ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Σχολή Χρηματοοικονομικής και Στατιστικής

Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ

ΣΤΗΝ **ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

ΣΤΑΤΙΣΤΙΚΑ ΜΕΤΡΑ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟΙ ΔΕΙΚΤΕΣ

Αγγελική Μαρία Ι. Οικονόμου

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και

Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς

ως μέρος των απαιτήσεων για την απόκτηση του

Μεταπτυχιακού Διπλώματος Ειδίκευσης στην

Εφαρμοσμένη Στατιστική

Πειραιάς

Ιούλιος 2022

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

Σχολή Χρηματοοικονομικής και Στατιστικής

Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ

ΣΤΗΝ **ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

ΣΤΑΤΙΣΤΙΚΑ ΜΕΤΡΑ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟΙ ΔΕΙΚΤΕΣ

Αγγελική Μαρία Ι. Οικονόμου

Διπλωματική Εργασία

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

Πειραιάς

Ιούλιος 2022

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

Τα μέλη της Επιτροπής ήταν:

- Ψαρράκος Γεώργιος (Επιβλέπων)

-Πολίτης Κωνσταντίνος

-Τζαβελάς Γεώργιος

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

UNIVERSITY OF PIRAEUS

School of Finance and Statistics

Department of Statistics and Insurance Science

POSTGRADUATE PROGRAM IN

APPLIED STATISTICS

STATISTICAL INFORMATION MEASURES AND FINANCIAL INDICES

By

Angeliki Maria I. Oikonomou

MSc Dissertation

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

Piraeus, Greece

July 2022

Στην αδερφή μου

Ματίνα

Ευχαριστίες

Ευχαριστώ πολύ τον επιβλέποντα αναπληρωτή καθηγητή κύριο Ψαρράκο Γεώργιο για την καθοδήγηση του σε όλη τη διάρκεια της εκπόνησης της παρούσας διπλωματικής εργασίας, όπως επίσης για τις εύστοχες υποδείξεις που μου έκανε για να επιτύχω το καλύτερο δυνατό αποτέλεσμα. Επιπλέον θα ήθελα να ευχαριστήσω και τους καθηγητές κύριο Τζαβελά Γεώργιο και κύριο Πολίτη Κωνσταντίνο για την πολύτιμη βοήθειά τους.

Περίληψη

Στην εργασία αυτή μελετώνται διεξοδικά δυο στατιστικά μέτρα πληροφορίας, η αθροιστική υπολειπόμενη εντροπία (CRE) και η αθροιστική παρελθοντική εντροπία (CPE) .Στο πεδίο της χρηματοοικονομικής θεωρίας, τα μέτρα αυτά παίζουν σημαντικό ρόλο στην εύρεση των δεικτών Gini και Bonferroni.Οι δείκτες αυτοί χρησιμοποιούνται για την μέτρηση της οικονομικής ανισότητας σε μια κοινωνία. Σκοπός λοιπόν της εργασίας είναι αρχικά η μελέτη των ιδιοτήτων και των εκτιμήσεων των CPE και CRE και στη συνέχεια η μελέτη των δεικτών που αναφέρθηκαν παραπάνω.

Abstract

In this dissertation two statistical measures of information are thoroughly studied, the cumulative residual entropy (CRE) and the cumulative past entropy (CPE).In the field of financial theory, these measures play an important role in finding the Gini and Bonferroni indices. These indicators are used to measure economic inequality in society. Therefore, the purpose of this paper is initially to study the properties and estimates of CPE and CRE and then to study the indicators mentioned above.

ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη

Abstract

1.Εντροπία

2. Δείκτης Gini

ΚΕΦΑΛΑΙΟ 1
Εντροπία

1.1 Σύντομη ιστορική αναδρομή

Στην παράγραφο αυτή θα αναφερθούμε στην έννοια της πληροφορίας και στη συνέχεια στην έννοια της εντροπίας. Πληροφορία είναι το στοιχείο γνώσης που μετουσιώνει και δίνει αξία στα πράγματα τα οποία αποκτούν έτσι σημασία. Στις αρχές με μέσα του 20ου αιώνα, η πληροφορία ήταν έννοια αφηρημένη και ποιοτική. Ωστόσο, η ραγδαία ανάπτυξη της μετάδοσης πληροφορίας οδήγησε στην αντίληψη και μελέτη της πληροφορίας υπό ένα νέο πρίσμα, αυτό της ποσοτικής και μαθηματικής της υπόστασης. Στην σημερινή κοινωνία, την κοινωνία της πληροφορίας, η κωδικοποίηση και η μετάδοση πληροφοριών διαδραματίζουν κεντρικό ρόλο στην κατανόηση και την περιγραφή της συμπεριφοράς των βιολογικών και μηχανικών συστημάτων αλλά και της χρηματοοικονομικής αγοράς. Ο τομέας της Χρηματοοικονομικής επιστήμης έχει ως κύριο αντικείμενο την ανάλυση επενδύσεων, τη διαχείριση χαρτοφυλακίων και κεφαλαίων, την επιχειρηματική στρατηγική, την επιχειρησιακή οργάνωση αλλά και την λογιστική. Ας δούμε λοιπόν ένα παράδειγμα:

- Ο επενδυτής χρειάζεται πληροφορίες από τον χρηματιστή (Συμβουλή).
- Ο χρηματιστής τις δίνει βασιζόμενος σε πληροφορίες που κατέχει (Γνώση).
- Ο χρηματιστής πληροφορεί τον επενδυτή για τις μετοχές μιας εταιρείας (Ενημέρωση).
- Ο επενδυτής αξιολογεί τις πληροφορίες για να πουλήσει / αγοράσει (Δεδομένα).

Η πληροφορία είναι **Μέτρο της αβεβαιότητας.**

- Ο επενδυτής μειώνει την αβεβαιότητα του συμβουλευόμενος τον χρηματιστή.
- Ο χρηματιστής γνωρίζει με βεβαιότητα τα χρηματιστηριακά θέματα .

Η ανάγκη για την ποσοτικοποίηση της αβεβαιότητας που διέπει ένα γεγονός οδήγησε σε χρήση μέτρων εντροπίας, δανεισμένα από τον τομέα της Θεωρίας Πληροφορίας. Η παρούσα διπλωματική εργασία θα μελετήσει την αθροιστική παρελθοντική εντροπία και την αθροιστική υπολειπόμενη εντροπία ως μέτρο πληροφορίας ή μέτρο αβεβαιότητας στην Χρηματοοικονομική επιστήμη. Η εντροπία στη θεωρία πληροφορίας είναι το «μέτρο αβεβαιότητας» που διακατέχει ένα σύστημα. Ο όρος εντροπία χρησιμοποιήθηκε αρχικά στη θερμοδυναμική. Στη θεωρία πληροφορίας εισήχθη από τον Shannon το 1948 και γι' αυτό ονομάζεται και *εντροπία του Shannon*. Η χρήση του ίδιου όρου με τη θερμοδυναμική εντροπία, παρότι μπορεί να προκαλέσει σύγχυση, υιοθετήθηκε από τον Shannon μετά και από παρότρυνση ενός άλλου σπουδαίου μαθηματικού, του Von Neumann, ο οποίος φέρεται ότι είχε πει στον Shannon:

«Πρέπει να το ονομάσεις εντροπία για δύο λόγους: Πρώτον, η συνάρτηση αυτή χρησιμοποιείται ήδη στη θερμοδυναμική με το ίδιο όνομα. Δεύτερο, και σημαντικότερο, ο περισσότερος κόσμος δεν γνωρίζει τι πραγματικά είναι η εντροπία, και αν χρησιμοποιείς τον όρο εντροπία σε ένα αντεπιχείρημα θα κερδίζεις πάντα…»

1.2 Η εντροπία του Shannon

Σε αυτήν την παράγραφο θα ασχοληθούμε με την εντροπία του Shannon. Αρχικά θα ορίσουμε την παραπάνω εντροπία τόσο για τις διακριτές όσο και τις συνεχείς κατανομές. Στη συνέχεια θα σχολιάσουμε κάποια μειονεκτήματα της διαφορικής εντροπίας του Shannon.

Ορισμός 1.1

Έστω X μια τυχαία μεταβλητή, διακριτή με συνάρτηση πιθανότητας ή συνεχής με συνάρτηση πυκνότητας πιθανότητας $f(x)$. Η αναμενόμενη αβεβαιότητα, ή εντροπία (entropy) της X (συμβολικά $H(X)$), όπως ορίστηκε από τον Shannon (1948) δίνεται από τις σχέσεις:

$$H(X) = -\sum_i p_i \, ln p_i, \qquad (1.1)$$

όπου υποθέτουμε ότι η X είναι μια τυχαία μεταβλητή, διακριτή με συνάρτηση πιθανότητας $p_i = P(X = i)$.

Αν η X είναι συνεχής τυχαία μεταβλητή η εντροπία δίνεται από τον τύπο:

$$H(X) = -\int_0^{\infty} f(x) ln f(x) dx \,. \tag{1.2}$$

Η περίπτωση της συνεχούς εντροπίας αναφέρεται επίσης και ως διαφορική εντροπία. Η συνάρτηση ολοκλήρωσης στο δεξί μέλος της σχέσης (1.2) εξαρτάται μόνο από το x μόνο μέσω του $f(x)$ καθιστώντας έτσι το $H(X)$ ανεξάρτητο από τη μετατόπιση. Με άλλα λόγια το $H(X)$ είναι χωρίς θέση, με την έννοια ότι η τυχαία μεταβλητή X έχει την ίδια διαφορική εντροπία με την $X + b$, για κάθε $b > 0$.

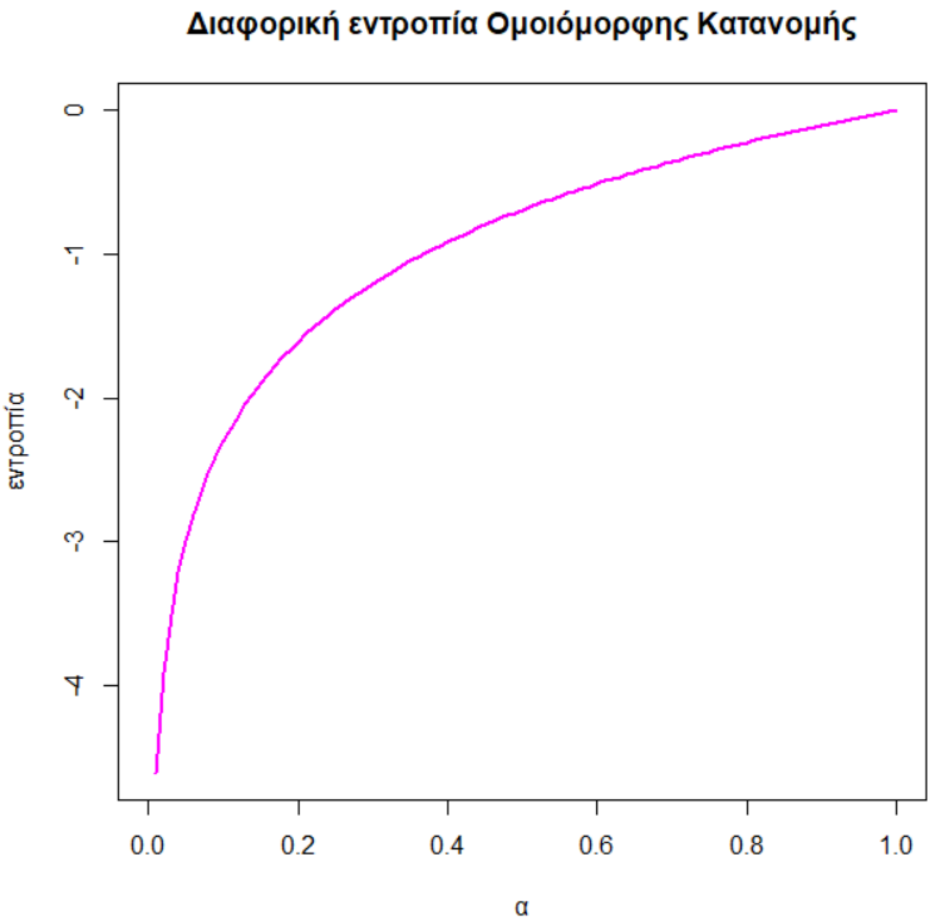
Παρακάτω θα μελετήσουμε κάποιες σημαντικές ιδιότητες της εντροπίας αυτής αλλά και κάποια μειονεκτήματά της. Για διακριτές κατανομές η εντροπία ως μέτρο πληροφορίας έχει ορισμένες σημαντικές ιδιότητες. Η εντροπία είναι ίση με το μηδέν αν και μόνο αν η τυχαία μεταβλητή X είναι σταθερή. Σε κάθε άλλη περίπτωση είναι θετική. Το μέτρο αυτό ποσοτικοποιεί την αναμενόμενη αβεβαιότητα που σχετίζεται με το αποτέλεσμα ενός πειράματος τύχης, ή ισοδύναμα παρέχει πληροφορία για την προβλεψιμότητα ενός πιθανού αποτελέσματος του δειγματικού χώρου $\boldsymbol{X}$.

Παρόλα αυτά η διαφορική εντροπία του Shannon έχει ορισμένα μειονεκτήματα για τις συνεχείς κατανομές (Schroeder 2004). Αρχικά ορίζεται μόνο για κατανομές με πυκνότητα, ενώ δεν μπορεί να υπάρξει ορισμός της εντροπίας για μια μικτή κατανομή. Ένα άλλο σημαντικό πρόβλημα της είναι ότι είναι ''ασυνεπής'' , με την έννοια ότι μπορεί να πάρει και αρνητικές τιμές. Αυτό συμβαίνει στην περίπτωση της ομοιόμορφης κατανομής. Ας θεωρήσουμε ότι η τυχαία μεταβλητή X ακολουθεί την ομοιόμορφη κατανομή στο $(0, \alpha)$, με συνάρτηση πυκνότητας πιθανότητας $f(x) = \frac{1}{\alpha}$ για $0 < x < \alpha$.Τότε από την (1.2) προκύπτει ότι η διαφορική εντροπία της $U(0, \alpha)$ ισούται με

$H(X) = -\int_0^{\alpha} \frac{1}{\alpha} \cdot ln\frac{1}{\alpha}\, dx = ln\alpha.$ Παρατηρούμε ότι ανάλογα με την τιμή της παραμέτρου α έχουμε 3 περιπτώσεις. Πιο συγκεκριμένα:

- για $\alpha = 1$, η $H(X)$ ισούται με το μηδέν , δηλαδή $H(X)=0$,
- για $\alpha < 1$, η $H(X)$ παίρνει αρνητικές τιμές , δηλαδή $H(X)<0$,
- και τέλος όταν $\alpha > 1$, η $H(X)$ παίρνει θετικές τιμές, δηλαδή $H(X)>0$.

Παρακάτω δίνουμε την σχηματική απεικόνιση της παραπάνω σχέσης για $0 < \alpha < 1$:



Σχήμα 1.1 : Διαφορική εντροπία της Ομοιόμορφης κατανομής για $0 < \alpha < 1$.

Τέλος, αξίζει να αναφέρουμε ότι στην βιβλιογραφία έχουν παρουσιαστεί αρκετές εναλλακτικές για την εντροπία μιας συνεχούς κατανομής. Πιο συγκεκριμένα ο Schroeder (2004) πρότεινε ένα μέτρο που σε αντίθεση με την εντροπία μπορεί εύκολα να επεκταθεί στις συνεχείς κατανομές και σε αντίθεση με την διαφορική εντροπία είναι πάντα θετική και αμετάβλητη σε σχέση με τους γραμμικούς συνδυασμούς των μεταβλητών. Άλλα μέτρα αβεβαιότητας ως κατάλληλες γενικεύσεις ή τροποποιήσεις της κλασικής εντροπίας έχουν προταθεί , όπως η σταθμισμένη εντροπία (Di Crescenzo and Longobardi, 2006), η οποία ορίζεται ως:

$$H^{w}(X) = -E[X \ln f(X)] = -\int_0^{\infty} x f(x) \ln f(x) dx, \tag{1.3}$$

και είναι ένα μέτρο πληροφοριών που εκχωρεί μεγαλύτερα βάρη σε μεγαλύτερες τιμές του X.

1.3 Αθροιστική υπολειπόμενη εντροπία (CRE)

Σε αυτήν την παράγραφο θα ορίσουμε την αθροιστική υπολειπόμενη εντροπία και στη συνέχεια θα δούμε πως υπολογίζεται για κάποιες γνωστές κατανομές.

Προκειμένου να ξεπεραστούν τα παραπάνω προβλήματα , οι Rao et al.(2004) όρισαν ένα νέο μέτρο πληροφορίας που ονομάζεται αθροιστική υπολειπόμενη εντροπία (Cumulative Residual Entropy-CRE) το οποίο συμβολίζεται με $\mathcal{E}(X)$

$$\mathcal{E}(X) = - \int_0^\infty \bar{F}(x) \ln \bar{F}(x)\, dx, \qquad (1.4)$$

όπου $\bar{F}(x) = P(X > x)$ η συνάρτηση επιβίωσης της τυχαίας μεταβλητής X. Ας δούμε την CRE για κάποιες γνωστές κατανομές.

1.Κατανομή Pareto με παραμέτρους $\alpha > 1$ και $\beta > 0$.

Αν X τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha > 1, \beta > 0$ και έχει συνάρτηση επιβίωσης $\bar{F}(x) = \left(\frac{\beta}{\beta+x}\right)^{\alpha}$, $x \geq 0$, τότε με βάση τον τύπο (1.4) λαμβάνουμε:

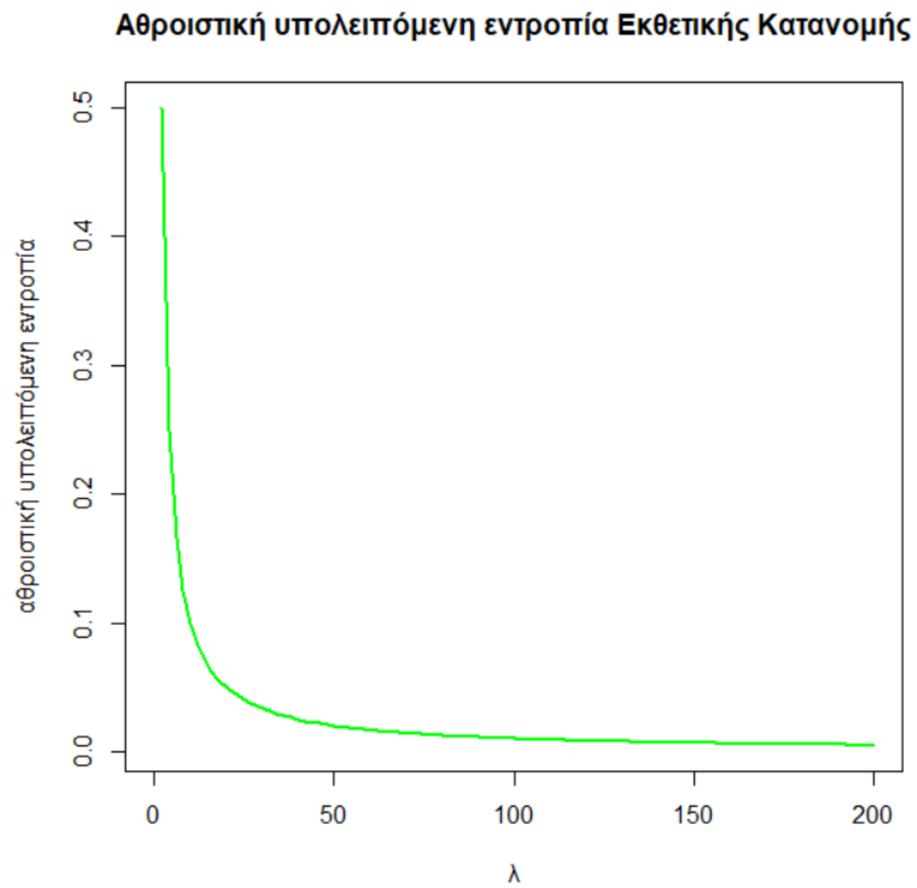
$$\mathcal{E}(X) = - \int_0^\infty \left(\frac{\beta}{\beta+x}\right)^{\alpha} \ln \left(\frac{\beta}{\beta+x}\right)^{\alpha} dx = \frac{\alpha\beta}{(\alpha-1)^2}$$

2. Εκθετική κατανομή με παράμετρο $\lambda > 0$.

Αν X τυχαία μεταβλητή η οποία ακολουθεί την εκθετική κατανομή με παράμετρο $\lambda > 0$ και έχει συνάρτηση επιβίωσης $\bar{F}(x) = e^{-\lambda x}$, $x \geq 0$, τότε με βάση τον τύπο (1.4) προκύπτει ότι :

$$\mathcal{E}(X) = - \int_0^\infty e^{-\lambda x} \ln\left(e^{-\lambda x}\right) dx = 1/\lambda.$$

Παρακάτω δίνουμε και την σχηματική απεικόνιση της παραπάνω εντροπίας για τις όλες τις τιμές της παραμέτρου λ στο διάστημα (0, 200).



Σχήμα 1.2 : Αθροιστική υπολειπόμενη εντροπία της Εκθετικής κατανομής όταν $\mathbf{0 < \lambda < 200}$.

3.Δυναμοκατανομή Power(α, β).

Αν X τυχαία μεταβλητή η οποία ακολουθεί την δυναμοκατανομή Power(α, β) με παραμέτρους $\alpha, \beta > 0$ και έχει συνάρτηση επιβίωσης $\overline{F}\,(x) = \left(\frac{\beta - x}{\beta}\right)^{\alpha}$, $0 \leq x \leq \beta$, τότε με βάση τον τύπο (1.4) προκύπτει ότι η αθροιστική υπολειπόμενη εντροπία ισούται με :

$$\mathcal{E}(X) = -\int_0^{\infty} \left(\frac{\beta - x}{\beta}\right)^{\alpha} ln\left(\left(\frac{\beta - x}{\beta}\right)^{\alpha}\right) dx = \frac{\alpha\beta}{(\alpha+1)^2}\,.$$

1.4 Αθροιστική υπολειπόμενη εντροπία και μέσος υπολειπόμενος χρόνος ζωής

Σε αυτήν την παράγραφο θα δούμε πώς συνδέεται ο μέσος υπολειπόμενος χρόνος ζωής ενός συστήματος με την αθροιστική υπολειπομένη εντροπία. Στη συνέχεια θα παρουσιάσουμε τον εύκολο τρόπο υπολογισμού της αθροιστικής υπολειπόμενης εντροπίας όταν γνωρίζουμε τον μέσο υπολειπόμενο χρόνο ζωής. Η αθροιστική υπολειπόμενη εντροπία είναι πάντα μη αρνητική και είναι ιδιαίτερα κατάλληλη για να περιγράψει τις πληροφορίες σε προβλήματα που σχετίζονται με τις ιδιότητες γήρανσης της θεωρίας αξιοπιστίας με βάση τη συνάρτηση μέσης υπολειπόμενης ζωής (Asadi and Zohrevand, 2007).

1.4.1 Ο μέσος υπολειπόμενος χρόνος ζωής

Αν υποθέσουμε ότι η τυχαία μεταβλητή X, παριστά το χρόνο ζωής μιας μονάδας , η οποία έχει επιβιώσει στο χρονικό διάστημα $[0, t]$, με $t \geq 0$, τότε η δεσμευμένη τυχαία μεταβλητή $X_t = X - t \mid X > t$, η οποία εκφράζει το χρόνο ζωής της από τη στιγμή t και έπειτα και καλείται υπολειπόμενος χρόνος ζωής της μονάδας ηλικίας t. Η μέση τιμή της X_t δίνεται από τον τύπο:

$$m(t)=E(X_t)=\int_0^{\infty} \frac{\bar{F}(x)}{\bar{F}(t)} dx . \quad (1.5)$$

Θεώρημα 1.1: Για την μη αρνητική, συνεχή τυχαία μεταβλητή X με μέσο υπολειπόμενο χρόνο ζωής $m(t)$ και αθροιστική υπολειπόμενη εντροπία $\mathcal{E}(X)$ τέτοια ώστε $\mathcal{E}(X)<\infty$, ισχύει :

$$\mathcal{E}(X)=E(m(X)) \quad (1.6)$$

Απόδειξη.

Το δεύτερο μέλος της (1.6) εκφράζει τον αναμενόμενο μέσο υπολειπόμενο χρόνο ζωής της X, δηλαδή

$$E(m(X)) = \int_0^{\infty} m(t) f(t) dt.$$

Αντικαθιστώντας την $m(t)$ από την (1.5) , λαμβάνουμε

$$E(m(X)) = \int_0^{\infty} \frac{\int_t^{+\infty} \bar{F}(x)dx}{\bar{F}(t)} f(t)dt$$

Αλλάζοντας τη σειρά ολοκλήρωσης , παίρνουμε:

$$E(m(X)) = \int_0^{\infty} \int_0^{x} \frac{f(t)}{\bar{F}(t)} dt \, \bar{F}(x) \, dx$$

$$= \int_0^{\infty} \int_0^{x} \lambda(t) \, dt \, \bar{F}(x) dx$$

$$= \int_0^{\infty} (-ln \, \bar{F}(x)) \bar{F}(x) \, dx = \mathcal{E}(X),$$

το οποίο ολοκληρώνει την απόδειξη. ■

Με τον παραπάνω τύπο (1.6), είναι σαφές ότι αν γνωρίζουμε τον μέσο υπολειπόμενο χρόνο ζωής μπορούμε να υπολογίσουμε πολύ εύκολα την αθροιστική υπολειπόμενη εντροπία. Ας δούμε λοιπόν πως μπορούμε να εφαρμόσουμε τη σχέση (1.6) για κάποιες γνωστές κατανομές.

Κατανομή Pareto με παραμέτρους $\boldsymbol{\alpha > 1, \beta > 0}$

Για την κατανομή Pareto με παραμέτρους $\alpha > 1, \beta > 0$, βρίσκουμε μέσο υπολειπόμενο χρόνο ζωής $m(t) = \frac{t+\beta}{\alpha-1}$ $, t \geq 0$. Συνεπώς με την βοήθεια της (1.4) παίρνουμε ότι η αθροιστική υπολειπόμενη εντροπία είναι:

$$\mathcal{E}(X) = E\left(\frac{X+\beta}{\alpha-1}\right) = \frac{E(X)+\beta}{\alpha-1} = \frac{\frac{\beta}{\alpha-1}+\beta}{\alpha-1} = \frac{\alpha\beta}{(\alpha-1)^2},$$

όπου $E(X)$ είναι η μέση τιμή της Pareto.

