**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ**

**ΑΝΑΛΟΓΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗΣ ΚΙΝΔΥΝΟΥ**

# «ΘΕΩΡΙΑ ΠΛΗΡΟΦΟΡΙΑΣ, ΜΕΤΡΑ ΑΒΕΒΑΙΟΤΗΤΑΣ ΕΝΤΡΟΠΙΑΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΣΤΟΝ ΑΝΑΛΟΓΙΣΜΟ»

**ΑΛΕΞΑΝΔΡΑ ΤΙΠΑΠΙΔΗ ΚΑΡΑΜΑΪΛΗ**

**Διπλωματική Εργασία**

**που υποβλήθηκε στο τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου**

**ΠΕΙΡΑΙΑΣ**

**ΙΟΥΝΙΟΣ 2018**

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίσθηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Αναλογιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. … συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου.

Τα μέλη της Επιτροπής ήταν:

-Επίκ. Καθηγητής Πιτσέλης Γεώργιος (Επιβλέπων)

-Αναπλ. Καθηγητής Κωνσταντίνος Πολίτης

-Αναπλ. Καθηγητής Σεβρόγλου Βασίλειος

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

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

UNIVERSITY OF PIRAEUS

DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

MASTER OF SCIENCE IN ACTUARIAL SCIENCE AND RISK MANAGEMENT

# "ENTROPY"

ALEXANDRA TIPAPIDI KARAMAILI

MSc Dissertation submitted to the Department of Statistics and Insurance Science of the University of Piraeus in partial fulfillment of the requirements for the degree of Master of Science in Actuarial Science and Risk Management

PIRAEUS

JUNE 2018

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

*…as nobody knows what entropy is, whenever you use the term you will always be at an advantage*.

(John von Neumann)

*Στους  γονείς μου*

*Κωνσταντίνο και Κυριακούλα*

*και στην αγαπημένη μου γιαγιά Θηρεσία*

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Ευχαριστίες

Ευχαριστώ τον Επιβλέποντα Καθηγητή της Διπλωματικής μου Εργασίας κ. Πιτσέλη Γεώργιο για την υπομονή και τη σημαντική καθοδήγηση που μου παρείχε μέχρι την περάτωση της εργασίας.

Θα ήθελα ακόμη να ευχαριστήσω τους καθηγητές κ. Κωνσταντίνο Πολίτη και κ. Βασίλειο Σεβρόγλου για τη συμμετοχή στην παρουσίαση και αξιολόγηση της διπλωματικής μου εργασίας.

Τέλος θέλω να απευθύνω ένα μεγάλο ευχαριστώ στην οικογένειά μου και στους φίλους μου για τη στήριξη που μου παρείχαν.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# ΠΕΡΙΛΗΨΗ

Ο τομέας της Αναλογιστικής Επιστήμης έχει ως κύριο αντικείμενο την ανάλυση και αποτίμηση των οικονομικών συνεπειών που αφορούν κινδύνους και αβέβαια μελλοντικά γεγονότα. Η ανάγκη για την ποσοτικοποίηση της αβεβαιότητας που διέπει ένα γεγονός οδήγησε σε χρήση μέτρων εντροπίας.

Στην παρούσα διπλωματική εργασία θα αναλύσουμε ορισμένα μέτρα αβεβαιότητας εντροπίας και την εφαρμογή τους στον Αναλογισμό.

Συγκεκριμένα, στο πρώτο κεφάλαιο γίνεται μια αναδρομή στη Θεωρία Πληροφορίας και στην επέκτασή της στην Επιστήμη του Αναλογισμού. Κατόπιν, στο δεύτερο κεφάλαιο ορίζεται η εντροπία του Shannon και ορισμένες ιδιότητες που διέπουν το μέτρο, και εξετάζεται η εφαρμογή του σε Μοντέλα Απώλειας υπό την επίδραση πληθωρισμού.

Στο τρίτο κεφάλαιο ορίζεται η Απόκλιση Kullback–Leibler και παρουσιάζεται η εφαρμογή του γραμμικού προβλήματος Ελάχιστης Διαχωριστικής Πληροφορίας (*Minimum Discrimination Information)* στις υποθέσεις ομοιόμορφης κατανομής θανάτων *(UDD)* και σταθερής έντασης θνησιμότητας.

Στο τέταρτο και πέμπτο κεφάλαιο της εργασίας παρουσιάζονται τα μέτρα Υπολειπόμενης και Παρελθοντικής εντροπίας *(Residual and Past Entropy)* αντίστοιχα και η εφαρμογή τους σε περικομμένα ή λογοκριμένα δείγματα απωλειών και σε Μοντέλα Επιβίωσης.

Τέλος, στο έκτο και τελευταίο κατά σειρά κεφάλαιο εξετάζονται τα μέτρα Αθροιστικών Εντροπιών, Αθροιστικής Υπολειπόμενης, Δυναμικής Αθροιστικής Υπολειπόμενης και Διπλά Περικομμένης Εντροπίας *(Cumulative Residual, Dynamic Cumulative Residual and Doubly Truncated Cumulative Residual Entropy)* ως προς τις ιδιότητές τους και την κατηγοριοποίηση κατανομών σε κλάσεις γήρανσης.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# ABSTRACT


The main objective of the field of Actuarial Science is the analysis and valuation of economic consequences that refer to risks and uncertain future events. The necessity to quantify the uncertainty conditions of an event has led to the use of measures of entropy and uncertainty.

In the following thesis, various measures of entropy are to be analyzed, and their applications in Actuarial Science are to be examined.

Specifically, the first chapter contains a summary of the history of Information Theory and its extension in Actuarial Science. After that, in the second chapter we will define Shannon’s Entropy and some features of this measure. We will also examine its application in Loss Models, under the influence of inflation.

In the third chapter, the Kullback-Leibler Divergence is defined, and the application of the Minimum Discrimination Information linear problem in the assumption of Uniform Distribution of Deaths *(UDD)* and Constant Force of Mortality is described.

The fourth and fifth chapter of this thesis explore the measures of Residual and Past Entropy respectively, and also their application in truncated or censored loss data, as well as in Survival Models.

Finally, in the sixth chapter we will examine the measures of Cumulative Entropies; Cumulative Residual Entropy, Dynamic Cumulative Residual and Doubly Truncated Cumulative Residual Entropy, in relation to their properties, and the categorization of distributions in aging classes.


ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Περιεχόμενα

## Κεφάλαιο 1ο



## Κεφάλαιο 2ο



## Κεφάλαιο 3ο



## Κεφάλαιο 4ο





ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

## Κεφάλαιο 5$^{ο}$



## Κεφάλαιο 6$^{ο}$



## Κεφάλαιο 7$^{ο}$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Κεφάλαιο 1ο

## 1.1. Σύντομη ιστορική αναδρομή

Η έννοια της πληροφορίας πριν τα μέσα του 20ου αιώνα αποτελούσε κατά βάση ποιοτική και αφηρημένη έννοια. Ωστόσο, η ραγδαία ανάπτυξη της μετάδοσης πληροφορίας (information transmission) οδήγησε στην αντίληψη και μελέτη της πληροφορίας υπό ένα νέο πρίσμα, αυτό της ποσοτικής και μαθηματικής της υπόστασης.

Πρώτος ο Samuel Morse το 1837 ανέπτυξε το αλφάβητο Morse για τη μετάδοση μηνυμάτων με βέλτιστο τρόπο σε μεγάλη απόσταση, ενώ κατασκεύασε και την πρώτη τηλεγραφική μηχανή. Κατόπιν, ο Graham Bell το 1903 ανακάλυψε τον πολλαπλό τηλέγραφο, ενώ το 1878 είχε ήδη κατοχυρώσει την ευρεσιτεχνία της συσκευής που μεταδίδει ήχο και φωνή τηλεγραφικώς. Τέλος, ο Thomas Edison το 1874 εισήγαγε την τετραπλή κωδικοποίηση (quadruplex) για να αυξήσει το ρυθμό μετάδοσης της πληροφορίας. Τα επιτεύγματα αυτά, όπως ήταν επόμενο, ώθησαν την επιστημονική κοινότητα στη μελέτη βέλτιστων τρόπων αναμετάδοσης της πληροφορίας και συνετέλεσαν στη θεμελίωση ενός νέου τομέα της Θεωρίας Επικοινωνίας, αυτού της Θεωρίας Πληροφορίας.

Η έννοια της ποσοτικοποίησης της πληροφορίας απασχόλησε αρχικά το Hartley κατά τη μελέτη των τηλεγραφικών επικοινωνιών το 1928. Σύμφωνα με την πρόταση που διατυπώθηκε, η «ποσότητα» πληροφορίας διαμορφώνεται από τη διαδοχική επιλογή λέξεων από ένα δεδομένο αλφάβητο N-το πλήθος συμβόλων. Υποθέτοντας λοιπόν ότι σχηματίζονται λέξεις ή μηνύματα k θέσεων από ένα αλφάβητο $|N|$: συμβόλων, μπορούν να εμφανιστούν $|N|^k$ διαφορετικές λέξεις, υπό την προϋπόθεση ότι όλα τα στοιχεία του αλφάβητου έχουν την ίδια πιθανότητα εμφάνισης. Σε αυτή την περίπτωση η ποσότητα πληροφορίας που εξάγεται είναι:

$$I(N^k) = \log|N|^k = k\log|N|,\ \acute{o}\pi o\upsilon\ |N|\, o\ \pi\lambda\eta\theta\acute{\alpha}\rho\iota\theta\mu o\varsigma\ \tau o\upsilon\ N. \tag{1.1}$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Η μονάδα μέτρησης είναι εξαρτώμενη από τη βάση του λογάριθμου. Σε δεκαδική βάση, η μονάδα ποσότητας πληροφορίας είναι η decit (decimal unit), ή αλλιώς μονάδα Hartley, όπως ονομάστηκε προς τιμήν του, και για το φυσικό λογάριθμο η μονάδα nat (natural unit). Ωστόσο, η επικρατέστερη μονάδα μέτρησης είναι το bit, η λογαριθμική συνάρτηση με βάση 2, εξαιτίας της χρήσης του δυαδικού συστήματος αρίθμησης στους υπολογιστές.

Κατόπιν, ο Claude E. Shannon το 1948, με τη δημοσίευση του άρθρου "*A Mathematical Theory of Communication*" στο περιοδικό *Bell System Technical Journal*, επέκτεινε την έννοια της πληροφορίας του Hartley για σύνολα που δεν έχουν απαραίτητα την ίδια πιθανότητα εμφάνισης και θεμελίωσε τις βασικές μαθηματικές αρχές που διέπουν το μέτρο.

Το μέτρο πληροφορίας όπως ορίστηκε από το Shannon, δίνεται στη σχέση:

$$I(X) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i . \tag{1.2}$$

Όπως μπορεί να παρατηρήσει κανείς, η πληροφορία του Shannon εκφράζει την αναμενόμενη αβεβαιότητα του μέτρου του Hartley, περιλαμβάνοντας και ενδεχόμενα όχι κατ' ανάγκη ισοπίθανα. Φυσικά, σε περίπτωση που τα ενδεχόμενα έχουν ίσες πιθανότητες εμφάνισης οδηγούμαστε στην πληροφορία του Hartley.

Η συμφωνία του μέτρου του Shannon με το μέτρο της εντροπίας του Boltzmann, όπως ορίστηκε στον τομέα της Θερμοδυναμικής, συντέλεσε στο να ονομαστεί το πρώτο Εντροπία του Shannon.

## 1.2. Λίγα λόγια για το μέτρο της Εντροπίας

Τον όρο της εντροπίας πρώτος διατύπωσε ο φυσικός Clausius το 1850 για να περιγράψει το ποσοστό της θερμότητας που πρέπει να εισαχθεί σε ένα κλειστό σύστημα προκειμένου αυτό να έλθει σε μια δεδομένη κατάσταση.

Ο όρος της εντροπίας είναι απόλυτα συνυφασμένος με το $2^{ο}$ Θερμοδυναμικό νόμο, κατά τον οποίο είναι αδύνατη η παραγωγή ενέργειας εκ του μηδενός. Σε οποιαδήποτε διεργασία, ένα μέρος της διαθέσιμης ενέργειας μετατρέπεται σε θερμότητα ή

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

χρησιμοποιείται για την εσωτερική αναδιάταξη χημικών ενώσεων ή για άλλες αλλαγές εντροπίας. Σε μία μεταβολή ενός μονωμένου συστήματος η εντροπία αυξάνει πάντοτε και η μεταβολή της ισούται με το μέτρο της θερμικής ενέργειας που δε μπορεί πια να χρησιμοποιηθεί για την παραγωγή έργου.

Αργότερα, κατά τη διάρκεια της δεκαετίας του 1870, ο L. E. Boltzmann όρισε την εντροπία σε μικροσκοπικό επίπεδο, ως το πλήθος των μικροκαταστάσεων ενός φυσικού συστήματος το οποίο αντιστοιχεί σε δεδομένη μακροκατάσταση, συνδέοντας το $2^{ο}$ Θερμοδυναμικό νόμο με τη Θεωρία Πιθανοτήτων. Η εντροπία απέκτησε τη φυσική ερμηνεία της αταξίας που διέπει ένα σύστημα και απεδείχθη ότι αύξηση της εντροπίας συνεπάγεται και αύξηση της αταξίας του συστήματος. Ο τύπος του Boltzmann για την εντροπία ορίζεται ως

$$S = k_B \ln \Omega, \tag{1.3}$$

όπου $k_B$:1,381$\times 10^{-23} J/_K$, η σταθερά Boltzmann και $\Omega$ είναι ο αριθμός των μικροσκοπικών καταστάσεων στις οποίες μπορεί να βρεθεί ένα σύστημα.

Η επέκταση της έννοιας της εντροπίας στον τομέα Θεωρίας Πληροφορίας ήρθε από τον Claude Shannon με αφορμή τη μελέτη της «χωρητικότητας» ενός καναλιού μέσα από το οποίο μεταφέρεται ένα μήνυμα. Στόχος ήταν η εύρεση ενός μέτρου που μεγιστοποιεί την απόδοσης της πληροφορίας μέσω του διαύλου, λαμβάνοντας υπόψη το «θόρυβο» ή «αταξία».

Τα αποτελέσματα της μελέτης οδήγησαν σε ένα μέτρο ισοδύναμο με αυτό του Boltzmann και προτάθηκε να περιγράφεται από τον όρο εντροπία.

Η παρούσα διπλωματική εργασία θα μελετήσει την εντροπία ως μέτρο πληροφορίας ή μέτρο αβεβαιότητας σε θέματα Αναλογισμού.

## 1.3. Η Θεωρία Πληροφορίας στον Αναλογισμό

Η συνεισφορά της επιστήμης της Θεωρίας Πληροφορίας επεκτάθηκε σε πληθώρα επιστημονικών πεδίων, όπως της Στατιστικής, της Βιολογίας και μεταξύ άλλων στον τομέα της Αναλογιστικής Επιστήμης.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Ο τομέας της Αναλογιστικής Επιστήμης έχει ως κύριο αντικείμενο την ανάλυση και αποτίμηση των οικονομικών συνεπειών που αφορούν κινδύνους και αβέβαια μελλοντικά γεγονότα. Η ανάγκη για ποσοτικοποίηση της αβεβαιότητας που διέπει ένα γεγονός, προκειμένου να εξασφαλιστεί η ευρωστία του ασφαλιστικού φορέα, οδηγεί σε αναπόφευκτη χρήση μέτρων εντροπίας αβεβαιότητας, δανεισμένα από τον τομέα της Θεωρίας Πληροφορίας.

Οι δύο βασικοί κλάδοι δραστηριοποίησης ενός Αναλογιστή είναι ο κλάδος ασφαλίσεων ζωής (Life Insurance) και ο κλάδος που δεν αφορά τις ασφαλίσεις ζωής (Non-Life Insurance). Ο Αναλογιστής αναλαμβάνει τον κίνδυνο ασφάλισης ενός προσώπου ή ιδιοκτησίας (Life Insurance, Non-Life Insurance) και χρησιμοποιεί κατάλληλα μαθηματικά μοντέλα, ώστε ο κίνδυνος να αποτιμάται σε επίπεδο ασφαλίστρου. Για τον προσδιορισμό της βέλτιστης αποτίμησης του ασφαλίστρου που υποχρεούται να καταβάλλει ένας ασφαλισμένος, είναι αναγκαία η εκτίμηση βασικών παραμέτρων όπως συναρτήσεων επιβίωσης, θανάτου ή αναπηρίας, πινάκων θνησιμότητας κλπ. από διαθέσιμα δεδομένα στατιστικά στοιχεία (Life Insurance). Ανάλογα, στον κλάδο Non-Life Insurance αναγκαία είναι η εκτίμηση κατανομών ζημιάς σε συμβόλαια μερικής κάλυψης (αφαιρετέο ποσό, όριο ιδίας κράτησης).

Και στις δύο περιπτώσεις τα δεδομένα που έχει στη διάθεσή του ο Αναλογιστής συνήθως είναι περικομμένα ή λογοκριμένα. Αυτό μεταφράζεται στη μεν πρώτη περίπτωση σε δεδομένα που δεν αναφέρονται ή δεν καταγράφονται καν ή είναι εκτός διαστήματος μελέτης, ενώ στη δεύτερη σε δεδομένα για τα οποία έχουμε εν μέρει γνώση.

Παραδείγματος χάρη, ζημιές που παρατηρούνται σε ένα χαρτοφυλάκιο κάτω από κάποιο ποσό μπορεί να μην αναφέρονται καν (περικοπή), ενώ μπορεί να αναφέρονται αλλά να παρατηρούνται ως μηδενικές τιμές (λογοκρισία). Στη δεύτερη περίπτωση έχουμε γνώση ότι έλαβαν χώρα, ωστόσο δεν έχουμε γνώση για το ακριβές ποσό της ζημίας.

Εκτιμήσεις μοντέλων που προέρχονται από τέτοια δείγματα συνήθως εμπεριέχουν μεγάλη αβεβαιότητα η οποία θα πρέπει να ποσοτικοποιηθεί.

Όπως θα δούμε παρακάτω, εντροπίες υπολειπόμενης και παρελθοντικής διάρκειας ζωής *(Residual and Past Entropy),* χρησιμοποιούνται για τον υπολογισμό

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

αβεβαιότητας που εμπεριέχεται σε περικομμένα ή λογοκριμένα δείγματα ζημιών (Kluggman et al., 2008), όταν έχει εφαρμοστεί αφαιρετέο ποσό ή όριο ιδίας κράτησης (Sachlas and Papaioannou, 2012), σε Μοντέλα Απώλειας (*Loss Models).*

Αντίστοιχα τα μοντέλα αυτά, στην Ανάλυση Επιβίωσης *(Survival Analysis)* βρίσκουν τόπο εφαρμογής σε μοντέλα κινδύνων (Sachlas and Papaioannou, 2012).

Επεκτάσεις των εντροπιών υπολειπόμενης και παρελθοντικής διάρκειας ζωής, όπως η Αθροιστική Υπολειπόμενη και Παρελθοντική Εντροπία (*Cumulative Residual and Past Entropy),* η Δυναμική Αθροιστική Υπολειπόμενη και Παρελθοντική Εντροπία (*Dynamic Cumulative Residual and Past Entropy),* και η Διπλά Περικομμένη Αθροιστική Υπολειπόμενη και Παρελθοντική Εντροπία *(Doubly Truncated Cumulative (Interval) Residual and Past Entropy),* χρησιμοποιούνται για τον προσδιορισμό κατανομών ζωής και την ένταξη των τελευταίων σε κλάσεις γήρανσης βάσει της μονοτονίας των συναρτήσεων κινδύνων και του μέσου υπολειπόμενου χρόνου.

Προβλήματα βελτιστοποίησης όπως αυτό της αρχής ελάχιστης διασταυρούμενης πληροφορίας (*Minimum Discrimination Information -MDI)* ή ισοδύναμα της αρχής μέγιστης εντροπίας *(Maximum Entropy Principle -MEP),* βρίσκουν τόπο εφαρμογής σε ένα ευρύτερο φάσμα αναλογιστικών θεμάτων.

Χαρακτηριστικά, όπως προτείνεται από τον Brockett (1991), η *MEP* μπορεί να χρησιμοποιηθεί για να την προσαρμογή μιας τυπικής κατανομής, ώστε να λαμβάνει υπόψη τα ιδιαίτερα χαρακτηριστικά ενός ασφαλισμένου. Επιπλέον μπορεί να χρησιμοποιηθεί για τη προσαρμογή πινάκων θνησιμότητας, ώστε αυτοί να αναπαριστούν γνωστή πληροφορία για τους ασφαλισμένους, ενώ είναι όσο το δυνατόν πιο όμοιοι με τους τυπικούς πίνακες (Brockett and Cox, 1984).

Στην παρούσα διπλωματική εργασία αρχικά θα παρουσιαστούν κάποιες βασικές ιδιότητες του μέτρου εντροπίας του Shannon και η εφαρμογή του μέτρου σε Μοντέλα Απώλειας υπό την επίδραση του πληθωρισμού. Στη συνέχεια θα παρουσιαστεί η απόκλιση Kullback–Leibler και τα προβλήματα βελτιστοποίησης που προκύπτουν βάσει των δύο μέτρων, *MEP* και *MDI,* στην επίλυση θεμάτων αναλογισμού (Brockett, 1991). Κατόπιν θα εξεταστούν κάποιες ιδιότητες των μέτρων των υπολειπόμενων και παρελθοντικών εντροπιών (*Residual and Past Entropy)* και η

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ισοδύναμη έκφρασή τους σε Μοντέλα Απώλειας υπό την επίδραση αφαιρετέου ποσού και ορίου ιδίας κράτησης, αντίστοιχα. (Sachlas and Papaioannou, 2012). Τέλος, θα μελετηθούν, η υπολειπόμενη εντροπία και οι Αθροιστικές Υπολειπόμενες Εντροπίες, *Cumulative Residual, Dynamic Cumulative Residual* και *(Doubly Truncated (Iinterval) Cumulative Residual)* ως προς τον προσδιορισμού κατανομών και ως προς την κατηγοριοποίηση των τελευταίων σε νέες κλάσεις γήρανσης.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Κεφάλαιο 2ο

## 2.1. Η εντροπία του Shannon

**Ορισμός 2.1.1**

Έστω $X$ μια τυχαία μεταβλητή, διακριτή με συνάρτηση πιθανότητας $p_i = P(X = i)$, 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:

$$H(X) = -\sum_i p_i \log p_i, \tag{2.1}$$

αν η $X$ είναι διακριτή τυχαία μεταβλητή και

$$H(X) = -\int f(x) \log f(x), \tag{2.2}$$

αν η $X$ είναι συνεχής τυχαία μεταβλητή.

Η περίπτωση της συνεχούς εντροπίας αναφέρεται επίσης και ως διαφορική εντροπία και αποτελεί επέκταση τις εντροπίας του Shannon, για συνεχείς τυχαίες μεταβλητές, όταν υπάρχει η συνάρτηση πυκνότητας πιθανότητας.

Είναι εύκολο να αποδειχθεί ότι οι παραπάνω σχέσεις ισοδύναμα γράφονται:

$$H(x) = E\left(\log\left(\frac{1}{p_i}\right)\right) \tag{2.3}$$

ή

$$H(x) = E\left(\log\left(\frac{1}{f(x)}\right)\right) \tag{2.4}$$

για τη διακριτή ή συνεχή περίπτωση αντίστοιχα.

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 $X$.

