπλεξης, εκδόσεις Τυπωθήτω-Δαρδανός, Αθήνα.
- Βασδέκης Β. :** Πανεπιστημιακές Σημειώσεις στις Επαναλαμβανόμενες Μετρήσεις, Τμήμα Στατιστικής, Οικονομικό Πανεπιστήμιο Αθηνών
- Γείτονα Μ., Κυριόπουλος Γ., Καραλής Γ. ,(1997).** Η Άτυπη Συμμετοχή στο Κόστος των Υπηρεσιών Υγείας από Ίδιες Πληρωμές και η Παραοικονομική δραστηριότητα, Κυριόπουλος Γ., Σισσούρας Α., (επιμ.) , Ενιαίος Φορέας Υγείας: Αναγκαιότητα ή Αυταπάτη, Θεμέλιο, Αθήνα.
- Κοντοζαμάνης Β. (2000):** Μεθοδολογικά προβλήματα στον υπολογισμό της φαρμακευτικής δαπάνης στην Ελλάδα. Ίδρυμα Οικονομικών & Βιομηχανικών Ερευνών- Τμήμα Οικονομικών της Υγείας. Άρθρο στο Ιατρικό Βήμα
- Κωλέτση-Κουνάρη Χ.,(1992).** Η Διανομή των Οδοντιατρικών Φροντίδων, στο **Κωλέτση-Κουνάρη Χ., Κυριόπουλος Γ., (επιμ. )** , Τα Οικονομικά της Οδοντιατρικής Φροντίδας στην Ελλάδα, Κέντρο Κοινωνικών Επιστημών της Υγείας & Ελληνική Εταιρεία Κοινωνικής Οδοντιατρικής, Αθήνα.
- Κυριόπουλος Γ. και Σουλιώτης (2002),** ‘Οι Δαπάνες Υγείας στην Ελλάδα: Μεθοδολογικά προβλήματα στη μέτρηση και συνέπειες για τις πολιτικές υγείας’ Εκδόσεις Παπαζήση, Αθήνα.
- Καλοκαιρινού Α. Κοδέλλα Κ, Τολικά Φ. (2003),** Θεωρητικές προσεγγίσεις της

Διαπολιτισμικής Νοσηλευτικής. Νοσηλευτική 2003, 42 (4): 451-456.

**Κυριόπουλος Γ., (1985).** Κοινωνιολογία-Ψυχολογία, Λίτσας, Αθήνα.

**Λεμονίδου Χ, Πατηράκη- Κουρμπάνη Ε.,(2002),** Θεμελιώδεις Αρχές της Νοσηλευτικής: Η Επιστήμη και η Τέχνη της Νοσηλευτικής Φροντίδας. Τόμος Ι, Ιατρικές Εκδόσεις Πασχαλίδη, Αθήνα.

**Μαρούδιας Ν.,(2001).** Ανθρώπινοι Πόροι, Ιατρικό Προσωπικό στις Υπηρεσίες Υγείας στην Ελλάδα, Πραγματικότητα και Προοπτικές, Υπουργείο Υγείας Κύπρου & ΕΣΔΥ, Κύπρος .

**Μοραΐτης Ε., Γεωργούση Ε., Ζηλίδης Χ., Θεοδώρου Μ., Πολύζος Ν. , (1995).** Μελέτη για την Οργάνωση και Λειτουργία Ολοκληρωμένου Συστήματος Ιατρικής Φροντίδας, Υπουργείο Υγείας και Πρόνοιας, Αθήνα

**Μπούτσικας Μ.Υ. (2004),** Σημειώσεις μαθήματος «Στατιστικά προγράμματα»: Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς

**Νέττλετον Σ. (2002),** Κοινωνιολογία της υγείας και της ασθένειας, Τυπωθήτω-Δαρδανός, Αθήνα.

**Ξεκαλάκη Ε. (2001).** Μη Παραμετρική Στατιστική. Αθήνα

**Παυλόπουλος Βασίλης (2008):** Εισαγωγή στην Ανάλυση Συνδιακύμανσης (Analysis of Covariance, ANCOVA). Τομέας Ψυχολογίας, Πανεπιστήμιο Αθηνών

**Ποταμιάνος Γρ. (1995),** Δοκίμια στην Ψυχολογία της υγείας, εκδόσεις ΕΛΛΗΝΙΚΑ ΓΡΑΜΜΑΤΑ, Αθήνα.

**Πουρναράς Ν. Τσόμπανου Μ, Σουρτζή Π., (2004)** Εκπαιδευτικές ανάγκες φοιτητών Νοσηλευτικής στα πλαίσια της Διαπολιτισμικής Φροντίδας Υγείας. Νοσηλευτική, 43 (2): 185-194.