Παρατηρούμε ότι η αθροιστική υπολειπόμενη εντροπία της Pareto μπορεί να γραφτεί ως εξής:

$$\mathcal{E}(X) = \frac{\alpha}{(\alpha-1)} \cdot \frac{\beta}{\alpha-1} = \frac{\alpha}{(\alpha-1)} \cdot E(X).$$

Συνεπώς η αθροιστική υπολειπόμενη εντροπία είναι μεγαλύτερη από την μέση τιμή. Θυμίζουμε εδώ ότι η κατανομή Pareto είναι μια από τις πιο σημαντικές κατανομές , εξαιτίας της βαριάς δεξιάς ουράς που έχει. Συνάρτηση πυκνότητας πιθανότητας της Pareto(α, β) είναι η:

$$f(x) = \frac{\alpha\beta^{\alpha}}{(x+\beta)^{\alpha+1}} \ , x \ \geq 0,$$

συνάρτηση επιβίωσης είναι η:

$$\bar{F}(x) = \left(\frac{\beta}{x+\beta}\right)^{\alpha} \ , x \geq 0,$$

ενώ η μέση τιμή της ισούται με:

$$E(X)=\begin{cases} \frac{\beta}{\alpha-1} \ , \text{αν } \alpha > 1 \\ \infty \ , \text{αν } 0 < \alpha \leq 1 \end{cases}.$$

2. Εκθετική κατανομή με παράμετρο λ.

Ο μέσος υπολειπόμενος χρόνος ζωής της εκθετικής κατανομής με παράμετρο $\lambda>0$ δίνεται από την σχέση:

$$m(t)= \frac{1}{\lambda}, \ \lambda>0.$$

Τότε από το Θεώρημα 1.1 ,

$$\mathcal{E}\,(X)=E\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda}.$$

Παρατηρούμε ότι η αθροιστική υπολειπόμενη εντροπία της εκθετικής κατανομής είναι ίση με την μέση τιμή.

3. Δυναμοκατανομή Power(α, β).

Ο μέσος υπολειπόμενος χρόνος ζωής της δυναμοκατανομής δίνεται από την σχέση:

$$m(t) = \frac{\beta-t}{\alpha+1}.$$

Τότε,

$$\mathcal{E}\,(X) = E\left(\frac{\beta-t}{\alpha+1}\right) = \frac{\beta-E(X)}{\alpha+1} = \frac{\beta-\frac{\beta}{\alpha+1}}{\alpha+1} = \frac{\alpha\beta}{(\alpha+1)^2}.$$

Όπου $E(X)$ είναι η μέση τιμή της δυναμοκατανομής Power(α, β). Παρατηρούμε ότι η αθροιστική υπολειπόμενη εντροπία της δυναμοκατανομής μπορεί να γραφτεί ως εξής:

$$\mathcal{E}(X) = \frac{\alpha}{(\alpha+1)} \cdot \frac{\beta}{\alpha+1} = \frac{\alpha}{(\alpha+1)} \cdot E(X).$$

Συνεπώς η αθροιστική υπολειπόμενη εντροπία είναι μικρότερη από την μέση τιμή.

Θυμίζουμε εδώ ότι η Power(α, β) , όπου $\alpha, \beta > 0$ έχει συνάρτηση πυκνότητας την :

$$f(x) = \frac{\alpha(\beta-x)^{\alpha-1}}{\beta^{\alpha}}, \ 0< x <\beta,$$

συνάρτηση επιβίωσης την:

$$\bar{F}(x) = \left(\frac{\beta-x}{\beta}\right)^{\alpha}, 0 \leq x \leq \beta.$$

Τέλος η μέση τιμή της ισούται με:

$$E(X) = \frac{\beta}{\alpha+1}.$$

1.5 Αθροιστική παρελθοντική εντροπία (CPE)

Σε αυτήν την παράγραφο θα ορίσουμε την αθροιστική παρελθοντική εντροπία και στη συνέχεια θα δούμε πως υπολογίζεται για κάποιες γνωστές κατανομές. Οι Di Crescenzo και Longobardi (2009) χρησιμοποιώντας τη συνάρτηση κατανομής όρισαν ένα ακόμα στατιστικό μέτρο πληροφορίας, το οποίο συμβολίζουμε με $\mathcal{CE}(X)$ και ονομάζεται αθροιστική παρελθοντική εντροπία (cumulative past entropy, CPE). Πιο συγκεκριμένα, για μη αρνητική συνεχή τυχαία μεταβλητή X η CPE δίνεται από τον παρακάτω τύπο:

$$\mathcal{CE}(X) = -\int_0^{\infty} F(x) lnF(x) dx, \tag{1.7}$$

όπου $F(X)=P(X \leq x)$ η συνάρτηση κατανομής της X. Η αθροιστική παρελθοντική εντροπία παίρνει τιμές στο διάστημα [0, ∞) ,δηλαδή ισχύει $0 \leq \mathcal{CE}(X) < \infty$. Επίσης $\mathcal{CE}(X) = 0$ αν και μόνο αν η τυχαία μεταβλητή X είναι σταθερή.

Παρατήρηση 1.5: Αξίζει να τονίσουμε εδώ ότι η αθροιστική παρελθοντική εντροπία $\mathcal{CE}(X)$ και η αθροιστική υπολειπόμενη εντροπία $\mathcal{E}\,(X)$ είναι ίδιες όταν η κατανομή της τυχαία μεταβλητής X είναι συμμετρική ως προς τη μέση τιμή της μ=E(X) ,όπου μ πεπερασμένη , δηλαδή αν ισχύει $F\,(\mu + X)$= 1- $F(\mu - X)$ για κάθε $x\epsilon R$.

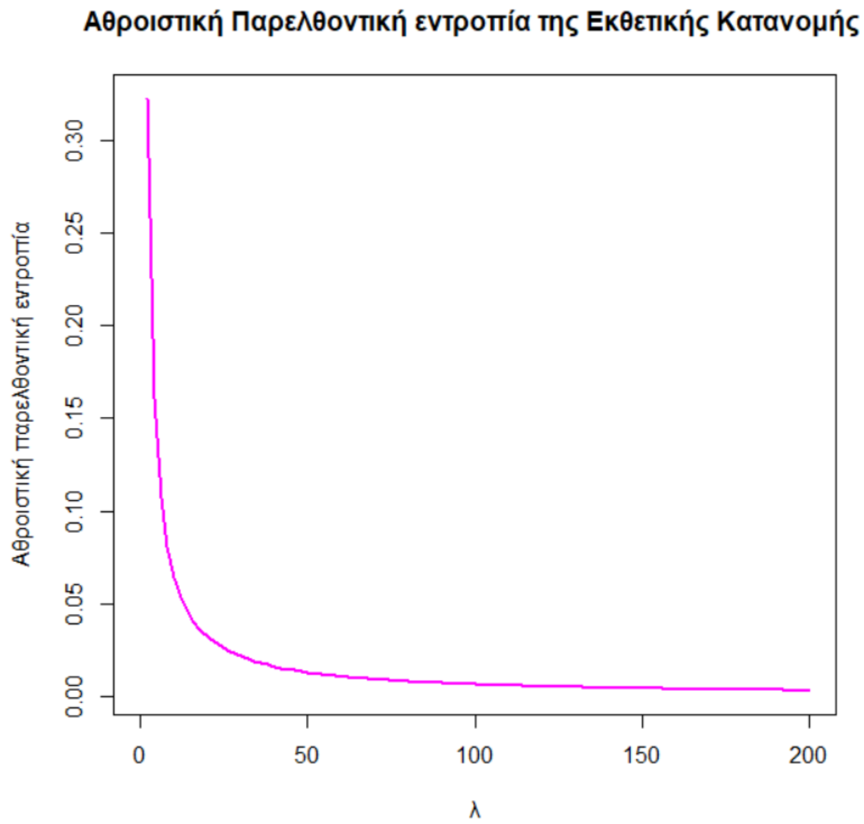
Ας δούμε την CPE για κάποιες γνωστές κατανομές.

1. Εκθετική κατανομή με παράμετρο λ>0.

Αν X τυχαία μεταβλητή η οποία ακολουθεί την εκθετική κατανομή με παράμετρο λ >0 και έχει συνάρτηση επιβίωσης $F(x)$ =1 $-$ e$^{-\lambda x}$, x≥0 , τότε με βάση τον τύπο (1.7) λαμβάνουμε ότι :

$$\mathcal{CE}(X) = - \int_0^{\infty} (1 - \mathrm{e}^{-\lambda x})\, ln\big(1 - \mathrm{e}^{-\lambda x}\big)\, dx = \frac{\pi^2 - 6}{6\lambda}.$$

Παρακάτω δίνουμε και την σχηματική απεικόνιση της παραπάνω εντροπίας για τις διάφορες τιμές της παραμέτρου λ στο διάστημα (0,100).



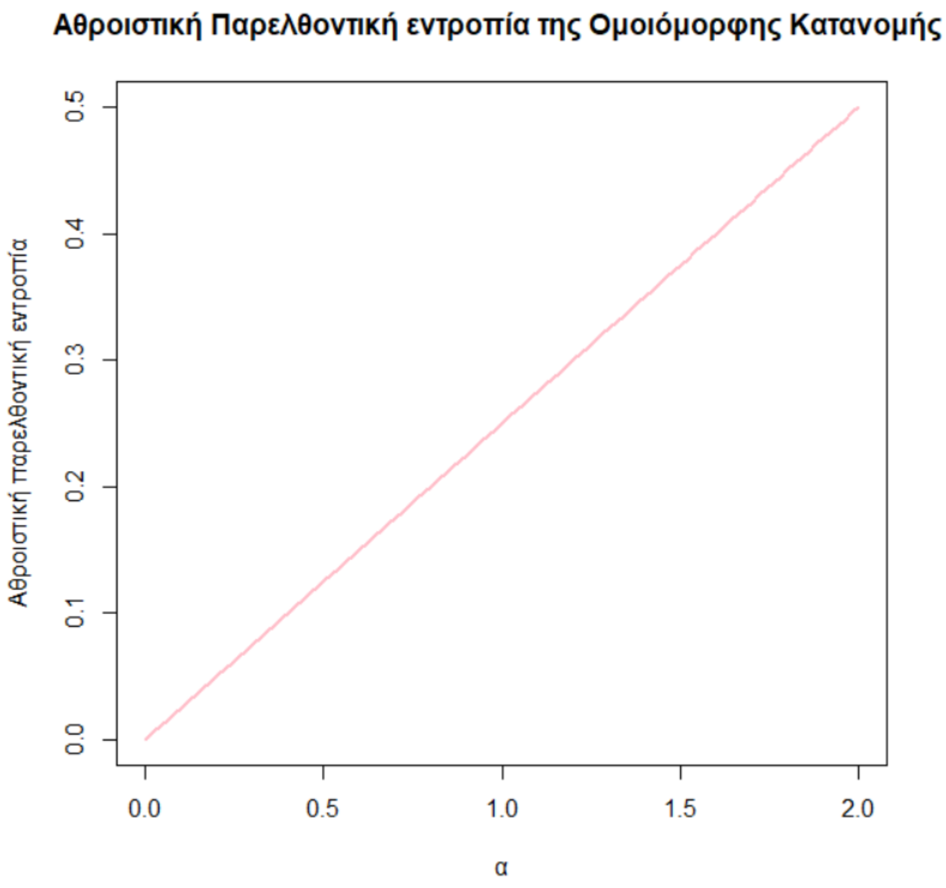
Σχήμα 1.3: Αθροιστική παρελθοντική εντροπία της Εκθετικής κατανομής όταν $\mathbf{0} < \lambda < 200$.

2. Ομοιόμορφη κατανομή $U(0, \alpha)$.

Ας θεωρήσουμε ότι η τυχαία μεταβλητή X ακολουθεί την ομοιόμορφη κατανομή στο (0, α) , με συνάρτηση πυκνότητας πιθανότητας $f(x) = \frac{1}{\alpha}$ για $0< x <\alpha$ και συνάρτηση κατανομής $F(x)=\frac{x-\alpha}{\alpha}$ τότε με βάση τον τύπο (1.7) :

$$\mathcal{CE}(X) = - \int_0^{\alpha} \left(\frac{t-\alpha}{\alpha}\right) ln\left(\frac{t-\alpha}{\alpha}\right) dt = \frac{\alpha}{4}.$$

Παρακάτω δίνουμε και την σχηματική απεικόνιση της παραπάνω εντροπίας όταν η παράμετρος α παίρνει τιμές στο (0 , 2).



Σχήμα 1.4: Αθροιστική παρελθοντική εντροπία της Ομοιόμορφης Κατανομής όταν $0 < \alpha < 2$.

Παρατηρούμε ότι η ομοιόμορφη κατανομή έχει ίδια αθροιστική παρελθοντική εντροπία και ίδια αθροιστική υπολειπόμενη εντροπία , δηλαδή $\mathcal{CE}(X)= \mathcal{E}(X)$,μιας και η ομοιόμορφη κατανομή είναι συμμετρική ως προς τη μέση της τιμή. (βλέπε Παρατήρηση 1.5)

Τέλος , αξίζει να αναφέρουμε ότι στην βιβλιογραφία έχουν παρουσιαστεί αρκετές εναλλακτικές για την αθροιστική παρελθοντική εντροπία μιας συνεχούς κατανομής. Πιο συγκεκριμένα το 2011 οι Misagh et al. όρισαν την σταθμισμένη εντροπία

$$\mathrm{CE}^{\mathrm{w}}(X) = -\int_0^{\infty} xF_{\mathrm{X}}(x)\ln F_{\mathrm{X}}(x)dx \ ,$$

και είναι ένα μέτρο πληροφοριών που εκχωρεί μεγαλύτερα βάρη σε μεγαλύτερες τιμές του X.

1.6 Ιδιότητες της αθροιστικής παρελθοντικής εντροπίας

Στην παράγραφο αυτή θα παρουσιάσουμε κάποιες χρήσιμες ιδιότητες της αθροιστικής παρελθοντικής εντροπίας. Όλες οι παρακάτω σχέσεις που θα αναφέρουμε μελετήθηκαν και αποδείχθηκαν από τους Di Crescenzo and Longobardi (2009).

Πρόταση 1.1: Αν $Y=\alpha X + b$, με $\alpha > 0$ $και$ $b \geq 0$, τότε

$$\mathcal{CE}(Y) = \alpha\mathcal{CE}(X) \tag{1.8}$$

Απόδειξη.

Από τις πιθανότητες γνωρίζουμε ότι:

$$P(Y \leq x) = P\left(X \leq \frac{x-b}{\alpha}\right) = F\left(\frac{x-b}{\alpha}\right)$$

$$\mathcal{CE}(Y) = -\int_{\mathrm{b}}^{\infty} F\left(\frac{x-b}{\alpha}\right) lnF\left(\frac{x-b}{\alpha}\right) dx \tag{1.9}$$

Θέτουμε όπου $y = \frac{x-b}{\alpha}$, έτσι αν παραγωγίσουμε παίρνουμε ότι $d\mathrm{y}=\frac{1}{\alpha}dx$.

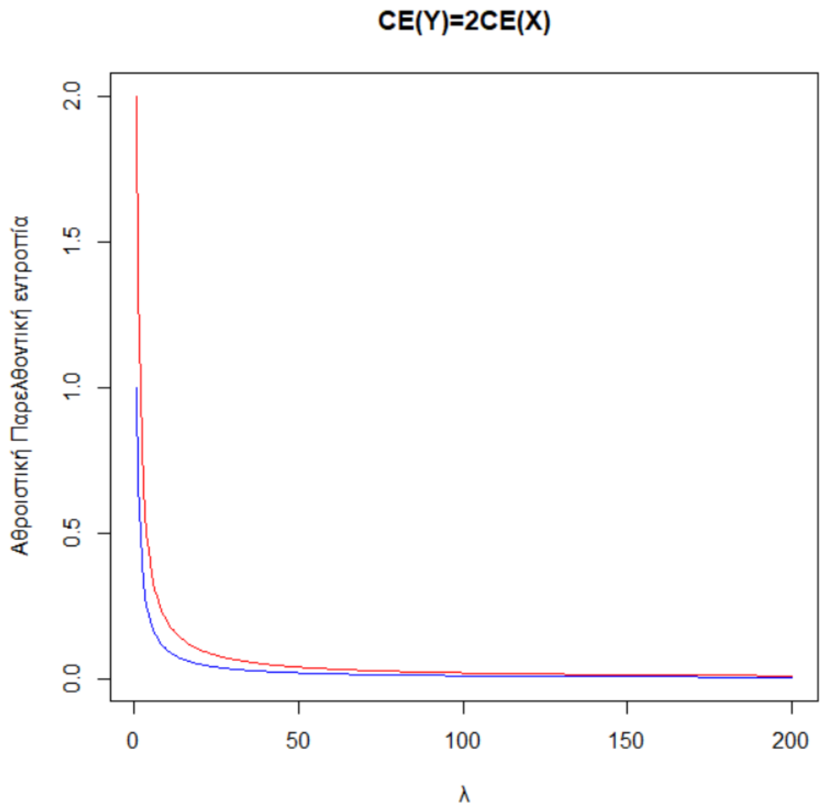
Συνεπώς η (1.9) γίνεται μέσω της $dy = \frac{1}{\alpha}dx$

$$\mathcal{CE}\,(\mathrm{Y}) = -\int_0^{\infty} F(y)lnF(y)adx$$

$$= -\alpha\int_0^{\infty} F(x)lnF(x)dx = \alpha\mathcal{CE}\,(\mathrm{x}).$$

Η απόδειξη ολοκληρώθηκε.■

Παρακάτω παραθέτουμε ένα γράφημα των τυχαίων μεταβλητών X και Y , όπου Y=2X +3 και X ~Exp(λ).



Σχήμα 1.5: Αθροιστική παρελθοντική εντροπία των τυχαίων μεταβλητών X και Y

Πρόταση 1.2: Για μη αρνητική συνεχή τυχαία μεταβλητή X με αθροιστική παρελθοντική εντροπία $\mathcal{CE}$ (X)<∞ , έχουμε ότι :

$$\mathcal{CE}(X) = E[\mathrm{T}^{(2)}(X)], \tag{1.10}$$

όπου

$$\mathrm{T}^{(2)}(x) = - \int_x^\infty ln\mathrm{F}(z)dz \;,\; x \geq 0. \tag{1.11}$$

Απόδειξη.

Από την σχέση (1.7) και αλλάζοντας τη σειρά ολοκλήρωσης (Θεώρημα Fubini) έχουμε ότι

$$\mathcal{CE}(X) = -\int_0^\infty [\int_0^x f(t)dt] lnF(x)dx = -\int_0^\infty f(t)[\int_t^\infty lnF(x)dx]dt,$$

από την οποία προκύπτει η (1.10) μέσα από την χρήση της (1.11)■

Πρόταση 1.3: Έστω X τυχαία μεταβλητή η οποία παίρνει τιμές στο $[0, b]$, με b πεπερασμένο. Τότε,

$$\mathcal{CE}(X) \leq [b - E(X)]\left|log(1 - \frac{E(X)}{b})\right| \quad (1.12)$$

Απόδειξη.

Χρησιμοποιώντας την log-sum ανισότητα (Rao ,2005) παίρνουμε ότι

$$\int_0^b F(x)\log F(x)dx \geq \int_0^b F(x)dx \log\frac{\int_0^b F(x)dx}{b}.$$

Μέσα από την σχέση (1.7) προκύπτει το ζητούμενο.■

Πρόταση 1.4: Για μη αρνητική συνεχή τυχαία μεταβλητή X με $E(X)<\infty$, τότε έχουμε :

$$\mathcal{CE}(X) \leq E(X). \quad (1.13)$$

Πρόταση 1.5 : Η αθροιστική υπολειπόμενη εντροπία ικανοποιεί την κυρτότητα ενός μέτρου κινδύνου (convexity of a risk measure)

Απόδειξη.

Για κάθε λ $\epsilon[0,1]$ και X , $Y \epsilon X$, σύμφωνα με τη υποπροσθετικότητα της αθροιστικής υπολειπόμενης εντροπίας , έχουμε:

$$\mathcal{CE}(\lambda X + (1-\lambda)Y) \leq \mathcal{CE}(\lambda X) + \mathcal{CE}((1-\lambda)Y). \quad (1.14)$$

Τώρα σύμφωνα με την θετική ομοιογένεια της αθροιστικής υπολειπόμενης εντροπίας , έχουμε:

$$\mathcal{CE}\,(\lambda X) + \mathcal{CE}\,((1\text{-}\lambda)\,Y) = \lambda\mathcal{CE}\,(X) + (1\text{-}\lambda)\mathcal{CE}\,(Y) \qquad (1.15)$$

Από τα παραπάνω προκύπτει το ζητούμενο.■

Σε αυτό το σημείο θα κάνουμε μια αναφορά στον χρηματοοικονομικό δείκτη Gini,στον οποίο θα αναφερθούμε διεξοδικά στα Κεφάλαια 3 και 4 της παρούσας εργασίας.

Ορισμός 1.2: Ένα διάσημο μέτρο της οικονομικής ανισότητας είναι ο δείκτης Gini , ο οποίος ορίζεται ως εξής για μια μη αρνητική τυχαία μεταβλητή με συνάρτηση επιβίωσης $\bar{F}(x)$:

$$\text{Gini}(X) = \frac{D_G(X)}{2E(X)}, \qquad (1.16)$$

όπου $D_G(X) = \int_0^\infty 2\bar{F}\,(x)(1-\bar{F}(x))dx$ είναι η μέση διαφορά Gini ως μέτρο διασποράς.

Πρόταση 1.6: Έστω X μη αρνητική τυχαία μεταβλητή. Τότε ισχύει ότι :

$$\mathcal{CE}\,(X) \geq \int_0^\infty F(x)\bar{\mathrm{F}}\,(\mathrm{x})\,d\mathrm{x}. \qquad (1.17)$$

Απόδειξη.

Θυμίζουμε ότι $-logu \geq 1-u$ για κάθε u>0 και από την σχέση (1.7) έχουμε το ζητούμενο.■

Πρόταση 1.7: Έστω X και Y μη αρνητικές και ανεξάρτητες τυχαίες μεταβλητές. Τότε ισχύει ότι

$$max\{\,\mathcal{CE}\,(X), \mathcal{CE}\,(Y)\} \leq \mathcal{CE}\,(X+Y). \qquad (1.18)$$

Εφαρμογή της Πρότασης 1.7:

Η παρακάτω εφαρμογή αποδείχθηκε από τους Crescenzo and Longobardi (2009).Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές, οι οποίες ακολουθούν την ομοιόμορφη κατανομή $U(0,1)$. Δηλαδή $X \sim U(0,1)$, $Y \sim U(0,1)$.Τότε έχουμε ότι η $F(x)$= 0 αν x<0. Επίσης $F(x) = \frac{x^2}{2}$, για $0 \leq x \leq 1$.Τώρα για $1 \leq x \leq 2$ παίρνουμε ότι $F(x) = 1 - \frac{(1-x)^2}{2}$.Τέλος για $x>2$ λαμβάνουμε ότι η $F(x)=1$.

Από την σχέση (1.7) βρίσκουμε ότι $\mathcal{CE}(X+Y)=0.3645$. Είδαμε στην Παράγραφο 1.5 ότι η αθροιστική παρελθοντική εντροπία της ομοιόμορφης κατανομής στο $[0, b]$ ισούται με $\mathcal{CE}(X)= \frac{b}{4}$. Εδώ b=1 συνεπώς $\mathcal{CE}(X) = \frac{1}{4}$. Τελικά παρατηρούμε ότι ικανοποιείται η σχέση $\max\{\mathcal{CE}(X), \mathcal{CE}(Y)\} \leq \mathcal{CE}(X+Y)$.

Πρόταση 1.8 : Αν $X_1, \ldots, X_n$ είναι μη αρνητικές i. i.d. τυχαίες μεταβλητές , τότε :

$$\mathcal{CE}(nX_1) \geq \mathcal{CE}(max\{X_1, \ldots, X_n\}). \tag{1.19}$$

1.7 Από κοινού αθροιστική παρελθοντική εντροπία

Σε αυτήν παράγραφο θα αναφερθούμε στην από κοινού αθροιστική παρελθοντική εντροπία μεταξύ 2 τυχαίων μεταβλητών X και Y.Το μέτρο της αθροιστικής παρελθοντικής εντροπίας μπορεί να χρησιμοποιηθεί και για τον ορισμό μετρήσεων πληροφορίας οι οποίες αναδεικνύουν τις σχέσεις μεταξύ δύο τυχαίων μεταβλητών X και Y .

Ορισμός 1.3: Η από κοινού αθροιστική παρελθοντική εντροπία για 2 τυχαίες μεταβλητές μετρά τη συνολική πληροφορία των X, Y και δίνεται ως εξής:

$$\mathcal{CE}(X,Y)= -\int_0^\infty \int_0^\infty F(x,y) ln F(x,y) dxdy. \tag{1.20}$$

Αν οι μεταβλητές X και Y είναι ανεξάρτητες τότε παίρνουμε:

$$\mathcal{CE}(X,Y)= [\int_0^\infty F_Y(y)dy]\, \mathcal{CE}(X)+[\int_0^\infty F_X(x)dx]\, \mathcal{CE}(Y). \tag{1.21}$$

Πρόταση 1.9: Αν X και Y είναι ανεξάρτητες , μη αρνητικές τυχαίες μεταβλητές με στηρίγματα $[0, b_X]$ και $[0, b_Y]$ αντιστοίχως με b_X και b_Y πεπερασμένα τότε

$$\mathcal{CE}\,(X, Y) = [\,b_Y \;-\; E(Y)]\,\mathcal{CE}\,(X) + [b_X - E(X)]\,\mathcal{CE}\,(Y) \qquad (1.22)$$

Αν οι μεταβλητές X και Y είναι ανεξάρτητες και οι 2 έχουν στήριγμα στο $[0, b]$ και μέση τιμή μ τότε :

$$\mathcal{CE}\,(X, Y) = (b - \mu)][\,\mathcal{CE}\,(X) + \mathcal{CE}\,(Y)] \qquad (1.23)$$

1.8 Αθροιστική παρελθοντική εντροπία και μέσος χρόνος αδράνειας

Σε αυτήν την παράγραφο θα δούμε πώς συνδέεται ο μέσος χρόνος αδράνειας ενός συστήματος με την αθροιστική παρελθοντική εντροπία. Για την τυχαία διάρκεια ζωής X ενός συστήματος ,υπενθυμίζουμε ότι η σχέση $[t - X|\, X \le t]$, περιγράφει τον χρόνο αδράνειας ενός συστήματος.