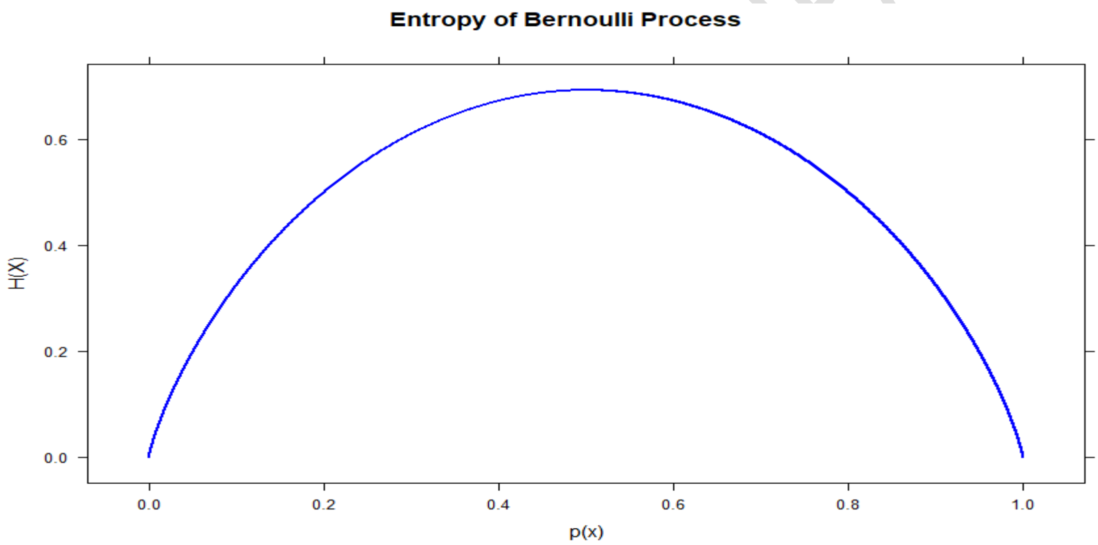
**Παράδειγμα:**

**Η εντροπία της κατανομής Bernoulli**

Έστω ότι η τυχαία μεταβλητή $X$ ακολουθεί τη Bernoulli με πιθανότητα επιτυχίας $p$. Τότε η εντροπία της προκύπτει ως:

$$H(p) = -p \log p - (1-p) \log(1-p). \tag{2.5}$$

Η σχηματική απεικόνιση της παραπάνω σχέσης είναι:




*Σχήμα 1: Γραφική απεικόνιση της εντροπίας μιας Bernoulli κατανομής*

Όπως είναι εύκολο να διαπιστωθεί από το παραπάνω σχήμα, η εντροπία εμφανίζει ολικό μέγιστο στην τιμή $p = 1/2$. (π.χ. αμερόληπτο νόμισμα).

Όσο μεγαλύτερη είναι η εντροπία, τόσο λιγότερη πληροφορία παρέχεται από μια παρατήρηση της $X$. Επομένως, βάσει του παραδείγματος που δόθηκε, η εντροπία για την τιμή $p = 1/2$ παρέχει τη μέγιστη δυνατή πληροφορία. Επιπλέον, στο βέβαιο γεγονός $p = 1$ η πληροφορία που παρέχεται είναι μηδενική.

Παρατηρούμε πως το μέτρο της εντροπίας είναι συναρτησιακό (functional) της συνάρτησης πιθανότητας της τυχαίας μεταβλητής $X$, μιας και η τιμή της δεν εξαρτάται από τις τιμές των παρατηρήσεων αλλά από τις αντίστοιχες πιθανότητες εμφάνισης.

## Ιδιότητες της εντροπίας ως μέτρο πληροφορίας

Για διακριτές κατανομές, όπως θα δούμε παρακάτω, η εντροπία ως μέτρο αβεβαιότητας (πληροφορίας) παρουσιάζει ορισμένες σημαντικές ιδιότητες, όπως η μη-αρνητικότητα και η ισότητα με το μηδέν σε περίπτωση βέβαιου γεγονότος.

Επιγραμματικά αναφέρονται κάποιες βασικές ιδιότητες της διακριτής περίπτωσης:

I. $H(X)$: συνεχής συνάρτηση των $p_i$.

II. $H(X)$:συμμετρική ως προς $p_i$.

III. $0 \leq H(X) \leq \log n$, με $\log n = H(Uniform)$.
Η εντροπία είναι μη-αρνητική, με την ισότητα με το μηδέν να προκύπτει αν και μόνο αν η $X$ εκφράζει το βέβαιο γεγονός (με πιθανότητα $p = 1$). Επίσης είναι άνω φραγμένη από την εντροπία της ομοιόμορφης κατανομής. Η ισότητα προκύπτει όταν τα ενδεχόμενα $\{x_1, \dots x_n\} \in X$ είναι ισοπίθανα με πιθανότητα εμφάνισης $p_i = \frac{1}{n}$, $\forall i = 1, .. n$. Η περίπτωση της ομοιόμορφης κατανομής των παρατηρήσεων της Χ αναφέρεται και ως μέγιστη εντροπία.

IV. $H(X, Y) \leq H(X) + H(Y)$ (Αρχή υποπροσθετικότητας),
$H(X, Y) = H(X) + H(Y)$ αν $X, Y$ ανεξάρτητες τυχαίες μεταβλητές $H(X, Y)$ καλείται από κοινού εντροπία και παρουσιάζεται σε αναλυτική μορφή σε ακόλουθη παράγραφο.

*V.* $H(X|Y) \leq H(X)$. Αυτό σημαίνει ότι η εκ των προτέρων γνώση της τυχαίας μεταβλητής $Y$ μειώνει την αβεβαιότητα.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Ως προς τη συνεχή περίπτωση, είναι σημαντικό να αναφέρουμε ότι παρουσιάζει ορισμένα μειονεκτήματα, καθώς ορίζεται μόνο για κατανομές με πυκνότητα, δεν ορίζεται για μικτού τύπου κατανομές και επιπλέον μπορεί να πάρει αρνητικές τιμές. Χαρακτηριστικό παράδειγμα αποτελεί η περίπτωση της ομοιόμορφης κατανομής που ορίζεται στο διάστημα $(0, a)$. Εύκολα αποδεικνύεται ότι η εντροπία της $X$, με $X \sim U(0, a)$, για $a < 1$ προκύπτει αρνητική. Οι υπόλοιπες ιδιότητες της διακριτής περίπτωσης επεκτείνονται στη συνεχή.

## 2.2. Από κοινού εντροπία και Υπό συνθήκη εντροπία

Το μέτρο της εντροπίας μπορεί να χρησιμοποιηθεί και για τον ορισμό μετρήσεων πληροφορίας οι οποίες αναδεικνύουν τις σχέσεις μεταξύ δύο τυχαίων μεταβλητών $X$ και $Y$.

I. Από κοινού ή συνδετική εντροπία *(joint entropy)* η οποία μετρά τη συνολική πληροφορία των $X, Y$.
II. Την υπό συνθήκη εντροπία *(conditional entropy)* η οποία εκφράζει τη συνολική πληροφορία της $X$ όταν είναι γνωστή η $Y$ και αντίστροφα.

### 2.2.1. Από κοινού εντροπία (joint entropy)

**Ορισμός 2.2.1**

Έστω X, Y τυχαίες μεταβλητές με από κοινού συνάρτηση μάζας πιθανότητας $p(x, y)$, για $x \in X$ και $y \in Y$. Τότε, η από κοινού εντροπία ορίζεται ως:

$$H(X, Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) \tag{2.6}$$

ή ισοδύναμα,

$$H(X, Y) = -E(\log p(X, Y)) = E\left(\log\left(\frac{1}{p(X, Y)}\right)\right). \tag{2.7}$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Για τη συνεχή περίπτωση ορίζεται ομοίως:

$$H(X,Y) = -\int_X \int_Y p_{x,y}\,(x,y)\log p_{x,y}(x,y)\,dx\,dy, \qquad (2.8)$$

με $p(x,y)$: από κοινού συνάρτηση πυκνότητας πιθανότητας.

### 2.2.2. Υπό συνθήκη εντροπία (conditional entropy)

**Ορισμός 2.2.2**

Η υπό συνθήκη εντροπία $H(X|Y)$ είναι η μέση τιμή των $H(X|Y=y)$

$$H(X|Y=y) = -\sum_{x\in X} p(X|Y=y)\log p(X|Y=y). \qquad (2.9)$$

Πράγματι,

$$H(X|Y) = -\sum_{y\in Y} p(y)H(X|Y=y) = -\sum_{y\in Y} p(y)\sum_{x\in X} p(x|y)\log p(x|y)$$

$$= -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(x|y) = -E(\log p(X|Y)). \qquad (2.10)$$

Κατ' αντιστοιχία, στη συνεχή περίπτωση ορίζεται:

$$H(X|Y) = -\int_X \int_Y p_{X,Y}\,(x,y)\log p_{X|Y}(X|Y)\,dxdy. \qquad (2.11)$$

## 2.3. Η σχέση που συνδέει την εντροπία με τις από κοινού και υπό συνθήκη.

**Θεώρημα 2.3.1 (Κανόνας Αλυσίδας)**

$$H(X,Y) = H(Y) + H(X|Y) = H(X) + H(Y|X). \qquad (2.12)$$

(Παρουσιάζεται απόδειξη μόνο για τη διακριτή περίπτωση)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Απόδειξη:**

$$H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(x,y) = -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log[p(y)p(x|y)]$$

$$= -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(y) - \sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(x|y)$$

$$= -\sum_{y\in Y} p(y)\log p(y) - \sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(x|y) = H(Y) + H(X|Y).$$

Ομοίως αποδεικνύεται και η δεύτερη ισότητα.

**Πόρισμα 2.3.1**

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z). \tag{2.13}$$

## 2.4. Η Διαφορική Εντροπία του Shannon και η επίδραση του πληθωρισμού σε Μοντέλα Απώλειας

Ο πληθωρισμός είναι το φαινόμενο της συνεχούς αύξησης του γενικού επιπέδου τιμών, αγαθών και υπηρεσιών, σε όλο το εύρος μιας οικονομίας, με αποτέλεσμα τη μείωση της αγοραστικής αξίας του χρήματος. Η ποσοστιαία αύξηση του δείκτη τιμών έναντι εκείνου της προηγούμενης περιόδου καλείται ποσοστό πληθωρισμού.

Από Αναλογιστική άποψη μας ενδιαφέρει η επίπτωση στις τιμές των ζημιών ενός χαρτοφυλακίου, και η αβεβαιότητα που συνοδεύει την αύξηση των τιμών. Τα υπάρχοντα οικονομικά ή αναλογιστικά μοντέλα θα πρέπει να τροποποιούνται κατάλληλα προκειμένου να εκφράζουν το τρέχον επίπεδο των απωλειών, εφόσον έχουν εκτιμηθεί από παρατηρήσεις που συλλέχθηκαν στο παρελθόν.

Σε αυτή την παράγραφο θα μελετηθεί η επίδραση του πληθωρισμού στην διαφορική εντροπία.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Ορισμός 2.4.1**

Έστω $X$ η τυχαία μεταβλητή, η οποία εκφράζει τις απώλειες ενός έτους. Τότε, η τυχαία μεταβλητή που θα εκφράζει τις απώλειες μετά από ένα έτος και υπό την επίδραση του πληθωρισμού θα είναι η

$$X(r) = (1 + r)X, \tag{2.14}$$

με $r$ τέτοιο ώστε *0>r>1*, ο ετήσιος ρυθμός πληθωρισμού.

Με άλλα λόγια, αν *X* είναι η τιμή της τρέχουσας απώλειας και $r$ ο ετήσιος ρυθμός πληθωρισμού τότε η τιμή της απώλειας μετά το πέρας ενός έτους θα είναι $X(r) = (1 + r)X$.

Η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής $X(r)$ είναι:

$$f_{X(r)}(z) = \frac{1}{1 + r} f_X\left(\frac{z}{1 + r}\right), \ \ z \in \mathbb{R}, \tag{2.15}$$

όπου $f_X$: συνάρτηση πυκνότητας πιθανότητας της X.

Η συνάρτηση κατανομής της $X(r)$ θα είναι:

$$F_{X(r)}(x) = Pr(X(r) \leq x) = Pr\left(\frac{X(r)}{1 + r} \leq \frac{x}{1 + r}\right) = F_X\left(\frac{x}{1 + r}\right) \tag{2.16}$$

με $F_X(\cdot)$: συνάρτηση κατανομής της *X*.

Η σχέση μεταξύ της εντροπίας της μεταβλητής $X$ και $X(r)$ δίνεται στο παρακάτω λήμμα.

**Λήμμα 2.4.1**

Η εντροπία της τυχαίας μεταβλητής $X(r) = (1 + r)X$ η οποία εκφράζει τις απώλειες μετά από ένα χρόνο και υπό την επίδραση πληθωρισμού, με ετήσιο ρυθμό πληθωρισμού $r$: *rϵ(0,1)* δίνεται από τη σχέση:

$$H\big(X(r)\big) = H(X) + ln(1 + r), \tag{2.17}$$

όπου $H(X)$ η εντροπία της *X*.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Απόδειξη:**

Η απόδειξη θα γίνει με χρήση του Θεωρήματος 9.6.4 των Cover and Thomas (1991, σελίδα 233).

**Θεώρημα 2.4.1**

$$H(aX) = H(X) + \log|a|. \tag{2.18}$$

Έστω $(1+r) = a,$ με $r \in (0,1)$ ο ετήσιος ρυθμός πληθωρισμού και έστω η τυχαία μεταβλητή $Y = aX = g(x)$. Τότε, η $Y = aX$ έχει μοναδική λύση την $X = u(Y) = \frac{Y}{a}$.

$$f_Y(y) = f_X\big(u(Y)\big)|u'(Y)| = \frac{1}{|a|} f_X\left(\frac{Y}{a}\right) \qquad \text{και } dY = adX$$

$$H(aX) = -\int f_y(y) \log f_y(y) dy = -\int \frac{1}{|a|} f_x\left(\frac{y}{a}\right) log\left(\frac{1}{|a|} f_x\left(\frac{y}{a}\right)\right) adx.$$

Όμως $a > 1$ , επομένως,

$$H(aX) = -\int f_x(x) log\left(\frac{1}{a} f_x(x)\right) dx = -\int f_x(x) log\left(\frac{f_x(x)}{a}\right) dx$$

$$= -\int f_x(x) log\big(f_x(x)\big) dx + \int f_x(x) log\, adx$$
$$= H(X) + \log a \int f_x(x)\, dx = H(X) + \log a = H(X) + \log(1+r).$$

Η επέκταση για $k$ έτη είναι προφανής.

**Πρόταση 2.4.1**

Για $r > 0$, η διαφορική εντροπία της τυχαίας μεταβλητής $X(r)$ είναι πάντα μεγαλύτερη από την εντροπία της τυχαίας μεταβλητής $X$ και αύξουσα συνάρτηση ως προς $r$.

**Απόδειξη:**

Η απόδειξη είναι προφανής εφόσον ισχύει ότι $ln(1+r) \;> 0, \quad \forall\, r \epsilon (0,1)$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Όπως έχει αναφερθεί και στα προηγούμενα κεφάλαια, η εντροπία είναι ταυτόσημη με την αναμενόμενη αβεβαιότητα, στα πλαίσια ενός πειράματος τύχης. Η αύξηση της εντροπίας συνεπάγεται αύξηση της αβεβαιότητας. Στην παραπάνω πρόταση γίνεται σαφές ότι στην περίπτωση κατανομών απώλειας, υπό την επίδραση πληθωρισμού, η αβεβαιότητα θα είναι πάντα μεγαλύτερη συγκριτικά με τις κατανομές απώλειας άνευ πληθωρισμού.

Επιπλέον η εντροπία είναι αύξουσα συνάρτηση ως προς το ρυθμό του πληθωρισμού, άρα, ισοδύναμα, όσο αυξάνει ο ρυθμός $r$ αυξάνεται και η εντροπία (αβεβαιότητα).

Συμπερασματικά, εφόσον η εντροπία αποτελεί περιγραφικό μέτρο της μεταβλητότητας μιας κατανομής, όσο αυξάνει ο πληθωρισμός, τόσο αυξάνει και η μεταβλητότητα των απωλειών.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Κεφάλαιο 3ο

## 3.1. Σχετική εντροπία και Kullback–Leibler απόκλιση

**Ορισμός 3.1.1** (διακριτή περίπτωση)

Έστω $X, Y$ τυχαίες μεταβλητές αντίστοιχες συναρτήσεις μάζας πιθανότητας $p(x)$, $q(x), x \in X$. Η Σχετική εντροπία *(relative entropy)* ή αλλιώς Kullback-Leibler απόκλιση, μεταξύ των $p(x)$, $q(x), x \in X$ ορίζεται ως:

$$D_{KL}(p\|q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)} = E\left(\log \frac{p(x)}{q(x)}\right). \tag{3.1}$$

(Θεωρούμε πως $0\log\frac{0}{0} = 0 \log\frac{0}{q} = 0, \ \ p\log\frac{1}{0} = \infty$).

**Ορισμός 3.1.2** (συνεχής περίπτωση)

Έστω $X, Y$ τυχαίες μεταβλητές με αντίστοιχες συναρτήσεις πυκνότητας πιθανότητας $f_X$, $f_Y$. Η Σχετική εντροπία *(relative entropy)* ή αλλιώς Kullback-Leibler απόκλιση, μεταξύ των $f_X(x)$, $f_Y(x)$, ορίζεται ως:

$$D_{KL}(f_X\|f_Y) = \int_0^\infty f_X(x) \log\left(\frac{f_X(x)}{f_Y(x)}\right) dt, \tag{3.2}$$

με στήριγμα $\mathcal{S} = \{X, Y: f_X > 0, f_Y > 0\}$.

Θεωρούμε $0\log 0 = 0$.

Η $D_{KL}(f_X\|f_Y)$ είναι καλώς ορισμένη αν και μόνο αν για $f_Y(x) = 0$ έχουμε $f_X(x) = 0$.

Η απόκλιση Kullback-Leibler εκφράζει το ποσό της πληροφορίας που χάνεται όταν μια συνάρτηση πιθανότητας χρησιμοποιείται προκειμένου να εκτιμηθεί μια άλλη.

Βάσει των ορισμών που παρατέθηκαν, οι συναρτήσεις $q(x)$, $(f_Y)$ χρησιμοποιούνται για να εκτιμήσουν τις $p(x)$, $(f_X)$ (διακριτή και συνεχής περίπτωση αντίστοιχα).

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Αναφέρεται συχνά και ως διασταυρούμενη εντροπία *(cross entropy)* ή πληροφορία διαχωρισμού *(discrimination entropy).*

## 3.2. Ιδιότητες της Kullback–Leibler απόκλισης ως μέτρο πληροφορίας

### Εισαγωγικές έννοιες-θεωρήματα:

Στο σημείο αυτό παραθέτουμε κάποιες βασικές έννοιες και θεωρήματα του Απειροστικού Λογισμού που θα χρησιμοποιηθούν για τις αποδείξεις των ιδιοτήτων της Kullback-Leibler απόκλισης.

**Ορισμός 3.2.1**

Έστω $f(x)$ πραγματική συνάρτηση, ορισμένη διάστημα $[a, b]$.

Τότε $f$ καλείται κυρτή *(convex)* στο διάστημα $(a, b)$ αν $\forall x_1, x_2 \in [a, b]$ και $\forall \lambda :$ $0 \leq \lambda \leq 1$ ισχύει

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2). \qquad (3.3)$$

Η $f(x)$ καλείται αυστηρώς κυρτή αν η ισότητα ισχύει αυστηρά για $x_1 \neq x_2$.

Η $f(x)$ καλείται κοίλη *(concave)* αν η - $f(x)$ είναι κυρτή.

### Ανισότητα Jensen

**Θεώρημα 3.2.2**

Έστω $f(x)$: $[a, b] \to \mathbb{R}$ κυρτή συνάρτηση και $p_i \geq 0 \; για \; i = 0,1, \dots n$ ,τέτοια ώστε:

$$\sum_{i=1}^{n} p_i = 1.$$

Τότε, $\forall x_j \in [a, b]$ , $j = 1,2, \dots n$ ισχύει:



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i=1}^{n} p_i f(x_i). \quad (3.4)$$

Για $n = 2$ προκύπτει ο ορισμός της κυρτής συνάρτησης.

Ισοδύναμα,

αν $f(x)$: κυρτή και $x \in \{x_1, x_2, \ldots, x_n\}$ τυχαία μεταβλητή με μάζες πιθανότητας $p(x_i)$ με $\sum_{i=1}^{n} p(x_i) = 1$ τότε:

$$f\big(E(x)\big) \leq E\big(f(x)\big). \quad (3.5)$$

Επιπλέον, αν $f(x)$ αυστηρά κυρτή τότε $X = E(X)$ με πιθανότητα 1, δηλαδή $X$: σταθερή (Cover and Thomas, 1991).

## Log-sum ανισότητα:

**Θεώρημα 3.2.3**

Έστω $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \geq 0$. Τότε,

$$\sum_{i=1}^{n} a_i \log\frac{1}{b_i} \leq \left(\sum_{i=1}^{n} a_i\right) \log\frac{(\sum_{i=1}^{n} a_i)}{(\sum_{i=1}^{n} b_i)} \quad (3.6)$$

ή

$$\int P(X) \log\frac{P(X)}{Q(X)} dx \geq \left(\int P(X) dx\right) \log\left(\frac{\int P(X) dx}{\int Q(X) dx}\right). \quad (3.7)$$

## Ιδιότητες:

I. Η απόκλιση K-L δεν είναι συμμετρική, ισοδύναμα $D(p\|q) = D(q\|p)$ δεν ισχύει πάντα.

Ωστόσο δύο συμμετρικές εκδοχές του μέτρου αποτελούν:

i. $J(f_X, f_Y) = D_{KL}(f_X\|f_Y) + D_{KL}(f_Y\|f_X)$ (Jeffreys,1948).
ii. $\delta(f_X, f_Y) = \min\{D_{KL}(f_X\|f_Y) D_{KL}(f_Y\|f_X)\}$,

ως μέτρο εσωτερικής πληροφορίας (*intrinsic information*) μεταξύ των $f_Y$ , $f_X$ (Bernardo and Rueda, 2002).

II. Η απόκλιση K-L είναι πάντα μη- αρνητική.

**Θεώρημα 3.2.4** *Ανισότητα Πληροφορίας (information inequality)- (Gibb's inequality)*

Έστω $p(x)$ και $q(x), x \in X$ συναρτήσεις μάζας πιθανότητας.