Στατιστική μελέτη του Επίκουρου Καθηγητή Προπονητικής Τμήμα Επιστήμης Φυσικής Αγωγής και Αθλητισμού του Δημοκρίτειου Πανεπιστήμιο Θράκης Αντώνιο Καμπά

<http://www.sciencetech.gr/arxeia/ARTHRA/EXERCISE/FYSIKH%20DRASTHRIOTHTA%20KAI%20KATHISTIKH.pdf>

Στοιχεία από το Γενικό Νοσοκομείο Κατερίνης <http://www.gnkaterini.gr/home/images/1vathmia.doc>

Σχολή Διοίκησης Δημόσιας Υγείας [http://www.nsph.gr/files/011\\_Ygeias\\_Paidiou/Epidimiologiki\\_epitirisi\\_mathimata/Digmatolipsia.pdf](http://www.nsph.gr/files/011_Ygeias_Paidiou/Epidimiologiki_epitirisi_mathimata/Digmatolipsia.pdf)

**Τούντας Γ. και συν. (2007):** Η υγεία του Ελληνικού πληθυσμού. Κέντρο Μελετών Υπηρεσιών Υγείας. Εργαστήριο Υγιεινής και Επιδημιολογία. Ιατρική Σχολή Πανεπιστημίου Αθηνών.

**Τριχόπουλος Δ. , (1982).** Επιδημιολογία , Αρχές, Μέθοδοι, Εφαρμογές, Παρισιάνος, Αθήνα.

**Φωκιανός Κωνσταντίνος (2011).** Σημειώσεις Διαλέξεων στην Στατιστική. Τμήμα Μαθηματικών και Στατιστικής, Πανεπιστήμιο Κύπρου

**Χάλκος Γεώργιος Εμμ.,** Οικονομετρία, Θεωρία και Πράξη, Οδηγίες χρήσης σε ENIEWS, MINITAB, SPSS & EXCEL, Β. Γκιούρδας Εκδοτική, Αθήνα, 2006.

## ΠΑΡΑΡΤΗΜΑ 1 ΠΙΝΑΚΕΣ

### Πίνακας 1

Πίνακας απεικόνισης για διάφορες τιμές που παίρνουν τα  $\alpha$  και η τα  $\alpha$ -ποσοστιαία σημεία της  $X_n^2$ .

<b>n</b>	<b><math>\alpha=0.995</math></b>	<b><math>\alpha=0.99</math></b>	<b><math>\alpha=0.975</math></b>	<b><math>\alpha=0.95</math></b>	<b><math>\alpha= 0.05</math></b>	<b><math>\alpha=0.025</math></b>	<b><math>\alpha= 0.01</math></b>	<b><math>\alpha=0.005</math></b>
<b>1</b>	0.000	0.000	0.001	0.004	3.841	5.024	6.635	7.879
<b>2</b>	0.010	0.020	0.051	0.103	5.991	7.378	9.210	10.597
<b>3</b>	0.072	0.115	0.216	0.352	7.815	9.348	11.345	12.838
<b>4</b>	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
<b>5</b>	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
<b>6</b>	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
<b>7</b>	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
<b>8</b>	1.344	1.647	2.180	2.733	15.507	17.535	20.090	21.955
<b>9</b>	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
<b>10</b>	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
<b>11</b>	2.603	3.053	3.816	4.575	19.675	21,92	24.725	26.757
<b>12</b>	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
<b>13</b>	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
<b>14</b>	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
<b>15</b>	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
<b>16</b>	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
<b>17</b>	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
<b>18</b>	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
<b>19</b>	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
<b>20</b>	7.434	8.260	9.591	10.851	31.414	34.170	37.566	39.997
<b>21</b>	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
<b>22</b>	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
<b>23</b>	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
<b>24</b>	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
<b>25</b>	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
<b>26</b>	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
<b>27</b>	11.808	12.878	14.573	16.151	40.113	43.194	46.963	49.645
<b>28</b>	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.994
<b>29</b>	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.335
<b>30</b>	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672
<b>40</b>	20.706	22.164	244.331	26.509	55.756	59.342	63.691	66.766
<b>50</b>	27.991	29.708	323.574	34.764	67.505	71.420	76.154	79.490
<b>60</b>	35.535	37.485	404.817	43.188	79.082	83.298	88.379	91.952
<b>70</b>	43.275	45.442	487.576	51.739	90.531	95.023	100.425	104.215
<b>80</b>	51.172	53.540	571.532	60.392	101.879	106.629	112.329	116.321
<b>90</b>	59.196	61.754	656.466	69.126	113.145	118.136	124.116	128.299
<b>100</b>	67.328	70.065	742.219	77.930	124.342	129.561	135.807	140.169