Ο μέσος χρόνος αδράνειας ενός συστήματος δίνεται από τον παρακάτω τύπο:

$$\tilde{\mu}(t) \;\; = E[t - X|\, X \le t\,] = \frac{1}{F(t)} \int_0^t F(x)dx \qquad (1.24)$$

Θεώρημα 1.2: Για μη αρνητική τυχαία μεταβλητή X με μέσο χρόνο αδράνειας $\tilde{\mu}(t)$ και αθροιστική παρελθοντική εντροπία $\mathcal{CE}\,(X) < \infty$ έχουμε:

$$\mathcal{CE}\,(X) = E[\tilde{\mu}(X)] \qquad (1.25)$$

Εφαρμογή του Θεωρήματος 1.2:

Έστω X τυχαία μεταβλητή με συνάρτηση κατανομής $F(x) = x^3$, $0 \le x \le 2$

Βρίσκουμε ότι ο μέσος χρόνος αδράνειας ισούται με $\tilde{\mu}(t) = \frac{t}{3}$, $0 \le t \le 2$

Επομένως με βάση την σχέση (1.26) η αθροιστική παρελθοντική εντροπία ισούται με

$$\mathcal{CE}(X) = E\left[\frac{x}{3}\right] = \frac{E[x]}{3} = \frac{\frac{32}{5}}{3} = \frac{32}{15} = 2{,}1$$

1.9 Σχέση αθροιστικής παρελθοντικής εντροπίας και αθροιστικής υπολειπόμενης εντροπίας

Σε αυτήν την παράγραφο θα μελετήσουμε την σύνδεση της αθροιστικής υπολειπόμενης εντροπίας $\mathcal{E}(X)$ με την αθροιστική παρελθοντική εντροπία $\mathcal{CE}(X)$ Από την σχέση (1.3) και (1.5) συνεπάγεται ότι η αθροιστική παρελθοντική εντροπία και η αθροιστική υπολειπόμενη εντροπία σχετίζονται ως εξής (Di Crescenzo και Longobardi (2012)):

$$\mathcal{E}(X) + \mathcal{CE}(X) = \int_{-\infty}^{\infty} h(x)dx \ , \tag{1.26}$$

όπου,

$$h(x) = -[F(x)\log F(x) + \bar{F}(x)\log\bar{F}(x)],\ x \in R \tag{1.27}$$

είναι η μερική εντροπία της τυχαίας μεταβλητής X (Bowden, 2010).

1.10 Κανονικοποιημένη αθροιστική παρελθοντική εντροπία

Σε αυτήν την παράγραφο θα ορίσουμε την κανονικοποιημένη αθροιστική παρελθοντική εντροπία, ενώ θα δούμε πως συνδέεται με την καμπύλη Bonferroni,την οποία θα μελετήσουμε στο Κεφάλαιο 4 της παρούσας εργασίας.

Για μη αρνητική τυχαία μεταβλητή X η οποία έχει πεπερασμένη μέση τιμή , οι Di Crescenzo και Longobardi (2009) όρισαν την κανονικοποιημένη αθροιστική παρελθοντική εντροπία ως εξής:

$$\mathcal{NCE}(X) = \frac{\mathrm{CE(X)}}{E(X)} = -\frac{1}{E(X)}\int_{0}^{\infty} F(x)lnF(x)dx \tag{1.28}$$

Θυμίζουμε ότι $0 \leq \mathcal{CE}(X) \leq E(X)$ (ενότητα 1.6), επομένως η $\mathcal{NCE}(X)$ παίρνει τιμές στο [0,1].

Ας δούμε την κανονικοποιημένη αθροιστική παρελθοντική εντροπία για κάποιες γνωστές κατανομές.

1.Εκθετική κατανομή με παράμετρο λ>0

Αν η τυχαία μεταβλητή X ακολουθεί την εκθετική κατανομή με παράμετρο λ>0 ,

έχουμε $E(X)=\frac{1}{\lambda}$ και είδαμε στην Παράγραφο 1.5 ότι $\mathcal{CE}\,(X) = \frac{\pi^2-6}{6\lambda}$.

Συνεπώς διαιρώντας κατά μέλη τις σχέσεις $\mathcal{CE}\,(X)$ και $E(X)$ παίρνουμε ότι :

$$\mathcal{NCE}\,(X) = \frac{\pi^2-6}{6}$$

2.Ομοιόμορφη κατανομή $\boldsymbol{U(0,b)}$

Αν η τυχαία μεταβλητή X ακολουθεί την ομοιόμορφη κατανομή $U(0,b)$ έχουμε $E(X)$=b/2 και είδαμε στην Παράγραφο 1.5 ότι $\mathcal{CE}\,(X)$ =b/4.Συνεπώς διαιρώντας κατά μέλη τις σχέσεις $\mathcal{CE}$ (X) και $E(X)$ παίρνουμε ότι :

$$\mathcal{NCE}\,(X)=\frac{1}{2}$$

1.11 Εμπειρική αθροιστική παρελθοντική εντροπία

Σε αυτήν την παράγραφο θα ορίσουμε έναν μη παραμετρικό εκτιμητή της αθροιστικής παρελθοντικής εντροπίας την εμπειρική αθροιστική παρελθοντική εντροπία. Στη συνέχεια θα δούμε τον τρόπο υπολογισμού της μέσης τιμής της εμπειρικής αθροιστικής παρελθοντικής εντροπίας για την ομοιόμορφη κατανομή.

Ορισμός 1.4: Έστω $X_1, X_2,\ldots, X_n$ ένα τυχαίο δείγμα μη αρνητικών , συνεχών i.i.d τυχαίων μεταβλητών .Ένας κατάλληλος εκτιμητής της αθροιστικής παρελθοντικής εντροπίας $\mathcal{CE}\,(X)$ είναι η εμπειρική αθροιστική παρελθοντική εντροπία , η οποία προτάθηκε από τους (Di Crescenzo και Longobardi (2009) και ορίζεται ως εξής:

$$\mathcal{CE}\,(\hat{F}_n) = -\int_0^\infty \hat{F}_n(x) ln\, \hat{F}_n(x)dx \qquad (1.29)$$

όπου η εμπειρική κατανομή του δείγματος ισούται με:

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \le x\}, x \in R. \qquad (1.30)$$

Διατάσσουμε το δείγμα (order statistics) ως εξής $X_{(1)}<\ldots<X_{(n)}$ και θεωρούμε ότι $U_1=X_{(1)}$, $U_{(i)}=X_{(i)}-X_{(i-1)}$, $i=1,2,..,n$. Έτσι η εμπειρική αθροιστική παρελθοντική εντροπία γράφεται ως:

$$\mathcal{CE}\,(\hat{F}_n) = -\sum_{j=1}^{n-1} U_{j+1}\,\frac{j}{n}\log\frac{j}{n}\,, \tag{1.31}$$

και συγκλίνει στην αθροιστική παρελθοντική εντροπία της X

$$\mathcal{CE}\,(\hat{F}_n) \to \mathcal{CE}\,(X)\,,\ \text{όσο}\ n \to +\infty. \tag{1.32}$$

Παρακάτω θα υπολογίσουμε την εμπειρική αθροιστική παρελθοντική εντροπία της ομοιόμορφης κατανομής $U\sim(0,1)$. Έστω $X_1, X_2,\ldots,X_n$ ένα τυχαίο δείγμα μη αρνητικών ,συνεχών τυχαίων μεταβλητών που ακολουθούν την $U\sim(0,1)$. Διατάσσουμε το δείγμα (order statistics) ως εξής $X_{(1)}<\ldots<X_{(n)}$ και θεωρούμε ότι $U_1=X_{(1)}$, $U_{(i)}=X_{(i)}-X_{(i-1)}$, $i=1,2,..,n$. Η συνάρτηση πυκνότητας πιθανότητας των $X_{(i)}$ είναι η :

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!\,(n-i)!}\,f_X(x)[F_X(x)\,]^{i-1}[\,\bar{F}_X(x)]^{n-i}$$

Ειδικά για την ομοιόμορφη $U(0,1)$ έχουμε $f_X(x)=1$, $F_X(x)=x$, $\bar{F}_X(x) = 1-x$ συνεπώς η συνάρτηση πυκνότητας πιθανότητας γίνεται:

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!\,(n-i)!}\cdot 1\cdot x^{i-1}(1-x)^{n-i}\,, 0\le x\le 1.$$

Παρατηρούμε ότι η τυχαία μεταβλητή $X_{(i)}\sim Beta(\,i, n+1-i)$. Συνεπώς

$$E(X_{(i)}) = \int_0^1 xf(x)dx = \frac{i}{n+1}\,.$$

Επομένως η μέση τιμή της $E(\,U_{(i)}\,)=E(\,X_{(i)}\,)-E(X_{(i-1)}\,) = \frac{i}{n+1}-\frac{i-1}{n+1}=\frac{1}{n+1}$ είναι ανεξάρτητο του i. Από την σχέση παίρνουμε ότι η μέση τιμή της εμπειρικής αθροιστικής παρελθοντικής εντροπίας είναι ίση με :

$$E(\mathcal{CE}(\hat{F}_n)) = \frac{\zeta'(-1)-\zeta^{(1,0)}(-1,n)}{n(n+1)} + \frac{(n-1)logn}{2(n+1)},$$

Όπου η $\zeta'(s)$ η παράγωγος της συνάρτησης ζήτα Riemann και ορίζεται ως εξής :

$$\zeta'(s) = -\sum_{k=1}^{\infty} \frac{lnk}{k^s}$$

$\zeta^{(1,0)}(s,\alpha)$ είναι η μερική παράγωγος της γενικευμένης συνάρτησης ζήτα Riemann σε σχέση με το πρώτο όρισμα.

1.12 Η εντροπία του Tsallis

Σε αυτήν την παράγραφο θα μελετήσουμε την εντροπία του Tsallis , την οποία παρουσίασε ο Tsallis το 1988, στη δημοσίευσή του "Πιθανή γενίκευση των στατιστικών Boltzmann-Gibbs" στο περιοδικό *Journal of Statistical Physics*. .

Ορισμός 1.5: Για μη αρνητική συνεχή τυχαία μεταβλητή X με συνάρτηση πυκνότητας πιθανότητας $f(x)$ η εντροπία του Tsallis τάξης α δίνεται από την σχέση:

$$T_\alpha(X) = \frac{1}{\alpha-1}\left(1 - \int_0^\infty f^a(x)dx\right); \alpha \neq 1, \alpha > 0. \qquad (1.33)$$

Παρατηρούμε ότι όσο το $\alpha \to 1$ τότε η $T_\alpha(X)$ “reduces to” στην διαφορική εντροπία του Shannon (σχέση 1.2 , δείτε την ενότητα 1.2).Ένας άλλος τρόπος να εκφράσουμε την σχέση (1.33) είναι ο εξής:

$$T_\alpha(X) = \frac{1}{\alpha-1}\int_0^\infty (f(x) - f^a(x))dx \ ; \alpha \neq 1, \alpha > 0. \qquad (1.34)$$

1.13 Αθροιστική υπολειπόμενη εντροπία του Tsallis

Σε αυτήν την παράγραφο θα αναφερθούμε στην αθροιστική υπολειπόμενη εντροπία του Tsallis. Με βάση την σχέση (1.34) οι Sati και Gupta πρότειναν την αθροιστική υπολειπόμενη εντροπία του Tsallis τάξης α (CRTE) , η οποία για μια μη αρνητική συνεχή τυχαία μεταβλητή δίνεται από τον τύπο:

$$\eta_\alpha(X) = \frac{1}{\alpha-1}(1- \int_0^\infty \bar{F}(x)^a dx) \ ; \ \alpha \neq 1, \alpha > 0, \qquad (1.35)$$

όπου $\bar{F}(x)$ η συνάρτηση επιβίωσης της τυχαίας μεταβλητής X.

1.14 Αθροιστική παρελθοντική εντροπία του Tsallis

Σε αυτήν την παράγραφο θα αναφερθούμε στην αθροιστική παρελθοντική εντροπία του Tsallis.

Ορισμός 1.6: Για μη αρνητική συνεχή τυχαία μεταβλητή X με συνάρτηση κατανομής $F(x)$ η αθροιστική παρελθοντική εντροπία του Tsallis (CTE) τάξης α :

$$C\xi_{\alpha}(X) = \frac{1}{\alpha-1}(\int_0^{\infty}\big(F(x) - F^{\alpha}(x)\big)dx) \ ; \alpha \neq 1, \alpha > 0 \qquad (1.36)$$

Παρατηρούμε ότι όσο το $\alpha \rightarrow 1$ τότε η $C\xi_{\alpha}(X)$ "τείνει" στην αθροιστική παρελθοντική εντροπία $CE(X)$ (σχέση 1.7 , δείτε την Ενότητα 1.5).

Θεώρημα 1.3: $C\xi_{\alpha}(X)$= 0 αν και μόνο αν η τυχαία μεταβλητή X είναι σταθερή , ενώ $C\xi_{\alpha}(X)$>0 για κάθε μη αρνητική συνεχή τυχαία μεταβλητή X.

Απόδειξη.

Έστω $0 < \alpha < 1$. Τότε έχουμε ότι $F(x) \leq F^{\alpha}(x)$ συνεπώς από την σχέση 1.36 παίρνουμε ότι $C\xi_{\alpha}(X) \geq 0$. Αν $\alpha > 1$, $F(x) \geq F^{\alpha}(x)$ και έτσι $C\xi_{\alpha}(X) \geq 0$. Αν η τυχαία μεταβλητή X είναι σταθερή , τότε $C\xi_{\alpha}(X)$= 0.

Αντίστροφα, αν $C\xi_{\alpha}(X)$= 0 , τότε το 2ο μέλος της σχέσης (1.37) γίνεται ίσο με μηδέν. Δηλαδή $(\int_0^{\infty}\big(F(x) - F^{\alpha}(x)\big)dx)$ =0 επειδή $\alpha \neq 1$. Η συνάρτηση μέσα στο ολοκλήρωμα είναι μη αρνητική για κάθε x , σύμφωνα με την τιμή του α. Έτσι παίρνουμε ότι $F(x)\big(1 - F^{\alpha-1}(x)\big) = 0$. Συνεπώς έχουμε ότι $F(x)$ = 0 ή $F(x)$ = 1, άρα η τυχαία μεταβλητή X είναι σταθερή. Η απόδειξη ολοκληρώθηκε.■

Θα μελετήσουμε τώρα την σύνδεση της αθροιστικής παρελθοντικής εντροπίας του Tsallis με τον μέσο χρόνο αδράνειας ενός συστήματος.

Λήμμα 1.1: Για μη αρνητική συνεχή τυχαία μεταβλητή X με συνάρτηση κατανομής $F(x)$, τότε έχουμε ότι:

$$C\xi_\alpha(X) = E[\tilde{\mu}_X(x)(F(X))^{\alpha-1}], \tag{1.37}$$

όπου με $\tilde{\mu}_X(x)$ συμβολίζουμε τον μέσο χρόνο αδράνειας ενός συστήματος (σχέση 1.25).

Από το Λήμμα 1.1, αν $\alpha \geq 1$ τότε $C\xi_\alpha(X) \leq E[\tilde{\mu}_X(X)]$. Τώρα αν $0 < \alpha \leq 1$, παίρνουμε ότι $C\xi_\alpha(X) \geq E[\tilde{\mu}_X(X)]$. Παρακάτω θα αναφερθούμε στον τρόπο με τον οποίο σχετίζεται η αθροιστική εντροπία του Tsallis με την αθροιστική παρελθοντική εντροπία.

Θεώρημα 1.4: Για μη αρνητική συνεχή τυχαία μεταβλητή X με συνάρτηση πυκνότητας πιθανότητας $f(x)$, αν $\alpha \geq 1$ $(0 < \alpha \leq 1)$,τότε :

$$C\xi_\alpha(X) \leq (\geq)\ \mathcal{CE}(X). \tag{1.38}$$

Απόδειξη.

Αν $\alpha > 1$ $(0 < \alpha < 1)$ μπορούμε να γράψουμε :

$$\begin{aligned} C\xi_\alpha(X) &= \frac{1}{\alpha-1}\left[\int_0^\infty (F(x) - F^a(x))dx\right] \\ &= \frac{1}{a-1}\left[\int_0^\infty F(x)(1 - F^{a-1}(x))dx\right] \\ &\leq (\geq) \frac{1}{\alpha-1}\left[-F(x)\left(lnF^{a-1}(x)\right)dx\right] \\ &= \mathcal{CE}(X). \end{aligned}$$

Η ανισότητα προκύπτει από την χρήση της σχέσης , $1 - u \leq -lnu,\ όταν\ u > 0$. Η απόδειξη ολοκληρώθηκε.■

Μια εναλλακτική απόδειξη της παραπάνω σχέσης , θα μπορούσε να προκύψει από το Λήμμα 1.1 , όπου αποδείξαμε $C\xi_\alpha(X) \geq (\leq) E[\tilde{\mu}_X(X)]$. Θυμίζουμε εδώ ότι οι Di Crescenzo and Longobardi (2009) , απέδειξαν ότι $\mathcal{CE}(X) = E[\tilde{\mu}_X(X)]$ (σχέση 1.26). Επομένως παίρνουμε το ζητούμενο.■

Εφαρμογή του Θεωρήματος 1.4: Αν X ακολουθεί την ομοιόμορφη $U[0, b]$ τότε για $0 \leq t \leq b$ τότε με βάση τον τύπο (1.7) έχουμε:

$$\mathcal{CE}(X) = - \int_0^b \left(\frac{\mathrm{t-b}}{\mathrm{b}}\right) ln\left(\frac{\mathrm{t-b}}{\mathrm{b}}\right) dt = \frac{\mathrm{b}}{4}\,.$$

Επίσης με βάση τον τύπο (2.7) υπολογίζουμε την αθροιστική παρελθοντική εντροπία του Tsallis:

$$\mathrm{C}\xi_{\alpha}(X) = \frac{b}{2(a+1)}\,.$$

Παρατηρούμε ότι για $\alpha > 1$ έχουμε $\mathrm{C}\xi_{\alpha}(X) \leq \mathcal{CE}(X)$, ενώ για $0 < \alpha < 1$ παίρνουμε $\mathrm{C}\xi_{\alpha}(X) \geq \mathcal{CE}(X)$. Συνεπώς επαληθεύεται το Θεώρημα 1.4 ■

Κλείνοντας το Κεφάλαιο 1 της εργασίας θα κάνουμε μια σύντομη περίληψη. Αρχικά μελετήσαμε την έννοια της εντροπίας του Shannon, ενώ στην συνέχεια παρουσιάσαμε την αθροιστική υπολειπόμενη εντροπία(CRE), όπως επίσης και την αθροιστική παρελθοντική εντροπία (CPE). Ασχοληθήκαμε διεξοδικά με τα παραπάνω στατιστικά μέτρα πληροφορίας, είδαμε τις ιδιότητες τους και τις εκτιμήσεις τους. Μελετήσαμε ακόμη την εντροπία του Tsallis και την αθροιστική υπολειπόμενη εντροπία του Tsallis. Όλα αυτά τα μέτρα παίζουν σημαντικό ρόλο στην εύρεση των δεικτών Gini και Bonferroni, δύο από των σημαντικότερων χρηματοοικονομικών δεικτών. Στο Κεφάλαιο 2 θα αναφερθούμε στον δείκτη Gini πολύ αναλυτικά και θα καταλάβουμε την χρησιμότητά του Κεφαλαίου 1 στον υπολογισμό του.

ΚΕΦΑΛΑΙΟ 2

Ο δείκτης Gini

2.1 Ο δείκτης Gini και η οικονομική ανισότητα

Το ζήτημα της ανισότητας προσλαμβάνει όλο και μεγαλύτερη σημασία τα τελευταία χρόνια. Οι επιπτώσεις της οικονομικής κρίσης στην Ευρώπη υπήρξαν έντονες, ανατρέποντας χρόνια σύγκλισης των συνθηκών διαβίωσης και ασκώντας σημαντική πίεση στα συστήματα κοινωνικής προστασίας. Οι ανισότητες αυξήθηκαν στα περισσότερα κράτη μέλη, προκαλώντας προβληματισμό για τη βιωσιμότητα της ανάπτυξης και για την κοινωνική συνοχή. Μια άνιση κοινωνία συνδέεται με μια κοινωνία με μικρή οικονομική ευημερία και ένα φτωχό κράτος πρόνοιας.

Ο δείκτης Gini ή ο συντελεστής Gini είναι ένα οικονομικό μέτρο που χρησιμοποιείται για τον υπολογισμό της εισοδηματικής ανισότητας που υπάρχει μεταξύ των πολιτών μιας περιοχής, συνήθως μιας χώρας. Αναπτύχθηκε από τον Ιταλό στατιστικολόγο και κοινωνιολόγο Corrado Gini στις αρχές του 1900, ενώ στη συνέχεια προτάθηκαν αρκετές τροποποιήσεις του συγκεκριμένου δείκτη από τους Dalton(1920), Atkinson(1970), Sheshinski(1972), όπως και από αρκετούς άλλους. Ο συντελεστής Gini είναι ο πλέον συνηθισμένος δείκτης μέτρησης της εισοδηματικής ανισότητας διεθνώς. Μετρά την κατανομή του εισοδήματος συγκρίνοντας την εισοδηματική κατάσταση κάθε νοικοκυριού με την εισοδηματική κατάσταση όλων των άλλων νοικοκυριών.

Σε αυτό το σημείο θα δώσουμε τον ορισμό της εισοδηματικής ανισότητας σύμφωνα με την ΕΕ: Η εισοδηματική ανισότητα αφορά τον τρόπο με τον οποίο κατανέμεται στον πληθυσμό το εισόδημα που κερδίζεται σε μια οικονομία. Υπολογίζεται συνήθως σε επίπεδο νοικοκυριού(δηλαδή, προσθέτοντας το εισόδημα όλων των μελών του νοικοκυριού) και σταθμίζεται ως προς τον αριθμό των μελών του νοικοκυριού και την ηλικία τους.

Η τιμή του δείκτη Gini κυμαίνεται μεταξύ 0 και 1, με το μηδέν να είναι η μέγιστη ισότητα (όλοι οι πολίτες έχουν το ίδιο εισόδημα) και 1 να είναι η μέγιστη ανισότητα (όλα τα έσοδα ανήκουν σε έναν μόνο πολίτη). Όσο υψηλότερος είναι ο δείκτης Gini, τόσο μεγαλύτερη είναι η εισοδηματική ανισότητα στον πληθυσμό. Δηλαδή, μερικές χρεώνουν περισσότερο από τον υπόλοιπο πληθυσμό. Συνιστάται μια χώρα ή περιοχή να μην έχει συντελεστές Gini κοντά σε έναν, καθώς αυτό θα σήμαινε ότι η κοινωνία, μιλώντας νομισματικά, είναι πολύ άνιση.

Στον παρακάτω πίνακα, ο οποίος προέρχεται από την ΕΛΣΤΑΤ, παρουσιάζονται οι τιμές του δείκτη Gini σε ευρωπαϊκές χώρες για το 2019. Παρατηρούμε ότι για την Ελλάδα ο συντελεστής Gini εκτιμήθηκε το 2019 σε 31,0%, η μικρότερη τιμή των τελευταίων χρόνων. Το ποσοστό αυτό ερμηνεύεται ως εξής: επιλέγοντας 2 τυχαία άτομα του πληθυσμού, το εισόδημά τους θα διαφέρει κατά 31% του μέσου ισοδύναμου διαθέσιμου εισοδήματος. Ακόμα το 2014 ο δείκτης Gini στην χώρα μας ήταν ίσος με 34.5% ο μεγαλύτερος των τελευταίων 12 ετών. Από τον παρακάτω πίνακα διαπιστώνουμε ότι η Φιλανδία το 2019 είχε τον μικρότερο δείκτη Gini, 26,2%, από όλες τις χώρες που δίνονται παρακάτω. Αυτό σημαίνει ότι στη Φιλανδία καταγράφεται μικρότερη οικονομική ανισότητα.