Τότε,

$$D_{KL}(p\|q) \geq 0 \tag{3.8}$$

και $D_{KL}(p\|q) = 0 \; αν \; και \; μόνο \; αν \; p(x) = q(x), \; \forall x \epsilon X.$

**Απόδειξη:**

Έστω $\mathcal{S} = \{x/ \, p(x) > 0\} \subset X$, το στήριγμα της $p(x)$ και $\sum_{x \in X} q(x) = 1$ .Τότε,

$$-D(p\|q) = -\sum_{x \in \mathcal{S}} p(x) \log \frac{p(x)}{q(x)} = \sum_{x \in \mathcal{S}} p(x) \log \frac{q(x)}{p(x)}$$

$$\leq \log \sum_{x \in \mathcal{S}} p(x) \frac{q(x)}{p(x)} \qquad (logt: αυστηρά \; κοίλη \; \forall t)$$

$$= \log \sum_{x \in \mathcal{S}} q(x) \leq \log \sum_{x \in X} q(x) = \log 1 = 0.$$

Η ισότητα ισχύει αν και μόνο αν $\frac{q(x)}{p(x)} = 1 \; \forall x \in X.$

Άρα $D_{KL}(p\|q) = 0, \leftrightarrow p(x) = q(x) \forall x \in X.$

**Θεώρημα 3.2.5** *Ανισότητα Πληροφορίας (information inequality)-(Gibb's inequality)*

Έστω $f_1(x), f_2(x)$ με $x \in X$ δύο συναρτήσεις πιθανότητας. Τότε,

$$D_{KL}(f_1\|f_2) \geq 0 \tag{3.9}$$

με $D_{KL}(f_1\|f_2) = 0$ ανν και μόνο αν $f_1 = f_2 , \quad \forall x \in X.$

**Απόδειξη:**



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Έστω $\mathcal{S} = \{x/f_1(x) > 0\} \subset X$, το στήριγμα της $f_1(x)$ και $\int_X f_2(x)dx = 1$ .Τότε,

$$-D_{KL}(f_1\|f_2) = -\int_A f_1(x)\ln\frac{f_2(x)}{f_1(x)}dx$$

$$= \int_A f_1(x)\ln\frac{f_2(x)}{f_1(x)}dx \leq \ln\int_A f_1(x)\frac{f_2(x)}{f_1(x)}dx$$

$$= \ln\int_A f_2(x)dx \leq \ln\int_X f_2(x)dx = \ln 1 = 0.$$

Η συνάρτηση $\ln t$ είναι αυστηρά κοίλη συνάρτηση του $t$ επομένως η $t\ln t$ αυστηρά κυρτή συνάρτηση του $t$ και η ισότητα ισχύει αν και μόνο αν $\frac{f_2(x)}{f_1(x)} = 1,\ \forall x \in X$, δηλαδή αν $f_1(x)$=$f_2(x)$, $\forall x \in X$ (Cover and Thomas, 1991).

III. Η απόκλιση K-L είναι κυρτή συνάρτηση.

**Λήμμα 3.2.1**

Έστω $p_1, q_1$ και $p_2, q_2$ μάζες πιθανότητας μιας τυχαίας μεταβλητής $X$ και $\forall\lambda \in (0,1)$

$$p = \lambda p_1 + (1-\lambda)p_2\,,$$
$$q = \lambda q_1 + (1-\lambda)q_2.$$

Τότε,

$$D_{KL}(p\|q) \leq \lambda D_{KL}(p_1\|q_1) + (1-\lambda)D_{KL}(p_2\|q_2). \tag{3.10}$$

**Απόδειξη:**

Έστω $a_1(x) = \lambda p_1(x)$, $a_2(x) = (1-\lambda)p_2(x)$, $b_1(x)\lambda q_1(x)$, $b_2(x) = (1-\lambda)q_2(x)$.

Τότε,

$$D_{KL}(p\|q) = \sum_x \big(\lambda p_1(x) + (1-\lambda)p_2(x)\big)\log\frac{\big(\lambda p_1(x) + (1-\lambda)p_2(x)\big)}{(\lambda q_1(x) + (1-\lambda)q_2(x))}$$

$$= \sum_x (a_1(x) + a_2(x))\log\frac{a_1(x) + a_2(x)}{b_1(x) + b_2(x)}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$\leq \sum_{x} a_1(x)\log\frac{a_1(x)}{b_1(x)} + a_2(x)\log\frac{a_2(x)}{b_2(x)} \quad (log - sum\ ανισότητα)$$

$$= \sum_{x} \lambda p_1(x)\log\frac{\lambda p_1(x)}{\lambda q_1(x)} + (1-\lambda)p_2(x)\log\frac{(1-\lambda)p_2(x)}{(1-\lambda)q_2(x)}$$

$$= \lambda D_{KL}(p_1\|q_1) + (1-\lambda)D_{KL}(p_2\|q_2).$$

**Ένα κάτω φράγμα για την απόκλιση Kullback-Leibler**

**Λήμμα 3.2.2** *(Divergence and total Variation)*

Για οποιεσδήποτε $p, q$ συναρτήσεις ισχύει:

$$D_{KL}(p\|q) \geq \frac{1}{2\ln 2}\|p-q\|_1^2. \tag{3.11}$$

## 3.3. Kullback-Leibler απόκλιση και Μέγιστη Εντροπία

**Θεώρημα 3.3.1**

$H(X) \leq \log|X|$, όπου $|X|$ το πλήθος των στοιχείων του πεδίου τιμών της $X$ με ισότητα αν και μόνο αν η $X$ είναι ομοιόμορφα κατανεμημένη επί του $|X|$.

**Απόδειξη:**

Έστω $u(x) = \frac{1}{|X|}$ η ομοιόμορφη συνάρτηση μάζας πιθανότητας επί του $X$ και $p(x)$ συνάρτηση μάζας πιθανότητας της $X$. Τότε,

$$D_{KL}(p\|u) = \sum_{x\in X} p(x)\log\frac{p(x)}{u(x)} = \sum_{x\in X} p(x)\log\frac{p(x)}{\frac{1}{|X|}} = \log\ |X| - H(X). \tag{3.12}$$

Λόγω της μη-αρνητικότητας του μέτρου έπεται ότι:

$$0 \leq D_{KL}(p\|u) = \log\ |X| - H(X),$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ή ισοδύναμα,

$$H(X) \leq \log|X|. \tag{3.13}$$

Το παραπάνω θεώρημα μας λέει ότι η μέγιστη εντροπία επιτυγχάνεται όταν η τυχαία μεταβλητή $X$ ακολουθεί την ομοιόμορφη κατανομή.

## 3.4. Υπό συνθήκη απόκλιση

**Ορισμός 3.4.1**

Έστω οι από κοινού κατανομές πιθανότητας $p(x,y)$ και $q(x,y)$ των δύο τυχαίων μεταβλητών $X, Y$. Τότε η υπό συνθήκη απόκλιση μεταξύ των δεσμευμένων κατανομών $p(y|x)$ και $q(y|x)$ δίνεται από τη σχέση:

$$D_{KL}\big(p(y|x)\|q(y|x)\big) = \sum_{x\in X} p(x) \sum_{y\in Y} p(y|x) \log\frac{p(y|x)}{q(y|x)}. \tag{3.14}$$

**Λήμμα 3.4.2**

$$D_{KL}\big(p(x,y)\|q(x,y)\big) = D_{KL}\big(p(x)\|q(x)\big) + D_{KL}\big(p(y|x)\|q(y|x)\big). \tag{3.15}$$

**Απόδειξη:**

$$\begin{aligned}
D_{KL}\big(p(x,y)\|q(x,y)\big) &= \sum_x \sum_y p(x,y) \log\frac{p(x,y)}{q(x,y)} \\
&= \sum_x \sum_y p(x)p(y|x) \log\frac{p(x)p(y|x)}{q(x)q(y|x)} \\
&= \sum_x \sum_y p(x)p(y|x) \log\frac{p(x)}{q(x)} + \sum_x \sum_y p(x)p(y|x) \log\frac{p(y|x)}{q(y|x)} \\
&= \sum_x p(x) \log\frac{p(x)}{q(x)} \sum_y p(y|x) + \sum_x p(x) \sum_y p(y|x) \log\frac{p(y|x)}{q(y|x)}
\end{aligned}$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$= D_{KL}\big(p(x)\|q(x)\big) + D_{KL}\big(p(y|x)\|q(y|x)\big).$$

## 3.5. Αμοιβαία πληροφορία

**Ορισμός 3.5.1**

Έστω $X, Y$ διακριτές με από κοινού συνάρτηση μάζας πιθανότητας $p(x, y)$ και περιθώριες $p(x)$, $p(y)$. Τότε ως αμοιβαία πληροφορία (*mutual information*) ή συμβολικά $I(X;Y)$ ορίζεται η σχετική εντροπία μεταξύ της από κοινού κατανομής και του γινομένου των περιθωρίων, δηλαδή

$$I(X;Y) = D_{KL}\big(p(x,y)\|p(x)p(y)\big)$$

$$= E\left(\log\frac{p(x,y)}{p(x)p(y)}\right) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}. \tag{3.16}$$

Αντίστοιχα, για τη συνεχή περίπτωση έχουμε:

**Ορισμός 3.5.2**

Έστω οι δύο τυχαίες μεταβλητές $X$ και $Y$ με από κοινού συνάρτηση πυκνότητας πιθανότητας $p(x, y)$ και περιθώριες συναρτήσεις πυκνότητας πιθανότητας $f(x)$, $g(y)$ αντίστοιχα. Τότε η σχετική εντροπία μεταξύ της από κοινού κατανομής και του γινομένου των περιθωρίων κατανομών (αμοιβαία πληροφορία), ορίζεται ως:

$$I(X;Y) = \iint p(x,y)\ln\frac{p(x,y)}{f(x)g(y)}dxdy = D_{KL}\big(p(x,y)\|f(x)g(y)\big). \tag{3.17}$$

Το παραπάνω μέτρο ερμηνεύεται ως η πληροφορία που παρέχει η μια μεταβλητή για την άλλη (Soofi, 1994).

## 3.6. Σχέσεις που συνδέουν την αμοιβαία πληροφορία με την εντροπία

**Θεώρημα 3.6.1** (*Mutual information and entropy)*

Ισχύουν οι παρακάτω σχέσεις:



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

1. $I(X;Y) = H(X) - H(X|Y)$
2. $I(X;Y) = H(Y) - H(Y|X)$
3. $I(X;Y) = H(X) + H(Y) - H(X,Y)$
4. $I(X;Y) = D(Y\|X)$
5. $I(X;X) = H(X)$

**Λήμμα 3.6.1** *(Chain Rule)*

$$D_{KL}(f_1(X,Y)\|f_2(X,Y)) = D_{KL}\big(f_1(X)||f_2(X)\big) + D_{KL}\big(f_1(Y)||f_2(Y)\big). \tag{3.18}$$

Το μέτρο Kullback-Leibler, όπως ορίστηκε παραπάνω, δεν αποτελεί μετρική, καθώς δεν είναι συμμετρικό και επιπλέον δεν ικανοποιεί την τριγωνική ανισότητα. Επομένως δε μπορεί να θεωρηθεί καθαρή απόσταση. Στην ουσία πρόκειται για μέτρο κατευθυνόμενης απόκλισης.

## 3.7. (MDI) Το Κανονικό Πρόβλημα Ελάχιστης Διαχωρίσιμης Πληροφορίας

Έστω η απόκλιση Kullback-Leibler

$$D_{KL}(p\|q) = \begin{cases} \displaystyle\sum_{x\in X} p(x)\log\frac{p(x)}{q(x)}, \\ \displaystyle\int_0^{\infty} p(x)\log\left(\frac{p(x)}{q(x)}\right)dt, \end{cases}$$

για τη διακριτή και συνεχή περίπτωση αντίστοιχα και έστω ότι μας ενδιαφέρει η εύρεση συνάρτησης $\boldsymbol{p}$ η οποία να βρίσκεται όσο το δυνατόν πιο «κοντά» στη γνωστή συνάρτηση $\boldsymbol{q}$, ή ισοδύναμα, η ελαχιστοποίηση της «απόστασης» $D_{KL}(p\|q)$ ως προς $\boldsymbol{p}$. Τότε, η ελάχιστη (βέλτιστη) κατανομή $\widehat{\boldsymbol{p}}$ που προκύπτει από τη διαδικασία ελαχιστοποίησης, καλείται Ελάχιστη Διαχωριστική Πληροφορία (Minimum Discrimination Information - MDI).

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Η μη-αρνητικότητα του μέτρου, όπως αναφέρεται στο θεώρημα Ανισότητας Πληροφορίας, εξασφαλίζει ότι η εκτιμήτρια κατανομή $\widehat{\boldsymbol{p}}$ που θα προκύψει από τη διαδικασία της ελαχιστοποίησης υπό κάποιους γραμμικούς περιορισμούς είναι όσο το δυνατόν πιο «κοντά» (ελάχιστα διαχωρίσιμη) στην $\boldsymbol{q}$, ενώ ταυτόχρονα ικανοποιεί περιορισμούς που ίσως η $\boldsymbol{q}$ δεν ικανοποιεί (Brockett, 1991).

**Συνεχής περίπτωση:**

$$Minimize D_{KL}(\boldsymbol{p}\|\boldsymbol{q}) = \int_{-\infty}^{+\infty} p(x) \ln \frac{p(x)}{q(x)} dx \tag{3.19}$$

υπό τους περιορισμούς

$$E\left[a_j(X)\right] = \theta_j, j = 1,2, \dots, m, \tag{3.20}$$

έτσι ώστε

$$1 = \theta_0 = \int_{-\infty}^{+\infty} p(x)\, dx \tag{3.21}$$

και

$$\theta_j = \int_{-\infty}^{+\infty} a_j(x) p(x)\, dx\,, \quad j = 1,2, \dots, m. \tag{3.22}$$

Με την επιλογή $a_j(x) = x^j$, εισάγονται στο πρόβλημα ελαχιστοποίησης περιορισμοί της μορφής ροπών, όπως για παράδειγμα μέσες τιμές και διακυμάνσεις της συνάρτησης $p$.

Επίσης, αν η $a_j(x)$ έχει τη μορφή δείκτριας συνάρτησης, τότε εισάγονται περιορισμοί που αφορούν πιθανότητες, όπως πιθανότητες επιβίωσης.

**Διακριτή περίπτωση:**



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$Minimize D_{KL}(\boldsymbol{p}\|\boldsymbol{q}) = \sum_{i=1}^{n} p_i \ln\frac{p_i}{q_i} \tag{3.23}$$

υπό τους περιορισμούς:

$$\sum_{i=1}^{n} p_i = 1 \tag{3.24}$$

$$\sum_{i=1}^{n} a_{ij} p_i = \theta_j, \quad j = 1,2,\dots,m. \tag{3.25}$$

$$p_i \geq 0, \forall i = 1,2,\dots,n. \tag{3.26}$$

Ισοδύναμα, υπό τη μορφή πινάκων, οι περιορισμοί εκφράζονται ως:

$$\boldsymbol{Ap} = \boldsymbol{\theta}, \qquad \boldsymbol{p} \geq \boldsymbol{0}, \tag{3.27}$$

όπου $\boldsymbol{A}$ είναι ένας πίνακας $(m+1)\times n$ με στοιχεία της πρώτης γραμμής ίσα με τη μονάδα και $\theta_0 = 1$ προκειμένου να ικανοποιείται ο περιορισμός $\sum_{i=1}^{n} p_i = 1$, ώστε η $\boldsymbol{p}$ να αποτελεί κατανομή πιθανότητας.

Το παραπάνω πρόβλημα βελτιστοποίησης (συνεχής και διακριτή περίπτωση) λύνεται με χρήση πολλαπλασιαστών Lagrange.

Ο Brockett (1991), αναφέρει πως στην περίπτωση που έχουμε πιθανότητες θνησιμότητας οι οποίες δεν αθροίζουν στη μονάδα, όταν παραδείγματος χάρη χρησιμοποιούμε τμήμα του πίνακα, μπορούμε να διαγράψουμε τον πρώτο περιορισμό χωρίς να δημιουργείται πρόβλημα.

Στο παραπάνω πρόβλημα ελαχιστοποίησης, οι πιθανότητες $q_i$ και οι τιμές $\theta_j$ είναι σταθερές. Επιπλέον μπορεί να προσδιορίζονται εξωγενώς ή μέσω του δείγματος.

Η ελαχιστοποίηση αφορά τον έλεγχο της μηδενικής υπόθεσης $H_0: p = q$, δηλαδή ότι οι παρατηρούμενες και αναμενόμενες κατανομές είναι στατιστικώς μη διαχωρίσιμες. Σε περίπτωση που η υπόθεση $H_0$ δεν είναι αληθής τότε το ασυμπτωτικό όριο της $D_{KL}(\boldsymbol{p}\|\boldsymbol{q})$ είναι η «απόσταση» μεταξύ όλων των κυρτών κατανομών $\boldsymbol{p}$ που ικανοποιούν τους περιορισμούς και της υποθετικής γνωστής συνάρτησης $\boldsymbol{q}$ (Brockett, 1991).

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Αν θεωρήσουμε $\widehat{\boldsymbol{p}}$ βέλτιστο διάνυσμα, τότε $D_{KL}(\,\widehat{\boldsymbol{p}}\|\boldsymbol{q})$ είναι η ελάχιστη υπό τους περιορισμούς λύση του προβλήματος. Αν θεωρήσουμε ότι η παρατηρούμενη κατανομή πιθανότητας προέρχεται από ένα τυχαίο διάνυσμα μεγέθους N από την κατανομή πιθανότητας $\boldsymbol{q}$ τότε η ποσότητα 2N $D(\,\widehat{\boldsymbol{p}}\|\boldsymbol{q})$ ακολουθεί ασυμπτωτικά $x^2$ κατανομή με βαθμούς ελευθερίας που εξαρτώνται από το μέγεθος των διανυσμάτων $n$ και τον αριθμό των γραμμικώς ανεξάρτητων περιορισμών (Kullback, 1959).

## 3.8. Η Αρχή της Μέγιστης Εντροπίας ως ισοδύναμη του MDI–(MEP)

Μια κοινή πρακτική που ακολουθείται όταν δεν διατίθεται γνωστή κατανομή της $\boldsymbol{q}$, είναι η υπόθεση της διακριτής ομοιόμορφης κατανομής της $\boldsymbol{q}$, με $qi=\frac{1}{n}$,$\forall i$=1,2,…,$n$.

Τότε η Kullback-Leibler απόκλιση, όπως αποδείχθηκε στο Θεώρημα 3.3.1 προκύπτει ως:

$$D_{KL}(\,\boldsymbol{p}\|\boldsymbol{q}) = -H(\boldsymbol{p}) + \ln n\,,$$

όπου $H(\boldsymbol{p})$ η εντροπία της $\boldsymbol{p}$.

Από την σχέση γίνεται σαφές ότι το πρόβλημα ελαχιστοποίησης της απόκλισης Kullback-Leibler είναι ισοδύναμο με τη μεγιστοποίηση της εντροπίας, δηλαδή

$$Minimize \sum_{i=1}^{n} p_i \ln\frac{p_i}{q_i} \Leftrightarrow Maximize \sum_{i=1}^{n} p_i \ln\frac{1}{p_i}, \quad \text{όταν} \quad q_i = \frac{1}{n}, \forall i = 1,2,\dots,n. \tag{3.28}$$

Ο Brockett (1991) ωστόσο αναφέρει ότι το μοντέλο MDI παρέχει μεγαλύτερη ευελιξία, καθώς η συνάρτηση $\boldsymbol{q}$ μπορεί να είναι οποιαδήποτε γνωστή κατανομή και όχι αποκλειστικά η ομοιόμορφη. Το μοντέλο μέγιστης εντροπίας μπορεί να θεωρηθεί ότι επιλέγει την πιο «αβέβαιη» υπό τους περιορισμούς κατανομή.

## 3.9. Εφαρμογές σε θέματα Αναλογισμού

Στο σημείο αυτό θα παρουσιάσουμε τη γενική διαδικασία ελαχιστοποίησης της απόκλισης Kullback-Leibler και κατόπιν θα παρουσιαστεί εφαρμογή σε δύο γνωστές Αναλογιστικές υποθέσεις, αυτή της ομοιόμορφης κατανομής θανάτων (UDD) και της σταθερής έντασης θνησιμότητας.

### Γενική διαδικασία επίλυσης

Όπως ήδη έχει αναφερθεί η διαδικασία του προβλήματος ελαχιστοποίησης της $D_{KL}(\boldsymbol{p}\|\boldsymbol{q})$ χρησιμοποιεί πολλαπλασιαστές Lagrange προκειμένου να εξάγει την ελάχιστη τιμή της $\boldsymbol{p}$, $\widehat{\boldsymbol{p}}$.

Τότε για κάθε περιορισμό $\theta_j = \int_{-\infty}^{+\infty} a_j(x)p(x)\,dx$, όπου $a_j(x) = x^j$, απαιτείται η εισαγωγή $z_j$ πολλαπλασιαστών Lagrange, και το πρόβλημα ανάγεται στη μορφή:

$$MinimizeD_{KL}(\boldsymbol{p}\|\boldsymbol{q}) =$$

$$Minimize[\int_{-\infty}^{+\infty} p(x)\ln\frac{p(x)}{q(x)}d - z_0\int_{-\infty}^{+\infty} p(x)\,d - z_1\int_{-\infty}^{+\infty} a_1(x)p(x)\,dx - \cdots - z_m\int_{-\infty}^{+\infty} a_m(x)p(x)\,dx\,]. \tag{3.29}$$

Πολλαπλασιάζοντας την παραπάνω σχέση με (-1) οδηγούμαστε στην ισοδύναμη έκφραση:

$$Maximize[\int_{-\infty}^{+\infty} p(x)\ln\frac{q(x)}{p(x)}d + z_0\int_{-\infty}^{+\infty} p(x)\,d + z_1\int_{-\infty}^{+\infty} a_1(x)p(x)\,dx\, +$$

$$\ldots + z_m\int_{-\infty}^{+\infty} a_m(x)p(x)\,dx\,] = \int_{-\infty}^{+\infty} p(x)\left\{\ln\frac{q(x)}{p(x)} + \sum_{j=1}^{m} z_j a_j(x)p(x)\right\}dx$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$= \int_{-\infty}^{+\infty} p(x) \left\{ \ln \frac{q(x) e^{\sum_{j=1}^{m} z_j a_j(x) p(x)}}{p(x)} \right\} dx$$

$$\leq \int_{-\infty}^{+\infty} p(x) \left\{ \ln \frac{q(x) e^{\sum_{j=1}^{m} z_j a_j(x) p(x)}}{p(x)} - 1 \right\} dx,$$

από την οποία έπεται ότι η ισότητα με το άνω φράγμα ισχύει για:

$$z = \frac{q(x) e^{\sum_{j=1}^{m} z_j a_j(x) p(x)}}{p(x)} = 1 \tag{3.30}$$

ή

$$p(x) = q(x) e^{\sum_{j=1}^{m} z_j a_j(x) p(x)}. \tag{3.31}$$

Όπως γίνεται εμφανές, η λύση παραπέμπει σε εκθετικές οικογένειες κατανομών.

## I. Η υπόθεση της ομοιόμορφης κατανομής θανάτων (UDD)

Έστω ότι η μοναδική πληροφορία που διαθέτουμε είναι η πιθανότητα θανάτου σε ένα διάστημα της μορφής $(a, b)$.

Τότε, η πληροφορία αυτή μεταφράζεται στον περιορισμό:

$$\theta_0 = \int_{-\infty}^{+\infty} a_0(x) p(x)\, dx$$

με

$$a_0 = \begin{cases} 0, & x < a \\ 1, & a \leq x \leq b \\ 0, & x \geq b \end{cases}$$

όπου η $x$ η μεταβλητή που εκφράζει την ηλικία και $\theta_0$ η γνωστή σε μας ηλικία θανάτου.

Τότε η σχέση (3.30) αντιστοιχεί την βέλτιστη εκτίμηση στην ομοιόμορφη κατανομή. Όπως είναι προφανές, η ομοιόμορφη κατανομή θανάτων παρέχει την ελάχιστη πληροφορία (μέγιστη εντροπία).

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

## II. Η υπόθεση σταθερής έντασης θνησιμότητας

Έστω ότι γνωρίζουμε την πιθανότητα θανάτου και τη μέση διάρκεια ζωής στο διάστημα $(a, b)$.