## Πίνακας 2

Στον παρακάτω πίνακα βλέπουμε για διάφορες τιμές που παίρνουν τα  $\alpha$  και η τα  $\alpha$ -ποσοστιαία σημεία της  $t_n$ .

<b>n</b>	<b><math>\alpha=0.1</math></b>	<b><math>\alpha=0.05</math></b>	<b><math>\alpha=0.025</math></b>	<b><math>\alpha=0.01</math></b>	<b><math>\alpha=0.005</math></b>
<b>2</b>	1.886	2.920	4.303	6.965	9.925
<b>3</b>	1.638	2.353	3.182	4.541	5.841
<b>4</b>	1.533	2.132	2.776	3.747	4.604
<b>5</b>	1.476	2.015	2.571	3.365	4.032
<b>6</b>	1.440	1.943	2.447	3.143	3.707
<b>7</b>	1.415	1.895	2.365	2.998	3.499
<b>8</b>	1.397	1.860	2.306	2.896	3.355
<b>9</b>	1.383	1.833	2.262	2.821	3.250
<b>10</b>	1.372	1.812	2.228	2.764	3.169
<b>11</b>	1.363	1.796	2.201	2.718	3.106
<b>12</b>	1.356	1.782	2.179	2.681	3.055
<b>13</b>	1.350	1.771	2.160	2.650	3.012
<b>14</b>	1.345	1.761	2.145	2.624	2.977
<b>15</b>	1.341	1.753	2.131	2.602	2.947
<b>16</b>	1.337	1.746	2.120	2.583	2.921
<b>17</b>	1.333	1.740	2.110	2.567	2.898
<b>18</b>	1.330	1.734	2.101	2.552	2.878
<b>19</b>	1.328	1.729	2.093	2.539	2.861
<b>20</b>	1.325	1.725	2.086	2.528	2.845
<b>21</b>	1.323	1.721	2.080	2.518	2.831
<b>22</b>	1.321	1.717	2.074	2.508	2.819
<b>23</b>	1.319	1.714	2.069	2.500	2.807
<b>24</b>	1.318	1.711	2.064	2.492	2.797
<b>25</b>	1.316	1.708	2.060	2.485	2.787
$\infty$	1.282	1.645	1.960	2.326	2.576

**Πίνακας 3**

**Πίνακας α-ποσοστιαίων σημείων της κατανομής F με n και m βαθμούς ελευθερίας για α=0,05**

m\n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	INF
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18,5	19	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,7	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6	5,96	5,91	5,86	5,8	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,5	4,46	4,43	4,4	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,1	4,06	4	3,94	3,87	3,84	3,81	3,77	3,74	3,7	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,3	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,5	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,9	2,86	2,83	2,79	2,75	2,71
10	4,96	4,1	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,7	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,2	3,09	3,01	2,95	2,9	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,4
12	4,75	3,89	3,49	3,26	3,11	3	2,91	2,85	2,8	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,3
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,6	2,53	2,46	2,42	2,38	2,34	2,3	2,25	2,21
14	4,6	3,74	3,34	3,11	2,96	2,85	2,76	2,7	2,65	2,6	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,9	2,79	2,71	2,64	2,59	2,54	2,48	2,4	2,33	2,29	2,25	2,2	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,2	2,96	2,81	2,7	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,1	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,9	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,1	2,87	2,71	2,6	2,51	2,45	2,39	2,35	2,28	2,2	2,12	2,08	2,04	1,99	1,95	1,9	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,1	2,05	2,01	1,96	1,92	1,87	1,81
22	4,3	3,44	3,05	2,82	2,66	2,55	2,46	2,4	2,34	2,3	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,8	2,64	2,53	2,44	2,37	2,32	2,27	2,2	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,4	3,01	2,78	2,62	2,51	2,42	2,36	2,3	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,6	2,49	2,4	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,9	1,85	1,8	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,2	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,2	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,7	2,55	2,43	2,35	2,28	2,22	2,18	2,1	2,03	1,94	1,9	1,85	1,81	1,75	1,7	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4	3,15	2,76	2,53	2,37	2,25	2,17	2,1	2,04	1,99	1,92	1,84	1,75	1,7	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,5	1,43	1,35	1,25
inf	3,84	3	2,6	2,37	2,21	2,1	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1