Πίνακας 5

Δείκτης άνισης κατανομής εισοδήματος (συντελεστής Gini) σε ευρωπαϊκές χώρες με διαθέσιμα στοιχεία για το 2019: 2008-2019

%

Χώρες	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Βουλγαρία	35,9	33,4	33,2	35,0	33,6	35,4	35,4	37,0	38,3	40,2	39,6	40,8
Λετονία	37,5	37,5	35,9	35,1	35,7	35,2	35,5	35,4	34,5	34,5	35,6	35,2
Ρουμανία	35,9	34,5	33,5	33,5	34,0	34,6	35,0	37,4	34,7	33,1	35,1	34,8
Ελλάδα	33,4	33,1	32,9	33,5	34,3	34,4	34,5	34,2	34,3	33,4	32,3	31,0
Πολωνία	32,0	31,4	31,1	31,1	30,9	30,7	30,8	30,6	29,8	29,2	27,8	28,5
Μάλτα	28,1	27,4	28,6	27,2	27,1	28,0	27,7	28,1	28,6	28,2	28,7	28,0
Ουγγαρία	25,2	24,7	24,1	26,9	27,2	28,3	28,6	28,2	28,2	28,1	28,7	28,0
Δανία	25,1	26,9	26,9	26,6	26,5	26,8	27,7	27,4	27,7	27,6	27,9	27,5
Αυστρία	27,5	26,4	26,6	26,3	26,5	25,9	25,9	26,2	26,3	26,0	26,8	27,5
Φινλανδία	26,3	25,9	25,4	25,8	25,9	25,4	25,6	25,2	25,4	25,3	25,9	26,2

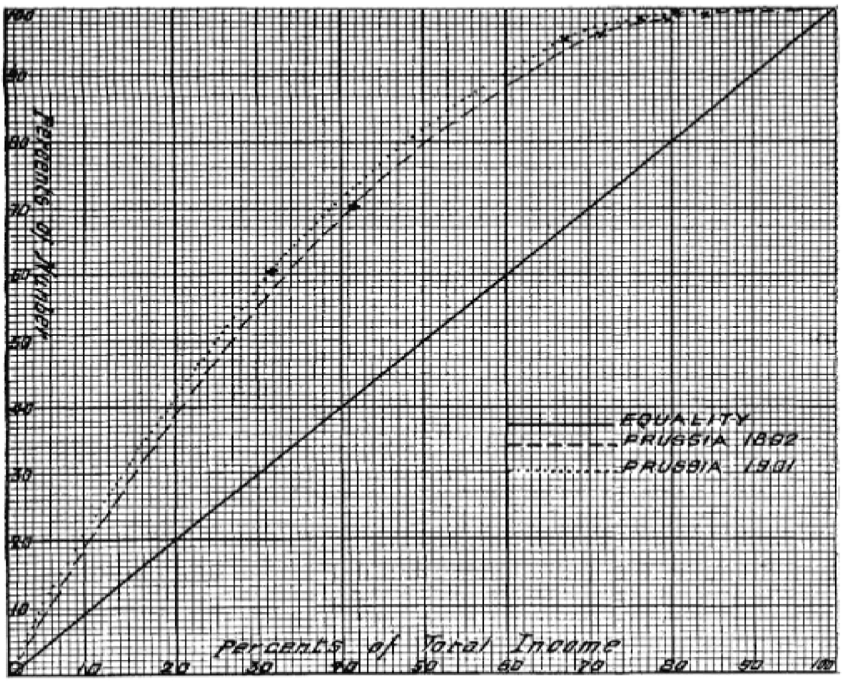
(Πηγή ΕΛΣΤΑΤ)

2.2 Η καμπύλη Lorenz και η σχέση της με τον δείκτη Gini

Από τον περασμένο αιώνα, η ανάλυση και η γραφική αναπαράσταση της ανισότητας διαδραματίζουν πολύ σημαντικό ρόλο στα οικονομικά. Στη βιβλιογραφία έχουν προταθεί και έχουν αναπτυχθεί πολλές καμπύλες για να απλοποιηθεί η περιγραφή της ανισότητας. Μια από αυτές είναι και η καμπύλη Lorenz. Στην συγκεκριμένη παράγραφο θα μελετήσουμε την σχέση της καμπύλης Lorenz με τον δείκτη Gini.

Το 1905 ο Max O. Lorenz παρουσίασε την καμπύλη Lorenz, η οποία μέχρι και σήμερα θεωρείται ένα από τα πιο δημοφιλή γραφήματα για την αναπαράσταση της εισοδηματικής κατανομής. Έχουν δοθεί πολλοί ισοδύναμοι ορισμοί της, από τους Pietra (1915) αλλά και τον Gastwirth (1972) .

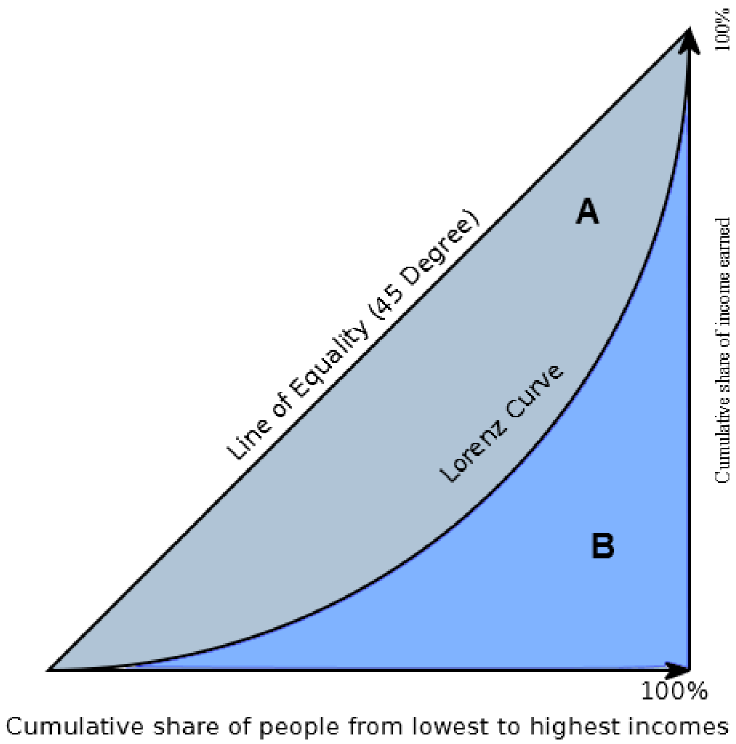
Σχήμα 2.1: Η αυθεντική καμπύλη Lorenz (1905)



Πηγή :American Statistical Association June 1905 JSTOR

Για την κατασκευή της χρησιμοποιούνται σχεδόν πάντα εισοδηματικά δεδομένα. Είναι ένα πολύ εύκολα ερμηνεύσιμο γράφημα καθώς με τη βοήθειά του μπορούμε να έχουμε μια σαφή εικόνα για την αναπαράσταση και τη σύγκριση της οικονομικής ανισότητας. Ένα πολύ σημαντικό πλεονέκτημα της καμπύλης αυτής είναι ότι μπορεί να χρησιμοποιηθεί και στη σύγκριση περισσότερων πληθυσμών. Η καμπύλη Lorenz δείχνει την ανισότητα που ενδεχομένως υπάρχει στην κατανομή μιας μεταβλητής, όπου η μεταβλητή αυτή είναι το εισόδημα ή ο πλούτος. Θα πρέπει να σημειωθεί ότι ο εντοπισμός ανισότητας στην καμπύλη Lorenz δεν συνεπάγεται ότι μπορούμε να εντοπίσουμε τα αίτια της ύπαρξης αυτής της εισοδηματικής ανισότητας μεταξύ των ατόμων που μελετάμε. Προκειμένου λοιπόν να εντοπίσουμε τα αίτια της ανισότητας αν υφίσταται , πρέπει να αποφανθούν κατάλληλα καταρτισμένοι επιστήμονες, όπως οικονομολόγοι, κοινωνιολόγοι και μαθηματικοί. Με άλλα λόγια η καμπύλη Lorenz μας βοηθάει στον εντοπισμό της ανισότητας και όχι στην επίλυση της.

Στο παρακάτω σχήμα απεικονίζεται η μορφή με την οποία συναντάμε την καμπύλη Lorenz σήμερα. Αξίζει να τονίσουμε ότι πρόκειται για μια κυρτή καμπύλη. Στον οριζόντιο άξονα εμφανίζεται το αθροιστικό ποσοστό του πληθυσμού, ενώ στον κατακόρυφο άξονα έχουμε το αθροιστικό ποσοστό του εισοδήματος.



Σχήμα 2.2:Μορφή Καμπύλης Lorenz

Πηγή:Wikipedia

Η ευθεία με κλίση 45 μοιρών θα έδειχνε την απόλυτα ίση κατανομή του εισοδήματος, δηλαδή ο πληθυσμός και ο πλούτος κατανέμονται εξίσου, επομένως αυτή τη στιγμή η τιμή του δείκτη Gini ισούται με μηδέν.

Η περιοχή Α του διαγράμματος μεταξύ της τέλειας ισότητας και της καμπύλης του Lorenz δείχνει την ανισότητα. Η περιοχή Β του παραπάνω σχήματος χρησιμοποιείται για τον υπολογισμό του δείκτη Gini.

Μαζί με την καμπύλη Lorenz μπορούμε να υπολογίσουμε τον δείκτη Gini, ο οποίος δίνεται από την σχέση:

$$\text{Gini} = \frac{A}{\mathrm{A+B}} \tag{2.1}$$

Ο αριθμητής Α είναι το εμβαδό ανάμεσα στην καμπύλη Lorenz και τη διαγώνιο, ενώ ο παρονομαστής Β είναι το εμβαδό κάτω από τη διαγώνιο. Όμως το Α+Β ισούται με $\frac{1}{2}$ τετραγωνικές μονάδες, καθώς το Α+Β είναι το εμβαδόν ενός ορθογωνίου τριγώνου με βάση 1 και ύψος 1 όπως φαίνεται και στο παραπάνω σχήμα. Θυμίζουμε ότι το Εμβαδόν τριγώνου δίνεται από τον τύπο $\frac{\text{βάση}\cdot\text{ύψος}}{2}$. Έτσι τελικά ο δείκτης Gini θα ισούται με:

$$\text{Gini} = \frac{\mathrm{A}}{\frac{1}{2}} = 2\mathrm{A}.$$

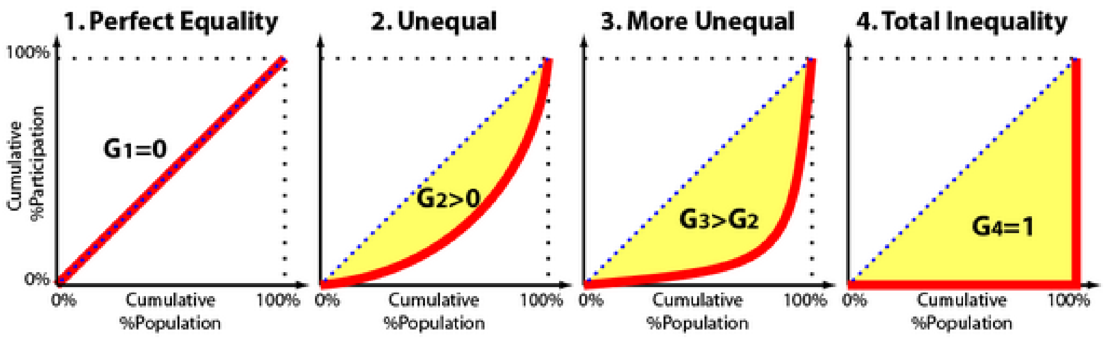
Με αυτόν τον τρόπο λαμβάνουμε τον συντελεστή ή με τη σειρά πολλαπλασιάζοντας το αποτέλεσμα με 100, λαμβάνουμε το ποσοστό. Η παραπάνω διαδικασία αποτελεί την γεωμετρική προσέγγιση του δείκτη Gini με τη βοήθεια της καμπύλης Lorenz, η οποία αποδείχθηκε από τον Raymond L.Lows (1984).

Κάποιες παρατηρήσεις σχετικά με την σχέση του δείκτη Gini με την καμπύλη Lorenz που προκύπτουν από τον παραπάνω τύπο είναι οι εξής:

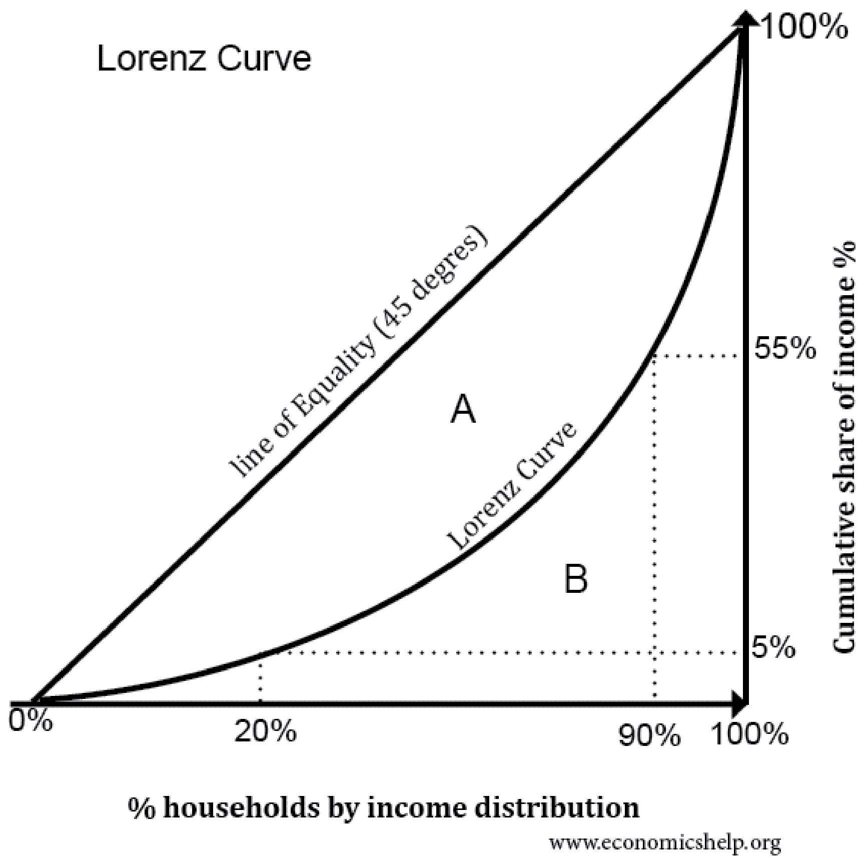
- Όσο πιο κοντά είναι η καμπύλη Lorenz στη γραμμή ισότητας, τόσο μικρότερη είναι η περιοχή Α. Συνεπώς και ο συντελεστής Gini θα είναι χαμηλός.

- Εάν υπάρχει υψηλός βαθμός ανισότητας, τότε η περιοχή A θα είναι το μεγαλύτερο ποσοστό της συνολικής επιφάνειας.
- Μια αύξηση του συντελεστή Gini δείχνει αύξηση της ανισότητας, με άλλα λόγια η καμπύλη Lorenz είναι πιο μακριά από τη γραμμή ισότητας,

Σχήμα 2.3: Δείκτης Gini και Καμπύλη Lorenz

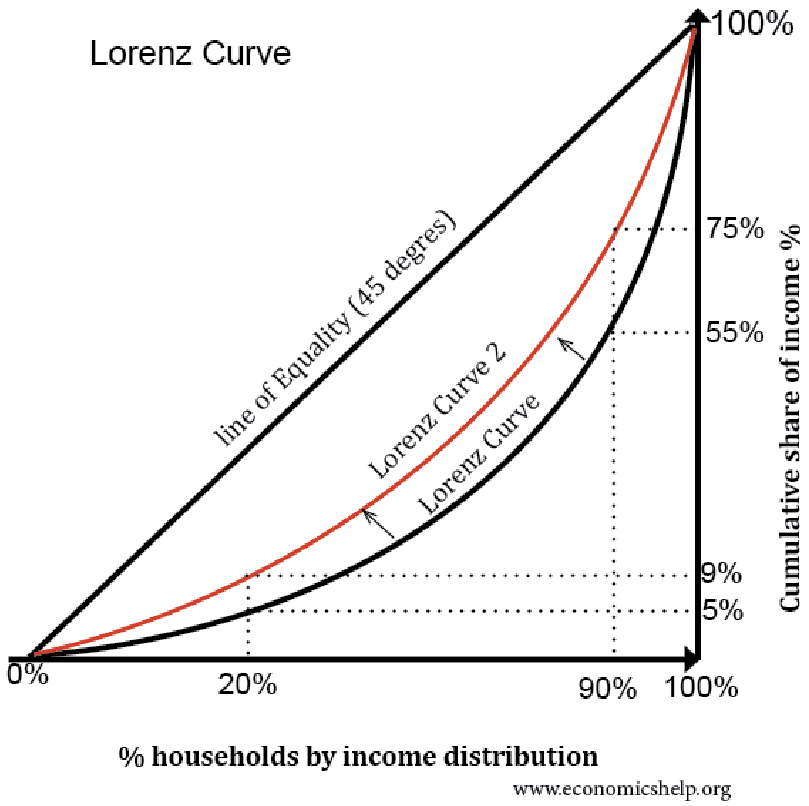


Ερμηνεία της Καμπύλης Lorenz



Από την παραπάνω καμπύλη Lorenz παρατηρούμε ότι το 20% του φτωχότερου πληθυσμού των νοικοκυριών έχει το 5% του συνολικού εισοδήματος της χώρας. Το 90% του φτωχότερου πληθυσμού των νοικοκυριών έχει το 55% του συνολικού εισοδήματος της χώρας. Αυτό σημαίνει πως το πλουσιότερο 10% των εισοδηματιών κερδίζει το 45% του συνολικού εισοδήματος.

Μετατόπιση στην Καμπύλη Lorenz



Στο σημείο αυτό θα ερμηνεύσουμε την μετατόπιση της καμπύλης Lorenz.Στο παραπάνω σχήμα παρατηρούμε ότι η καμπύλη Lorenz έχει πλησιάσει τη γραμμή της ισότητας, πράγμα που σημαίνει ότι υπάρχει μείωση της ανισότητας. Ας ερμηνεύσουμε λοιπόν το παραπάνω γράφημα . Από την παραπάνω καμπύλη Lorenz παρατηρούμε ότι το 20% του φτωχότερου πληθυσμού των νοικοκυριών έχει τώρα το 9% του συνολικού εισοδήματος της χώρας (σε σχέση με το 5% που είχε πριν). Το 90% του φτωχότερου πληθυσμού των νοικοκυριών έχει το 75% του συνολικού εισοδήματος της χώρας. Αυτό σημαίνει πως το πλουσιότερο 10% των εισοδηματιών που κέρδιζε το 45% του συνολικού εισοδήματος, τώρα έχει μόνο το 25% του συνολικού εισοδήματος. Θα μπορούσε βέβαια η κόκκινη γραμμή να είναι κάτω από την μαύρη γραμμή. Στην περίπτωση αυτή θα είχαμε μεγαλύτερη εισοδηματική ανισότητα καθώς το 10% του πλουσιότερου πληθυσμού θα καταλάμβανε ακόμα μεγαλύτερο μέρος του πλούτου της χώρας.

Στάδια υπολογισμού του δείκτη Gini

Για να υπολογιστεί ο δείκτης Gini πρέπει να ακολουθήσουμε κάποια βήματα.

- Αρχικά στο πρώτο βήμα, ταξινομούμε τα άτομα με βάση το εισόδημά τους κατά αύξουσα σειρά, από το φτωχότερο στο πλουσιότερο. Έτσι αν στην ερευνά μας έχουμε 100 άτομα, θα συμβολίσουμε με 1 το άτομο με το χαμηλότερο εισόδημα ενώ το άτομο 100 θα είναι εκείνο με το υψηλότερο εισόδημα.
- Στη συνέχεια, στο δεύτερο στάδιο θα υπολογίσουμε την αθροιστική κατανομή εισοδήματος.
- Στο τρίτο βήμα, θα βρούμε το αθροιστικό ποσοστό του εισοδήματος κάνοντας την διαίρεση κάθε αθροιστικού εισοδήματος με το συνολικό εισόδημα.
- Στο τέταρτο βήμα υπολογίζουμε το αθροιστικό ποσοστό του πληθυσμού.
- Τέλος στο πέμπτο βήμα θα υπολογίσουμε το δείκτη Gini από την καμπύλη Lorenz. Σχηματίζονται πολύγωνα, των οποίων θα υπολογίσουμε τα εμβαδά. Το πρώτο πολύγωνο είναι πάντα τρίγωνο, ενώ τα υπόλοιπα είναι τραπέζια. Όταν βρούμε όλα τα εμβαδά στο τέλος τα αθροίζουμε, μιας και αυτό το άθροισμα αποτελεί το εμβαδόν της περιοχής κάτω από την καμπύλη Lorenz (Εμβαδό Β). Τότε ο δείκτης Gini δίνεται από τον τύπο:

$$\text{Gini} = 1-2B \qquad \textbf{(2.2)}$$

Ας δούμε όμως αυτήν την διαδικασία με ένα παράδειγμα, ενώ στο Κεφάλαιο 4 της εργασίας θα υπολογίσουμε τον δείκτη Gini με αυτήν την μέθοδο.

Παράδειγμα 2.1

Σε ένα πληθυσμό 20 ατόμων θα υπολογίσουμε τον δείκτη Gini . Ξέρουμε ότι 5 άτομα κερδίζουν 50 ευρώ το μήνα το καθένα, 10 άτομα κερδίζουν 100 ευρώ το μήνα το καθένα, ενώ τα υπόλοιπα 5 άτομα κερδίζουν 300 ευρώ το μήνα το καθένα. Τα δεδομένα μας παρουσιάζονται και στον παρακάτω πίνακα.

Πληθυσμός	5	10	5
Μηνιαίο εισόδημα για ένα άτομο	50€	100€	300€
Συνολικός πληθυσμός	20		

Αρχικά θα υπολογίσουμε το συνολικό εισόδημα ,το ποσοστό του εισοδήματος και το ποσοστό του πληθυσμού.

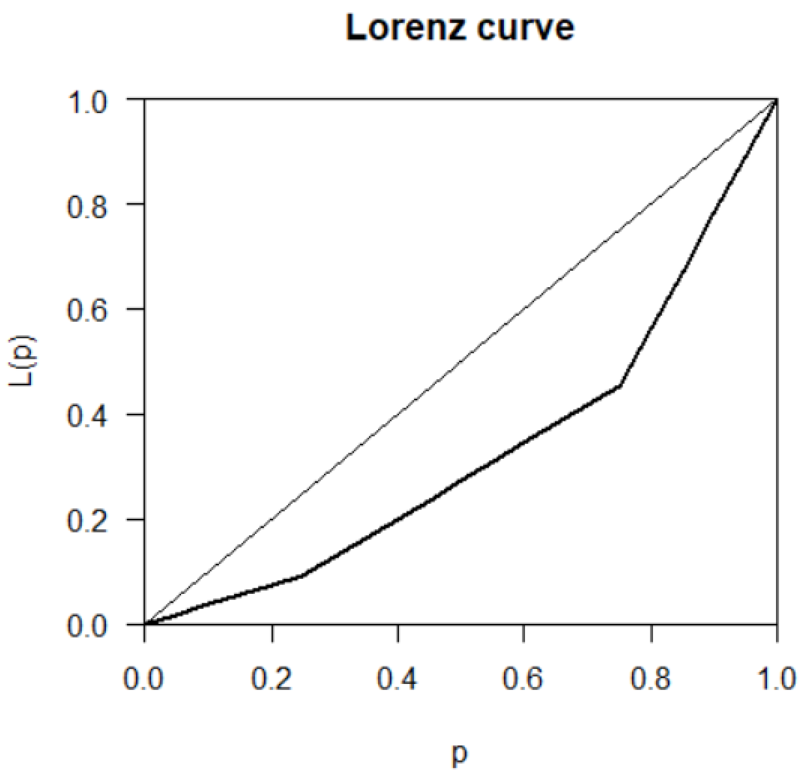
Πληθυσμός	Εισόδημα	Συνολικό εισόδημα	Ποσοστό του εισοδήματος	Ποσοστό του πληθυσμού
5	50	250	0,09	0,25
10	100	1000	0,36	0,5
5	300	1500	0,55	0,25
20		2750		

Στη συνέχεια με τη βοήθεια του παραπάνω πίνακα θα υπολογίσουμε το αθροιστικό ποσοστό του εισοδήματος και το αθροιστικό ποσοστό του πληθυσμού.

Πληθυσμός	Εισόδημα	Αθροιστικό Ποσοστό του εισοδήματος	Ποσοστό του πληθυσμού
5	50	0,09	0,25
10	100	0,45	0,75
5	300	1	1

Θα κατασκευάσουμε την καμπύλη Lorenz με βάση τα παραπάνω αποτελέσματα. Στον οριζόντιο άξονα απεικονίζεται το αθροιστικό ποσοστό του πληθυσμού, ενώ στον κατακόρυφο άξονα έχουμε το αθροιστικό ποσοστό του εισοδήματος. Ο πληθυσμός έχει χωριστεί σε 3 μέρη. Έχουμε 3 σημεία τα οποία θα ενώσουμε για να προκύψει η καμπύλη, τα οποία είναι τα εξής: (0.25, 0.09), (0.75, 0.45),(1,1).Η παραπάνω καμπύλη φαίνεται στο γράφημα που ακολουθεί. Σχηματίζεται ένα ορθογώνιο και 2 τραπέζια. Το εμβαδόν της περιοχής κάτω από την καμπύλη (Εμβαδό Β) ισούται με:

$$\frac{0.25 \cdot 0.09}{2} + \frac{(0.09+0.45) \cdot 0.50}{2} + \frac{(0.45+1) \cdot 0.25}{2} = 0.3275$$



Σχήμα 2.4: Καμπύλη Lorenz του πληθυσμού του παραδείγματος 2.1

Ο δείκτης Gini ισούται με:

$$\text{Gini}=1-2 \cdot 0.3275 \cong 0.34$$

Η παραπάνω τιμή μας δείχνει πως υπάρχει οικονομική ανισότητα στα παραπάνω άτομα και ερμηνεύεται ως εξής: επιλέγοντας 2 τυχαία άτομα του πληθυσμού, το εισόδημά τους θα διαφέρει κατά 34% του μέσου ισοδύναμου διαθέσιμου εισοδήματος

Κάποιες άλλες παρατηρήσεις που προκύπτουν από την παραπάνω καμπύλη είναι ότι το υψηλότερο εισόδημα αντιστοιχεί στο ανώτερο 25% του πληθυσμού. Επίσης λαμβάνουμε ότι το 25% του φτωχότερου πληθυσμού των νοικοκυριών έχει το 9% του συνολικού εισοδήματος .

Τόσο ο δείκτης Gini όσο και η καμπύλη Lorenz αναπτύσσονται ως μέθοδοι για τον εντοπισμό των ανισοτήτων μεταξύ του πληθυσμού μιας περιοχής (έθνος, πολιτεία, τοποθεσία κ.λπ.), κατανοώντας ότι όσο περισσότερη δικαιοσύνη υπάρχει μεταξύ των κατοίκων, τόσο μεγαλύτερη είναι η προσέγγιση της καμπύλης με μια τέλεια γραμμή, ενώ το αντίθετό της, μια μεγάλη ανισότητα μεταξύ του πληθυσμού μιας περιοχής, η καμπύλη γίνεται όλο και πιο έντονη.

2.3 Μαθηματική μορφή της καμπύλης Lorenz

Σε αυτήν την παράγραφο θα μελετήσουμε τη μαθηματική μορφή της καμπύλης Lorenz η οποία προτάθηκε από τον Joseph L.Gastwirth(1972), όπως επίσης θα αναφερθούμε στα χαρακτηριστικά της αλλά και στην ερμηνεία της.