Τότε εισάγεται στο πρόβλημα ένας επιπλέον περιορισμός της μορφής:

$$\theta_1 = \int_{-\infty}^{+\infty} a_1(x) p(x)\, dx$$

με

$$a_1 = \begin{cases} 0, & x < a, \\ x, & a \leq x \leq b, \\ 0, & x \geq b. \end{cases}$$

Τότε από τη σχέση (3.30) προκύπτει:

$$p(x) = e^{\,z_0 + \,z_1}, \tag{3.32}$$

το οποίο ισοδυναμεί με την υπόθεση σταθερής έντασης θνησιμότητας. Έτσι, μπορεί να θεωρηθεί ότι η υπόθεση παρέχει την ελάχιστη δυνατή πληροφορία που μπορούμε να έχουμε γνωρίζοντας την πιθανότητα θανάτου και τη μέση διάρκεια ζωής σε ένα διάστημα $(a, b)$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Κεφάλαιο 4ο

## 4.1. Υπολειπόμενη Εντροπία (Residual Entropy)

Καθώς το μέτρο της εντροπίας του Shannon δεν ήταν εφαρμόσιμο σε συστήματα που έχουν επιβιώσει μέχρι κάποια μονάδα χρόνου, ήταν αναγκαία η εισαγωγή ενός νέου μέτρου για την εκτίμηση της αβεβαιότητας του «προσδόκιμου» ενός συστήματος που έχει «επιβιώσει» μέχρι την ηλικία $t$.

Η έννοια της υπολειπόμενης εντροπίας *(Residual Entropy)* εισήχθη από τους Ebrahimi and Pellerey (1995) και Ebrahimi (1996) και εκφράζει την αβεβαιότητα που εμπεριέχεται στον υπολειπόμενο χρόνο ζωής συστημάτων που έχουν επιβιώσει μέχρι το χρόνο $t$ , με $X_t = (X - t \,|\, X \geq t)$ και ορίστηκε ως:

$$H(X;t) = -\int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)} dx$$

$$= 1 - \frac{1}{\bar{F}(t)} \int_t^{\infty} f(x) \log h(x)\, dx = 1 - E(\log h(X) \,| X \geq t) \qquad (4.1)$$

$t \in D = \{x \in \mathbb{R} : \bar{F}(x) > 0\}$,

όπου $h(x) = \frac{f(x)}{\bar{F}(t)}$ η συνάρτηση κινδύνου της $X$ και $\bar{F}(X) = P(X > t)$ η συνάρτηση επιβίωσης της $X$.

Είναι προφανές ότι για $t = 0$ :

$$H(X;0) = -\int_0^{\infty} f(x) \log f(x)\, dx = H(X),$$

όπου $H(X)$ η εντροπία του Shannon.

Παρατηρούμε πως το νέο μέτρο αποτελεί έκφραση της υπό συνθήκη εντροπίας του Shannon (Ebrahimi, 1996). Στην πραγματικότητα, η υπολειπόμενη εντροπία μετρά την αναμενόμενη αβεβαιότητα που εμπεριέχεται στην υπό συνθήκη (δεσμευμένη) πυκνότητα της $(X - t)$ ,δοθέντος ότι $X > t$ .

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Ο Ebrahimi (1996) απέδειξε πως η υπολειπόμενη εντροπία προσδιορίζει μονοσήμαντα την κατανομή της $X$, όπως η συνάρτηση κινδύνου, και επιπλέον όρισε δύο νέες κλάσεις κατανομών, βάσει της μονοτονίας του μέτρου της υπολειπόμενης εντροπίας.

Στο πρώτο μέρος αυτής της ενότητας θα εξετάσουμε υπό ποιες προϋποθέσεις ισχύει ο χαρακτηρισμός των κατανομών και θα αναφέρουμε τις οικογένειες γήρανσης ως επέκταση των κλάσεων που ορίζει η συνάρτηση κινδύνου $h(t)$.

Κατόπιν θα δούμε πως το μέτρο επεκτείνεται στον τομέα των Μοντέλων Απώλειας, όπου τη θέση του υπολειπόμενου χρόνου καταλαμβάνει το αφαιρετέο ποσό $d$.

Στο εξής η μελέτη θα αναφέρεται μόνο σε συνεχείς, μη αρνητικές τυχαίες μεταβλητές.

## Εισαγωγικές έννοιες

**Ορισμός 4.1.1**

Συνάρτηση κινδύνου *(hazard rate)* ορίζεται ως η υπό συνθήκη πιθανότητα αποτυχίας στο διάστημα $(t, t+x]$ δοθέντος ότι δεν έχει υπάρξει αποτυχία μέχρι τη στιγμή $t$ και δίνεται από τη σχέση:

$$h(t) = \frac{\lim\limits_{x\to 0} \mathrm{P}(X \leq \mathrm{t} + X | X > \mathrm{t})}{x} \quad \text{, ή ισοδύναμα,}$$

$$h(t) = \frac{f(t)}{\bar{F}(t)} = -(\ln \bar{F}(t))', \quad t > 0, \tag{4.2}$$

όπου $\bar{F}(t) = P(X > t)$ η συνάρτηση επιβίωσης της $X$ και η $f(t) = -\bar{F}'(X)$.

**Ορισμός 4.1.2**

Μέσος υπολειπόμενος χρόνος *(mean residual life)* της τυχαίας μεταβλητής $X_t = (X - t \,|\, X \geq t)$ ορίζεται ως:

$$m(t) = E(X_t) = \int_0^{\infty} \bar{F}(X_t)dx = \int_0^{\infty} \frac{\bar{F}(x+t)}{\bar{F}(t)}dx = \int_t^{\infty} \frac{\bar{F}(x)}{\bar{F}(t)}dx\,, t \geq 0. \quad (4.3)$$

Προφανώς, $m(0) = E(X)$.

**Ορισμός 4.1.3**

Μια συνεχής τυχαία μεταβλητή (ή αντίστοιχα κατανομή) λέμε ότι έχει την ιδιότητα:

*i.* *IFR* (*Increasing Failure Rate)* αν η συνάρτηση κινδύνου $h(t)$ είναι αύξουσα ως προς $t$, και συμβολικά $X\in IFR$ ή $F\in IFR$.

*ii.* *DFR* (*Decreasing Failure Rate)* αν η συνάρτηση κινδύνου $h(t)$ είναι φθίνουσα ως προς $t$, και συμβολικά $X\in DFR$ ή $F\in DFR$.

**Ορισμός 4.1.4**

Μια συνεχής τυχαία μεταβλητή (ή αντίστοιχα κατανομή) λέμε ότι έχει την ιδιότητα:

*i.* *IMRL* (*Increasing Mean Residual Lifetime)* αν ο μέσος υπολειπόμενος χρόνος ζωής $m(t)$ είναι αύξουσα συνάρτηση ως προς $t$, και συμβολικά $X\in IMRL$ ή $F\in IMRL$.

*ii.* *DMRL (Decreasing Mean Residual Lifetime)* αν ο μέσος υπολειπόμενος χρόνος ζωής $m(t)$ είναι φθίνουσα συνάρτηση ως προς *t,* και συμβολικά $X\in$ *DMRL ή* $F\in$*DMRL.*

## 4.2. Υπολειπόμενη εντροπία και προσδιορισμός κατανομών

Στον τομέα της αξιοπιστίας και ανάλυσης επιβίωσης, όπου η τυχαία μεταβλητή $X$ περιγράφει το προσδόκιμο ζωής μιας μονάδας, τα μέτρα που προσδιορίζουν την κατανομή της τυχαίας μεταβλητής είναι η συνάρτηση κινδύνου (*hazard rate)* και ο μέσος υπολειπόμενος χρόνος ζωής *(mean residual life).* Μελετώντας τη μονοτονία των παραπάνω μέτρων, οι κατανομές κατατάσσονται σε κλάσεις με μονότονη γήρανση.

Στο σημείο αυτό θα εξετασθεί το πότε η υπολειπόμενη εντροπία μπορεί να χαρακτηρίσει μια κατανομή και έπειτα θα οριστούν δυο νέες κλάσεις κατανομών.

**Θεώρημα 4.2.1**

Έστω $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή, με απόλυτα συνεχή κατανομή $F(t)$ και αύξουσα υπολειπόμενη εντροπία $H(X;t)$. Τότε η $H(X;t)$ προσδιορίζει μονοσήμαντα την κατανομή $F(t)$.

**Απόδειξη:**

$$H(X;t) = -\int_t^{\infty} \frac{f(x)}{1-F(t)} \log \frac{f(x)}{1-F(t)} dx \quad \text{ή ισοδύναμα,}$$

$$\int_t^{\infty} f(x) \log f(x)\, dx = (1-F(t)) \log\big(1-F(t)\big) - (1-F(t))H(X;t).$$

Παραγωγίζοντας και τα δύο μέλη ως προς t έχουμε:

$$f(t) \log f(t) = f(t)[1 - H(X;t) + \log[(1-F(t)] + (1-F(t))H'(X;t) \text{ ή}$$

$$h(t)[\log h(t) + H(X;t) - 1] = H'(X;t), \tag{4.4}$$

όπου $h(x) = \dfrac{f(x)}{\bar{F}(x)}, \textit{συνάρτηση κινδύνου}.$

Έτσι, για t>0, η $h(t) = x$ είναι μια θετική λύση της παρακάτω εξίσωσης.

$$g(x) = x[H(X;t) -1 + \log x] - H'(X;t) = 0. \tag{4.5}$$

Παραγωγίζοντας και τα δύο μέλη ως προς x έχουμε:

$$g'(x) = H(X;t) + \log x. \tag{4.6}$$

Για $g'(x) = 0$ συνεπάγεται

$$x = e^{-H(X;t)}\ . \tag{4.7}$$

Έστω $x = e^{-RE(X;t)} = x_t$. Επίσης, $g(0) = -H'(X;t)$, $g(\infty) = \infty$, έτσι, η $g(x)$ αρχικά φθίνουσα και μετά αύξουσα ως προς $x$. Άρα παρουσιάζει τοπικό ελάχιστο (ολικό) στο $x_t = e^{-H(X;t)}$ που υποδηλώνει ότι η εξίσωση (4.5) έχει μοναδική θετική $h(t), \forall t$. Άρα, η $H(X;t)$ προσδιορίζει μονοσήμαντα την $h(t)$ και κατ' επέκταση την $F(t)$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Παρατήρηση:**

Αν η $H(X;t)$ φθίνουσα συνάρτηση του $t$ και $g(\,x_t) \neq 0$, τότε η (4.5) έχει δύο λύσεις για κάθε $t \geq 0$. Η μια είναι η είναι η συνάρτηση κινδύνου $h(t)$ και η άλλη θα μπορούσε να μην είναι κατάλληλη συνάρτηση κινδύνου. Στο παράδειγμα που ακολουθεί θα μελετηθεί η περίπτωση της κατανομής $Beta(\lambda,1)$, στο οποίο και οι δύο λύσεις που προκύπτουν αποτελούν συναρτήσεις κινδύνου.

**Παράδειγμα:**

Έστω $X$ συνεχής, μη αρνητική τυχαία μεταβλητή με συνάρτηση κατανομής $Beta(\lambda,1)$ και συνάρτηση πυκνότητας:

$$f(x) = \begin{cases} \lambda x^{\lambda-1}, & \lambda > 1, \quad 0 < x < 1 \\ 0, & \textit{διαφορετικά} \end{cases}$$

με συνάρτηση επιβίωσης της Χ, $\bar{F}(x) = P(X > t) = \int_x^1 f(u)\,du = 1 - x^{\lambda}$ και συνάρτηση κινδύνου:

$$h(x) = \frac{f(x)}{\bar{F}(x)} = \frac{\lambda x^{\lambda-1}}{1 - x^{\lambda}}.$$

Τότε, η $H(X;t)$ για την $Beta(\lambda,1)$ χρησιμοποιώντας την εξίσωση (4.5) είναι:

$$H(X;t) = \frac{\lambda - 1}{\lambda} + \log\left(\frac{1 - t^{\lambda}}{\lambda}\right) + \frac{(\lambda - 1)t^{\lambda}}{1 - t^{\lambda}} \log t \tag{4.8}$$

με ελάχιστο στο

$$x_t = e^{-H(X;t)} = \frac{\lambda}{1 - t^{\lambda}} t^{\frac{-(\lambda-1)t^{\lambda}}{1-t^{\lambda}}} exp\left\{-\frac{\lambda - 1}{\lambda}\right\}. \tag{4.9}$$

Έτσι,

$$\frac{x_t}{h(t)} = t^{\frac{-(\lambda-1)}{1-t^{\lambda}}} e^{\frac{1-\lambda}{\lambda}}. \tag{4.10}$$

Οπότε, $\forall t > 0$ η $g(x)$ έχει δύο θετικές λύσεις $h(t), h^*(t)$, με $h(t) < x_t < h^*(t)$, όπου $h(t)$ συνάρτηση κινδύνου της $Beta$ κατανομής. Ωστόσο και η λύση $h^*(t)$ αποτελεί συνάρτηση κινδύνου.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Πόρισμα 4.2.1**

Η ομοιόμορφη κατανομή στο $(a, b)$, με $a < b$ μπορεί να προσδιοριστεί από την φθίνουσα

$$H(X;t) = \log(b-t), \quad t < b. \tag{4.11}$$

**Απόδειξη:**

Για την περίπτωση της ομοιόμορφης στο $(a, b)$ με $a < b$ έχουμε:

$H(X;t) = \log(b-t), \quad t < b, t \geq 0$ η οποία είναι φθίνουσα ως προς $t$.

Επιπλέον, $H'(X;t) = -\frac{1}{b-t}$ και $x_t = e^{-RE(X;t)} = \frac{1}{b-t}$. Αντικαθιστώντας στην εξίσωση (4.5) προκύπτει ότι η $g(x) = 0$ έχει μοναδική λύση για $x = x_t$.

Όμως η $h(t) = \frac{1}{b-t} = x_t$ είναι η μοναδική λύση για την $g(x) = 0$.

Αυτό συνεπάγεται ότι η κατανομή είναι η ομοιόμορφη.

Οι Asadi και Ebrahimi (2000) προσδιόρισαν την κατανομή Pareto *(GPD)* μέσω της σχέσης $H(X;t) = c - \log h(t)$, όπου $c$:πραγματική σταθερά. Ο προσδιορισμός αυτός περιλαμβάνει την Εκθετική κατανομή για $c = 1$, την κατανομή Pareto για $c > 1$.

Επιπλέον οι Nair and Rajesh (1998) χαρακτήρισαν την Εκθετική κατανομή μέσω της σχέσης που δίνεται στο θεώρημα που ακολουθεί.

**Θεώρημα 4.2.1**

Έστω $X$ συνεχής, μη αρνητική τυχαία μεταβλητή με μέση τιμή $E(X) < \infty$.

Αν $H(f;t)$ η αθροιστική υπολειπόμενη εντροπία της $X$ και $m(t)$ ο μέσος υπολειπόμενος χρόνος ζωής, τότε:

$$H(f;t) - m(t) = H(f;0) + m(0). \tag{4.12}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

## 4.3. Η Υπολειπόμενη Εντροπία και κλάσεις κατανομών

Όπως ήδη αναφέραμε, η μονοτονία του μέτρου της συνάντησης κινδύνου, καθώς και του μέσου υπολειπόμενου χρόνου ζωής ορίζουν τις κλάσεις κατανομών $IFR$ $(DFR)$, $DMRL(IMRL)$ αντίστοιχα.
Ο Ebrahimi (1996) πρότεινε τις ακόλουθες κλάσεις κατανομών βάσει της μονοτονίας της $H(X;t)$.

**Ορισμός 4.3.1**

Μια τυχαία μεταβλητή $X$ λέγεται ότι είναι:

I. Φθίνουσα ως προς την υπολειπόμενη αβεβαιότητα (*DURL)* αν $H(X;t)$ φθίνουσα για $t > 0$.
II. Αύξουσα ως προς την υπολειπόμενη αβεβαιότητα (*IURL)* αν $H(X;t)$ αύξουσα για $t > 0$.

Η Εκθετική είναι η μόνη συνεχής κατανομή που ανήκει και στις δύο κλάσεις.

Ο Ebrahimi (1996) απέδειξε πως όταν μια τυχαία μεταβλητή ανήκει στην κλάση $IFR\,(DFR)$ τότε ανήκει και στην $DURL\,(IURL)$. Αργότερα, οι Ebrahimi and Kirmani (1996) επέκτειναν τα αποτελέσματα σε κλάσεις με σημείο αναφοράς τη μονοτονία του μέσου υπολειπόμενου χρόνου με τη Χ∈ $DMRL$ $(IMRL)$.

Ισχύει η διάταξη:
$$IFR\ (DFR) \Rightarrow DMRL\ (IMRL) \Rightarrow DURL\ (IURL). \qquad (4.13)$$

## 4.4. Υπολειπόμενη εντροπία και η επίδραση αφαιρετέου ποσού σε Μοντέλα Απώλειας

Με τον όρο αφαιρετέο ποσό (*deductible)* εννοούμε το μέρος της απώλειας ή απαίτησης που καλείται να καλύψει ο δικαιούχος και εκπίπτει από τις υποχρεώσεις του ασφαλιστή.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Έστω ότι οι απώλειες δεν καταγράφονται ή δεν αναφέρονται όταν είναι μικρότερες από ένα προκαθορισμένο ποσό. Σε αυτή την περίπτωση τα δεδομένα αναφέρονται ως αποκομμένα από κάτω *(truncated from below)* ή από αριστερά αποκομμένα *(left truncated)*. Ο πιο συνηθισμένος λόγος για τον οποίο συμβαίνει αυτό είναι η χρήση ενός αφαιρετέου ποσού ή απαλλαγής.

Στην περίπτωση των περικομμένων δεδομένων η ασφαλιστική δεν λαμβάνει γνώση τη ζημιά. Στη δεύτερη περίπτωση γνωρίζει το συμβάν, ωστόσο δε γνωρίζει το ακριβές ύψος της απώλειας.

**Ορισμός 4.4.1**

Έστω $X$ απόλυτα συνεχής, μη αρνητική τυχαία μεταβλητή η οποία εκφράζει τις απώλειες και έστω $d$ η τιμή του αφαιρετέου ποσού.

Τότε,σύμφωνα με τον ορισμό των Kluggman et al. (2008), η τυχαία μεταβλητή που εκφράζει τις απώλειες σε ένα ασφαλιστήριο συμβόλαιο το οποίο έχει εκδοθεί με χρήση αφαιρετέου ποσού δίνεται ως:

I. Κόστος ανά πληρωμή (per-payment):

$$Y_{p(d)} = \begin{cases} X, & X > d \\ \delta\varepsilon\nu\ o\rho\acute{\iota}\zeta\varepsilon\tau\alpha\iota, & X \leq d \end{cases}, \text{ ισοδύναμα, } Y_{p(d)} = X|X > d. \tag{4.14}$$

II. Κόστος ανά ζημιά (per-loss) ή (franchise deductible):

$$Y_{l(d)} = \begin{cases} X, & X > d \\ 0\,, & X \leq d. \end{cases} \tag{4.15}$$

Στην πρώτη περίπτωση έχουμε περικοπή από κάτω ενώ στη δεύτερη λογοκρισία από κάτω. Η διαφορά μεταξύ των δύο έγκειται στο ότι στη δεύτερη περίπτωση δίνεται μάζα πιθανότητας στο μηδέν (για $X \leq d$), κάτι που την καθιστά μικτού τύπου (διακριτή στο μηδέν και συνεχή αλλού).

Στην πραγματικότητα, η μεταβλητή του κόστους ανά πληρωμή είναι ισοδύναμη με αυτή του κόστους ανά ζημιά, δοθέντος ότι η τελευταία είναι θετική. Η μεταβλητή κόστους ανά ζημιά ωστόσο εκφράζει τις πραγματικές απώλειες του ασφαλιστή (Sachlas and Papaioannou, 2012).

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Στα λήμματα που ακολουθούν δίνονται οι σχέσεις μεταξύ της εντροπίας της τυχαίας μεταβλητής $X$ και των τυχαίων μεταβλητών που προκύπτουν ύστερα από περικοπή ή λογοκρισία από κάτω.

**Λήμμα 4.4.1**

Η εντροπία της μεταβλητής κόστους ανά πληρωμή με προκαθορισμένο αφαιρετέο ποσό $d$ ορίζεται ως $H(f,d)$ και δίνεται από τη σχέση:

$$H(f,d) = -\int_d^{\infty} \frac{f_X(x)}{\bar{F}_X(d)} \ln \frac{f_X(x)}{\bar{F}_X(d)} dx$$

$$= \frac{1}{\bar{F}_X(d)}\left[H(X) + \int_0^d f_X(x) \ln f_X(x)\, dx\right] + \ln \bar{F}_X(d)\,, \qquad (4.16)$$

όπου $H(X)$ είναι η εντροπία της τυχαίας μεταβλητής $X$, $\bar{F}_X(x)$= 1-$F_X(x)$ η συνάρτηση δεξιάς ουράς της $X$ και $d$ το αφαιρετέο ποσό.

**Απόδειξη:**

Είναι γνωστό ότι η συνάρτηση πυκνότητας πιθανότητας της $Y_{p(d)}$ δίνεται από τη σχέση:

$$f_{Yp(d)}(x) = \begin{cases} \dfrac{f_X(x)}{\bar{F}_X(d)}, & X > d, \\ \quad 0\,, & X \le d. \end{cases} \qquad . \qquad (4.17)$$

Συνεπώς η εντροπία της $Y_{p(d)}$ δίνεται από τη σχέση:

$$H(f,d) = -\frac{1}{\bar{F}_X(d)}\left[\int_0^d f_X(x) \ln f_X(x)\, dx - \int_d^{\infty} f_X(x) \ln \bar{F}_X(x)\, dx\right]$$

$$= \frac{1}{\bar{F}_X(d)}\left[H(X) + \int_0^d f_X(x) \ln f_X(x)\, dx + \bar{F}_X(d) \ln \bar{F}_X(d)\right]$$

$$= \frac{1}{\bar{F}_X(d)}\left[H(X) + \int_0^d f_X(x) \ln f_X(x)\, dx\right] + \ln \bar{F}_X(d),$$

όπου $\bar{F}_X(d) = 1 - F_X(d)$ η συνάρτηση επιβίωσης στο σημείο $d$.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Λήμμα 4.4.2**

Η εντροπία της μεταβλητής κόστους ανά ζημιά με προκαθορισμένο αφαιρετέο ποσό $d$, ορίζεται ως $H_l(f,d)$ και δίνεται από τη σχέση:

$$H_l(f,d) = -\int_d^{\infty} f_X(x)\ln f_X(x)\,dx \;-\; F_X(d)\ln F_X(d)$$

$$= H(X) + \int_0^{d} f_X(x)\ln f_X(x)\,dx \;-\; F_X(d)\ln F_X(d). \qquad (4.18)$$

**Απόδειξη:**

Η εντροπία της μεταβλητής κόστους ανά ζημιά $Y_{l(d)}$ προκύπτει προσθέτοντας το διακριτό τμήμα (μάζα πιθανότητας στο μηδέν). Σημειώνουμε ότι $P\big(\,Y_{l(d)} = 0\big) = F_X(d)$.

**Παρατηρήσεις :**

i. Οι $H_l(f,d)$, $H(f,d)$ μπορούν να πάρουν αρνητικές τιμές όπως επίσης και $\pm\infty$.

ii. Οι $H_l(f,d)$, $H(f,d)$ δεν είναι πάντα μεγαλύτερες από την εντροπία της $X$.