**Πίνακας 4**

**Πίνακας α-ποσοστιαίων σημείων της κατανομής F με n και m βαθμούς ελευθερίας για α=0,01**

m\n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	INF
1	4.052	5.000	5.403	5.625	5.764	5.859	5.928	5.981	6.022	6.056	6.106	6.157	6.209	6.235	6.261	6.287	6.313	6.339	6.366
2	98,5	99	99,17	99,25	99,3	99,33	99,36	99,37	99,39	99,4	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49	99,5
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,6	26,51	26,41	26,32	26,22	26,13
4	21,2	18	16,69	15,98	15,52	15,21	14,98	14,8	14,66	14,55	14,37	14,2	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,2	9,11	9,02
6	13,75	10,93	9,78	9,15	8,75	8,47	8,26	8,1	7,98	7,87	7,72	7,56	7,4	7,31	7,23	7,14	7,06	6,97	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,2	5,12	5,03	4,95	4,86
9	10,56	8,02	6,99	6,42	6,06	5,8	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,4	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,2	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,4	4,25	4,1	4,02	3,94	3,86	3,78	3,69	3,6
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,5	4,39	4,3	4,16	4,01	3,86	3,78	3,7	3,62	3,54	3,45	3,36
13	9,07	6,7	5,74	5,21	4,86	4,62	4,44	4,3	4,19	4,1	3,96	3,82	3,67	3,59	3,51	3,43	3,34	3,26	3,17
14	8,86	6,52	5,56	5,04	4,7	4,46	4,28	4,14	4,03	3,94	3,8	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4	3,9	3,81	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	8,53	6,23	5,29	4,77	4,44	4,2	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,1	3,02	2,93	2,85	2,75
17	8,4	6,11	5,19	4,67	4,34	4,1	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3	2,92	2,84	2,75	2,65
18	8,29	6,01	5,09	4,58	4,25	4,02	3,84	3,71	3,6	3,51	3,37	3,23	3,08	3	2,92	2,84	2,75	2,66	2,57
19	8,19	5,93	5,01	4,5	4,17	3,94	3,77	3,63	3,52	3,43	3,3	3,15	3	2,93	2,84	2,76	2,67	2,58	2,49
20	8,1	5,85	4,94	4,43	4,1	3,87	3,7	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,7	2,61	2,52	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,4	3,31	3,17	3,03	2,88	2,8	2,72	2,64	2,55	2,46	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,5	2,4	2,31
23	7,88	5,66	4,77	4,26	3,94	3,71	3,54	3,41	3,3	3,21	3,07	2,93	2,78	2,7	2,62	2,54	2,45	2,35	2,26
24	7,82	5,61	4,72	4,22	3,9	3,67	3,5	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,4	2,31	2,21
25	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,7	2,62	2,54	2,45	2,36	2,27	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,82	2,66	2,59	2,5	2,42	2,33	2,23	2,13
27	7,68	5,49	4,6	4,11	3,79	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,2	2,1
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,9	2,75	2,6	2,52	2,44	2,35	2,26	2,17	2,06
29	7,6	5,42	4,54	4,05	3,73	3,5	3,33	3,2	3,09	3,01	2,87	2,73	2,57	2,5	2,41	2,33	2,23	2,14	2,03
30	7,56	5,39	4,51	4,02	3,7	3,47	3,3	3,17	3,07	2,98	2,84	2,7	2,55	2,47	2,39	2,3	2,21	2,11	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,8	2,67	2,52	2,37	2,29	2,2	2,11	2,02	1,92	1,81
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,5	2,35	2,2	2,12	2,03	1,94	1,84	1,73	1,6
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,04	1,95	1,86	1,76	1,66	1,53	1,38
inf	6,64	4,61	3,78	3,32	3,02	2,8	2,64	2,51	2,41	2,32	2,19	2,04	1,88	1,79	1,7	1,59	1,47	1,33	1