Για τυχαία μεταβλητή X , η οποία εκφράζει το εισόδημα ενός πληθυσμού, με συνάρτηση πυκνότητας πιθανότητας $f(x)$**,** συνάρτηση κατανομής $F(x)$**,** ενώ με $F^{-1}(x)$ συμβολίζουμε την δεξιά αντίστροφη συνάρτηση της αθροιστικής συνάρτησης πιθανότητας, $F^{-1}(x)$ $=inf\{x: F(x) \geq t\}$. όπου $F^{-1}(0) = 0$. Τότε η καμπύλη Lorenz δίνεται από τον παρακάτω τύπο:

$$L(p) = \frac{1}{\mu} \int_0^p F^{-1}(x)dx \text{ , } p\epsilon[0,1], \qquad (2.3)$$

όπου $\mu = E(X)$ της τυχαίας μεταβλητής X. Αξίζει να τονίσουμε ότι η F μπορεί να είναι είτε διακριτή είτε συνεχής συνάρτηση κατανομής. Επίσης γίνεται φανερό ότι η καμπύλη Lorenz μπορεί να βρεθεί από την σχέση (2.3) μόνο αν είναι γνωστή η συνάρτηση κατανομής $F(x)$ της τυχαίας μεταβλητής X που ακολουθούν τα δεδομένα.

Παρακάτω θα αποδείξουμε τον τύπο της καμπύλης Lorenz για την εκθετική κατανομή $\lambda > 0$. Αρχικά θα υπολογίσουμε την συνάρτηση F^{-1}.

$$y = 1 - e^{-\lambda x} \leftrightarrow e^{-\lambda x} = 1 - y$$

$$\leftrightarrow -\lambda x = \ln(1-y)$$

$$\rightarrow x = \frac{-\ln(1-y)}{\lambda}$$

$$\rightarrow F^{-1}(x) = \frac{-ln(1-\mathrm{x})}{\lambda}$$

Τώρα θα υπολογίσουμε το ολοκλήρωμα :

$$\int_0^p F^{-1}(x)dx = \int_0^p \left(-\frac{\ln(1-\mathrm{x})}{\lambda}\right)dx$$

$$= -\frac{1}{\lambda}\int_0^p \ln(1-\mathrm{x})dx$$

$$= -\frac{1}{\lambda}\left[[x \cdot ln(1-x)]_0^p - \int_0^p xd(ln(1-x))\right]$$

$$= -\frac{1}{\lambda}\left[p \cdot ln(1-p) + \int_0^p x \cdot \frac{1}{1-x}dx\right]$$

$$= -\frac{1}{\lambda}\left[p \cdot ln(1-p) - [x + ln(1-x)]_0^p\right]$$

$$= -\frac{1}{\lambda}[p \cdot ln(1-p) - [p + ln(1-p)]$$

$$= -\frac{1}{\lambda}[p \cdot ln(1-p) - p - ln(1-p)]$$

$$= -\frac{1}{\lambda}[-p - (1-p)ln(1-p)]$$

Τώρα υπενθυμίζουμε ότι η μέση τιμή της Εκθετικής κατανομής με παράμετρο λ>0 είναι ίση με $\mu = E(X) = \frac{1}{\lambda}$. Επομένως με βάση τον τύπο (2.3) η καμπύλη Lorenz θα είναι ίση με:

$$L(p) = \frac{1}{\mu} \cdot \int_0^p F^{-1}(x)dx =$$

$$=\frac{1}{\frac{1}{\lambda}}\cdot\left(-\frac{1}{\lambda}[-p-(1-p)ln(1-p)]\right)$$

$$=p\ +(1-p)\cdot ln(1-p)$$

Ομοιόμορφη Κατανομή $\boldsymbol{U(a,b)}$

Θα αποδείξουμε τον τύπο της καμπύλης Lorenz για την ομοιόμορφη κατανομή με τιμές στο (a,b). Θυμίζουμε ότι η συνάρτηση πυκνότητας πιθανότητας είναι η $F(x)=\frac{x-\alpha}{b-\alpha}$ **.** Αρχικά θα υπολογίσουμε την συνάρτηση F^{-1}.

$$y=\frac{x-\alpha}{b-\alpha}\ \leftrightarrow x-\alpha=y\cdot(b-\alpha)$$

$$\leftrightarrow x=\alpha(1-y)+yb$$

$$\rightarrow\ F^{-1}(x)=\ \alpha(1-x)+xb$$

Τώρα θα υπολογίσουμε το ολοκλήρωμα :

$$\int_0^p F^{-1}(x)dx=\int_0^p(\ \alpha(1-x)+xb)dx$$

$$=\left[\alpha x-\frac{\alpha x^2}{2}+\frac{bx^2}{2}\right]\begin{matrix}p\\0\end{matrix}$$

$$=\alpha p+\frac{(b-\alpha)p^2}{2}$$

Θυμίζουμε ότι η μέση τιμή της Ομοιόμορφης Κατανομής $\boldsymbol{U(a,b)}$ είναι ίση με $\mu=E(X)=\frac{b+\alpha}{2}$. Επομένως με βάση τον τύπο (2.3) η καμπύλη Lorenz θα ισούται με:

$$L(p)=\frac{1}{\mu}\cdot\int_0^p F^{-1}(x)dx=$$

$$=\frac{1}{\frac{b+\alpha}{2}}\cdot\left(\alpha p+\frac{(b-a)p^2}{2}\right)$$

$$=\frac{2}{b+\alpha}\cdot\left(\alpha p+\frac{(b-\alpha)p^2}{2}\right)$$

$$=\ \frac{2\alpha p+(b-a)p^2}{\alpha+b}$$

Με την ίδια διαδικασία μπορούμε να υπολογίσουμε την καμπύλη Lorenz και για άλλες κατανομές. Στον παρακάτω πίνακα λοιπόν θα παρουσιάσουμε την μορφή που παίρνει ο τύπος της καμπύλης Lorenz για κάποιες από τις πιο γνωστές κατανομές.

Κατανομή	**Συνάρτηση**	**Καμπύλη Lorenz**
Εκθετική	$F(x) = 1 - \exp(-\lambda x)$, $x > 0$ και $\lambda > 0$	$L(p) = p + (1-p)\log(1-p)$
Frechet	$F(x) = \exp\left(-\left(\frac{x}{\lambda}\right)^{-\mathrm{k}}\right)$	$L(p) = \frac{\Gamma\left(1-\frac{1}{k}, (\frac{q}{\lambda})^{-k}\right)}{\Gamma\left(1-\frac{1}{k}\right)}$,όπου q=$\mathrm{F}^{-1}(p)$
Ομοιόμορφη	$F(x) = \frac{x-\alpha}{b-a}$, $\mu\varepsilon\ 0 < \alpha < x < b$	$L(p) = \frac{2\alpha p + (b-a)p^2}{\alpha + b}$
Pareto	$F(x) = 1 - \frac{k^\alpha}{x^\alpha}$, για $x \geq k > 0$ και $\alpha > 0$	$L(p) = 1 - (1-p)^{1-1/\alpha}$
Δυναμοκατανομή	$F(x) = x^\alpha$,για $0 < x < 1$ και $\alpha > 0$	$L(p) = p^{\ 1+1/\alpha}$
Weibull	$F(x) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^{\mathrm{k}}\right)$,για $x > 0\, , k > 0\ \kappa\alpha\iota\ \lambda > 0.$	$L(p) = \frac{\gamma(1+\frac{1}{k},\left(\frac{q}{\lambda}\right)^k)}{\Gamma(1+\frac{1}{k})}$, όπου $q = F^{-1}(p)$
Γενικευμένη εκθετική	$F(x) = 1 - \exp(-\lambda(x-a))$, $x > a$	$L(p) = p + (1+\lambda\alpha)^{-1}(1-p)\log(1-p)$

Πίνακας 1 : Καμπύλη Lorenz γνωστών κατανομών

Ορισμός 2.1: Για μια X συνεχή τυχαία μεταβλητή η οποία έχει καμπύλη Lorenz, $L(p)$**,** ο δείκτης Gini μπορεί να δοθεί από την παρακάτω σχέση:

$$G = 1 - 2\int_0^1 L(p)dp. \qquad (2.4)$$

Ο παραπάνω ορισμός μας δίνει πολύ γρήγορα τον δείκτη Gini αν γνωρίζουμε τον τύπο της καμπύλης Lorenz. Ας δούμε λοιπόν μια εφαρμογή του ορισμού 2.1.

Εφαρμογή του ορισμού 2.1 : Έστω ότι η κατανομή του εισοδήματος μιας χώρας περιγράφεται από την καμπύλη Lorenz , $L(p) = p^{4.3}$, $0 \leq p \leq 1$. Μπορούμε λοιπόν να εφαρμόσουμε τον τύπο (2.4) για να βρούμε την τιμή του δείκτη Gini. Έτσι λαμβάνουμε:

$$G = 1- 2\int_0^1 L(p)dp = 1- 2\int_0^1 p^{4.3}dp = 1 - \frac{\mathbf{2}}{\mathbf{5.3}} = 0.623$$

2.4 Βασικές ιδιότητες της καμπύλης Lorenz

Στο σημείο αυτό θα αναφερθούμε σε κάποιες βασικές ιδιότητες της καμπύλης Lorenz, οι οποίες παραμένουν ίδιες ανεξάρτητα από τα δεδομένα που διαθέτουμε. Αρχικά παίρνει τιμές από το 0 έως το 1, δηλαδή έχουμε ότι $0 \leq L(p) \leq 1$. Επίσης ισχύει ότι $L(0) = 0$ και $L(1) = 1$, συνεπώς οποιαδήποτε καμπύλη Lorenz περνά από τα σημεία με συντεταγμένες (0,0) και (1,1). Ακόμα, είναι αύξουσα και κυρτή.

Από το διάγραμμα της καμπύλης Lorenz μπορούν να προκύψουν ορισμένα μέτρα ανισότητας τα οποία έχουν εφαρμογή στις οικονομικές αναλύσεις. Ένα από αυτά είναι το μερίδιο εισοδήματος του p-ποσοστού των οικονομικά ασθενέστερων το οποίο δίνεται από το $L(p)$. Είναι προφανές ότι όσο αυξάνεται το μερίδιο εισοδήματος η κατανομή γίνεται λιγότερο άνιση για τους οικονομικά ασθενέστερους. Με άλλα λόγια το $L(p_0)$ αντιστοιχεί στο ποσοστό του εισοδήματος του φτωχότερου p_0 ποσοστό του πληθυσμού. Τέλος, από την καμπύλη Lorenz μπορούν να προκύψουν ορισμένα μέτρα κεντρικής τάσης, μέτρα μεταβλητότητας καθώς και μέτρα ασυμμετρίας.

2.5 Πλεονεκτήματα και μειονεκτήματά του δείκτη Gini ως μέτρο ανισότητας

Στην συγκεκριμένη παράγραφο θα αναφερθούμε τόσο στα πλεονεκτήματα όσο και στα μειονεκτήματα του δείκτη Gini.

Πλεονεκτήματα του δείκτη Gini

- Το κύριο πλεονέκτημα του δείκτη Gini είναι ότι αποτελεί ένα μέτρο της ανισότητας, όχι ένα μέτρο του μέσου εισοδήματος ή κάποιας άλλης μεταβλητής που δεν είναι αντιπροσωπευτική για το μεγαλύτερο μέρος του πληθυσμού, όπως είναι το Ακαθάριστο Εγχώριο Προϊόν.
- Χρησιμοποιείται για σύγκριση διαφόρων κατανομών εισοδημάτων διάφορων ομάδων πληθυσμού, συνεπώς είναι ένα πάρα πολύ καλό εργαλείο σε οικονομικές αναλύσεις.
- Ο δείκτης Gini είναι αρκετά απλός ώστε να μπορεί να συγκριθεί μεταξύ των χωρών και να ερμηνευθεί εύκολα. Τα στατιστικά στοιχεία για το Ακαθάριστο Εγχώριο Προϊόν συχνά επικρίνονται καθώς δεν αντιπροσωπεύουν αλλαγές για ολόκληρο τον πληθυσμό, σε αντίθεση με τον Gini ο οποίος δείχνει πως έχει αλλάξει το εισόδημα για τους φτωχούς και τους πλούσιους. Εάν ο δείκτης Gini και το Ακαθάριστο Εγχώριο Προϊόν αυξάνονται τότε η φτώχεια μπορεί να μην βελτιώνεται για τη συντριπτική πλειοψηφία του πληθυσμού.
- Ο συντελεστής Gini μπορεί να χρησιμοποιηθεί για να δείξει πως έχει αλλάξει η κατανομή του εισοδήματος σε μια χώρα σε μια χρονική περίοδο, επομένως είναι δυνατόν να δούμε αν η ανισότητα αυξάνεται η μειώνεται.

Ο συντελεστής Gini μετρά την ανισότητα διαμέσου της δημιουργίας ενός λόγου μεταξύ δύο μεγεθών. Είναι, δηλαδή, ένα απλός αριθμός, που συγκρίνει την εξέλιξη της ανισότητας σε μια χώρα ικανοποιώντας τα παρακάτω κριτήρια:

- **Ανωνυμία**: Δεν μπορούμε να συνάγουμε ποιοι κερδίζουν πολλά ή ποιοι κερδίζουν λίγα, με άλλα λόγια δεν γνωρίζουμε την ταυτότητα των εισοδηματιών.

- **Ανεξαρτησία κλίμακας**: Δεν λαμβάνει υπόψη του το μέγεθος της οικονομίας μιας χώρας, ή αν είναι πλούσια ή φτωχή.
- **Ανεξαρτησία** από το μέγεθος του πληθυσμού κάθε χώρας.
- **Αρχή μετακίνησης:** Αν το εισόδημα μεταφέρεται από ένα πλούσιο άτομο σε ένα φτωχό, η κατανομή του εισοδήματος ακόμα πιο ίση.
  (βλ. Υφαντόπουλος (1990))

Ο συντελεστής Gini λόγω του ότι αναφέρεται σε απόλυτες και όχι σε σχετικές διαφορές ενέχει ορισμένες αδυναμίες.

Μειονεκτήματα δείκτη Gini

- Ο δείκτης Gini όταν μετράται για μια μεγάλη γεωγραφική χώρα θα έχει γενικά ως αποτέλεσμα πολύ υψηλότερο συντελεστή από ότι έχει κάθε περιφέρεια ξεχωριστά. Έτσι οι συντελεστές κάθε χώρας ξεχωριστά μέσα στην Ευρωπαϊκή Ένωση είναι δύσκολο να συγκριθούν με αυτούς των Η.Π.Α.
- Δεύτερο μειονέκτημα είναι ο υπολογισμός του εισοδήματος ενός ατόμου χωρίς να γίνεται αναφορά σε όλη τη διάρκεια της ζωής του κατά την οποία αλλάζει το εισόδημα.
- Ο δείκτης Gini θα δώσει διαφορετικά αποτελέσματα όταν εφαρμόζεται σε άτομα αντί για νοικοκυριά. Όταν διαφορετικοί πληθυσμοί δεν μετρώνται με συνεπείς ορισμούς ,η σύγκριση δεν έχει νόημα.
- Υποστηρίζεται επίσης ότι ο δείκτης Gini είναι πιο ευαίσθητος στα εισοδήματα των μεσαίων κοινωνικών στρωμάτων.
- Όπως για όλα τα στατιστικά στοιχεία , θα υπάρχουν τυχαία σφάλματα στα δεδομένα. Η σημασία του συντελεστή αυτού μειώνεται καθώς τα δεδομένα γίνονται λιγότερο ακριβή. Επίσης, οι χώρες αυτές ενδέχεται να συλλέγουν δεδομένα με διαφορετικό τρόπο, γεγονός που καθιστά δύσκολη τη σύγκριση στατιστικών στοιχείων μεταξύ των χωρών.
- Η ερμηνεία του συντελεστή Gini είναι πολύ δύσκολη όταν οι καμπύλες Lorenz τέμνονται, κυρίως όταν μελετάμε διαχρονικές συγκρίσεις, γιατί δεν μας δείχνουν κατά πόσον οι μεταβιβάσεις εισοδήματος από υψηλά κλιμάκια προς τα κατώτερα μετέβαλλαν την αρχική κατανομή.

- Ένα ακόμη μειονέκτημα του δείκτη Gini είναι ότι δείχνει μόνο την κατανομή του εισοδήματος εντός μιας χώρας, χωρίς να γίνεται λόγος για το πόσο πλούσια είναι μια χώρα. Πιο συγκεκριμένα, σύμφωνα με τα στοιχεία του OECD, οι Η.Π.Α και η Λιθουανία έχουν τον ίδιο συντελεστή Gini , ο οποίος είναι περίπου στο 0.38. Όμως οι Η.Π.Α έχουν κατά κεφαλήν ΑΕΠ που είναι 3,7 φορές μεγαλύτερο από αυτό της Λιθουανίας, ενώ οι μισθοί είναι εξίσου δυσανάλογοι. Συνεπώς καθίσταται σημαντική έλλειψη του δείκτη Gini to γεγονός ότι δεν λαμβάνεται καθόλου υπόψη η ποιότητα ζωής ή γενική οικονομική ευημερία της κάθε χώρας .

2.6 Διαφορετικοί τρόποι υπολογισμού του δείκτη Gini

Αξίζει να τονίσουμε εδώ ότι υπάρχουν αρκετοί διαφορετικοί τρόποι υπολογισμού του δείκτη Gini.Παρακάτω θα δούμε αρχικά την προσέγγιση του δείκτη Gini με βάση την Μέση Διαφορά ,την οποία απέδειξαν οι Kendall and Stuart το 1958. Στη συνέχεια θα αναφερθούμε στην προσέγγιση του δείκτη Gini μέσα από την εντροπία του Tsallis , όπως επίσης και σε κάποιες άλλες μεθόδους.

Προσέγγιση του δείκτη gini με βάση την μέση διάφορα

Ορισμός 2.2: Έστω $y_{1,\dots,}y_n$ ανεξάρτητες τυχαίες μεταβλητές οι οποίες ακολουθούν μια κατανομή με συνάρτηση κατανομής $F(y)$. Ενδεικτικά μπορούν τα y_i να αντιπροσωπεύουν τις αθροιστικές αναλογίες (cumulative proportions) του εισοδήματος για έναν πληθυσμό.Η μέση διαφορά για μια συνεχή κατανομή δίνεται από τον τύπο:

$$\Delta = E\left|y_i - y_j\right| \tag{2.5}$$

Τώρα οι Kendall και Stuart(1958) όρισαν τον δείκτη Gini ως «το μισό του συντελεστή συγκέντρωσης», ο οποίος σήμερα είναι γνωστός ως η μέση σχετική διαφορά . Η μέση σχετική διαφορά για μια συνεχή κατανομή δίνεται από τον τύπο:

$$\frac{\Delta}{\mu}, \tag{2.6}$$

Όπου μ η μέση τιμή της F. Για μια διακριτή κατανομή τώρα λαμβάνουμε το Δ, την απόλυτη μέση διαφορά από την παρακάτω σχέση:

$$\Delta=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|y_i-y_j\right|, \tag{2.7}$$

όπου το n είναι ο αριθμός των αθροιστικών αναλογιών εισοδήματος, δηλαδή με άλλα λόγια το n ισούται με τον αριθμό των ατόμων που συμμετέχουν στην έρευνα.

Επομένως με βάση τα παραπάνω ο δείκτης Gini τόσο για μια συνεχή κατανομή όσο και για μια διακριτή κατανομή θα ισούται με :

$$\text{Gini} = \frac{\Delta}{2\mu} \tag{2.8}$$

Προσέγγιση του δείκτη Gini με βάση την αθροιστική παρελθοντική εντροπία του Tsallis $C\xi_\alpha(X)$

Σε αυτήν την παράγραφο θα δούμε πως σχετίζεται ο χρηματοοικονομικός δείκτης Gini με την αθροιστική παρελθοντική εντροπία του Tsallis τάξης α $C\xi_\alpha(X)$,την οποία ορίσαμε στο πρώτο κεφάλαιο της παρούσας εργασίας.

Παρατηρούμε ότι αν στο τύπο (2.4) θέσουμε α=2 ,τότε διαιρώντας με την μέση τιμή μ της τυχαίας μεταβλητής X , παίρνουμε τον δείκτη Gini , ένα μέτρο οικονομικής ανισότητας ,ο οποίος ορίζεται ως εξής:

$$Gini(X) = \frac{C\xi_2(X)}{\mu} \ . \tag{2.9}$$

Προσέγγιση του δείκτη Gini μέσα από την εμπειρική κατανομή

Σε ορισμένες περιπτώσεις, η εξίσωση (2.10) μπορεί να εφαρμοστεί για τον υπολογισμό του συντελεστή Gini χωρίς άμεση αναφορά στην καμπύλη Lorenz. Για παράδειγμα, σε περίπτωση που η κατανομή μας είναι διακριτή, δηλαδή αποτελείται από παρατηρήσεις $y_1, y_2, y_3, \ldots, y_n$ τις οποίες έχουμε διατάξει κατά αύξουσα σειρά

$y_1 \leq \cdots \leq y_n$ και με n συμβολίζεται ο αριθμός των ατόμων που συμμετέχουν στην έρευνα ο δείκτης Gini μπορεί να υπολογιστεί από τον παρακάτω τύπο:

$$G = \frac{1}{n}\left(n + 1 - 2\,\frac{\sum_{i=1}^{n}(n+1-i)y_i}{\sum_{i=1}^{n} y_i}\right) \tag{2.10}$$

Προσέγγιση του δείκτη Gini με βάση την συνδιασπορά

Μια άλλη διαφορετική προσέγγιση του δείκτη Gini βασίζεται στην συνδιασπορά (covariance) και την παρουσίασε πρώτη φορά ο Anand(1983). Πιο συγκεκριμένα απέδειξε ότι ο δείκτης Gini δίνεται από την ακόλουθη σχέση :

$$G = \frac{2\mathrm{cov}\,(y_i, i)}{N\mu},$$

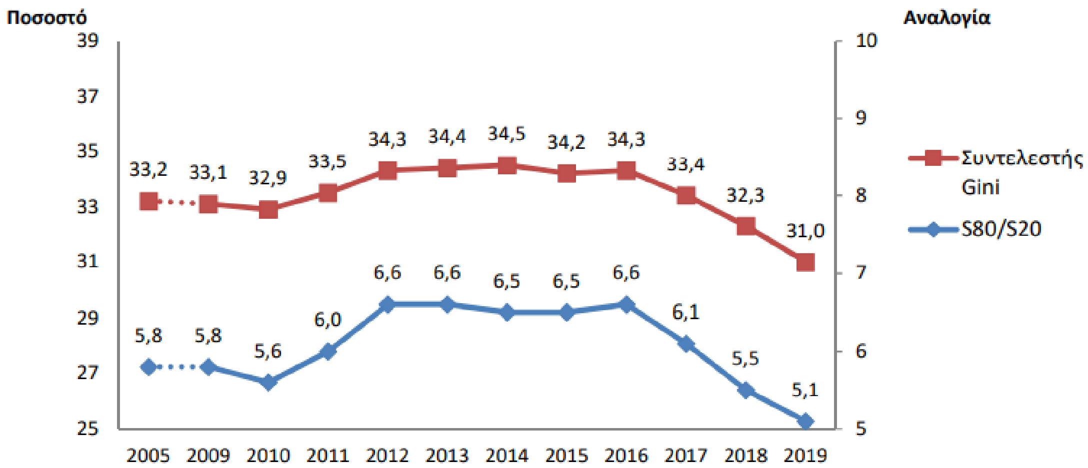
Όπου $cov(y_i, i)$ η συνδιασπορά του εισοδηματος y_i με την τάξη του i ατόμου και μ η μέση τιμή του συνολικού εισοδήματος. Ως τάξη θεωρείται ο αριθμός εκείνος από 1 έως και n που υποδηλώνει με το 1 το φτωχότερο άτομο και με n το πλουσιότερο. Τέλος το N αντιστοιχεί στο σύνολο των ατόμων.

2.7 Ο δείκτης Gini ανά τον κόσμο

Σε αυτήν την ενότητα θα μελετήσουμε την τιμή του δείκτη Gini ανά τον κόσμο. Μερικές από τις φτωχότερες χώρες του κόσμου έχουν μερικούς από τους υψηλότερους συντελεστές Gini, ενώ πολλοί από τους χαμηλότερους συντελεστές βρίσκονται σε πιο πλούσιες ευρωπαϊκές χώρες. Η Νότια Αφρική κατατάσσεται ως η χώρα με το χαμηλότερο επίπεδο εισοδήματος στον κόσμο, χάρη στον συντελεστή Gini 63% όταν μετρήθηκε τελευταία φορά το 2014. Αξίζει να επισημάνουμε ότι το 2005 ήταν ακόμη υψηλότερος 65%.Στη Νότια Αφρική , το πλουσιότερο 10% κατέχει μόλις το 7% του πλούτου. Επιπλέον, περισσότερο από το ήμισυ του πληθυσμού της Νότιας Αφρικής ζει σε συνθήκες φτώχειας.

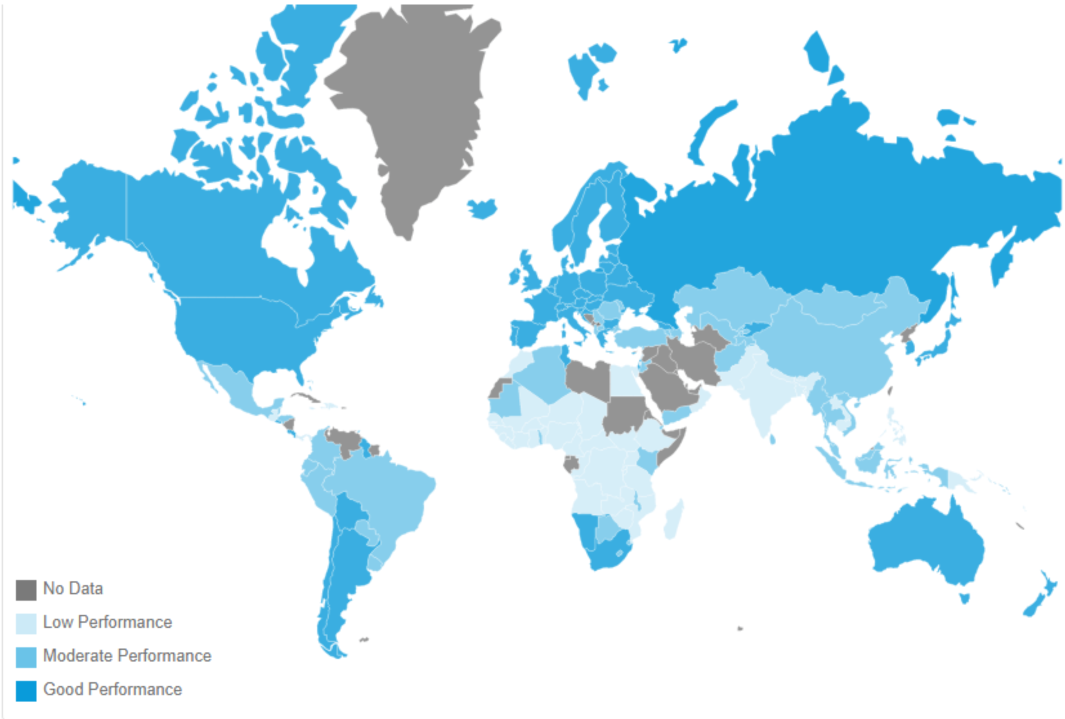
Οι Ηνωμένες Πολιτείες της Αμερικής έχουν συντελεστή Gini 41.1%. Το 2015, το 1% των μεγαλύτερων εισοδημάτων είχε κατά μέσο όρο 40 φορές μεγαλύτερο εισόδημα από το χαμηλότερο 90%. Στις Η.Π.Α , η φτώχεια είναι ένα αυξανόμενο ζήτημα καθώς εκτιμάται ότι είναι ένα 12,3%- 17,8 % που ζει κάτω από το επίπεδο της φτώχειας. Πολλοί από αυτούς τους χαμηλόμισθους εργάτες ζουν χωρίς ασφάλιση υγείας, σύνταξη και χωρίς σταθερό μεροκάματο.