**Εντροπία κόστους ανά πληρωμή με αφαιρετέο ποσό**

**Παραδείγματα:**

**1. Εκθετική κατανομή**

Η εντροπία για την μεταβλητή που εκφράζει το κόστος ανά πληρωμή υπό την επίδραση αφαιρετέου ποσού $d$ για την εκθετική κατανομή με μέση τιμή $1/\lambda$ είναι:

$$H(f,d) = 1 - \ln\lambda. \qquad (4.19)$$

**Απόδειξη:**

Για την εκθετική κατανομή με συνάρτηση πυκνότητας πιθανότητας $f_X(x) = \lambda e^{-\lambda x}$ έχουμε:



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$\int_0^d f_X(x)\ln f_X(x)\,dx = \ln\lambda F_X(d) - \lambda\int_0^d x f_X(x)\,dx$$

με

$$\int_0^d x f_X(x) = \int_0^d x\lambda e^{-\lambda x}\,dx = \frac{1-e^{-\lambda d}(1+\lambda d)}{\lambda}.$$

Επειδή $F_X(d) = 1 - e^{-\lambda d}$, έχουμε:

$$\int_0^d f_X(x)\ln f_X(x)\,dx = \ln\lambda\left(1-e^{-\lambda d}\right) - 1 + e^{-\lambda d}(1+\lambda d)$$

$$= \ln\lambda - 1 + e^{-\lambda d}(1+\lambda d - \ln\lambda).$$

Από τη σχέση $H(f,d) = \frac{1}{\bar{F}_X(d)}\left[H(X) + \int_0^d f_X(x)\ln f_X(x)\,dx + \bar{F}_X(d)\ln\bar{F}_X(d)\right]$ έχουμε:

$$\frac{1}{e^{-\lambda d}}\left[1 - \ln\lambda + \ln\lambda - 1 + e^{-\lambda d}(1+\lambda d - \ln\lambda) + \lambda d(1 - 1 + e^{-\lambda d}\right]$$

$$= \frac{1}{e^{-\lambda d}}e^{-\lambda d}(1-\ln\lambda) = 1 - \ln\lambda,$$

ίση με την εντροπία του Shannon για την εκθετική κατανομή.

Παρατηρούμε πως το αφαιρετέο ποσό δεν επηρεάζει την εντροπία κόστους ανά πληρωμή στην εκθετική κατανομή.

## 2. Κατανομή Weibull

Η $H(f,d)$ για την Weibull με αφαιρετέο ποσό $d$ είναι:

$$H(f,d) = 1 - \ln c - \ln\tau - (\tau-1)\ln d - (\tau-1)e^{cd^\tau}\frac{\Gamma(0,cd^\tau)}{\tau}. \qquad (4.20)$$

*i.* Για $\tau > 1$, προκύπτει $H(f,d) < H_X, \quad \forall c,d$

με $H(X) = 1 - \ln c - \ln\tau + \frac{\tau-1}{\tau}(\gamma + \ln c)$.

*ii.* Για $d > 0$,

a) $H(f,d)$ αύξουσα ως προς $d$ όταν $\tau < 1$.

b) $H(f,d)$ φθίνουσα ως προς $d$ όταν $\tau > 1$.

## 3. Κατανομή Pareto

Για την κατανομή Pareto με συνάρτηση πυκνότητας πιθανότητας $f_X(x) = \frac{a\lambda^a}{x^{a+1}}$, x≥ 0, έχουμε:

$$\int_0^d f_X(x)\ln f_X(x)\,dx = (\ln a + a\ln\lambda)\,F_X(d) - (\alpha+1)\int_0^d \ln(\lambda+x)\,f_X(x)dx,$$

όπου,

$$\int_0^d \ln(\lambda+x)\,f_X(x)dx = \int_0^d \ln(\lambda+x)\frac{a\lambda^a}{(\lambda+x)^{a+1}}dx$$

$$= \alpha\lambda^a\int_0^d \ln y y^{-(a+1)}dy$$

$$= \frac{1}{a}[1+\alpha\ln\lambda - \lambda^a(\lambda+d)^{-a}(1+a\ln(\lambda+d)].$$

Όμως, $F_X(d) = 1-\left(\frac{\lambda}{(\lambda+d)}\right)^\alpha$ επομένως,

$$\int_0^d f_X(x)\ln f_X(x)\,dx = \ln\alpha + \alpha\ln\lambda\left[1-\left(\frac{\lambda}{(\lambda+d)}\right)^\alpha\right]$$

$$-\frac{\alpha+1}{\alpha}\left[1+\alpha\ln\lambda - \frac{\lambda^a(1+a\ln(\lambda+d)}{(\lambda+d)^a}\right].$$

Όμως, από το Λήμμα 4.4.1 έχουμε:

$$H(f,d) = \frac{1}{\bar{F}_X(d)}\left[H(X) + \int_0^d f_X(x)\ln f_X(x)\,dx + \bar{F}_X(d)\ln\bar{F}_X(d)\right]$$

$$= \frac{1}{\left(\frac{\lambda}{(\lambda+d)}\right)^\alpha}\left\{1+\frac{1}{\alpha}+\ln\lambda-\ln\alpha+(\ln\alpha+\alpha\ln\lambda)\left(1-\left(\frac{\lambda}{\lambda+d}\right)^\alpha\right)\right.$$

$$-\frac{\alpha+1}{\alpha}\left[1+\alpha\ln\lambda-\frac{\lambda^a(1+a\ln(\lambda+d))}{(\lambda+d)^a}\right]$$

$$\left.+\left(\frac{\lambda}{(\lambda+d)}\right)^\alpha\ln\left(\frac{\lambda}{(\lambda+d)}\right)^\alpha\right\},$$

$$H(f,d) = \left(\frac{\lambda}{\lambda+d}\right)^{-a}\left(2+\frac{1}{a}+\ln\lambda\right)+1+\frac{1}{a}+\ln(\lambda+d)-\ln a \qquad (4.21)$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

με $H(f,d) > H(X)$ ,$\forall d, a$ και $\forall \lambda > 1$, όπου $H(X) = \left(1 + \frac{1}{a} + \ln\lambda - \ln a\right)$.

Επίσης , $H(f,d)$ αύξουσα συνάρτηση ως προς $d, \forall d$.

**Εντροπία κόστους ανά ζημιά με αφαιρετέο ποσό**

**Παραδείγματα**:

1. **Εκθετική κατανομή**

Η εντροπία κόστους ανά ζημιά υπό την υπό την επίδραση αφαιρετέου ποσού $d$ είναι:

$$H_l(f,d) = e^{-\lambda d}(1 + \lambda d - \ln\lambda) - \left(1 - e^{-\lambda d}\right)\ln\left(1 - e^{-\lambda d}\right). \tag{4.22}$$

Όπως παρατηρούμε σε αυτή την περίπτωση η $H_l(f,d)$ δεν είναι ανεξάρτητη του $d$.

2. **Κατανομή Weibull**

$$H_l(f,d) = e^{-cd^\tau}(1 - \ln c - \ln\tau - (\tau - 1)\ln d + cd^\tau) - (\tau - 1)\frac{\Gamma(0, cd^\tau)}{\tau}$$
$$-\left(1 - e^{-cd^\tau}\right)\ln\left(1 - e^{-cd^\tau}\right). \tag{4.23}$$

3. **Κατανομή Pareto**

$$H_l(f,d) = \left(\frac{\lambda}{\lambda + d}\right)^a \left(1 + \frac{1}{a} - \ln a - a\ln\lambda\right) + (a+1)\ln(\lambda + d)$$
$$-\left[1 - \left(\frac{\lambda}{\lambda + d}\right)^a\right]\ln\left[1 - \left(\frac{\lambda}{\lambda + d}\right)^a\right]. \tag{4.24}$$

Οι εντροπίες κόστους ανά πληρωμή και κόστους ανά ζημιά με αφαιρετέο ποσό είναι ιδιαιτέρως περίπλοκες ως προς τον υπολογισμό τους για κατανομές όπως η Gamma, Transformed Gamma, Lognormal κλπ. καθώς απαιτείται χρήση ανώτερων συναρτήσεων όπως η συνάρτηση Meijer, η συνάρτηση σφάλματος *(error function)* $\mathrm{Erf}(z)$ κλπ. (Abramowitz and Stegun, 1972).

Για τις κατανομές Burr και Generalized Pareto η εντροπία του Shannon δεν υπολογίζεται σε κλειστή μορφή, επομένως απαιτείται χρήση αριθμητικής μεθόδου για τον υπολογισμό τους.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Στον πίνακα που ακολουθεί δίνονται συνοπτικά οι εντροπίες κόστους ανά πληρωμή $H(f,d)$ για τις διάφορες κατανομές που αναφέρθηκαν παραπάνω.

| **ΚΑΤΑΝΟΜΗ** | **ΕΝΤΡΟΠΙΑ** $\boldsymbol{H(f,d)}$ |
|---|---|
| **Εκθετική** | $1-ln\,\lambda$ |
| **Pareto** | $1+\frac{1}{a}-\ln a+(1+2a)\ln(\lambda+d)$ |
| **Weibull** | $1+\ln c+\ln\tau-(\tau-1)\ln d-\frac{(\tau-1)e^{cd^{\tau}}\Gamma(0,cd^{\tau})}{\tau}$ |
| **Gamma** | $\frac{1}{\varphi(\alpha,\lambda d)}\Big\{\Gamma(\alpha+1)-\Gamma(\alpha+1,\lambda d)+(\alpha-1)\Gamma(\alpha,\lambda d)\ln d+(\alpha-1)G^{3,0}_{2,0}\left(\lambda d\,\middle\vert\,\begin{matrix}\boldsymbol{\alpha}\\ \boldsymbol{b}\end{matrix}\right)-\Gamma(\alpha)\left[\alpha+(\alpha-2)\ln\lambda+\ln\xi(\alpha,\lambda d)+\Gamma(\alpha,\lambda d)\ln\frac{\lambda^{\alpha}}{\xi(\alpha,\lambda d)}\right]\Big\}$ |
| **Truncated Gamma** | $\frac{1}{\varphi(\alpha,\lambda d^{\tau})}\Big\{\frac{\tau\alpha-1}{\tau}G^{3,0}_{2,0}\left(\lambda d^{\tau}\,\middle\vert\,\begin{matrix}\boldsymbol{\alpha}\\ \boldsymbol{b}\end{matrix}\right)+\Gamma(\alpha+1)+\Gamma(\alpha+1,\lambda d^{\tau})+(\tau\alpha-1)\Gamma(\alpha,\lambda d^{\tau})\ln d-\Gamma(\alpha)\left[\alpha+\ln\frac{\lambda^{\tau\alpha}\tau}{\xi(\alpha,\lambda d^{\tau})}+\Gamma(\alpha,\lambda d^{\tau})\ln\frac{\lambda^{\tau\alpha}\tau}{\xi(\alpha,\lambda d^{\tau})}\right]\Big\}$ |
| **Loggamma** | $\frac{1}{\lambda\varphi(\alpha,\lambda\ln d)}\Big\{\lambda(\alpha-1)G^{3,0}_{2,0}\left(\lambda\ln d\,\middle\vert\,\begin{matrix}\boldsymbol{\alpha}\\ \boldsymbol{b}\end{matrix}\right)+(\alpha-1)\lambda\Gamma(\alpha,\lambda\ln d)\ln(\ln d)-(\lambda+1)\Gamma(\alpha+1,\lambda\ln d)+\lambda\big(1-\Gamma(\alpha,\lambda\ln d)\big)[\Gamma(\alpha+1)\ln\lambda-\Gamma(\alpha)\ln\xi(\alpha,\lambda\ln d)]\Big\}$ |
| **Lognormal** | $\frac{1}{1-\Phi\left(\frac{\ln d-\mu}{\sigma}\right)}\Big\{\frac{1}{2}+\mu+\ln(\sigma\sqrt{2}\pi)-\left(\ln(\sigma\sqrt{2}\pi)+\frac{\mu^2}{\sigma^2}\right)\Phi\left(\left(\frac{\ln d-\mu}{\sigma}\right)\right)+\left(1-\Phi\left(\frac{\ln d-\mu}{\sigma}\right)\right)\ln\left(1-\Phi\left(\frac{\ln d-\mu}{\sigma}\right)\right)-\frac{1}{2\sigma^2}\left[Erfc\left(\frac{\mu-\ln d}{\sigma\sqrt{2}}\right)-\frac{\sigma(\mu+\ln d)}{\sqrt{2}\pi}e^{-\frac{(\mu-\ln d)^2}{2\sigma^2}}\right]+\left(\frac{\mu}{\sigma^2}-1\right)\left[-\frac{\sigma}{\sqrt{2}\pi}e^{-\frac{(\mu-\ln d)^2}{2\sigma^2}}+\frac{\mu}{2}-\frac{\mu}{2}Erf\left(\frac{\mu-\ln d}{\sigma\sqrt{2}}\right)\right]\Big\}$ |
| **Burr** | Όχι σε κλειστή μορφή |
| **Generalized Pareto** | Όχι σε κλειστή μορφή |

(Σαχλάς, 2010)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$\xi(\alpha,\chi) = \Gamma(\alpha)\big(1 - \Gamma(\alpha,\chi)\big)$ $\boldsymbol{b} = (0,0,a)$

$\varphi(\alpha,\chi) = \Gamma(\alpha)(\Gamma(\alpha,\chi) - 1)$ $\boldsymbol{\alpha} = (1,1)$

**Σχέση εντροπίας κόστους ανά ζημιά και κόστους ανά πληρωμή**

$$H_l(f,d) = \bar{F}_X(d)H(f,d) + H[F_X(d)\bar{F}_X(d)]. \tag{4.25}$$

**Πρόταση 4.4.1**

Αν $\int_d^{\infty} f_X(x)\ln f_X(x)\,dx > 0$, τότε $H(f,d) < H_l(f,d)$, $\forall d$.

**Απόδειξη:**

$$H_l(f,d) - H(f,d) = H(X) + \int_0^d f_X(x)\ln f_X(x)\,dx - F_X(d)\ln F_X(d) - \frac{1}{\bar{F}_X(d)}H_X$$

$$= -\frac{1}{\bar{F}_X(d)}\int_0^d f_X(x)\ln f_X(x)\,dx - \ln \bar{F}_X(d)$$

$$= -\frac{F_X(d)}{\bar{F}_X(d)}\left[H_X + \int_0^d f_X(x)\ln f_X(x)\,dx\right]$$

$$-F_X(d)\ln F_X(d) - \ln \bar{F}_X(d)$$

$$> -\frac{F_X(d)}{\bar{F}_X(d)}\left[H_X + \int_0^d f_X(x)\ln f_X(x)\,dx\right]$$

$$= \frac{F_X(d)}{\bar{F}_X(d)}\left[\int_d^{\infty} f_X(x)\ln f_X(x)\,dx\right].$$

**Ορισμός 4.4.2**

Οι υπολειπόμενες απώλειες μετά από *d (residual losses after d)* όπως περιγράφονται από την τυχαία μεταβλητή *X-d* (Kluggman et al., 2008) δίνονται από τις παρακάτω σχέσεις:

i. Ανά πληρωμή υπολειπόμενη ζημιά

$$Z_p(d) = \begin{cases} X - d, & X > d \\ \delta\varepsilon\nu\ o\rho\acute{\iota}\zeta\varepsilon\tau\alpha\iota, & X \le d \end{cases},$$

ή ισοδύναμα $Z_p(d) = X - d | X > d$. (4.26)

ii. Ανά κόστος υπολειπόμενη ζημιά

$$Z_l(d) = \begin{cases} X - d, & X > d \\ 0, & X \le d \end{cases},$$

ή ισοδύναμα $Z_l(d) = \max\{X - d, 0\}$. (4.27)

**Παρατηρήσεις:**

Η τυχαία μεταβλητή $X - d$ είναι η *retention loss* (Cox, 1991).

Η πρώτη περίπτωση, κατ' αντιστοιχία με τον αρχικό ορισμό, περιγράφει δεδομένα περικομμένα από κάτω, ενώ η δεύτερη δεδομένα λογοκριμένα από κάτω.

Στην πρώτη περίπτωση ο ασφαλιστής για ζημιές που ξεπερνούν το αφαιρετέο ποσό $d$ καταβάλλει ως αποζημίωση τη διαφορά του ποσού της ζημιάς μείον το αφαιρετέο ποσό, ενώ στη δεύτερη ο ασφαλιστής δεν καταβάλλει αποζημίωση για ζημιές μικρότερες του αφαιρετέου ποσού.

**Πρόταση 4.4.2**

Η $H(f, d)$ είναι ανεξάρτητη του $d$ αν και μόνο αν η συνάρτηση κινδύνου $h_X(x)$ είναι σταθερή.

**Απόδειξη:**

Έστω $h_X(x) = h$ σταθερή. Τότε αντικαθιστώντας στην παραπάνω εξίσωση παίρνουμε $H(f, d) = 1 - \ln h$ ανεξάρτητο του $d$.

Η έκφραση $H(f, d) = 1 - \ln h$ αποτελεί την εντροπία της εκθετικής κατανομής με παράμετρο $h$.

Αντίστροφα,έστω $H(f, d)$ ανεξάρτητη του $d$. Τότε παραγωγίζοντας τη σχέση:

$H(f, d) = 1 - \frac{1}{\bar{F}_X(d)} \int_d^{\infty} \ln\big(h_X(x)\big)\, f_X(x)\, dx$ ως προς $d$ έχουμε:



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$\frac{\partial}{\partial d} H(f,d) = h_X(d)[\ln h_X(x) + H(f,d) - 1].$$

$\frac{\partial}{\partial d} H(f,d) = 0$, ισοδύναμα, $h_X(d)[\ln h_X(x) + H(f,d) - 1] = 0.$

Επομένως $h_X(d) = e^{1-H(f,d)}$ η οποία είναι ανεξάρτητη του $d$.

Η παραπάνω πρόταση μας λέει πως αν η συνάρτηση κινδύνου είναι σταθερή η αβεβαιότητα που προκύπτει υπό την επίδραση αφαιρετέου ποσού δεν εξαρτάται από το αφαιρετέο ποσό. Αυτή είναι η περίπτωση της εκθετικής κατανομής.

Από το παραπάνω αποτέλεσμα συμπεραίνουμε ότι η αποκοπή από κάτω δεν επηρεάζει εντέλει την εντροπία της εκθετικής κατανομής, εφόσον είναι ανεξάρτητη από το αφαιρετέο ποσό $d$ (Sachlas and Papaioannou, 2009).

Στην ανάλυση επιβίωσης και στη θεωρία αξιοπιστίας, όπως αναφέραμε, η $H(f,d)$ χρησιμοποιείται ως μέτρο αβεβαιότητας της κατανομής του υπολειπόμενου χρόνου ζωής, δοθέντος ότι ένα σύστημα ή άτομο έχει επιβιώσει μέχρι την «ηλικία» $t$. Σε αυτή την ενότητα θα αναφερόμαστε σε αυτή ως εντροπία υπολειπόμενης ζημιάς. Η παράμετρος της ηλικίας έχει αντικατασταθεί από το αφαιρετέο ποσό.

Στη συνέχεια παρουσιάζονται κάποιες βασικές ιδιότητες της υπολειπόμενης εντροπίας υπό την επίδραση αφαιρετέου ποσού, κατά αντιστοιχία με τις ιδιότητες της υπολειπόμενης διάρκειας ζωής, σε σχέση με τη μέση υπολειπόμενη ζημιά (*Mean Residual Loss*) .

**Ορισμός 4.4.3**

Κατά αντιστοιχία με τον ορισμό του μέσου υπολειπόμενου χρόνου ζωής, η μέση υπολειπόμενη ζημιά δίνεται ως:

$$m(d) = E(X - d|X > d) = \begin{cases} \frac{\int_d^\infty \bar{F}_X(x)dx}{\bar{F}_X(d)}, & d < d^* \\ 0, & \text{διαφορετικά} \end{cases} \quad (4.28)$$

με $d^* = sup\{x: \bar{F}_X(x) > 0\}$.

Για τα μοντέλα που μελετάμε ισχύει ότι $d^* = \infty$.

Επιπρόσθετα σημειώνουμε ότι η $m(d)$ μπορεί να γραφεί ως:

$$m(d) = \int_d^{\infty} e^{-\int_d^y h_X(z)\,dz}\, dy = E\big(Y_{p(d)} - d\big) \tag{4.29}$$

με $h_X(x)$: συνάρτηση κινδύνου.

**Ιδιότητες**

I. Η υπολειπόμενη εντροπία των απωλειών, υπό την επίδραση αφαιρετέου ποσού $d, H(f,d)$, είναι φραγμένη από την μέση υπολειπόμενη απώλεια και ισχύει ότι

$$H(f,d) \leq 1 + \ln \mathrm{m}(d) \tag{4.30}$$

υπό την προϋπόθεση ότι η $\mathrm{m}(d) < \infty$.

Η παραπάνω ιδιότητα, όπως διατυπώθηκε από τον Ebrahimi (1996), αποδεικνύει ότι το φράγμα προκύπτει από την αρχή της μέγιστης εντροπίας για κατανομές στο (0,∞) με δεδομένη μέση τιμή (επίσης Asadi et. al., 2004).

II. Η υπολειπόμενη εντροπία των απωλειών ανά πληρωμή υπό την επίδραση αφαιρετέου ποσού $d$ , $H(f,d)$,
   i. είναι αύξουσα συνάρτηση ως προς $d$ όταν η συνάρτηση κινδύνου $h_X(x)$ φθίνουσα ως προς $x$.
   ii. είναι φθίνουσα συνάρτηση ως προς d όταν $h_X(x)$ αύξουσα ως προς $x$. (Ebrahimi, 1996).

III. Αν η μέση υπολειπόμενη απώλεια (*Mean residual loss*) είναι φθίνουσα ως προς $d$ τότε η $H(f,d)$ επίσης φθίνουσα ως προς $d$ (Ebrahimi et. al., 2007).

IV. Έστω δύο ασφαλιστήρια συμβόλαια με συναρτήσεις κινδύνου $h_1(x)$, $h_2(x)$, τέτοιες ώστε $h_2(x) = k(x)h_1(x)$. Έστω επίσης $h(x)$ αύξουσα συνάρτηση του $x$ και $0 < k(x) < 1$.Τότε αν το ασφαλιστήριο συμβόλαιο 1 έχει φθίνουσα μέση υπολειπόμενη απώλεια συνεπάγεται ότι και το συμβόλαιο 2 έχει φθίνουσα μέση υπολειπόμενη απώλεια και συνεπώς φθίνουσα $H(f,d)$ (Theorem 2.5 in Block et al., 1985).

V. Αν $X$ απόλυτα συνεχής τυχαία μεταβλητή και $H(f,d)$ αύξουσα ως προς $d$ τότε η $H(f,d)$ προσδιορίζει μονοσήμαντα τη συνάρτηση κατανομής της $X, F_X(x)$ (Belunze et.al., 2004).



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Στον πίνακα που ακολουθεί στην επόμενη σελίδα, δίνεται το άνω φράγμα για κάθε κατανομή απώλειας.

Το φράγμα για την κατανομή Gamma δίνεται συναρτήσει της Exponential Integral function $E_n(z) = \int_1^{\infty} \frac{e^{-zt}}{t^n} dt$.

Για την κατανομή Burr δε μπορεί να δοθεί αλγεβρικός τύπος υπολογισμού, ωστόσο δίνεται ένα κάτω όριο αυτής στο οποίο εμφανίζεται η Υπεργεωμετρική συνάρτηση *(Hypergeometric function).*

$F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}$ με $(x)_n = x(x+1) \dots (x+n-1)$ (Abramowitz and Stegun, 1972).

Για τις κατανομές Transformed Gamma, Loggamma, Lognormal και Generalized Pareto το άνω φράγμα δε μπορεί να υπολογιστεί σε κλειστή μορφή.

**Άνω φράγμα της εντροπίας υπό την επίδραση της από κάτω αποκοπή**

| **ΚΑΤΑΝΟΜΗ** | **ΦΡΑΓΜΑ** |
|---|---|
| **Εκθετική** | $1 - ln\, \lambda$ |
| **Pareto** | $1 - (1-\alpha)\ln(\lambda + d) - \ln(1-\alpha)$ |
| **Weibull** | $1 + cd^{\tau} + \ln\left[\frac{c^{-\frac{1}{\tau}}}{\tau} \Gamma\left(\frac{1}{\tau}, cd^{\tau}\right)\right]$ |
| **Gamma** | $1 + \ln\left(\frac{d(\lambda d)^{\alpha}}{\alpha}\right) + \ln\left[e^{-\lambda d} + (\alpha - \lambda d)E_{-a}(\lambda d)\right]$ |
| **Truncated Gamma** | Όχι σε κλειστή μορφή |
| **Loggamma** | Όχι σε κλειστή μορφή |
| **Lognormal** | Όχι σε κλειστή μορφή |
| **Burr** | $1 + \ln\left[\frac{d^{-(\tau\alpha - 1)}}{\tau\alpha - 1} F_1\left(a, \alpha - \frac{1}{\tau}; 1 + \alpha - \frac{1}{\tau}, -\lambda d^{-\tau}\right)\right]$ |
| **Generalized Pareto** | Όχι σε κλειστή μορφή |

(Σαχλάς, 2010)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Πρόταση 4.4.3**

Για την $H_l(f,d)$ ισχύει

$$H_l(f,d) \le \bar{F}_X(d)[1 + \ln \mathrm{m}(d)] + H[F_X(d), \bar{F}_X(d)]. \tag{4.31}$$

**Απόδειξη:**

Προκύπτει άμεσα από τις σχέσεις:

$$H_l(f,d) = \bar{F}_X(d) H(f,d) + H[F_X(d), \overline{F}_X(d)]$$

και

$$H(f,d) \le 1 + \ln \mathrm{m}(d).$$

## 4.5. Παράδειγμα: Εφαρμογή αφαιρετέου ποσού σε δείγμα ζημιών και υπολογισμός εντροπίας

Έστω οι ζημιές που κατέγραψε μια ασφαλιστική εταιρεία, για το έτος 2018, όπως παρουσιάζονται στον παρακάτω πίνακα, υπολογισμένες σε χιλιάδες Ευρώ:

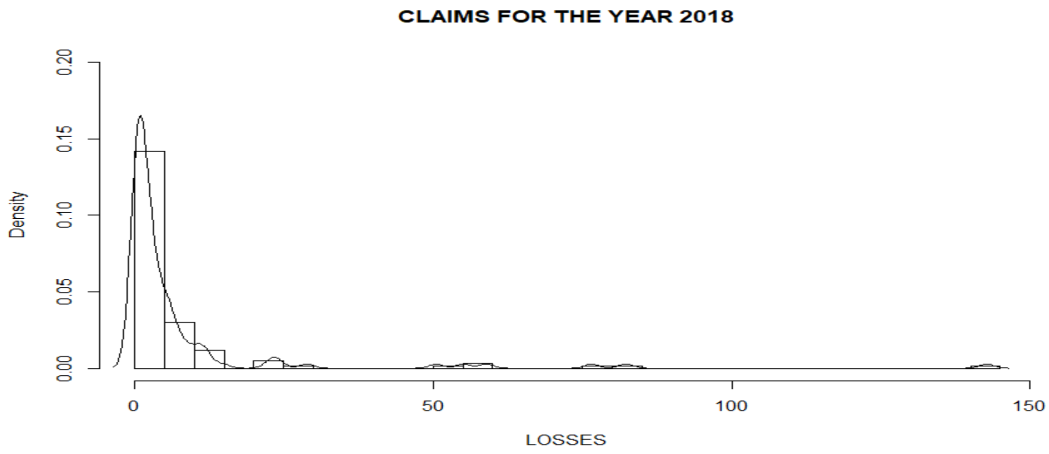
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.032 | 0.045 | 0.053 | 0.082 | 0.093 | 0.094 | 0.098 | 0.127 |
| 0.143 | 0.147 | 0.160 | 0.172 | 0.219 | 0.232 | 0.255 | 0.269 |
| 0.276 | 0.310 | 0.366 | 0.401 | 0.440 | 0.446 | 0.478 | 0.511 |
| 0.535 | 0.535 | 0.549 | 0.621 | 0.677 | 0.684 | 0.690 | 0.693 |
| 0.698 | 0.769 | 0.782 | 0.804 | 0.808 | 0.925 | 0.971 | 0.996 |
| 1.039 | 1.045 | 1.100 | 1.253 | 1.263 | 1.314 | 1.335 | 1.349 |
| 1.382 | 1.389 | 1.396 | 1.446 | 1.504 | 1.556 | 1.563 | 1.651 |
| 1.653 | 1.655 | 1.830 | 1.923 | 1.979 | 2.033 | 2.072 | 2.275 |
| 2.399 | 2.542 | 2.575 | 2.707 | 2.745 | 2.820 | 2.998 | 3.038 |
| 3.108 | 3.124 | 3.210 | 3.330 | 3.405 | 3.413 | 3.416 | 3.509 |
| 3.862 | 4.301 | 4.436 | 4.463 | 4.976 | 5.137 | 5.177 | 5.232 |
| 5.267 | 5.333 | 5.416 | 5.814 | 6.060 | 6.066 | 6.332 | 6.587 |
| 6.593 | 6.653 | 7.474 | 8.135 | 8.343 | 8.912 | 9.206 | 10.002 |
| 10.866 | 11.218 | 11.425 | 11.923 | 11.974 | 14.759 | 22.934 | 23.157 |
| 24.386 | 28.933 | 50.575 | 55.118 | 58.954 | 76.646 | 82.362 | 142.767 |

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

| Minimum | 1st Qu. | Median | Mean | 3rd Qu. | Maximum |
|---|---|---|---|---|---|
| 0.0320 | 0.6885 | 1.951 | 7.4190 | 5.3540 | 142.800 |

Όπως παρατηρούμε στον παραπάνω πίνακα, το μέσο κόστος υπολογίστηκε **7.4190** χιλ. €, ενώ το 25% των ζημιών είναι μικρότερες από **0.6885** χιλ. € και το 75% κάτω από **5.3540** χιλ. €.

Το ιστόγραμμα και η πυκνότητα του δείγματος παρουσιάζονται στο παρακάτω σχήμα που ακολουθεί στην επόμενη σελίδα.




*Σχήμα 2: Γραφική απεικόνιση ιστογράμματος και πυκνότητας ζημιών*

Στα δεδομένα προσαρμόστηκε κατανομή Weibull με εκτιμητές μεγίστης πιθανοφάνειας $\widehat{shape}$: 0.6022575 και $\widehat{scale}$ : 0.228384.
Για την εκτίμηση, καθώς και την κατασκευή του γραφήματος, χρησιμοποιήθηκε το στατιστικό πακέτο R και οι εντολές δίνονται στο παράρτημα κώδικα (σελ.104-105)

Θεωρήσαμε τα αφαιρετέα ποσά $\boldsymbol{d} = (0.5, 0.6, 0.7)$.
Τα αποτελέσματα παρουσιάζονται στον παρακάτω πίνακα:

| $\boldsymbol{d}$ | $\boldsymbol{H_X}$ | $\boldsymbol{H(f,d)}$ | $\boldsymbol{H_l(f,d)}$ |
|---|---|---|---|
| 0.5 | 3.508186 | 3.72552469 | 1.27314 |
| 0.6 | 3.508186 | 3.794337077 | 1.08533 |
| 0.7 | 3.508186 | 3.852517397 | 0.946654954 |

Στο παράδειγμα η $H(x)$ παριστά την εντροπία που εκτιμήθηκε στο δείγμα και οι $H(f,d)$, $H_l(f,d)$ τις εντροπίες κόστους ανά πληρωμή και κόστους ανά ζημιά αντίστοιχα. Παρατηρούμε πως η εντροπία κόστους ανά πληρωμή αυξάνει όσο αυξάνει το $d$.

Επιπλέον παρατηρούμε πως όσο αυξάνει η $H(f,d)$ μειώνεται η $H_l(f,d)$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Κεφάλαιο 5ο

## 5.1. Η Παρελθοντική Εντροπία (Past Entropy)

Στις προηγούμενες ενότητες αναλύθηκε η αβεβαιότητα που εμπεριέχεται στην υπολειπόμενη διάρκεια ζωής ενός συστήματος, και η επέκτασή της σε εφαρμογή αφαιρετέου ποσού σε Μοντέλα Απώλειας.

Στην ενότητα αυτή θα ασχοληθούμε με την παρελθοντική εντροπία *(Past Entropy)* όπως αυτή ορίστηκε από τους Di Crescenzo and Longobardi (2002), ως μέτρο αβεβαιότητας που αφορά τον πρότερο χρόνο ζωής ενός συστήματος και εκφράζεται από τη μεταβλητή $X_t^* = (t - X \,|\, X \leq t)$.

Θα εξεταστεί ομοίως το πότε μπορεί να χαρακτηρίσει μια κατανομή, για τη συνεχή περίπτωση μη αρνητικών τυχαίων μεταβλητών και θα δοθεί μια νέα κλάση κατανομών. Τέλος, θα εξεταστεί η ισοδύναμη έκφραση του μέτρου σε Μοντέλα Απώλειας υπό την επίδραση ορίου ιδίας κράτησης.

**Ορισμός 5.1.1**

$$\bar{H}(X;t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)}\, dx =$$

$$= 1 - \frac{1}{F(t)} \int_0^t f(x) \log rh(x)\, dx = 1 - E(\log h(X)\,|t \leq X), \quad (5.1)$$

όπου $rh(X) = \frac{f(x)}{F(t)}$, αντίστροφη συνάρτηση κινδύνου (*Reversed Hazard Rate).*

## 5.2. Η Παρελθοντική Εντροπία (Past Entropy) και προσδιορισμός κατανομών

Έστω,



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$\overline{H}(X;t) = 1 - \frac{1}{F(t)}\int_t^\infty f(x)\log rh(x)dx$$

Τότε παραγωγίζοντας και τα δύο μέλη ως προς $t$ έχουμε:

$$\frac{\partial}{\partial t}\,\overline{H}(X;t) - rh(t)[1 - \overline{H}(X;t) - \ln rh(t) = 0].$$

Επομένως, για σταθερό $t > 0$ η $rh(t)$ αποτελεί λύση της $g(x) = 0$, όπου

$$g(x) = \overline{H}'(X;t) - x[1 - \overline{H}(X;t) - \ln x] \tag{5.2}$$

Παραγωγίζοντας και τα δύο μέλη ως προς $x$ έχουμε:

$$g'(x) = \overline{H}(X;t) + \ln x. \tag{5.3}$$

Τότε για $g'(x) = 0$ έχουμε μοναδική λύση:

$$x_0 = \ exp\{-\overline{H}(X;t)\}. \tag{5.4}$$

**Θεώρημα 5.2.1**

Η ομοιόμορφη κατανομή στο διάστημα $(a,b), a < b$ χαρακτηρίζεται από την φθίνουσα

$$\overline{H}(X;t) = \ln(t-a), t > a. \tag{5.5}$$

**Απόδειξη**:

Έστω ότι η σχέση$\overline{H}\,(X;t) = \ln(t-a)$ ισχύει. Τότε, έχουμε $g(0) = \frac{\partial}{\partial t}\overline{H}\,(X;t) > 0$. Αποδεικνύεται πως η $g(x)$ είναι κυρτή συνάρτηση με ελάχιστο στο $x = x_0$. Οπότε η $g(x) = 0$ έχει λύση αν $g(x_0) = 0$. Χρησιμοποιώντας τη σχέση (5.5) παίρνουμε:

$$x_0 = exp\{-\overline{H}(X;t)\} = \left(\frac{1}{t-a}\right), t > a$$

και $g(x_0) = \frac{\partial}{\partial t}\,\overline{H}(X;t) - x_0[1 - \overline{H}(X;t) - \ln x_0] = 0.$

Επομένως η $g(x) = 0$ έχει μοναδική λύση στο $x = x_0$.

Από τη σχέση:
$g(x) = \overline{H}'(X;t) - x[1 - \overline{H}(X;t) - \ln x]$ προκύπτει ότι $rh(t)$ είναι λύση της.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Άρα αφού $x_0$ μοναδικό συνεπάγεται $rh(t) = x_0 = \left(\frac{1}{t-a}\right), t > a$.

Από αυτό προκύπτει ότι η κατανομή είναι η ομοιόμορφη στο $(a, b)$.

**Σημείωση:** Αν η $\overline{H}(X;t)$ αύξουσα συνάρτηση του $t$ και $g(\,x_0) < 0$, τότε η $g(x)$ έχει δύο λύσεις για κάθε $t \geq 0$. Η μια από αυτές τουλάχιστον θα είναι η αντίστροφη συνάρτηση κινδύνου $rh(t)$. Στο παράδειγμα που ακολουθεί θα μελετηθεί η περίπτωση της κατανομής $Beta(\lambda, 1)$

**Παράδειγμα:**

Έστω $X$ συνεχής, μη αρνητική τυχαία μεταβλητή με συνάρτηση κατανομής $Beta(\lambda, 1)$ και συνάρτηση πυκνότητας:

$$f(x) = \begin{cases} \lambda x^{\lambda-1}, & \lambda > 1, 0 < x < 1 \\ 0, & \text{διαφορετικά} \end{cases}$$

με συνάρτηση επιβίωσης της $X, \overline{F}(x) = P(X > t) = \int_x^1 f(u)\, du = 1 - x^\lambda$

και συνάρτηση αντίστροφου κινδύνου:
$rh(t) = \frac{f(t)}{\overline{F}(t)} = \frac{\lambda}{t}$, με $rh(t)$ αύξουσα για $t \in (0,1)$ και $\overline{H}(X;t) = \frac{\lambda-1}{\lambda} + \ln\left(\frac{t}{\lambda}\right)$.

Τότε

$$x_0 = exp\{-\overline{H}(X;t)\}$$

$$= exp\left(-\frac{\lambda-1}{\lambda} - \ln\left(\frac{t}{\lambda}\right)\right) = \frac{\lambda}{t} exp\left(\frac{1-\lambda}{\lambda}\right).$$

Τότε από τη σχέση $g(x) = \overline{H}'(X;t) - x[1 - \overline{H}(X;t) - \ln x]$ για $x_0$ έχουμε:

$$g(x_0\,) = \frac{1}{t}\left(1 - cexp\left(\frac{1-\lambda}{\lambda}\right)\right) < 0.$$

Επιπλέον, από την $x_0 = \frac{\lambda}{t} exp\left(\frac{1-\lambda}{\lambda}\right)$ προκύπτει $\frac{x_0}{rh(t)} = exp\left(\frac{1-\lambda}{\lambda}\right) > 1$ για $t \in (0,1)$.

Άρα, για κάθε $t \geq 0$ προκύπτουν δύο θετικές λύσεις με $rh(t) \leq x_0 \leq r'h(t)$ με $rh(t)$ επίσης αντίστροφη συνάρτηση κινδύνου (ικανοποιείται η συνθήκη :$\int_0^\infty r'h(t)dt = \infty$).

## 5.3. Παρελθοντική Εντροπία και κλάσεις κατανομών

**Ορισμός 5.3.1**

Η τυχαία μεταβλητή $X$ λέγεται ότι έχει αύξουσα αβεβαιότητα ζωής, $X \in IUL$ αν $\overline{H}(X;t)$ φθίνουσα ως προς $t$.

## 5.4. Παρελθοντική Εντροπία και επίδραση ορίου ιδίας κράτησης στα Μοντέλα Ζημιών

Όριο ιδίας κράτησης, σε ασφαλιστικούς όρους, ορίζεται ως το μέγιστο ποσό που έχει αναλάβει η ασφαλιστική εταιρία να καλύψει έναν δικαιούχο, μετά την επέλευση ενός ασφαλισμένου συμβάντος.

**Ορισμός 5.4.1**

Έστω ότι οι απώλειες $X$ δεν καταγράφονται όταν είναι ίσες ή μεγαλύτερες ενός καθορισμένου ποσού $u$, το οποίο καλείται όριο φερεγγυότητας ή όριο ιδίας κράτησης *(liability limit).* Σε αυτή την περίπτωση τα δεδομένα είναι αποκομμένα από πάνω, ή διαφορετικά δεξιά αποκομμένα *(truncated from above or right truncated).*

Τότε, η τυχαία μεταβλητή που περιγράφει τις από πάνω αποκομμένες απώλειες ορίζεται ως (Kluggman et al., 2008).

$$W(u) = \begin{cases} X, & X < u \\ \delta\varepsilon\nu\ o\rho\acute{\iota}\zeta\varepsilon\tau\alpha\iota, & X \geq u \end{cases} \text{ ή, } \qquad \text{ισοδύναμα, } W(u) = X|X < u. \qquad (5.6)$$

Η τυχαία μεταβλητή που περιγράφει τις ζημιές με λογοκρισία από πάνω *(Censored from above)* ορίζεται σύμφωνα με τους (Kluggman et al., 2008) ως:

$$V(u) = \begin{cases} X, & X < u \\ u, & X \geq u \end{cases} \text{ ή,} \qquad \text{ισοδύναμα } \; V(u) = \min\{X, u\}. \tag{5.7}$$

Σε αυτή την περίπτωση η λογοκριμένη από πάνω ή δεξιά μεταβλητή ορίζεται ως $V(u) = min\{X, u\}$, στην οποία αντιλαμβανόμαστε ότι αν ζημιά είναι $X \geq u$ τότε η ασφαλιστική καταβάλλει το ποσό $u$. Η τυχαία μεταβλητή $V(u)$ δεν είναι απολύτως συνεχής, αλλά μικτού τύπου, καθότι έχει μάζα πιθανότητας στο $u$. Αυτή συνιστά την περίπτωση κόστους ανά ζημία

**Λήμμα 5.4.1**

Η εντροπία κόστους ανά πληρωμή των απωλειών υπό την επίδραση ενός ορίου ιδίας κράτησης $u$, ορίζεται ως $\bar{H}(f, u)$ και δίνεται από τη σχέση:

$$\begin{aligned}\bar{H}(f,u) = & -\int_0^u \frac{f_X(x)}{\bar{F}_X(u)} \ln \frac{f_X(x)}{\bar{F}_X(u)} \\ = & \frac{1}{F_X(u)} \left[ H(X) + \int_u^{\infty} f_X(x) \ln f_X(x)\, dx \right] + \ln F_X(u). \end{aligned} \tag{5.8}$$

**Απόδειξη**:

Είναι γνωστό ότι η συνάρτηση πυκνότητας πιθανότητας της $W(u)$ δίνεται από τη σχέση:

$$f_{W(u)}(x) = \frac{f_X(x)}{F_X(u)}.$$

Συνεπώς, η εντροπία της μεταβλητής $W(u)$ δίνεται από τη σχέση:

$$\begin{aligned}\bar{H}(f,u) &= \int_0^u \frac{f_X(x)}{F_X(u)} \ln \frac{f_X(x)}{F_X(u)} dx \\ &= -\frac{1}{F_X(u)} \left[ \int_0^u f_X(x) \ln f_X(x)\, dx - \ln F_X(u) \int_0^u f_X(x)\, dx \right]. \end{aligned}$$

Όμως,



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$
\begin{aligned}
H(X) &= -\int_0^{\infty} f_X(x)\ln f_X(x)\,dx \\
&= -\left[\int_0^{u} f_X(x)\ln f_X(x)\,dx + \int_u^{\infty} f_X(x)\ln f_X(x)\,dx\right].
\end{aligned}
$$

Άρα,

$$
-\int_0^{u} f_X(x)\ln f_X(x)\,dx \;= H(X) + \int_u^{\infty} f_X(x)\ln f_X(x)\,dx\,.
$$

Αντικαθιστώντας στην σχέση έχουμε:

$$
\begin{aligned}
\overline{H}(f,u) &= \frac{1}{F_X(u)}\left[H(X) + \int_0^{u} f_X(x)\ln f_X(x)\,dx \;+ F_X(u)\ln F_X(u)\right] \\
&= \frac{1}{F_X(u)}\left[H(X) + \int_0^{u} f_X(x)\ln f_X(x)\,dx\right] + \ln F_X(u).
\end{aligned}
$$

Παρατηρούμε ότι, η παραπάνω έκφραση της εντροπίας με αποκοπή από πάνω είναι ταυτόσημη με αυτή που παρουσίασαν οι Di Crescenzo and Longobardi (2002) ως μέτρο παρελθοντικής εντροπίας *(Past lifetime Entropy).*

**Εντροπία κόστους ανά πληρωμή**

**1. Εκθετική κατανομή**

$$
\overline{H}(f,u) = \ln\left(\frac{1-e^{-\lambda u}}{\lambda}\right) + 1 - \frac{\left(\lambda u e^{-\lambda u}\right)}{(1-e^{-\lambda u})}. \tag{5.9}
$$

**2. Κατανομή Weibull**

$$
\begin{aligned}
\overline{H}(f,u) = {} & \frac{1-\ln c-\ln\tau+(\tau-1)(\gamma+\ln c\,+)+\dfrac{\Gamma(0,cd^{\tau})}{\tau}}{1-e^{-cu^{\tau}}} \\
& + \frac{e^{-cu}(\ln c+\ln\tau+(\tau-1)\ln u-1-cu^{\tau})}{1-e^{-cu^{\tau}}} \\
& + \ln(1-e^{-cu^{\tau}}).
\end{aligned} \tag{5.10}
$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**3. Κατανομή Pareto**

$$\overline{H}(f,u)=\frac{1+\frac{1}{a}+\ln\lambda-\ln a-\left(\frac{\lambda}{\lambda+u}\right)^{a}\left[\ln a+a\ln\lambda-\frac{a+1}{a}(1+a\ln(\lambda+u))\right]}{1-\left(\frac{\lambda}{\lambda+u}\right)^{a}}$$

$$+\ln\left[1-\left(\frac{\lambda}{\lambda+u}\right)^{a}\right]. \qquad (5.11)$$

Οι ζημιές που έλαβαν χώρα «πριν» από το όριο φερεγγυότητας (μικρότερες του ορίου) μπορούν ισοδύναμα να θεωρηθούν ζημιές που έλαβαν χώρα πριν από το σημείο $u$ (π. χ. «ηλικία»), παρελθοντικές ζημιές.

Όπως αναφέρθηκε και προηγουμένως, η $\overline{H}(f,u)$ είναι γνωστή ως Παρελθοντική Εντροπία (Past Lifetime Entropy) και ορίστηκε από τους Di Crescenzo and Longobardi (2002) ως:

$$\overline{H}(f,u)=1-\frac{1}{F(u)}\int_0^u f_X(x)\ln rh(x)\,dx,$$

όπου $rh(x)=\frac{f_X(x)}{F_X(x)}$ η αντίστροφη συνάρτηση κινδύνου.

Ιδιότητες (Di Crescenzo and Longobardi, 2002):

i. Ισχύει ότι

$$\overline{H}(f,u)\le\ln u, \qquad \forall u>0 \qquad (5.12)$$

ii. Αν οι απώλειες $X$ έχουν φθίνουσα συνάρτηση κινδύνου για κάθε $x$, τότε η $\overline{H}(f,u)$ είναι αύξουσα συνάρτηση για κάθε $u>0$.