Για την ΕΕ, δείκτης Gini κυμαίνεται γύρω στο 31%(2015), παρουσιάζοντας αύξηση από το 30,5% σε 31%. Με άλλα λόγια παρατηρείται αύξηση της εισοδηματικής ανισότητας στην ΕΕ. Τα πιο αναπτυγμένα Ευρωπαϊκά κράτη έχουν συντελεστή μεταξύ 0.24 και 0.36. Η χώρα μας έχει έναν από τους μεγαλύτερους δείκτες στην Ευρώπη. Ο δείκτης Gini της Ελλάδας μετά το 2008 αυξήθηκε ταχύτητα, φτάνοντας το 2012 στο 36,75%, ενώ μετά από μια μικρή μείωση το 2019 αυξήθηκε και πάλι. Δεδομένου της οικονομικής κρίσης που πέρασε η Ελλάδα ,είναι απόλυτα λογικό να υπάρχει εισοδηματική ανισότητα, μιας και τα περισσότερα εισοδήματα έχουν μειωθεί αισθητά. Άλλες χώρες με μεγάλο συντελεστή Gini είναι η Βουλγαρία, η Ρουμανία, η Πορτογαλία, η Λιθουανία και η Λετονία.



Σχήμα 2.5: Ο δείκτης Gini και ο δείκτης S80/S20 της Ελλάδας (2005, 2009-2019). Πηγή: ΕΛΣΤΑΤ

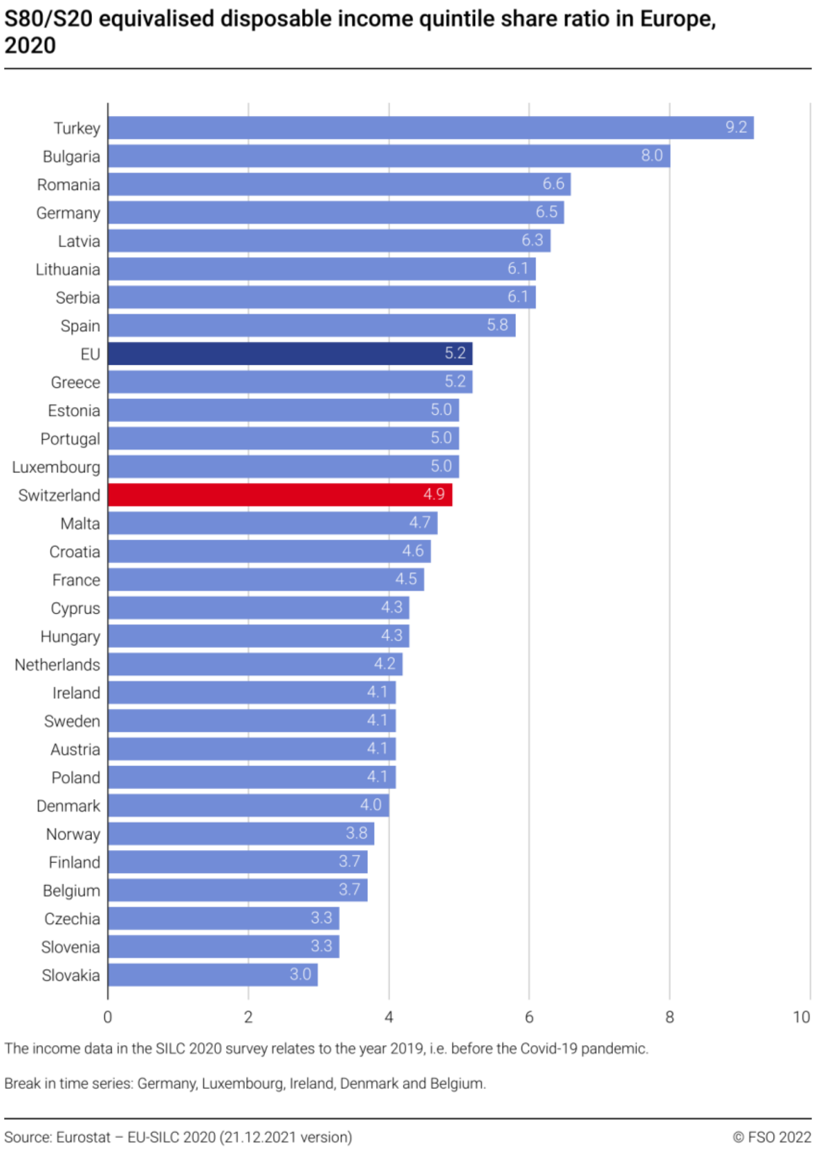
Ο συντελεστής Gini σημείωσε σταθερή ανάπτυξη κατά τον $19^{ο}$ και τον $20^{ο}$ αιώνα. Το 1820, ο παγκόσμιος συντελεστής Gini ήταν 0.50 και το 1980 και το 1992 ήταν 0.657. Σύμφωνα με την έκθεση της Παγκόσμιας Τράπεζας για τη Φτώχεια και την Κοινή Ευημερία 2020 , ο δείκτης Gini αυξάνεται περίπου 1,5 μονάδα στα 5 χρόνια μετά τις μεγάλες επιδημίες, όπως ο H1N1(2009), ο Έμπολα(2014) και ο Ζίκα(2016). Ενώ οι επιπτώσεις της πανδημίας COVID-19 εξακολουθούν να υπολογίζονται , οι πρώτες εκτιμήσεις προέβλεπαν αύξηση κατά 1,2- 1,9 ποσοστιαίες μονάδες ετησίως για το 2020 και το 2021, σηματοδοτώντας σημαντική αύξηση της εισοδηματικής ανισότητας. Επιπλέον για το 2022 με τον πόλεμο στην Ουκρανία αναμένεται εκτίναξη του δείκτη, μετά τις ραγδαίες αυξήσεις στις τιμές της ενέργειας.



Σχήμα 2.6: Ο δείκτης Gini ανά τον κόσμο

Πηγή :Google

Κλείνοντας την συγκεκριμένη παράγραφο , αξίζει να αναφερθούμε και στο γεγονός ότι η εισοδηματική ανισότητα μπορεί να μετρηθεί και με άλλους δείκτες. Ο συνηθέστερα χρησιμοποιούμενος δείκτης μετά τον δείκτη Gini είναι ο λόγος S80/S20. Ο συγκεκριμένος λόγος υπολογίζει το ετήσιο εισόδημα του 20% των πλουσιότερων νοικοκυριών σε σχέση με το 20% των φτωχότερων νοικοκυριών. Σε όλες τις ευρωπαϊκές μελέτες για την ανισότητα θα συναντήσουμε τον λόγο S80/S20 . Συνεπώς αν σε κάποιο άρθρο δούμε ότι ο λόγος S80/S20 είναι ίσος με 4,0 αυτό μεταφράζεται ως εξής : το ετήσιο εισόδημα του 20% των πλουσιότερων νοικοκυριών είναι τετραπλάσιο από το αντίστοιχο εισόδημα του 20% των φτωχότερων νοικοκυριών. Επιπλέον, υψηλές τιμές του δείκτη αντιστοιχούν σε μεγαλύτερη εισοδηματική ανισότητα, ενώ χαμηλές τιμές ισοδυναμούν με μικρότερη ανισότητα. Όταν S80/S20 ισούται με 1, υπάρχει τέλεια εισοδηματική ισότητα. Για την ΕΕ, ο λόγος αυτός κυμαίνεται γύρω στο 5,1(2015). Στο σχήμα που ακολουθεί παρουσιάζονται οι τιμές του δείκτη S80/S20 για όλα τα κράτη Ευρώπης το 2020. Η Ελλάδα έχει τιμή 5.2 ,συνεπώς το ετήσιο εισόδημα του 20% των πλουσιότερων νοικοκυριών είναι περίπου πενταπλάσιο από το αντίστοιχο εισόδημα του 20% των φτωχότερων νοικοκυριών. Η χώρα με τον μικρότερο δείκτη S80/S20 είναι η Σλοβακία.



Σχήμα 2.7: Τιμές του S80/S20 για την Ευρώπη.

Πηγή : Eurostat

ΚΕΦΑΛΑΙΟ 3

Δείκτης Bonferroni

3.1 Ο δείκτης Bonferroni

Ένας άλλος σημαντικός δείκτης ανισότητας εισοδηματικής είναι ο δείκτης Bonferroni. Ο Carlo Emilio Bonferroni (1930) πρότεινε τον δείκτη ανισότητας Bonferroni ως εναλλακτική του δείκτη Gini. Για περίπου μισό αιώνα, ο δείκτης Bonferroni παρέμεινε ξεχασμένος επειδή εξοστρακίστηκε από τον Gini και τους οπαδούς του.

Ορισμός 3.1: Ο δείκτης Bonferroni(1930) ορίζεται ως μια συνάρτηση των μερικών μέσων:

$$B = \frac{1}{N-1}\sum_{\mathrm{i}=1}^{N}\frac{(\mu-\mu_{\mathrm{i}})}{\mu} \qquad (3.1)$$

Όπου N ο πληθυσμός των ατόμων (με $N\geq 2$) και i είναι η κατάταξη εντός του παρατηρούμενου πληθυσμού κατά αύξουσα τιμή εισοδήματος, ενώ x_i είναι το εισόδημα του i-οστού ατόμου του πληθυσμού , ισχύει ότι $0\leq x_1 \leq \cdots \leq x_n$. Επιπλέον με $\sum_{i=1}^{N} x_i$ συμβολίζεται το συνολικό εισόδημα όλου του πληθυσμού ενώ από την άλλη με $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$ ο μέσος όρος του συνολικού εισοδήματος.

Ο παραπάνω τύπος μπορεί να γραφεί επίσης και ως εξής:

$$\mathrm{B} = \frac{1}{\mu}\left(\mu - \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \mu_i\right) \qquad (3.2)$$

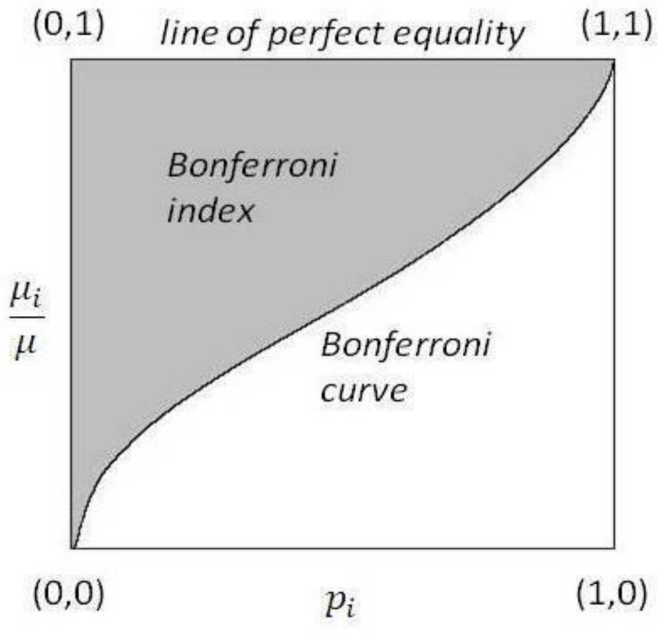
Ο δείκτης Bonferroni παίρνει τιμές μεταξύ 0 και 1, δηλαδή έχουμε ότι $0 \leq B\leq 1$. Επίσης, δίνει μεγαλύτερη βαρύτητα σε μονάδες με χαμηλότερο εισόδημα. Για το λόγο αυτό, ο δείκτης Bonferroni είναι πιο ευαίσθητος σε χαμηλότερα επίπεδα κατανομής. Ακόμα ο δείκτης Bonferroni είναι συμμετρικός.

3.2 Η καμπύλη Bonferroni

Σε αυτήν την παράγραφο θα μελετήσουμε την καμπύλη Bonferroni ,την οποία εισήγαγε πρώτος ο Bonferroni (1930), ενώ πολλοί ακόμη την έχουν μελετήσει όπως οι De Vergottini (1940) , Tarsitano(1990) και ο Zenga (2013).

Ορισμός 3.2: Έστω X μια μη αρνητική συνεχής τυχαία μεταβλητή με θετική και πεπερασμένη μέση τιμή μ και συνάρτηση κατανομής F. Η καμπύλη Bonferroni της X ορίζεται ως:

$$B(p)=\frac{1}{p\mu}\int_0^p F^{-1}(t)dt \text{ , για κάθε } p\in(0,1] \qquad (3.3)$$



Σχήμα 3.1: Η καμπύλη Bonferroni.

Πηγή: ResearchGate

Στον κατακόρυφο άξονα απεικονίζονται οι λόγοι $\frac{\mu_i}{\mu}$, ενώ στον οριζόντιο άξονα αναπαριστάνεται τα αθροιστικά μερίδια του πληθυσμού p_i . Η πολυγωνική γραμμή που ενώνει τα σημεία (p_i, $\frac{\mu_i}{\mu}$) στο παραπάνω γράφημα είναι η καμπύλη Bonferroni. Εάν όλα τα άτομα που συμμετέχουν στην έρευνα έχουν το ίδιο εισόδημα (δηλαδή ίσο με μ), η καμπύλη Bonferroni συμπίπτει με την γραμμή της τέλειας ισότητας, η οποία ενώνει τα σημεία (0,1) και (1,1). Αν μόνο ένα άτομο κατέχει το συνολικό εισόδημα , η καμπύλη Bonferroni είναι η διακεκομμένη γραμμή που ενώνει τα σημεία (0,0) , $(\frac{N-1}{N},0)$, $(\frac{N-1}{N},1)$.

Ακολούθως θα δούμε τον εύκολο τρόπο του υπολογισμού του δείκτη Bonferroni όταν γνωρίζουμε την καμπύλη Bonferroni. Ο δείκτης ανισότητας Bonferroni από γεωμετρική άποψη αντιπροσωπεύει την περιοχή μεταξύ της καμπύλη Bonferroni και της γραμμής της τέλειας ισότητας στο ενιαίο τετράγωνο.

Ορισμός 3.3: Έστω X μια μη αρνητική συνεχής τυχαία μεταβλητή με καμπύλη Bonferroni $B(p)$. Ο δείκτης Bonferroni δίνεται από τον τύπο:

$$B = 1-\int_0^1 B(p)dp \tag{3.4}$$

Εφαρμογή του ορισμού 3.3: Έστω ότι η κατανομή του εισοδήματος μιας χώρας περιγράφεται από την καμπύλη Bonferroni , $B(p) = p^2$, $0 \leq p \leq 1$. Μπορούμε λοιπόν να εφαρμόσουμε τον τύπο (3.4) για να βρούμε την τιμή του δείκτη Bonferroni. Έτσι λαμβάνουμε:

$$B = 1- \int_0^1 B(p)dp= 1- \int_0^1 p^2 dp=0.66$$

Η καμπύλη Bonferroni συγκρίνει τη μέση τιμή του εισοδήματος της κατώτερης ομάδας με το συνολικό μέσο εισόδημα όλου του πληθυσμού. Με άλλα λόγια, εάν η τυχαία μεταβλητή X αντιπροσωπεύει το εισόδημα και η $B(p)$ είναι η αντίστοιχη καμπύλη Bonferroni , τότε το $B(p_0) = B_0$ σημαίνει ότι το μέσο εισόδημα της αναλογίας <<κάτω>> του ποσοστού p_0 του πληθυσμού είναι ίσο με B_0 φορές το μέσο εισόδημα του ολόκληρου πληθυσμού.

Βασικές ιδιότητες της καμπύλης Bonferroni

Στο σημείο αυτό θα αναφερθούμε σε κάποιες βασικές ιδιότητες της καμπύλης Bonferroni, οι οποίες παραμένουν ίδιες ανεξάρτητα από τα δεδομένα που διαθέτουμε. Αρχικά παίρνει τιμές από το 0 έως το 1, ενώ επίσης ισχύει ότι $\mathrm{B}(1) = 1$ και B(0)=0. Πιο συγκεκριμένα ο δείκτης Bonferroni παίρνει την τιμή 1 όταν έστω ένα εισόδημα είναι θετικό. Είναι αύξουσα αλλά σε αντίθεση με την καμπύλη Lorenz, η καμπύλη Bonferroni δεν είναι απαραίτητα κυρτή. Η κυρτότητα της καμπύλης Bonferroni εξαρτάται από την κυρτότητα της κατανομής της τυχαίας μεταβλητής X. Εάν η τυχαία μεταβλητή X τείνει στην κατάσταση της ελάχιστης ανισότητας, τότε η αντίστοιχη καμπύλη Bonferroni τείνει να ισούται με 1, ενώ ο αντίστοιχος δείκτης Bonferroni θα είναι ίσος με 0.

Σύνδεση καμπύλης Bonferroni και καμπύλης Lorenz

Αν $L(p)$ η καμπύλη Lorenz της τυχαίας μεταβλητής X, τότε η καμπύλη Bonferroni δίνεται από την σχέση :

$$B(p)= \frac{L(p)}{p}\text{, για κάθε } p\epsilon\ (0,1] \tag{3.5}$$

Αν και από τον παραπάνω τύπο, φαίνεται ότι η καμπύλη Bonferroni αντιπροσωπεύει την ανισότητα σε έναν ισοδύναμο τρόπο με την καμπύλη Lorenz, εντούτοις οι πληροφορίες που δίνουν οι 2 καμπύλες είναι διαφορετικές. Οι τιμές της $L(p)$ είναι κλάσματα του συνολικού εισοδήματος, ενώ οι τιμές της $B(p)$ αναφέρονται σε σχετικά επίπεδα εισοδήματος.

3.3 Καμπύλη Bonferroni για γνωστές κατανομές

Σε αυτήν την ενότητα, θα υπολογίσουμε εύκολα και γρήγορα με τον τύπο την καμπύλη Bonferronni για κάποιες γνωστές κατανομές.

Εκθετική Κατανομή με παράμετρο $\lambda > 0$

Παρακάτω θα αποδείξουμε τον τύπο της καμπύλης Bonferroni για την εκθετική κατανομή $\lambda > 0$**.** Στην προηγούμενη παράγραφο είδαμε ότι η καμπύλη Lorenz της εκθετικής κατανομής ισούται με :

$$L(p) = p + (1-p)log(1-p).$$

Έτσι μέσα από τον τύπο (3.5) λαμβάνουμε ότι:

$$B(p) = \frac{L(p)}{p} = \frac{\mathrm{p}+(1-\mathrm{p})\log(1-\mathrm{p})}{\mathrm{p}} = 1 + \frac{(1-p)}{p}\log(1-p)$$

Ομοιόμορφη Κατανομή $U(0, \alpha)$

Θα αποδείξουμε τον τύπο της καμπύλης Bonferroni για την ομοιόμορφη κατανομή με τιμές στο $(\mathbf{0}, \boldsymbol{\alpha})$.Στην προηγούμενη παράγραφο είδαμε ότι η καμπύλη Lorenz της εκθετικής κατανομής ισούται με :

$$L(p) = \frac{2\alpha p + (b-a)p^2}{\alpha + b}$$

Έτσι μέσα από τον τύπο (3.5) λαμβάνουμε ότι:

$$B(p) = \frac{L(p)}{p} = \frac{2\alpha+(b-a)p}{\alpha+b}$$

Με την ίδια διαδικασία λοιπόν μπορούμε να βρούμε εύκολα την καμπύλη Bonferroni.Στον πίνακα που ακολουθεί παραθέτουμε την μορφή της καμπύλης αυτής για αρκετές γνωστές κατανομές, όπως υπολογίστηκε από τους Giorgi και Nadarahaj (2010).

Κατανομή	**Συνάρτηση**	**Καμπύλη Bonferroni**
Εκθετική	$F(x) = 1 - \exp(-\lambda x)$, $x > 0$ και $\lambda > 0$	$B(p) = 1 + \frac{(1-p)}{p}\log(1-p)$
Frechet	$F(x) = \exp\left(-\left(\frac{x}{\lambda}\right)^{-\mathrm{k}}\right)$	$B(p) = \frac{\Gamma\left(1-\frac{1}{k}, (\frac{q}{\lambda})^{-k}\right)}{p\Gamma\left(1-\frac{1}{k}\right)}$,όπου $q=F^{-1}(p)$
Ομοιόμορφη	$F(x) = \frac{x-\alpha}{b-\alpha}, \mu\varepsilon\ 0 < \alpha < x < b$	$B(p) = \frac{2\alpha + (b-a)p}{\alpha + b}$
Pareto	$F(x) = 1 - \frac{k^\alpha}{x^\alpha}$, για $x \geq k > 0$ και $\alpha > 0$	$B(p) = \frac{1 - (1-p)^{1-1/\alpha}}{p}$
Δυναμοκατανομή	$F(x) = x^\alpha$,για $0 < x < 1$ και $\alpha > 0$	$B(p) = p^{\ 1/\alpha}$
Weibull	$F(x) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^{\mathrm{k}}\right)$,για $x > 0\,, k > 0\ \kappa\alpha\iota\ \lambda > 0.$	$B(p) = \frac{\gamma(1+\frac{1}{k},\left(\frac{q}{\lambda}\right)^{k})}{p\Gamma(1+\frac{1}{k})}$, όπου $q = F^{-1}(p)$
Γενικευμένη εκθετική κατανομή	$F(x) = 1 - \exp(-\lambda(x-a))$, $x > a$	$B(p) = p + (1+\lambda\alpha)^{-1}(1-p)\log(1-p)$

Πίνακας 2: Καμπύλη Bonferroni γνωστών κατανομών

3.4 Σχέση Κανονικοποιημένης αθροιστικής παρελθοντικής εντροπίας και καμπύλης Bonferroni

Σε αυτήν την ενότητα θα δούμε τον τρόπο υπολογισμού της καμπύλης Bonferroni μέσα από την κανονικοποιημένη αθροιστική παρελθοντική εντροπία, την οποία μελετήσαμε στο πρώτο κεφάλαιο της παρούσας διπλωματικής εργασίας. Μια πολύ σημαντική ιδιότητα της κανονικοποιημένης αθροιστικής παρελθοντικής εντροπίας ,$\mathcal{NCE}$ (X), είναι η ακόλουθη:

Για μη αρνητική τυχαία μεταβλητή με πεπερασμένη μέση τιμή η κανονικοποιημένη αθροιστική παρελθοντική εντροπία εκφράζεται ως εξής:

$$\mathcal{NCE}(X)=1-E\{B_F[F(X)]\}\text{ , όπου } B_F \text{ η καμπύλη Bonferroni} \qquad (3.6)$$

3.5 Σύγκριση δείκτη Bonferroni και δείκτη Gini

Σε αυτήν την παράγραφο θα μελετήσουμε την σχέση του δείκτη Gini με τον δείκτη Bonferroni. Πιο συγκεκριμένα θα αποδείξουμε ότι ο δείκτης Bonferroni είναι πάντα πιο μεγάλος από τον δείκτη Gini.

Στο Κεφάλαιο 1 (βλέπε Παράγραφο 1.6) δείξαμε ότι ο δείκτη Gini δίνεται από την εξής σχέση:

$$\text{Gini } = \frac{\int_0^\infty F(x)(1-F(x))dx}{\mu} \qquad (1.16)$$

Αντίστοιχα ο Bonferroni πρότεινε ως μια εναλλακτική στον δείκτη Gini τον τύπο που ακολουθεί:

$$\text{Bonferroni}= \frac{-\int_0^\infty F(x)\ln(F(x))dx}{\mu} \qquad (3.7)$$

Θα συγκρίνουμε λοιπόν την σχέση (1.16) με την σχέση (3.7) με την βοήθεια της λογαριθμικής ανισότητας την οποία πρώτα θα αποδείξουμε:

$$t \le -ln(1-t)\text{ , όπου } t<1 \qquad (3.8)$$

Απόδειξη.

$$t\cdot 1 \ \le -ln(1-t) \leftrightarrow \text{ , } t<1$$

$$t\cdot lne \le -ln(1-t) \leftrightarrow$$

$$lne^t \le \ ln(1-t)^{-1} \leftrightarrow$$

$$lne^t \le ln\frac{1}{1-t} \leftrightarrow (ln \text{ γνησίως αύξουσα})$$

$$e^t \leq \frac{1}{1-t} \leftrightarrow (1-t>0)$$

$$e^t\,(1-t) \leq 1 \qquad (3.9)$$

Ορίζουμε την $f(t)$=$e^t(1-t)$ στο $(-\infty,1)$, η οποία είναι συνεχής ως πράξεις συνεχών στο πεδίο ορισμού της και παραγωγίσιμη ως πράξεις παραγωγίσιμων στο πεδίο ορισμού της. Παρατηρούμε ότι $f(0)$=1 (3.10). Άρα από (3.9) και (3.10) προκύπτει η παρακάτω σχέση:

$$f(t) \leq f(0)\,, \qquad (3.11)$$

που είναι ισοδύναμη με την ζητούμενη.