Για την τυχαία μεταβλητή $V(u)$, που όπως είδαμε εκφράζει τις ζημιές ανά κόστος υπό την επίδραση ορίου ιδίας κράτησης ισχύουν τα παρακάτω:



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Λήμμα 5.4.2**

Η εντροπία κόστους ανά ζημιά υπό την επίδραση ορίου ιδίας κράτησης ορίζεται ως:

$$\bar{H}_l(f,u) = -\int_0^u f_X(x)\ln f_X(x)\,dx \; - \; \bar{F}_X(u)\ln \bar{F}_X(u)$$

$$= H(X) + \int_u^{\infty} f_X(x)\ln f_X(x)\,dx \; - \bar{F}_X(u)\ln \bar{F}_X(u). \qquad (5.13)$$

**Εντροπία κόστους ανά ζημιά**

**Παραδείγματα:**

1. **Εκθετική κατανομή**

$$\bar{H}_l(f,u) = (1-\ln\lambda)\left(1-e^{-\lambda u}\right). \qquad (5.14)$$

2. **Κατανομή Weibull**

$$\bar{H}_l(f,d) = (1-\ln c - \ln\tau)(1-e^{-cu^{\tau}})$$
$$+\frac{\tau-1}{\tau}(\gamma + \Gamma(0,cd^{\tau}) + \ln c + e^{-cu^{\tau}}\tau\ln u) \qquad (5.15)$$

3. **Κατανομή Pareto**

$$\bar{H}_l(f,u) = 1+\frac{1}{a}+\ln\lambda-\ln a-\left(\frac{\lambda}{\lambda+u}\right)^a\left[1+\frac{1}{a}-\ln a+\ln(\lambda+u)\right]. \qquad (5.16)$$

## 5.5. Επίδραση αφαιρετέου ποσού και ορίου ιδίας κράτησης στα Μοντέλα Απώλειας

Σε αυτή την ενότητα, θεωρούμε ότι η ασφαλιστική δεν αναγνωρίζει ζημιές κάτω του ποσού απαλλαγής ή αφαιρετέου ποσού $d$ και επιπλέον ότι παρέχει μέγιστη ασφαλιστική κάλυψη για μια ζημιά ίση με το όριο φερεγγυότητας $u$.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Έστω λοιπόν $d$ το αφαιρετέο ποσό και $u$ το όριο φερεγγυότητας και προφανώς $u < d$.

Η μεταβλητή που περιγράφει τις ζημιές ορίζεται ως (Kluggman et al., 2008):

$$Y(d,u) = (X \wedge u) - (X \wedge d) = \begin{cases} 0\,, & X \le d \\ X-d, & d < X \le u \\ u-d, & X > u \end{cases} \tag{5.17}$$

με το σύμβολο «Λ» να απεικονίζει το minimum.

Θεωρούμε ότι το αφαιρετέο ποσό εφαρμόζεται κατόπιν εφαρμογής του ορίου ιδίας κράτησης, με αποτέλεσμα αν η ζημιά ξεπερνά το όριο $u$ τότε το μέγιστο ποσό της ασφαλιστικής κάλυψης να είναι $u-d$. Η τυχαία μεταβλητή $Y(d,u)$ είναι μικτού τύπου καθώς έχει μάζες πιθανότητας στα σημεία 0 και $u-d$ και απολύτως συνεχής στο διάστημα $(0, u-d)$.

Η συνάρτηση πιθανότητας της $Y(d,u)$ :

$$f_{Y(d,u)}(y) = \begin{cases} F_X(d), & y = 0, \\ f_X(y+d), & 0 < y < u-d, \\ \bar{F}_X(u), & y = u-d, \\ 0, & y > u-d. \end{cases}$$

**Λήμμα 5.5.1**

Η εντροπία κόστους ανά ζημιά υπό την επίδραση ενός αφαιρετέου ποσού $d$ και ορίου φερεγγυότητας $u$ ορίζεται ως $H_l(f,d,u)$ και δίνεται από τη σχέση:

$$H_l(f,d,u) = -F_X(d)\ln F_X(d) - \int_d^u f_X(x)\ln f_X(x)\,dx - \bar{F}_X(u)\ln \bar{F}_X(u)$$

$$= H(X) + \int_0^d f_X(x)\ln f_X(x)\,dx + \int_u^\infty f_X(x)\ln f_X(x)\,dx$$

$$-F_X(d)\ln F_X(d) - \bar{F}_X(u)\, ln\, \bar{F}_X(u), \qquad d < u. \tag{5.18}$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Απόδειξη:**

Όπως γνωρίζουμε η $H_l(f,d,u)$ είναι μικτού τύπου κατανομή, με μάζες πιθανότητας στα σημεία 0 και $u-d$ , $F_X(d)$, $\bar{F}_X(u)$ αντίστοιχα και συνεχής στο $(o,u-d)$.

Επομένως,

$$H_l(f,d,u) = -F_X(d)\ln F_X(d) + \int_0^{u-d} f_X(y+d)\ln f_X(y+d)\,dy - \bar{F}_X(u)\,ln\,\bar{F}_X(u). \tag{5.19}$$

Για το ολοκλήρωμα έχουμε:

$0 < y < u-d$, ισοδύναμα, $d < y+d < u$.

$x = y + d$ επομένως $d < x < u$ και $dx = dy$.

Αντικαθιστώντας στην (5.19) έχουμε:

$$H_l(f,d,u) = -F_X(d)\ln F_X(d) + \int_d^u f_X(x)\ln f_X(x)\,dy - \bar{F}_X(u)\,ln\,\bar{F}_X(u). \tag{5.20}$$

Όμως,

$$\int_0^\infty f_X(x)\ln f_X(x)\,dy = \int_0^d f_X(x)\ln f_X(x)\,dy + \int_d^u f_X(x)\ln f_X(x)\,dy + $$
$$+\int_u^\infty f_X(x)\ln f_X(x)\,dy$$

ή ισοδύναμα,

$$-\int_0^\infty f_X(x)\ln f_X(x)\,dy = -\int_0^d f_X(x)\ln f_X(x)\,dy - \int_d^u f_X(x)\ln f_X(x)\,dy + $$
$$-\int_u^\infty f_X(x)\ln f_X(x)\,dy.$$

Συνεπώς,

$$H(X) + \int_0^d f_X(x)\ln f_X(x)\,dy + \int_d^u f_X(x)\ln f_X(x)\,dy = -\int_u^\infty f_X(x)\ln f_X(x)\,dy. \tag{5.21}$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Αντικαθιστώντας τη σχέση (5.21) στη σχέση (5.20) προκύπτει το ζητούμενο.

**Πρόταση 5.5.1**

Η εντροπία κόστους ανά ζημιά με εφαρμογή αφαιρετέου ποσού $d$ και ορίου ιδίας κράτησης $u$, $H_l(f,d,u)$, σχετίζεται με τις $H_l(f,d)$ και $\bar{H}_l(f,u)$ μέσω της εξίσωσης:

$$H_l(f,d,u) = \bar{H}_l(f,u) + \int_0^d f_X(x)\ln f_X(x)\,dx - F_X(d)\ln F_X(d)$$

$$= \; H_l(f,d) + \int_u^\infty f_X(x)\ln f_X(x)\,dx - \bar{F}_X(u)\,ln\,\bar{F}_X(u), \;\; d < u. \qquad (5.2)$$

Η σχέση που συνδέει τις $H(X), \bar{H}(f,u)$, $H(f,d)$, υποθέτοντας ότι το όριο ιδίας κράτησης $u$ ισούται με το αφαιρετέο ποσό $d$, παρουσιάζεται στην παρακάτω πρόταση.

**Πρόταση 5.5.2**

$$H(X) = \bar{F}_X(d)H(f,d) + F_X(d)\bar{H}(f,u) + H[F_X(d)\bar{F}_X(d)] \qquad (5.23)$$

όπου $H[p,\bar{p}] = -p\ln p - \bar{p}\ln\bar{p}$, η εντροπία της κατανομής Bernoulli και προφανώς $\bar{p} = 1 - p$.

**Απόδειξη:**

Από τα λήμματα (αποκοπή από κάτω, αποκοπή από πάνω) έχουμε δείξει ότι:

$$H(f,d) = \frac{1}{\bar{F}_X(d)}\left[H_X + \int_0^d f_X(x)\ln f_X(x)\,dx + \bar{F}_X(d)\ln\bar{F}_X(d)\right]$$

και

$$\bar{H}(f,d) = \frac{1}{F_X(d)}\left[H_X + \int_d^\infty f_X(x)\ln f_X(x)\,dx + F_X(d)\ln F_X(d)\right].$$

Τότε, πολλαπλασιάζοντας τις αντίστοιχες εξισώσεις με τις συναρτήσεις δεξιάς ουράς και κατανομής της $X$ στο σημείο $d$ και προσθέτοντας κατά μέλη, έχουμε:

$$\bar{F}_X(d)H(f,d) + F_X(d)\bar{H}(f,d) = H(X) + \int_0^d f_X(x)\ln f_X(x)\,dx + \bar{F}_X(d)\ln\bar{F}_X(d)$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$+H(X) \ + \int_d^{\infty} f_X(x) \ln f_X(x)\, dx + F_X(d) \ln F_X(d)$$

$$= 2H(X) - +\bar{F}_X(d) \ln \bar{F}_X(d) + F_X(d) \ln F_X(d).$$

Όμως $\bar{F}_X(d) = 1 - F_X(d)$ επομένως $\bar{F}_X(d) \ln \bar{F}_X(d) + F_X(d) \ln F_X(d) = -H[F_X(d), \bar{F}_X(d)]$ , εντροπία μιας Bernoulli με πιθανότητα $F_X(d)$.

Άρα

$$\bar{F}_X(d) H(f, d) + F_X(d) \bar{H}(f, d) = 2H(X) - H(X) - H[F_X(d), \bar{F}_X(d)]$$

$$= H(X) - H[F_X(d), \bar{F}_X(d)].$$

Η τελευταία ισότητα μπορεί να γραφεί και ως:

$$H(X) = \bar{F}_X(d) H(f, d) + F_X(d) \bar{H}(f, d) + H[F_X(d), \bar{F}_X(d)]$$

και αποδεικνύεται το ζητούμενο.

Η παραπάνω σχέση υποδεικνύει ότι η αβεβαιότητα που εμπεριέχει μια ζημιά αποσυντίθεται σε τρείς συνιστώσες:

I. Την αβεβαιότητα για το αν η ζημιά έχει ξεπεράσει το αφαιρετέο ποσό $d$.
II. Την αβεβαιότητα για τη ζημιά εφόσον έχει ξεπεράσει το αφαιρετέο ποσό και αποτελεί κάλυψη της ασφαλιστικής .
III. Την αβεβαιότητα για τη ζημιά που καλύπτεται από την ασφαλιστική, δοθέντος ότι δεν έχει ξεπεράσει το όριο φερεγγυότητας $d$(Di Crescenzo and Longobardi, 2002).

**Παραδείγματα:**

1. **Εκθετική κατανομή**

$$H_l(f, d, u) = e^{-\lambda d}(1 + \lambda d - \ln \lambda) - e^{-\lambda u}(1 - \ln \lambda) - \left(1 - e^{-\lambda d}\right) \ln\left(1 - e^{-\lambda d}\right). \tag{5.24}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

2. **Κατανομή Weibull**

$$H_l(f,d,u) = e^{-cd^\tau}\big(1 - \ln c - \ln \tau - (\tau - 1)\ln d + cd^\tau + \ln\big(1 - e^{-cd^\tau}\big)\big)$$
$$- e^{-cu^\tau}(1 - \ln c - \ln \tau - (\tau - 1)\ln u)$$
$$+ \frac{(\tau - 1)}{\tau}[\Gamma(0, cu^\tau) - \Gamma(0, cd^\tau)] - \ln\big(1 - e^{-cd^\tau}\big), d < u. \qquad (5.25)$$

3. **Κατανομή Pareto**

$$H_l(f,d,u) = \left(\frac{\lambda}{\lambda + d}\right)^a \left(1 + \frac{1}{a} - \ln a - a\ln\lambda + (1 + a)\ln(\lambda + d)\right)$$
$$-\left(\frac{\lambda}{\lambda + d}\right)^a \left(1 + \frac{1}{a} - \ln a + \ln(\lambda + u)\right)$$
$$-\left[1 - \left(\frac{\lambda}{\lambda + d}\right)^a\right]\ln\left[1 - \left(\frac{\lambda}{\lambda + d}\right)^a\right], \quad d < u. \qquad (5.26)$$

**Πρόταση 5.5.3**

Ισχύουν οι παρακάτω σχέσεις:

**I.** $3H(X) = H_l(f,u) + \bar{H}_l(f,u) - H[F_X(d), \bar{F}_X(d)]$,

**II.** $\bar{H}_l(f,u) = F_X(u)\bar{H}(f,u) + H[F_X(d), \bar{F}_X(d)]$,

**III.** $H_l(f,u) \leq (\ln u)F_X(u) + H[F_X(u), \bar{F}_X(u)]$.

## 5.6. Υπολειπόμενη και Παρελθοντική εντροπία σε Μοντέλα Επιβίωσης

Έστω $X$ και $Y$ τυχαίες μεταβλητές με αντίστοιχες συναρτήσεις κατανομής $F$ και $G$. Έστω επίσης $f$ και $g$ οι αντίστοιχες συναρτήσεις πυκνότητας πιθανότητας των $X$ και $Y$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Μοντέλο αναλογικού κίνδυνου (Proportional Hazard Model –Cox)**

Έστω ότι οι $X$ και $Y$ ικανοποιούν το μοντέλο αναλογικών κινδύνων

$$\bar{G}(x) = \left(\bar{F}(x)\right)^{\theta}, \forall x > 0 \ \kappa\alpha\iota \ \forall \theta > 0 \tag{5.27}$$

$$\text{ή} \quad h_G(x) = \theta h_F(x).$$

Ισοδύναμα, ο κίνδυνος του $Y$ είναι ανάλογος του κινδύνου του $X$ (Cox, 1972).

**Θεώρημα 5.6.1**

Υπό την ισχύ του μοντέλου αναλογικού κινδύνου που παρουσιάστηκε στην παραπάνω εξίσωση, η υπολειπόμενη εντροπία $H(g,d)$ της τυχαίας μεταβλητής Y δίνεται από την παρακάτω σχέση:

$$H(g,d) = 1 - \frac{1}{\theta} - \ln\theta - \int_0^1 \ln\left(\frac{f(x)}{\bar{F}(d)}\right) dy^{\theta} \tag{5.28}$$

με $x = F^{-1}\left(1 - y\bar{F}(d)\right)$, *το οποίο προκύπτει από τη σχέση* $y = \frac{\bar{F}(x)}{\bar{F}(d)}$.

**Απόδειξη:**

$$H(g,d) = -\int_d^{\infty} \frac{g(x)}{\bar{G}(d)} \ln\frac{g(x)}{\bar{G}(d)} dx$$

$$= -\int_d^{\infty} \frac{\theta\left(\bar{F}(x)\right)^{\theta-1} f(x)}{\left(\bar{F}(d)\right)^{\theta}} \ln\frac{\theta\left(\bar{F}(x)\right)^{\theta-1} f(x)}{\left(\bar{F}(d)\right)^{\theta}} dx$$

$$= -\int_d^{\infty} \theta\left(\frac{\bar{F}(x)}{\bar{F}(d)}\right)^{\theta-1}\left(\frac{f(x)}{\bar{F}(d)}\right) \ln\left[\theta\left(\frac{\bar{F}(x)}{\bar{F}(d)}\right)^{\theta-1}\left(\frac{f(x)}{\bar{F}(d)}\right)\right] dx,$$

από το οποίο με αλλαγή της μεταβλητής $y = \frac{\bar{F}(x)}{\bar{F}(d)}$ παίρνουμε το ζητούμενο.

**Μοντέλο αντίστροφου αναλογικού κίνδυνου (Proportional Reversed Hazard Model)**

Έστω ότι οι $X$ και $Y$ ικανοποιούν το μοντέλο, τότε:

$$G(x) = \left(F(x)\right)^{\theta}, \qquad \forall x > 0 \; \kappa\alpha\iota \; \forall \theta > 0 \tag{5.29}$$

(Di Crescenzo (2000), Gupta and Gupta (2007), Sankaran and Gleeja (2008)).

**Θεώρημα 5.6.7**

Υπό την ισχύ του μοντέλου αναλογικού κινδύνου που παρουσιάστηκε στην παραπάνω εξίσωση, η παρελθοντική εντροπία $\bar{H}(g,u)$ της τυχαίας μεταβλητής Y δίνεται από την παρακάτω σχέση:

$$\bar{H}(g,u) = 1 - \frac{1}{\theta} - \ln\theta - \int_0^1 \ln\left(\frac{f(x)}{F(u)}\right) dy^{\theta} \tag{5.30}$$

με $x = F^{-1}\left(yF(u)\right)$, *το οποίο προκύπτει από τη σχέση* $y = \frac{F(x)}{F(u)}$.

**Απόδειξη:**

$$\bar{H}(g,u) = -\int_0^u \frac{g(x)}{G(u)} \ln\frac{g(x)}{G(u)} dx$$

$$= -\int_0^u \frac{\theta\left(F(x)\right)^{\theta-1} f(x)}{\left(F(u)\right)^{\theta}} \ln\frac{\theta\left(F(x)\right)^{\theta-1} f(x)}{\left(F(u)\right)^{\theta}} dx$$

$$= -\int_0^u \theta\left(\frac{F(x)}{F(u)}\right)^{\theta-1}\left(\frac{f(x)}{F(d)}\right)\ln\left[\theta\left(\frac{F(x)}{F(d)}\right)^{\theta-1}\left(\frac{f(x)}{F(u)}\right)\right] dx,$$

από το οποίο με αλλαγή της μεταβλητής $y = \frac{\bar{F}(x)}{\bar{F}(d)}$ παίρνουμε το ζητούμενο.

**Frailty Model**

Έστω ότι η τυχαία μεταβλητή $Y$ ακολουθεί ένα μοντέλο **(***Frailty)* έτσι ώστε

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$\bar{G}(y) = e^{-Q(\theta\Lambda(y))}, \ \ \theta > 0, \tag{5.31}$$

όπου $Q(\cdot)$ κοίλη και αύξουσα συνάρτηση με $Q(0) = 0$ και $Q(\infty)$=∞,

$\Lambda(x)$ : αναφορική ή βασική αθροιστική συνάρτηση κινδύνου (*baseline cumulative hazard function)* με $\Lambda(\infty)$=∞. (Sankaran and Gleeja (2008), Vonta (1996) and Vonta and Karagrigoriou (2010)).

Στη συνέχεια θα δοθούν σχέσεις για την υπολειπόμενη και παρελθοντική εντροπία της $Y$ κάνοντας χρήση της συνάρτησης $Q(\cdot)$ και της βασικής αθροιστικής συνάρτησης κινδύνου $\Lambda(x)$.

**Λήμμα 5.6.1**

Η υπολειπόμενη και παρελθοντική εντροπία της $Y$ υπό την επίδραση του Frailty Model που δόθηκε στην παραπάνω εξίσωση δίνονται αντίστοιχα από τις σχέσεις που ακολουθούν:

$$H(g,d) = \ln \bar{G}(d) - \frac{1}{\bar{G}(d)} \int_0^{\bar{G}(d)} \ln\left(-\frac{dz}{dy}\right) dz, \tag{5.32}$$

$$\bar{H}(g,u) = -\frac{\bar{G}(u)}{G(u)} \ln G(u) + \frac{1}{G(u)} \int_0^{\bar{G}(u)} \ln\left(-\frac{dz}{dy}\right) dz \tag{5.33}$$

αντίστοιχα, με $\bar{G}(d) = e^{-Q(\theta\Lambda(d))}$, $\bar{G}(u) = e^{-Q(\theta\Lambda(u))}$ και $\frac{dz}{dy} = -zQ'\big(\theta\Lambda(y)\big)\theta\lambda(y)$

και $y = \Lambda^{-1}\left(Q^{-1}\left(\frac{-\ln z}{\theta}\right)\right)$, το οποίο προκύπτει από τη σχέση $z = \bar{G}(y) = e^{-Q(\theta\Lambda(y))}$.

**Απόδειξη:**

Από τον ορισμό έχουμε:

$$H(g,d) = -\int_d^{\infty} \frac{g(x)}{\bar{G}(d)} \ln \frac{g(x)}{\bar{G}(d)} dx$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$= -\int_d^\infty \frac{e^{-Q(\theta\Lambda(d))}Q'(\theta\Lambda(y))\theta\lambda(y)}{\bar{G}(d)} \ln \frac{e^{-Q(\theta\Lambda(d))}Q'(\theta\Lambda(y))\theta\lambda(y)}{\bar{G}(d)} dy.$$

Κάνοντας αλλαγή μεταβλητών και θέτοντας σχέση $z = \bar{G}(y) = e^{-Q(\theta\Lambda(y))}$ και εφόσον $\frac{dz}{dy} = -zQ'(\theta\Lambda(y))\theta\lambda(y)$ και $\lim_{y\to\infty} e^{-Q(\theta\Lambda(y))} = 0$ παίρνουμε:

$$H(g,d) = -\frac{1}{\bar{G}(d)} \int_d^{\bar{G}(d)} \left[\left(-\frac{dz}{dy}\right) - \ln \bar{G}(d)\, dy\right]$$

$$= \ln \bar{G}(d) - \frac{1}{\bar{G}(d)} \int_0^{\bar{G}(d)} \ln\left(-\frac{dz}{dy}\right) dz.$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Κεφάλαιο 6ο

## Αθροιστικές Υπολειπόμενες Εντροπίες

Στα προηγούμενα κεφάλαια παρουσιάσαμε μέτρα εντροπίας που αποτελούν επεκτάσεις της εντροπίας του Shannon με τόπο εφαρμογής σε περικομμένα ή λογοκριμένα δείγματα.

Είδαμε ότι η Υπολειπόμενη και η Παρελθοντική εντροπία προσδιορίζουν τις συναρτήσεις κατανομών, και παρουσιάσαμε τις οικογένειες κατανομών που προκύπτουν από τη μονοτονία των μέτρων. Τα παραπάνω ωστόσο αφορούν κατανομές στις οποίες ορίζεται πυκνότητα.

Στο κεφάλαιο αυτό θα εξεταστούν μέτρα εντροπίας ως συναρτησιακά της συνάρτησης επιβίωσης μιας κατανομής.

Επιπλέον, όπως θα δούμε, δύναται να χρησιμοποιηθούν για το χαρακτηρισμό κατανομής του υπό εξέταση δείγματος. Τέλος, θα δειχθεί ότι οι Δυναμική και Διπλά Περικομμένη Αθροιστική Υπολειπόμενη Εντροπία ορίζουν κλάσεις κατανομών ζωής, βάσει μονοτονίας.

## 6.1. Αθροιστική Υπολειπόμενη Εντροπία (Cumulative Residual Entropy)

**Ορισμός 6.1.1**

Έστω $X$ μη αρνητική τυχαία μεταβλητή με $f(x)$ συνάρτηση πυκνότητας πιθανότητας και $\bar{F}(x)$ συνάρτηση επιβίωσης. Τότε η αθροιστική υπολειπόμενη εντροπία (*Cumulative Residual Entropy-CRE*) ή συμβολικά $\mathcal{E}(X)$ δίνεται από τον τύπο:

$$\mathcal{E}(X) = -\int_0^\infty \bar{F}(x) \log \bar{F}(x) dx = \int_0^\infty \bar{F}(x) \Lambda(x) dx, \tag{6.1}$$

όπου $\Lambda(x)$: αθροιστική συνάρτηση κινδύνου.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

## Σχέση Αθροιστικής Υπολειπόμενης Εντροπίας και Εντροπίας Shannon

**Θεώρημα 6.1.2**

$$\mathcal{E}(X) \geq Ce^{H(x)}, \tag{6.2}$$

με $C = e^{\int_0^1 \log(x|\log x|)dx} \cong 0.2.65$.