Μελέτη της f ως προς μονοτονία και ακρότατα.

$$f'(t) = [e^t(1-t)\,]' = -te^t \qquad (3.12)$$

Ρίζες της $f'(t)$ =0 στο (-$\infty, 1$) :

Μοναδική ρίζα η $t=0$

Πρόσημο της $f'(t)$ και μονοτονία της f :

- για $t<0 \rightarrow -te^t > 0 \rightarrow f'(t) > 0$ και αφού η f συνεχής είναι γνησίως αύξουσα στο (-∞,0].
- για $0<t<1 \rightarrow -te^t < 0 \rightarrow f'(t) < 0$ και αφού η f συνεχής είναι γνησίως φθίνουσα στο [0,1).

Άρα το $f(0)=1$ ολικό μέγιστο ,δηλαδή $f(t) \leq f(0)$ που είναι το ζητούμενο. Η απόδειξη ολοκληρώθηκε.■

Συνεχίζοντας τώρα στη σχέση 3.8 που ήδη αποδείξαμε θέτουμε όπου t την συνάρτηση επιβίωσης $\bar{F}(x)$ της τυχαίας μεταβλητής X.

Έτσι λαμβάνουμε ότι:

$$\bar{F}(x) \leq -ln(1-\bar{F}(x)) \leftrightarrow$$

$$1-F(x) \leq -ln(F(x)) \leftrightarrow$$

$$F(x) \cdot (1 - F(x)) \leq -F(x) lnF(x) \rightarrow$$

$$\int_0^\infty F(x) \cdot (1 - F(x)) \leq -\int_0^\infty F(x) lnF(x) \rightarrow$$

$$\frac{\int_0^\infty F(x)(1-F(x))dx}{\mu} \leq \frac{-\int_0^\infty F(x)\ln(F(x))dx}{\mu} \rightarrow$$

$$Gini \leq \ Bonferroni$$

Δηλαδή αποδείξαμε ότι ο δείκτης Gini είναι πάντα μικρότερος από τον δείκτη Bonferroni κάτι το οποίο θα εξετάσουμε αν ισχύει και στις εφαρμογές που ακολουθούν στο Κεφάλαιο 4 της εργασίας μας.

Στο συγκεκριμένο Κεφάλαιο παρουσιάσαμε έναν εξίσου σημαντικό δείκτη οικονομικής ανισότητας τον δείκτη Bonferroni. Μελετήσαμε τις ιδιότητές του , όπως επίσης ασχοληθήκαμε διεξοδικά και με την καμπύλη Bonferroni. Τέλος, συγκρίναμε τον δείκτη Gini με τον δείκτη Bonferroni. Στο επόμενο Κεφάλαιο που ακολουθεί θα δούμε στην πράξη την χρησιμότητα των όλων όσων έχουμε αναπτύξει μέχρι τώρα.

ΚΕΦΑΛΑΙΟ 4
Εφαρμογές του δείκτη Gini

4.1 Εφαρμογή του δείκτη Gini σε χρηματιστηριακά δεδομένα

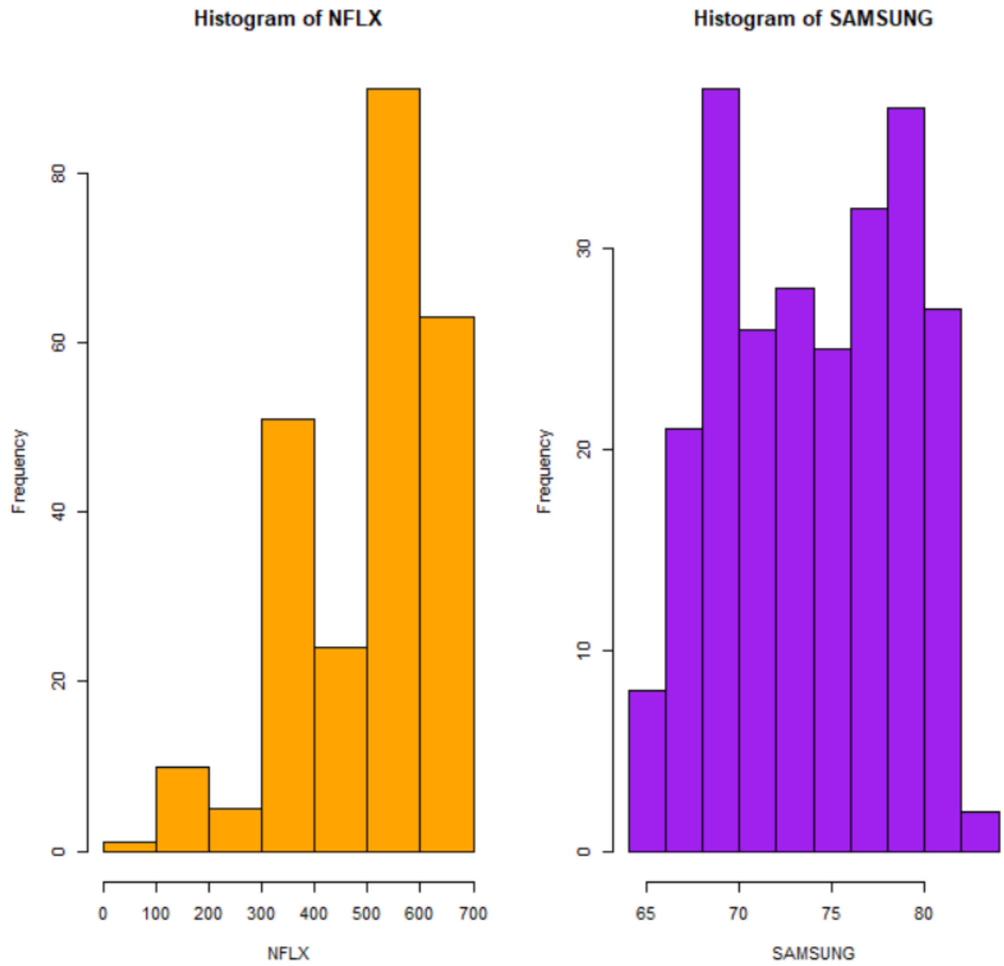
Με βάση τα ιστορικά δεδομένα τιμών των μετοχών (ημερήσιες τιμές κλεισίματος Μάιος 2021- Μάιος 2022) των εταιριών NFXL και Samsung Electronics Co. θα υπολογίσουμε τον δείκτη Gini, την καμπύλη Lorenz και την καμπύλη Bonferronni.Τα δεδομένα που θα χρησιμοποιήσουμε προέρχονται από το yahoo!finance.Η επεξεργασία των δεδομένων έγινε με τη γλώσσα προγραμματισμού R και με τη χρήση της βιβλιοθήκης ineq. Έχουμε 2 μεταβλητές την NFLX και την Samsung.

Αρχικά θα μελετήσουμε τις μετοχές της εταιρείας NFLX και στη συνέχεια τις μετοχές της Samsung. Παρακάτω βλέπουμε τις αποδόσεις των 2 μετοχών μας τις πρώτες 3 ημέρες.

	Date	**SAMSUNG**	**NFLX**
1	25/5/2021	80	503
2	26/5/2021	80	501
3	27/5/2021	80	502

Η μετοχή της NFLX έχει μέση τιμή 501.17 , ενώ η διασπορά της είναι ίση με 16722.88.
Η μετοχή της Samsung αντίστοιχα έχει μέση τιμή 74.50 και διασπορά 22.44.

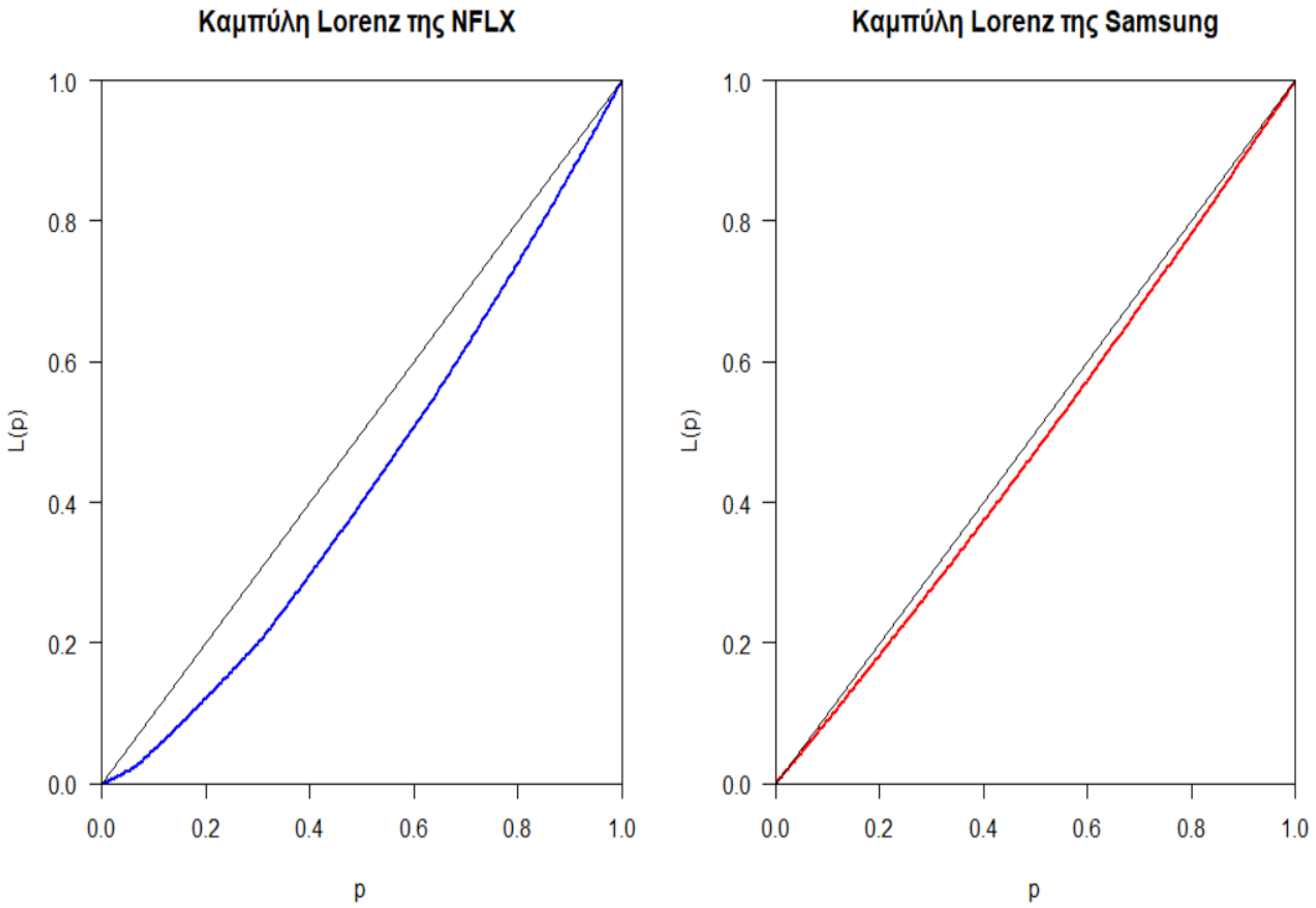
Ακολούθως, παραθέτουμε τα ιστογράμματα των εταιρειών μας. Και στα 2 ιστογράμματα δεν φαίνεται ότι θα μπορούσε να υπάρχει η κωνοειδής καμπύλη της κανονικής κατανομής , οπότε τα δεδομένα μας δεν θα μπορούσαν να ακολουθούν την κανονική κατανομή.



Σχήμα 4.1: Ιστόγραμμα των NFLX, Samsung.

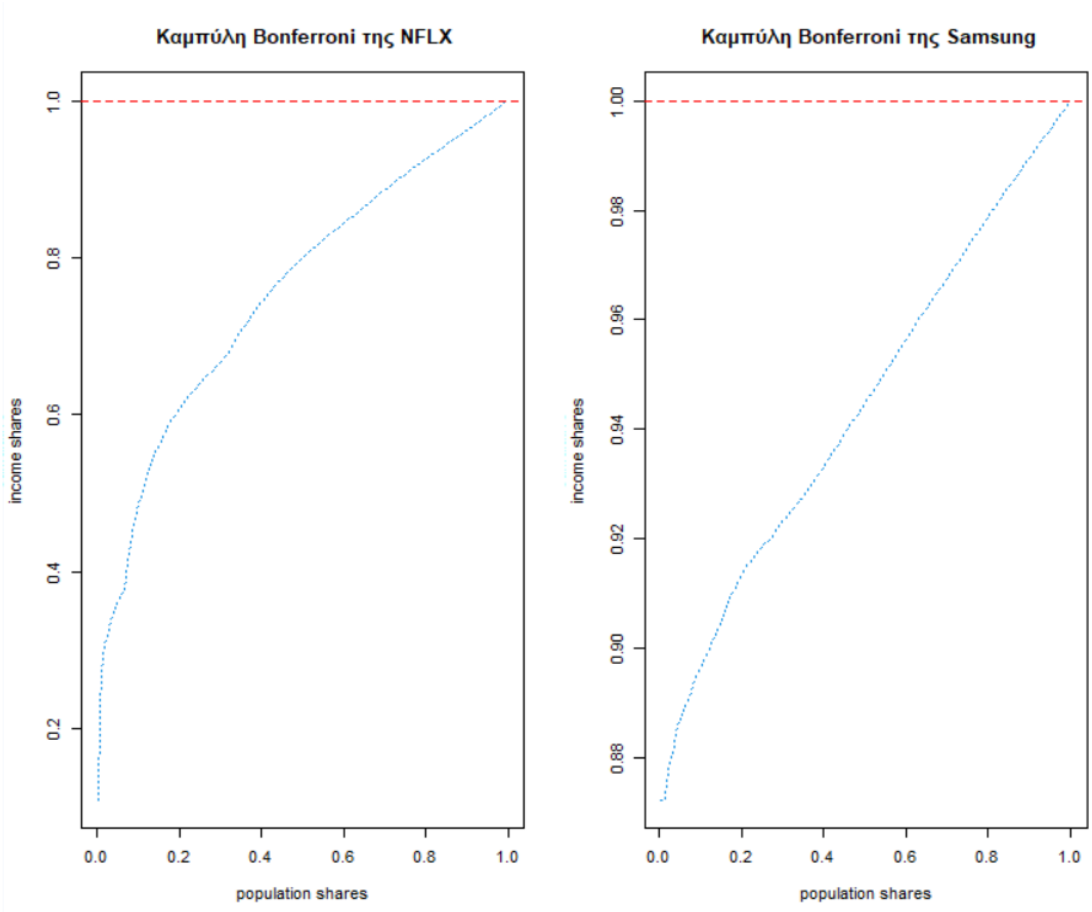
Στη συνέχεια δίνουμε τον δείκτη Gini για τις 2 εταιρείες. Ο δείκτης Gini της NFLX είναι 0.141 , ενώ ο δείκτης Gini της Samsung είναι 0.035. Το ποσοστό αυτό για την NFLX ερμηνεύεται ως εξής: επιλέγοντας 2 τυχαίες ημέρες κλεισίματος, η τιμή των μετοχών θα διαφέρει κατά 14%. Με τον ίδιο τρόπο για την Samsung λαμβάνουμε ότι: επιλέγοντας 2 τυχαίες ημέρες κλεισίματος, η τιμή των μετοχών θα διαφέρει κατά 3%.

Στο σημείο αυτό θα ερμηνεύσουμε την καμπύλη Lorenz των 2 εταιρειών. Στο παρακάτω σχήμα παρατηρούμε ότι η καμπύλη Lorenz για την εταιρεία Samsung (κόκκινη γραμμή) έχει πλησιάσει αρκετά τη γραμμή της ισότητας, σχεδόν συμπίπτει με αυτή , πράγμα που σημαίνει ότι δεν υπάρχει ανισότητα στις τιμές της μετοχής της συγκεκριμένης εταιρείας. Άλλωστε και ο δείκτης Gini είναι 0.03 σχεδόν ίσος με 0. Θυμίζουμε εδώ ότι η τιμή 0 για τον δείκτη Gini ισοδυναμεί με την μέγιστη ισότητα. Αντίθετα η καμπύλη Lorenz της NFLX(μπλε γραμμή) είναι πιο μακριά από την γραμμή της ισότητας. Με άλλα λόγια η μετοχή της Samsung έχει πιο σταθερή απόδοση σε σχέση με την μετοχή της NFLX.



Σχήμα 4.2: Καμπύλη Lorenz NFLX, SAMSUNG.

Στη συνέχεια θα μελετήσουμε τον δείκτη Bonferroni για τις 2 εταιρείες.Η τιμή του δείκτη Bonferroni της NFLX είναι 0.2572823, και της Samsung είναι 0.05569187, ενώ στο παρακάτω σχήμα έχουμε την καμπύλη Bonferroni των μετοχών της NFLX και της Samsung. Φαίνεται από το παρακάτω γράφημα, ότι η καμπύλη Bonferroni της Samsung τείνει να ισούται με 1, ενώ και ο αντίστοιχος δείκτης Bonferroni 0.055 είναι σχεδόν ίσος με μηδέν. Με άλλα λόγια συμπεραίνουμε ότι και οι 2 δείκτες έχουν παρόμοια αποτελέσματα.



Σχήμα 4.3: Καμπύλη Bonferroni NFLX, SAMSUNG.

4.2 Εφαρμογές του δείκτη Gini σε οικονομικά δεδομένα

4.2.1 1η Εφαρμογή του δείκτη Gini σε οικονομικά δεδομένα

Σε αυτήν την παράγραφο θα υπολογίσουμε τον δείκτη Gini των αποδοχών των εργαζομένων στον ιδιωτικό τομέα της Ελλάδας. Τα δεδομένα μας τα οποία παρουσιάζονται στον παρακάτω πίνακα έχουν προκύψει από την επεξεργασία των αναλυτικών περιοδικών δηλώσεων που υποβάλλουν όλοι εργοδότες στην ΕΡΓΑΝΗ. Στην Ελλάδα του 2022 2.278.394 άτομα είναι υπάλληλοι του ιδιωτικού τομέα, ενώ υπάρχουν 291.808 ιδιωτικές επιχειρήσεις. Από τον πίνακα που ακολουθεί συμπεραίνουμε ότι όσο πιο λίγα άτομα απασχολεί μια επιχείρηση, τόσο λιγότερες είναι και οι πιθανότητες ο εργαζόμενος να πάρει έναν μεγάλο μισθό.

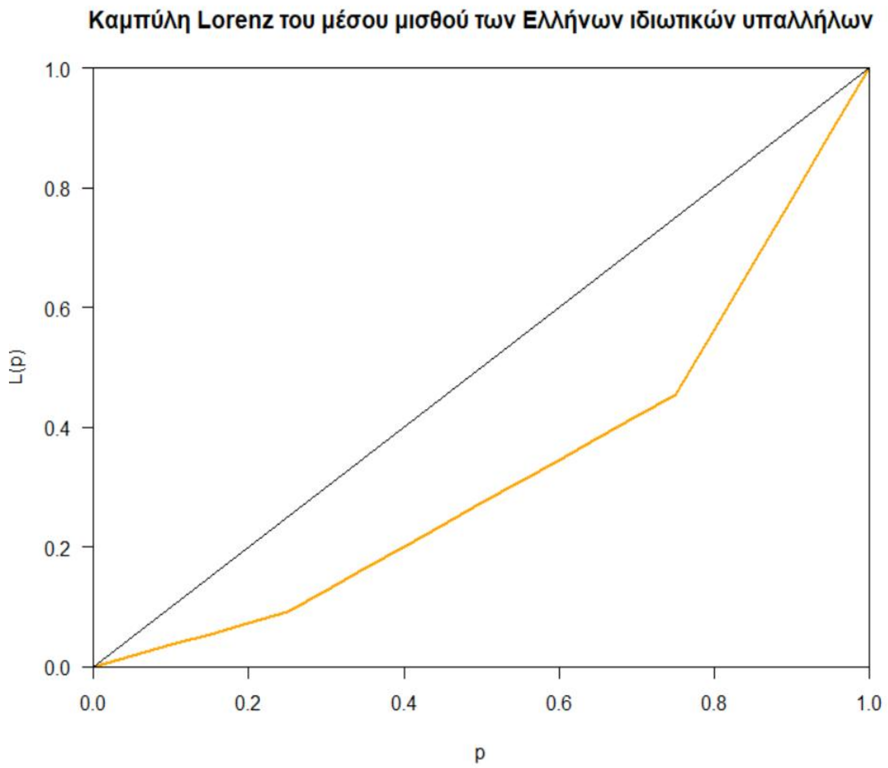
Για παράδειγμα, ο μέσος μισθός στις επιχειρήσεις που απασχολούν έως 10 άτομα προσωπικό διαμορφώνεται μόλις στα 666 ευρώ μεικτά. Στις λίγο μεγαλύτερες επιχειρήσεις στις οποίες εργάζονται από 11 έως 50 άτομα, ο μέσος μισθός ανέρχεται στα 863 ευρώ, ενώ μέσο μισθό άνω των 1.500 ευρώ προσφέρουν μόνο οι εταιρείες στις οποίες εργάζονται από 50 έως 250 άτομα.

Πλήθος εργαζομένων	Αριθμός επιχειρήσεων	Αριθμός εργαζομένων	Μέσος μισθός σε Ευρώ(€)
0-10	257.650	696.219	666,13
11-50	29.439	594.191	863,04
51-250	4004	399.524	1542,65
251-500	413	142.593	1415,34
501-1000	183	128.373	1346,43
1001-1500	55	66.810	1.418,16
1501-2000	20	33.479	1730,25
2001-2500	13	28.873	1492,65
2501-3000	6	16.334	1186,17
3001-3500	6	19.211	1221,20
3501-4000	4	14.906	707,85
>4001	15	137.881	1464,12
ΣΥΝΟΛΟ		**2.278.394**	

Πίνακας 3: Μέσος μισθός Ελλήνων ιδιωτικών υπαλλήλων

Πηγή : Εργάνη

Καταχωρώντας τα δεδομένα στην γλώσσα προγραμματισμού R και με τη βοήθεια της βιβλιοθήκης ineq λαμβάνουμε ότι ο δείκτης Gini ισούται με 0.188 και βλέπουμε και την καμπύλη Lorenz στο σχήμα που έπεται. Τα αποτελέσματα ερμηνεύονται ως εξής: επιλέγοντας 2 τυχαία άτομα του πληθυσμού, ο μισθός τους θα διαφέρει κατά 18% του μέσου ισοδύναμου διαθέσιμου εισοδήματος. Επιπλέον και από την καμπύλη Lorenz φαίνεται να μην υπάρχει ισότητα στην μισθοδοσία των ιδιωτικών υπαλλήλων της χώρας μας. Η διαγώνιος αντιπροσωπεύει μια τέλεια ισότητα, μιας και σε κάθε ποσοστό p του πληθυσμού αντιστοιχεί ποσοστό p του συνολικού εισοδήματος



Σχήμα 4.4: Καμπύλη Lorenz του μέσου μισθού των Ελλήνων ιδιωτικών υπαλλήλων.

Ο κώδικας στη γλώσσα R που χρησιμοποιήθηκε στην συγκεκριμένη εφαρμογή είναι ο παρακάτω:

```
> library(ineq)

> ineq(d, type="Gini")

[1] 0.1887564 # δείκτης Gini

> plot(Lc(d), col="orange", lwd=2, main="Καμπύλη Lorenz του μέσου μισθού των Ελλήνων ιδιωτικών υπαλλήλων"
```

4.2.2 2η Εφαρμογή του δείκτη Gini σε οικονομικά δεδομένα

Σε αυτήν την παράγραφο θα υπολογίσουμε τον δείκτη Gini των αποδοχών των Ελλήνων συνταξιούχων. Τα δεδομένα μας παρουσιάζονται στον παρακάτω πίνακα και προέρχονται από την έκθεση του συστήματος « ΗΛΙΟΣ» για τον Μάρτιο του 2019 και δημοσιεύτηκαν στην εφημερίδα Έθνος. Ο αριθμός των συνταξιούχων της χώρας ανέρχεται σε 2.501.662 και για τις συντάξεις πληρώνονται κάθε μήνα 2,24 δις. Ευρώ. Οι άνδρες συνταξιούχοι είναι περισσότεροι σε σχέση με τις γυναίκες και λαμβάνουν υψηλότερη σύνταξη. Οι συντάξεις που θα μελετήσουμε εμείς αφορούν στις συντάξεις γήρατος, θανάτου, αναπηρίας και ανασφάλιστων υπερηλίκων ΟΠΕΚΑ. Αξίζει να αναφέρουμε ότι η πλειοψηφία των συνταξιούχων ανδρών είναι ηλικίας 71-75 (20,4%) και των συνταξιούχων γυναικών (17,3%) είναι ηλικίας μεταξύ 71-75 ετών.

Ηλικία	**Πλήθος συνταξιούχων**	**Μηναίο Ποσό**	**Μέσο εισόδημα από συντάξεις (€)**
<=25	28.332	9.634.707,77	340,06
25-50	49.618	29.588.881,36	596,33
51-55	55.016	44.296.999,85	805,17
56-60	132.128	123.196.037,81	932,40
61-65	275.125	279.629.582,03	1016,37
66-70	399.250	400.598.228,00	1003,38
71-75	469.191	422.485.738,54	900,46
76-80	368.630	297.947.533,89	808,26
81-85	361.686	270.426.109,80	747,68
86-90	244.620	170.072.857,98	695,25
91-95	97.464	64.651.227,30	663,33
>95	20.521	13.145.165,25	640,57
Απροσδιόριστη	81	53.883,34	664,61
Σύνολο	2.501.662	2.125.727.902,92	849,73

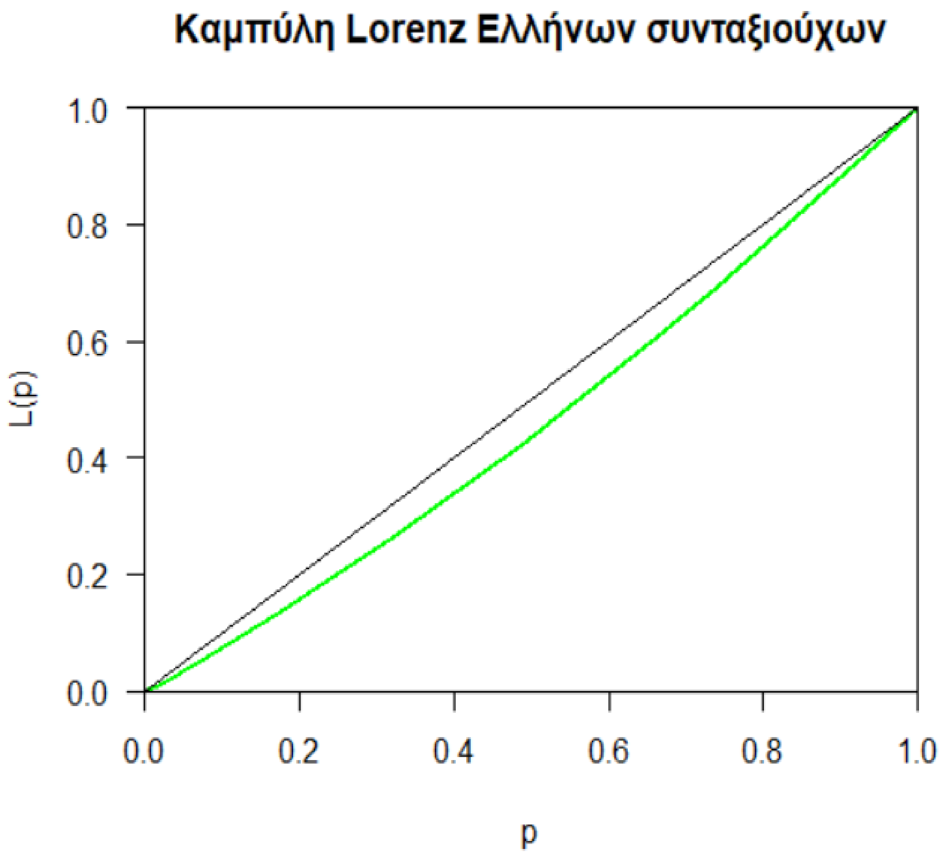
Πίνακας 4: Πλήθος Συνταξιούχων και μέσο μηνιαίο εισόδημα ανά Ηλικία

Πηγή: ΗΔΙΚΑ

Ο κώδικας στη γλώσσα R που χρησιμοποιήθηκε είναι ο εξής:

```
z<-c( rep(340.06,28332), rep(596.33,49618),rep(805.17,55016),rep(932.40,132128),
rep(1016.37,275125), rep(1003.38,399250), rep(900.46,469191),
rep(808.26,368630),rep(747.68,361686),rep(695.25,244620),rep(663.33,97464),
rep(664.61,81), rep(640.57,20521))  # διάνυσμα των τιμών

> mean(z) # Μέση τιμή των συντάξεων

849.7266

> var(z)

17730.66  #διακύμανση των συντάξεων

> library(ineq)  #χρήση της βιβλιοθήκης ineq

> ineq(z, type="Gini")

[1] 0.08645784  # δείκτης Gini

>plot(Lc(z), col="green", lwd=2, main="Καμπύλη Lorenz Ελλήνων συνταξιούχων")
```

Καταχωρώντας τα δεδομένα στην γλώσσα προγραμματισμού R αρχικά υπολογίζουμε ότι η μέση μηναία απολαβή ενός Έλληνα συνταξιούχου είναι περίπου 850€. Με τη βοήθεια της βιβλιοθήκης ineq λαμβάνουμε ότι ο δείκτης Gini ισούται με 0.08 και βλέπουμε και την καμπύλη Lorenz στο σχήμα που έπεται. Τα αποτελέσματα ερμηνεύονται ως εξής: επιλέγοντας 2 τυχαίους συνταξιούχους του πληθυσμού, η σύνταξη τους θα διαφέρει κατά 8% του μέσου ισοδύναμου διαθέσιμου εισοδήματος.



Σχήμα 4.5:Καμπύλη Lorenz μέσου μηνιαίου εισοδήματος Ελλήνων Συνταξιούχων.

Επιπλέον και από την καμπύλη Lorenz φαίνεται να υπάρχει ισότητα στις συντάξεις της χώρας μας. Άλλωστε και η καμπύλη Lorenz είναι πάρα πολύ κοντά στην γραμμή ισότητας, εξού και η πολύ μικρή τιμή του δείκτη Gini.

4.2.3 3η Εφαρμογή του δείκτη Gini σε οικονομικά δεδομένα

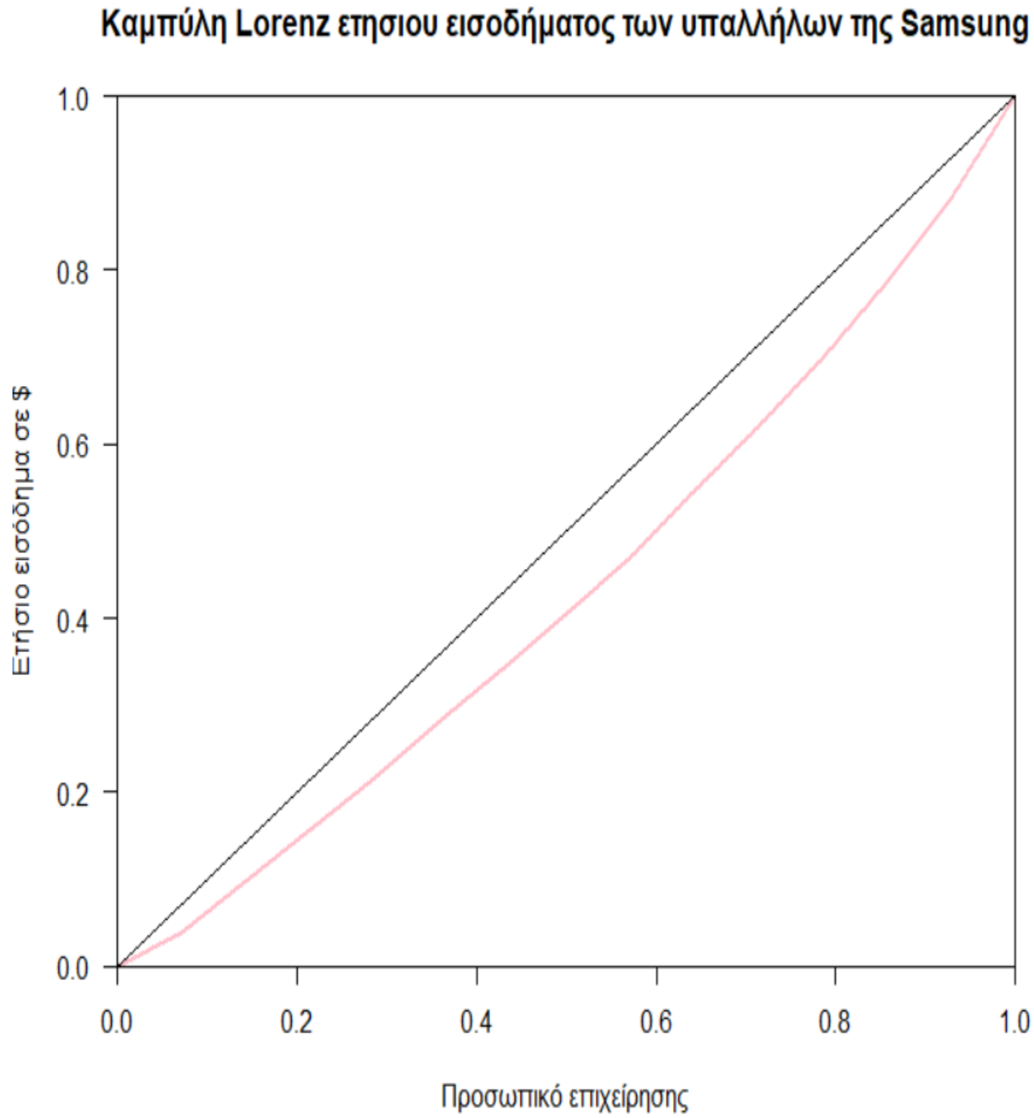
Στην συγκεκριμένη παράγραφο θα συγκρίνουμε με τη βοήθεια της καμπύλης Lorenz το ετήσιο εισόδημα των υπαλλήλων της Samsung και της Xiaomi Technologies για ένα συγκεκριμένο αντίστοιχο τμήμα των δύο εταιρειών. Οι συγκεκριμένες πολυεθνικές εταιρείες έχουν παρόμοια δράση μιας και πουλούν παρόμοια προϊόντα. Τα δεδομένα προέρχονται από το COMPARABLY και παρουσιάζονται στον παρακάτω πίνακα.

Θέση Εργασίας	Αμοιβές Samsung σε δολάρια	Αμοιβές Xiaomi Technologies σε δολάρια
Υπάλληλος Γραφείου, Admin	76000	58000
Business Development	116.000	146000
Communications	235000	99000
Customer Support	123000	63000
Design	134000	103000
Engineering	158000	135000
Finance	118000	105000
HR	149000	114000
IT	124000	114000
Legal	148000	188000
Marketing	122000	106000
Operations	119000	96000
Product	187000	146000
Sales	169000	140000

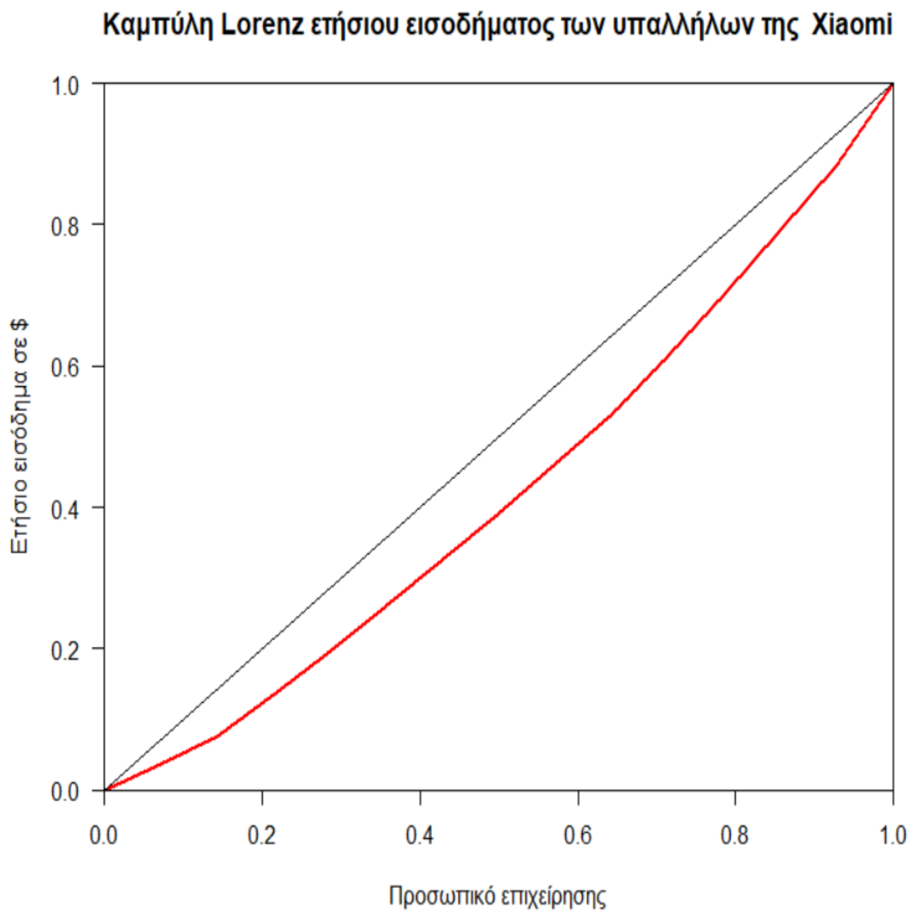
Πίνακας 5 : Ετήσιο εισόδημα στις εταιρείες Samsung και Xiaomi Technologies

Πηγή : Comparably

Καταχωρώντας τα δεδομένα στη γλώσσα R και με τη χρήση της βιβλιοθήκης ineq λαμβάνουμε τα ακόλουθα γραφήματα. Το πρώτο γράφημα παριστάνει την καμπύλη Lorenz για την Samsung , ενώ το δεύτερο παριστάνει την καμπύλη Lorenz της Xiaomi Technologies συγκριτικά με την καμπύλη τέλειας ισότητας. Παρατηρούμε ότι δεν υπάρχει σημαντική εισοδηματική ανισότητα ανάμεσα στους δυο ομίλους, μιας και οι 2 καμπύλες είναι πολύ κοντά στην γραμμή ισότητας. Επιπρόσθετα, οι τιμές των δεικτών Gini είναι πολύ κοντά , μιας και έχουμε ότι ο δείκτης Gini της Samsung ισούται με 0.13 , ενώ της Xiaomi 0.15 αντίστοιχα. Τα αποτελέσματα ερμηνεύονται ως εξής: επιλέγοντας 2 τυχαίους υπαλλήλους της Samsung, το ετήσιο εισόδημά τους θα διαφέρει κατά 13% του μέσου ισοδύναμου διαθέσιμου εισοδήματος. Για 2 υπαλλήλους της Xiaomi, το ετήσιο εισόδημά τους θα διαφέρει κατά 15% του μέσου ισοδύναμου διαθέσιμου εισοδήματος. Με άλλα λόγια οι απολαβές είναι αρκετά κοντά στις εταιρείες αυτές. Η μόνη αισθητή διαφορά είναι ότι ο μισθός για το επικοινωνιακό τμήμα της Xiaomi ανέρχεται σε 99.000$, σε αντίθεση με της Samsung που αντιστοιχεί σε 235.000$.

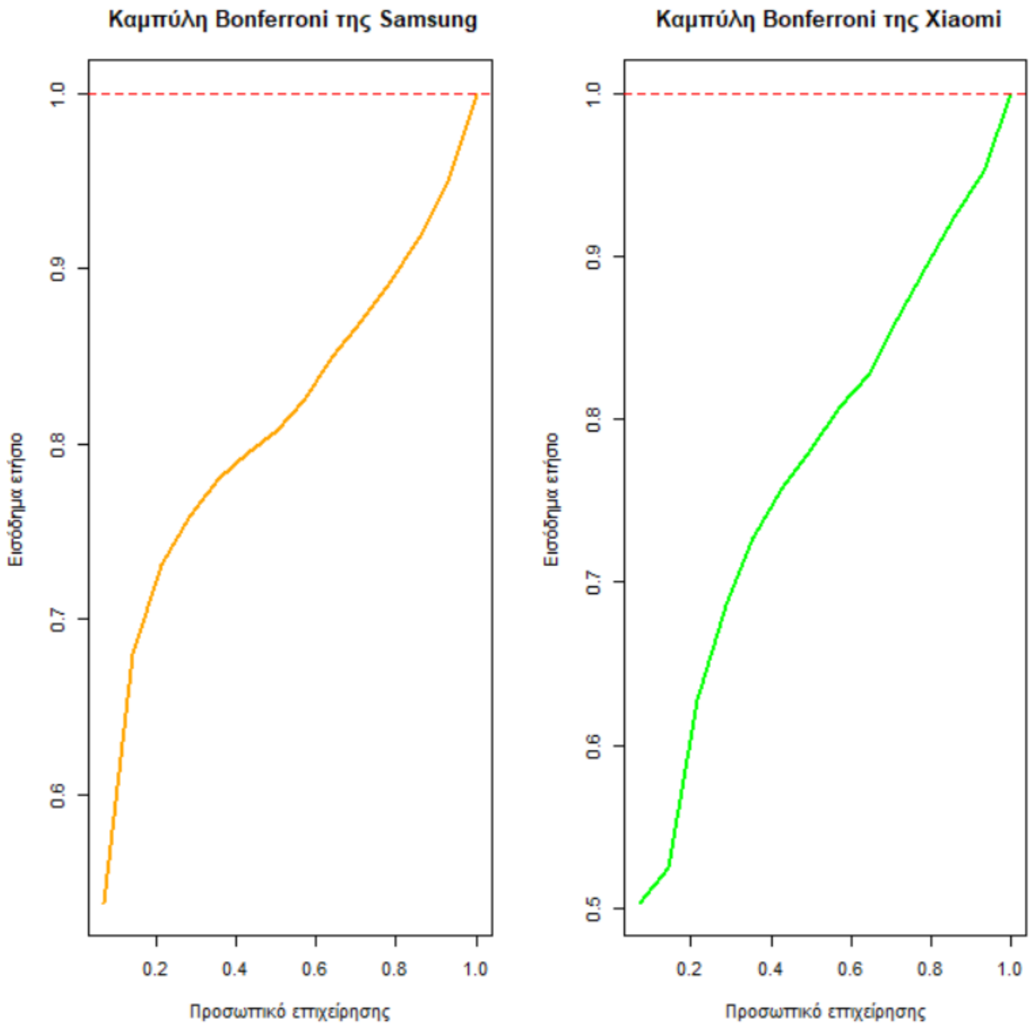


Σχήμα 4.6 : Καμπύλη Lorenz ετήσιου εισοδήματος των υπαλλήλων της Samsung.



Σχήμα 4.7 : Καμπύλη Lorenz ετήσιου εισοδήματος των υπαλλήλων της Xiaomi.

Στη συνέχεια θα παραθέσουμε τον δείκτη και την καμπύλη Bonferroni των 2 εταιρειών. Με τη βοήθεια της γλώσσας R παίρνουμε ότι ο δείκτης Bonferroni για την Samsung είναι 0.18 , ενώ για την Xiaomi είναι 0.22. Η Xiaomi έχει μεγαλύτερη τιμή δείκτη Bonferroni γεγονός που μας δείχνει ότι η Samsung έχει μικρότερη εισοδηματική ανισότητα σε σχέση με την Xiaomi. Ακολούθως παραθέτουμε και τις καμπύλες Bonferroni. Τα γραφήματα είναι αρκετά όμοια για τις 2 εταιρείες μιας και οι τιμές του δείκτη Bonferroni είναι πολύ κοντά. Συνοψίζοντας και οι 2 δείκτες μας οδηγούν σε κοινά συμπεράσματα. Επιπρόσθετα αξίζει να επισημάνουμε ότι ο δείκτης Bonferroni είναι πιο μεγάλος από τον δείκτη Gini τόσο για την Samsung όσο και για την Xiaomi κάτι το οποίο αποδείξαμε ότι ισχύει πάντα στο Κεφάλαιο 3 της παρούσας εργασίας. (βλέπε Παράγραφο 3.6)



Σχήμα 4.8: Καμπύλη Bonferroni SAMSUNG και Xiaomi.

Ο κώδικας στην γλώσσα R για να πάρουμε τον δείκτη Gini και την καμπύλη Lorenz ήταν ο εξής:

>Samsung=c(76000,116000,235000,123000,134000,158000,118000,149000,124000,148000,122000,119000,187000,169000)

> library(ineq)

> ineq(Samsung, type="Gini")

[1] 0.1372237

> plot(Lc(Samsung), col="pink", lwd=2, main="Καμπύλη Lorenz ετησιου εισοδήματος των υπαλλήλων της Samsung",)

Για την Xiaomi αντίστοιχα:

>Xiaomi<-c(58000,146000,99000,63000,103000,135000,105000, 114000,114000,188000,106000,96000,146000,140000)

> library(ineq)

> ineq(Xiaomi, type="Gini")

[1] 0.1565406

> plot(Lc(Xiaomi), col="red", lwd=2, main="Καμπύλη Lorenz ετησιου εισοδήματος των υπαλλήλων της Xiaomi" , xlab="Προσωπικό επιχείρησης", ylab="Ετήσιο εισόδημα σε $")

Πριν κλείσουμε το τελευταίο Κεφάλαιο της εργασίας μας αξίζει να επισημάνουμε ότι στο παράρτημα που ακολουθεί θα υπάρχουν οι συνδέσμοι για την εύρεση των δεδομένων που χρησιμοποιήθηκαν στις παραπάνω εφαρμογές. Όλα τα δεδομένα μελετήθηκαν με την χρήση της γλώσσας προγραμματισμού R και τη βιβλιοθήκη ineq. Οι εφαρμογές που υπάρχουν στο συγκεκριμένο Κεφάλαιο τονίζουν την χρησιμότητα των όσων μελετήσαμε στην παρούσα εργασία. Οι χρηματοοικονομικοί δείκτες Gini και Bonferroni είναι πολύτιμα εργαλεία στα χέρια των οικονομολόγων και των στατιστικών προκειμένου να ερμηνεύσουν εύκολα την οικονομική ανισότητα.

ΣΥΜΠΕΡΑΣΜΑΤΑ ΕΡΓΑΣΙΑΣ

Μετά την ολοκλήρωση της παρούσας διπλωματικής εργασίας καταλήξαμε στα εξής:

1.Μελετήσαμε διεξοδικά τα στατιστικά μέτρα πληροφορίας :

- αθροιστική υπολειπόμενη εντροπία
- Αθροιστική παρελθοντικη εντροπία
- Κανονικοποιημένη αθροιστική παρελθοντική εντροπία
- Εντροπία Tsallis
- Αθροιστική παρελθοντική εντροπία Tsallis.

2. **Διαπιστώσαμε πως** μπορούμε με την βοήθειά τους να υπολογίζουμε τον δείκτη Gini και τον δείκτη Bonferroni.

3.**Ο υπολογισμός των δεικτών αυτών** είναι απαραίτητα εργαλεία για την χρηματοοικονομική ανάλυση και σύγκριση οικονομικών μεγεθών, που στη συγκεκριμένη εργασία επαληθεύτηκαν μέσα από την εφαρμογή τους σε πραγματικά δεδόμενα.

4.**Οι χρηματοοικονομικοί δείκτες που μελετήθηκαν** είναι πολύτιμα εργαλεία στα χέρια των στατιστικών για τον εντοπισμό εισοδηματικής ανισότητας σε μια κοινωνία. Όμως για την αντιμετώπιση και πλήρη εξάλειψή της απαιτείται η συνεργασία όλων των κοινωνικών δομών, αν φυσικά αυτό είναι εφικτό σε μια σύγχρονη κοινωνία.

ΠΑΡΑΡΤΗΜΑ

Τα δεδομένα για την εφαρμογή 4.1 μπορούν να βρεθούν αν ακολουθήσετε τους παρακάτω συνδέσμους.

https://finance.yahoo.com/quote/NFLX/history?p=NFLX

https://finance.yahoo.com/quote/SMSN.IL/history?p=SMSN.IL

Τα δεδομένα για την εφαρμογή 4.2.1 μπορούν να βρεθούν αν ακολουθήσετε τον εξής σύνδεσμο.

https://www.kathimerini.gr/economy/561739024/mikres-epicheiriseis-ligoi-ergazomenoi-chamiloi-misthoi/

Τα δεδομένα για την εφαρμογή 4.2.3 μπορούν να βρεθούν αν ακολουθήσετε τους παρακάτω συνδέσμους.

- https://www.comparably.com/companies/xiaomi-technology/salaries
- https://www.comparably.com/companies/samsung/salaries

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνόγλωση

1. Υφαντόπουλος, Γ. (1990) Στατιστικά Μέτρα για την Οικονομική Ανισότητα. *Επιθεώρηση Κοινωνικών Ερευνών* (73Α) ,38-78

Ξενόγλωσση

1. Aaberge , R. (2007) . Gini’s nuclear family. *The Journal of Economic Inequality*, 5, 305-322.
2. Afify, A. and Cordeiro, G. and Butt, S. and Ortega, E. and Suzuki , A. (2017) . A new lifetime model with variable shapes for the hazard rate. *Brazilian Journal of Probability and Statistics* 31, 516-541
3. Anand, S. and Kanbur, S. (1993) . Inequality and development A critique, *Journal of Development Economics* 41, 19-43
4. Arcagni, A. and Porro, F. (2014) .The Graphical Representation of Inequality Revista . *Current Topics in Statistical Graphics* 37,419-436
5. Asadi , M. and Zohrevand , Y. (2007). On the dynamic cumulative residual entropy . *Jounal of Statistical Planning and Inference* 137, 1931-1941.
6. Bonferroni C.E. (1930). *Elmenti di statistica generale*. Libreria Seber, Firenze
7. Cali, C. and Longobardi, M. and Ahmadi, J. (2017). Some properties of cumulative Tsallis entropy. *Physica A* 486 ,1012-1021
8. Di Crescenzo , A. and Longobardi, M. (2009). On cumulative entropies . *Journal of Statistical Planning and Inference* 139 , 4072-4087

9. Fjaerli, E. and Aaberge, R. (1999). Tax Reforms, Dividend Policy and Trends in Income Inequality Empirical Evidence based on Norwegian Data *Discussion Papers*, 1-19

10. Gastwirth, J. L. (1972). The Estimation of the Lorenz curve and Gini Index. *The Review of Economics and Statistics*, LIV,306-316

11. Gastwirth, J. L and Modarres, R. and Bura, E. (2005). The use of the Lorenz curve, Gini index and related measures of relative inequality and uniformity in securities law. *International Journal of Statistics*, LXIII, 451-469

12. Giorgi, M, and Nadarajah, S. (2010). Bonferroni and Gini indices for various parametric families of distributions *International Journal of Statistics* LXVIII, 23-46

13. Kendall, M. and Stuart, A. (1958). The Advanced Theory of Statistics 1, Butler and Tanner, London

14. Navarro, J. and Aguila ,Y. and Asadi , M. (2010). Some new results on the cumulative residual entropy. *Journal of Statistical Planning and Inference* 140 , 310-322

15. Psarrakos, G. and J.Navarro (2013), Generalized cumulative residual entropy and record values. *Metrika* 76,623-640

16. Tahmasebi , S. and Parsa, H. (2019). Notes on Cumulative Entropy as a Risk Measure. *Stohastics and Quality Control* 34(1) , 1-7,

17. Tarsitano, A. (1990) . The Bonferroni index of income inequality . *Income and Wealth Distribution, Inequality and Poverty* , 228-242

18. Tsallis, C. (1988). Possible generalization of Boltzmann- Gibbs statistic. *Journal of Statistical Physics* 52, 479-487

Διαδικτυακές πηγές

1. BoyceWire
   https://boycewire.com/what-is-the-gini-coefficient/

2. Cleartax

   https://cleartax.in/g/terms/lorenz-curve

3. ECONOMICS HELP
   https://www.economicshelp.org/blog/glossary/lorenz-curve/

4. ECONOMY PEDIA
   https://el.economy-pedia.com/11031530-gini-index

5. Investopedia
   https://www.investopedia.com/terms/g/gini-index.asp

6. Towards Data Science
   https://towardsdatascience.com/clearly-explained-gini-coefficient-and-lorenz-curve-fe6f5dcdc07

7.ΟΙΚΟΝΟΜΙΑ ΟΙΚΟΝΟΜΙΚΑ

https://www.economiafinanzas.com/el/indice-de-gini/