**Απόδειξη:**

Έστω $\bar{F}(x) = P(X > x) = \int_x^\infty f_X(t)dt.$ Χρησιμοποιώντας τη Log-sum ανισότητα παίρνουμε:

$$\int_0^\infty f_X(x) \log \frac{f_X(x)}{\bar{F}(x)|\log \bar{F}(x)|} dx \geq \log \frac{1}{\int_0^\infty \bar{F}(x)|\log \bar{F}(x)|dx} = \log \frac{1}{\mathcal{E}(X)}.$$

Αν $\int \bar{F}(x) \log \bar{F}(x) = \infty$ τότε προφανώς ισχύει.

Για το αριστερό μέρος της παραπάνω εξίσωσης:

$$-H(x) - \int_0^\infty f_X(x) \log \bar{F}(x)|\log \bar{F}(x)|dx,$$

έτσι ώστε,

$$H(x) + \int_0^\infty f_X(x) \log \bar{F}(x)|\log \bar{F}(x)|dx \leq \log \mathcal{E}(X).$$

Κάνοντας αλλαγή μεταβλητών παίρνουμε:

$$\int_0^\infty f_X(x) \log \bar{F}(x)|\log \bar{F}(x)|dx = \int_0^1 \log(x|\log x|)dx.$$

Επομένως,

$$H(x) + \int_0^1 \log(x|\log x|)dx \leq \log \mathcal{E}(X).$$

Εφαρμόζοντας την εκθετική κατανομή και στα δύο μέλη παίρνουμε το ζητούμενο.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

## Σχέση Υπολειπόμενης Αθροιστικής και Αμοιβαίας Εντροπίας

**Πρόταση 6.1.3**

$$\mathcal{E}(X,\Upsilon) \geq 2Ce^{\left(\frac{H(X,\Upsilon)}{2}\right)}. \tag{6.3}$$

**Απόδειξη:**

Κάνοντας χρήση της προηγούμενης πρότασης, της κυρτότητας της $e^x$ και της ανισότητας Jensen προκύπτει:

$$\mathcal{E}(X,\Upsilon) = \mathcal{E}(X) + E\big(\mathcal{E}(\Upsilon|\, X)\big) \geq Ce^{H(X)} + Ce^{E\left(H(X|Y)\right)} \geq 2Ce^{\left(\frac{H(X)+E\left(H(X|Y)\right)}{2}\right)}$$

$$= 2Ce^{\left(\frac{H(X,\Upsilon)}{2}\right)}.$$

Στη δεύτερη ανισότητα χρησιμοποιήθηκε η σχέση $2e^{\left(\frac{(x+y)}{2}\right)} \leq e^x + e^y$.

Η πρόταση που ακολουθεί συνδέει τη διαφορική εντροπία του Shannon με την αθροιστική υπολειπόμενη εντροπία.

**Πρόταση 6.1.4**

Έστω $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή με συνεχή κατανομή. Τότε υπάρχει συνάρτηση $\varphi = \varphi_X$ τέτοια ώστε η διαφορική εντροπία του Shannon $H(\Upsilon)$ της $\Upsilon = \varphi(X)$ συνδέεται με την $\mathcal{E}(X)$ με την παρακάτω σχέση:

$$H(\Upsilon) = \frac{\mathcal{E}(X)}{E(X)} - \frac{1}{E(X)}\log\frac{1}{E(X)}. \tag{6.4}$$

Στην ενότητα που ακολουθεί, το Θεώρημα 6.2.1 (Asadi and Zohrevand, 2007) δίνει τη σχέση μεταξύ της αθροιστικής υπολειπόμενης εντροπίας $\mathcal{E}(X)$ και του μέσου υπολειπόμενου χρόνου ζωής.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

## 6.2. Αθροιστική εντροπία και μέσος υπολειπόμενος χρόνος ζωής

**Θεώρημα 6.2.1**

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

$$\mathcal{E}(X) = E\big(m(X)\big). \tag{6.5}$$

**Απόδειξη:**

Το δεύτερο μέλος της εξίσωσης εκφράζει τον αναμενόμενο μέσο υπολειπόμενο χρόνο ζωής $X$, δηλαδή

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

Αλλάζοντας τα όρια ολοκλήρωσης έχουμε:

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

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

με $h(t) = \frac{f(t)}{\bar{F}(t)}$: συνάρτηση κινδύνου.

**Παραδείγματα**

**I. Ομοιόμορφη κατανομή**

Έστω $X$ ομοιόμορφη στο $(a, b)$, για $a > 0$.

Ο μέσος υπολειπόμενος χρόνος ζωής της υπολογίζεται ως:

$$m(t) = \frac{a - t}{2}.$$

Τότε, με χρήση του θεωρήματος προκύπτει:



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$\mathcal{E}(X) = E\left(\frac{a-t}{2}\right) = \int_0^{\alpha}\left(\frac{a-x}{2}\right)\frac{1}{a}dx = \frac{1}{2a}\left[ax - \frac{x^2}{2}\right]_0^{a} = \frac{a}{4}. \tag{6.6}$$

### II. Εκθετική κατανομή

Ο μέσος υπολειπόμενος χρόνος της εκθετικής δίνεται από τη σχέση:

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

Τότε,

$$\mathcal{E}(X) = E\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda}. \tag{6.7}$$

### III. Κατανομή Pareto

Ο μέσος υπολειπόμενος χρόνος της κατανομής Pareto δίνεται από τη σχέσηQ

$$m(t) = \frac{t+b}{a-1}, \ \ \alpha > 1, b > 0.$$

Τότε,

$$\mathcal{E}(X) = E\left(\frac{X+b}{a-1}\right) = \frac{E(X)}{a-1} = \frac{\frac{b}{a-1}+b}{a-1} = \frac{ab}{(a-1)^2}. \tag{6.8}$$

## 6.3. Δυναμική Αθροιστική Υπολειπόμενη Εντροπία

**Ορισμός 6.3.1**

Η Δυναμική αθροιστική υπολειπόμενη εντροπία του χρόνου ζωής της τυχαίας μεταβλητής $X_t$ (*Dynamic Cumulative Residual Entropy -DCRE)* ή ισοδύναμα, αθροιστική υπολειπόμενη εντροπία, ορίζεται από την παρακάτω σχέση:

$$DCRE(X;t) = \int_t^{\infty}\bar{F}_{x_t}(x)\log\bar{F}_{x_t}(x)\,dx = -\int_t^{\infty}\frac{\bar{F}(x)}{F(t)}\log\frac{\bar{F}(x)}{\bar{F}(t)}dx$$

$$= -\frac{1}{\bar{F}(t)}\int_t^{\infty}\bar{F}\,(x)\log\bar{F}(x)dx + \log\bar{F}(t)\int_t^{\infty}\frac{\bar{F}(x)}{\bar{F}(t)}dx$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$= -\frac{1}{\bar{F}(t)}\int_t^{\infty}\bar{F}\,(x)\log\bar{F}(x)dx + m(t)\log\bar{F}(t). \tag{6.9}$$

Για $t = 0$, έπεται $DCRE(X; 0) = \mathcal{E}(X)$.

## 6.4 Δυναμική Αθροιστική Υπολειπόμενη Εντροπία και χαρακτηρισμός κατανομών

**Θεώρημα 6.4.1**

Έστω $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή με μέσο υπολειπόμενο χρόνο ζωής και δυναμική αθροιστική υπολειπόμενη εντροπία $DCRE(X; t)$, με $DCRE(X; t) < \infty$.

Τότε,

$$DCRE(X; t) = E(m(x)(X)|X > t), \forall t \geq 0. \tag{6.10}$$

**Απόδειξη:**

Το 2ο μέλος της σχέσης παριστά τη δεσμευμένη μέση τιμή του μέσου υπολειπόμενου χρόνου ζωής της τυχαίας μεταβλητής $t$, δοθέντος ότι έχει υπερβεί την ηλικία $t$.

Επομένως έχουμε:

$$E(m(X)|X > t) = \frac{1}{\bar{F}(t)}\int_t^{\infty} m(x)f\,(x)dx$$

με

$$m(t) = \int_t^{\infty}\frac{\bar{F}(x)}{\bar{F}(t)}dx.$$

Οπότε αντικαθιστώντας παίρνουμε:

$$E(m(X)|X > t) = \frac{1}{\bar{F}(t)}\int_t^{\infty}\frac{\int_x^{\infty}\bar{F}(u)\,du}{\bar{F}(x)}f\,(x)dx.$$

Αλλάζοντας τη σειρά ολοκλήρωσης για $t < x < u$

$$E(m(X)|X > t) = \frac{1}{\bar{F}(t)}\int_t^{\infty}\left(\int_t^{u}\frac{f(x)}{\bar{F}(x)}dx\right)\bar{F}(u)du$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$= \frac{1}{\bar{F}(t)} \int_t^{\infty} \left( \int_0^u h(x)dx - \int_0^t h(x)\, dx \right) \bar{F}(u)du$$

$$= \int_t^{\infty} (-\log \bar{F}(u) + \log \bar{F}(t)) \frac{\bar{F}(u)}{\bar{F}(t)} du$$

$$= -\int_t^{\infty} \frac{\bar{F}(x)}{F(t)} \log \frac{\bar{F}(x)}{\bar{F}(t)} dx.$$

Το προηγούμενο Θεώρημα αποτελεί γενίκευση του Θεωρήματος

$$\mathcal{E}(X) = E\big(m(X)\big). \tag{6.11}$$

Προφανώς, για $t = 0$ , $DCRE(X;t) = \mathcal{E}(X) = E(m(X)|X > 0) = E\big(m(X)\big)$.

**Παραδείγματα**

**I. Ομοιόμορφη κατανομή**

Έστω $X$ ομοιόμορφη στο $(a, b)$, για $a > 0$.

Τότε, με χρήση του θεωρήματος,

$$DCRE(X;t) = E\left(\frac{a-t}{2} |X \geq t\right) = \int_0^{\alpha} \left(\frac{a-x}{2}\right) \frac{1}{a} dx = \frac{1}{2a} \left[ax - \frac{x^2}{2}\right]_0^{a} = \frac{a}{4}. \tag{6.12}$$

**II. Εκθετική κατανομή**

Έστω $X \sim Exp(\lambda),\ \lambda \geq 0$.

Τότε,

$$DCRE(X;t) = E\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda}. \tag{6.13}$$

**III. Κατανομή Pareto**

Έστω $X \sim Pareto(a, b),\ a > 0, b > 0$.

$$DCRE(X;t) = E\left(\frac{X+b}{a-1} |X \geq t\right) = \frac{1}{a-1} \frac{1}{\frac{b^a}{(t+b)^a}} \int_t^{\infty} (x+b) \frac{\alpha b^a}{(x+b)^{a+1}} dx$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$= \frac{a(t+b)}{(a-1)^2}. \tag{6.14}$$

**IV. Δυναμοκατανομή,** $a > 0, b > 0$

$$DCRE(X;t) = E\left(\frac{b-X}{a+1}|t \geq 0\right) = \frac{1}{a-1}\frac{1}{\left(\frac{b-t}{b}\right)^a}\int_t^b (b-x)\frac{a(b-x)^{a-1}}{b^a}dx$$

$$= \frac{a(b-t)}{(a+1)^2}. \tag{6.15}$$

**Χαρακτηρισμός κατανομής**

**Λήμμα 6.4.1**

Έστω $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή με μέσο υπολειπόμενο χρόνο ζωής $m(t)$ και μέση τιμή.

Η $m(t)$ είναι γραμμική συνάρτηση της μορφής:

$$m(t) = (c-1)t + \mu, \tag{6.16}$$

αν και μόνο η $X$ ακολουθεί μια από τις παρακάτω κατανομές:

i. Εκθετική, όταν $c = 1$,
ii. Pareto, όταν $c > 1$.

**Θεώρημα 6.4.2**

Έστω $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή με μέσο υπολειπόμενο χρόνο ζωής $m(t)$ και δυναμική αθροιστική υπολειπόμενη εντροπία $DCRE(X;t)$.Τότε,

$$DCRE(X;t) = cm(t), \tag{6.17}$$

αν και μόνο αν η $X$ ακολουθεί μια από τις παρακάτω κατανομές:

i. Εκθετική, όταν $c = 1$,
ii. Pareto, όταν $c > 1$.

**Απόδειξη:**

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Έστω

$$DCRE(X;t) = cm(t).$$

Τότε από τον ορισμό έχουμε:

$$-\frac{1}{\bar{F}(t)}\int_t^{\infty} \bar{F}\,(x) \log \bar{F}(x) dx + m(t) \log \bar{F}(t) = cm(t).$$

Παραγωγίζοντας και τα δύο μέλη ως προς $t$ προκύπτει:

$$c'm(t) = m'(t) \log \bar{F}(t) - m(t)h(t) + \log \bar{F}(t)$$

$$-h(t)\frac{1}{\bar{F}(x)}\int_t^{\infty} \bar{F}\,(x) \log \bar{F}(x) dx\,, \mu\varepsilon\ h(t)\text{: } \sigma\upsilon\nu\acute{\alpha}\rho\tau\eta\sigma\eta\ \kappa\iota\nu\delta\acute{\upsilon}\nu o\upsilon.$$

$$m'(t) \log \bar{F}(t) - m(t)h(t) + \log \bar{F}(t) + h(t)[cm(t) - m(t) \log \bar{F}(t)].$$

Έχοντας λάβει υπόψη ότι:

$$\frac{d}{dt}\int_t^{\infty} \bar{F}\,(x) \log \bar{F}(x) dx = -\frac{d}{dt}\int_0^{t} \bar{F}\,(x) \log \bar{F}(x) dx = -\bar{F}(t) \log \bar{F}(t), \qquad (6.18)$$

$$\frac{d}{dt}\frac{1}{\bar{F}(t)} = \frac{f(t)}{\bar{F}^2(t)} = \frac{h(t)}{\bar{F}(t)}.$$

Όμως $m'(t) + 1 = h(t)m(t)$. Αντικαθιστώντας στην (6.18) έχουμε:

$$c(m(t)h(t) - 1) = (m(t)h(t) - 1) \log \bar{F}(t) - m(t)h(t) + \log \bar{F}(t).$$

Επομένως για $\forall\ x \geq 0$ έχουμε:

$$h(x)m(x) = c.$$

Ισοδύναμα,

$$m'(x) = c - 1,$$

$$m'(t) + 1 = m(t)h(t).$$

Ολοκληρώνοντας και τα δύο μέλη ως προς $x$ στο $(0, t)$, προκύπτει η γραμμική συνάρτηση της $m(t)$:

$$m(t) = (c-1)t + m(0) = (c-1)t + \mu, \qquad \text{όπου } \mu\text{: } \mu\acute{\varepsilon}\sigma\eta\ \tau\iota\mu\acute{\eta}.$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Τότε, από το λήμμα έπεται το συμπέρασμα.

i. Για την Εκθετική, έχουμε $DCRE(X;t) = m(t)$ επομένως ισχύει για $c = 1$.
ii. Για την Pareto, έχουμε $DCRE(X;t) = \frac{a(t+b)}{(a-1)^2} = \frac{\alpha}{\alpha-1} m(t)$. Άρα $c = \frac{\alpha}{\alpha-1} > 0$.

## 6.5 Μονοτονία Δυναμικής Αθροιστικής Υπολειπόμενης Εντροπίας και κλάσεις κατανομών

Έστω οι οικογένειες κατανομών με μονότονη γήρανση (Asadi and Zohrevand, 2007).

**Ορισμός 6.5.1**

Η $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή (ή αντίστοιχα κατανομή $F$) θα λέμε ότι:

I. Έχει την ιδιότητα *IDCRE (Increasing Dynamic Cumulative Entropy)* αν η $DCRE(X;t)$ αύξουσα ως προς $t$, και ισχύει $X \in IDCRE$ (ή αντίστοιχα $F \in IDCRE$).

II. Έχει την ιδιότητα *DDCRE (Decreasing Dynamic Cumulative Entropy)* αν η $DCRE(X;t)$ φθίνουσα ως προς $t$, και ισχύει $X \in DDCRE$ (ή αντίστοιχα $F \in DDCRE$).

**Λήμμα 6.5.1**

Έστω $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή, $DCRE(X;t)$: $DCRE(X;t) < \infty$.

Τότε,

$$DCRE'^{(X;t)} = h(t)[DCRE - m(t)], \qquad (6.19)$$

όπου $h(t)$: συνάρτηση κινδύνου και $m(t)$: μέσος υπολειπόμενος χρόνος ζωής.

Το επόμενο θεώρημα, όπως διατυπώθηκε από τους Asadi and Zohrevand (2007), δίνει την ικανή και αναγκαία συνθήκη ώστε η δυναμική αθροιστική υπολειπόμενη εντροπία να είναι μονότονη συνάρτηση του $t$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Θεώρημα 6.5.1**

Έστω $X$ συνεχής, μη-αρνητική τυχαία μεταβλητή με μέσο υπολειπόμενο χρόνο ζωής $m(t)$ και δυναμική αθροιστική υπολειπόμενη εντροπία $DCRE(X;t)$. Τότε:

I. Αν $m(t)$ αύξουσα συνάρτηση του $t$, τότε επίσης $DCRE(X;t)$ αύξουσα ως προς $t$.
I. Αν $m(t)$ φθίνουσα συνάρτηση του $t$, τότε επίσης $DCRE(X;t)$ φθίνουσα ως προς $t$.

**Απόδειξη:**

Έστω $m(t)$ αύξουσα συνάρτηση του $t$. Τότε, για $x \geq t$ συνεπάγεται

$$m(x) \geq m(t).$$

Επομένως,

$$\int_0^t \frac{m(x)f(t)}{\bar{F}(t)} dx \geq \frac{m(t)\int_0^t f(x)dx}{\bar{F}(t)}.$$

Όμως, $\int_0^t f(x)dx = 1$, οπότε,

$$\int_0^t \frac{m(x)f(t)}{\bar{F}(t)} dx \geq m(t).$$

Το πρώτο μέλος παριστάνει την δεσμευμένη μέση τιμή του υπολειπόμενου χρόνου ζωής της $X$ και επιπλέον έχει αποδειχθεί ότι $DCRE(X;t) = E(m(x)|X > t)$.

Επομένως, $DCRE(X;t) \geq m(t)$.

Άρα, $DCRE(X;t) - m(t) \geq 0, \quad t \geq 0$ και από το λήμμα έπεται ότι:

$$DCRE'(X;t) \geq 0, t \geq 0.$$

Άρα, $DCRE(X;t)$ γνησίως αύξουσα συνάρτηση του $t$. Ανάλογες σχέσεις ισχύουν και στην περίπτωση που η $m(t)$ είναι φθίνουσα ως προς $t$.

Όπως είδαμε σε προηγούμενη ενότητα, οι οικογένειες κατανομών με φθίνουσες (αύξουσες) συναρτήσεις κινδύνου ανήκουν στην οικογένεια κατανομών με αύξουσες (φθίνουσες) συναρτήσεις μέσων υπολειπόμενων χρόνων ζωής, ή ισοδύναμα

$$DFR\ (IFR) \Rightarrow IMRL(DMRL).$$

Το θεώρημα 6.5.1 υποδεικνύει τη διάταξη:

$$\begin{aligned} &DFR \Rightarrow IMRL \Rightarrow IDCRE, \\ &IFR \Rightarrow DMRL \Rightarrow DDCRE. \end{aligned} \tag{6.20}$$

## 6.6 Διπλά Περικομμένη Αθροιστική Υπολειπόμενη Εντροπία

Σε αυτή την ενότητα θα παρουσιαστεί ένα άλλο μέτρο εντροπίας, αυτό της διπλά περικομμένης αθροιστικής υπολειπόμενης εντροπίας *(Doubly truncated (interval) Cumulative Residual Entropy).*

Σε πολλές περιπτώσεις η μόνη πληροφορία που διαθέτουμε είναι μεταξύ δύο σημείων, επομένως τα στατιστικά μέτρα θα πρέπει να μελετηθούν υπό τη συνθήκη διπλά περικομμένων τυχαίων μεταβλητών.

Παραδείγματος χάρη, αν η τυχαία μεταβλητή X εκφράζει την «ηλικία» ενός μέρους, τότε η τυχαία μεταβλητή $X_{t1,t2} = (X - t_1 | t_1 \leq X \leq t_2)$ καλείται διπλά περικομμένη υπολειπόμενη διάρκεια ζωής.

Η τυχαία μεταβλητή $X_t = (X - t | X \geq t)$ αποτελεί ειδική περίπτωση της $X_{t1,t2}$ για $t_2 = \infty$.

Μια επέκταση της εντροπίας του Shannon βασίζεται στη διπλά περικομμένη τυχαία μεταβλητή και ορίζεται ως:

$$H(X; t_1, t_2) = -\int_{t_1}^{t_2} \frac{f(x)}{F(t_2) - F(t_1)} \log \frac{f(x)}{F(t_2) - F(t_1)} dx. \tag{6.21}$$

με $(t_1, t_2) \in D = \{(t_1, t_2) : F(t_1) < F(t_2)\}$ , $H(X; 0, \infty) = H(X)$ η εντροπία του Shannon, $H(X; t_1, \infty)$ η υπολειπόμενη εντροπία $RE(X; t)$ και $H(X; 0, t_2)$ η παρελθοντική εντροπία $PE(X; t)$.

Κατά αντιστοιχία θα παρουσιαστεί το μέτρο της Διπλά Περικομμένης Αθροιστικής Υπολειπόμενης Εντροπίας, όπως ορίστηκε και μελετήθηκε από τους M.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Khorashadizadeh et al. (2013). Αρχικά θα εξεταστεί ως προς τη σχέση του με τα μέτρα αξιοπιστίας (συνάρτηση κινδύνου, μέσος υπολειπόμενος χρόνος ζωής) και τέλος θα οριστούν δύο νέες κλάσεις κατανομών βάσει της μονοτονίας του μέτρου.

Να σημειωθεί πως το μέτρο διατηρεί όλες τις ιδιότητες της *CRE.*

**Ορισμός 6.6.1**

Έστω $X$ απόλυτα συνεχής τυχαία μεταβλητή και έστω $D = \{(x, y): F(x) < F(y)\}$. Τότε η διπλά περικομμένη αθροιστική υπολειπόμενη εντροπία *(Doubly truncated (interval) Cumulative Residual entropy –ICRE)* δίνεται από τη σχέση:

$$\mathrm{ICRE}(X; t_1, t_2) = -\int_{t_1}^{t_2} \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \log \frac{\bar{F}(x)}{\bar{F}(t_1) - \bar{F}(t_2)}\, dx, \quad (t_1, t_2) \in D$$

$$= -\frac{1}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} \bar{F}(x) \log \bar{F}(x)\, dx$$

$$+\log\big(\bar{F}(t_1) - \bar{F}(t_2)\big)\left[\mu(t_1, t_2) + (t_2 - t_1)\frac{\bar{F}(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)}\right], \quad (6.22)$$

όπου $m(t_1, t_2) = E(X - t_1 | t_1 \leq X \leq t_2)$ η διπλά περικομμένη μέση υπολειπόμενη διάρκεια ζωής.

Είναι προφανές ότι $\mathrm{ICRE}(X; 0, \infty) = \mathrm{CRE}(X)$ και $\mathrm{ICRE}(X; t_1, \infty)$ =δυναμική $\mathrm{DCRE}(X; t_1)$.

## 6.7 Μονοτονία της Διπλά Περικομμένης Αθροιστικής Υπολειπόμενης Εντροπίας και κλάσεις κατανομών

**Ορισμός 6.7.1**

Η τυχαία μεταβλητή $X$ έχει την ιδιότητα $DICRE$ (*Decreasing Interval Cumulative*



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

*Residual Entropy)* αν και μόνο αν ∀ σταθερό $t_2$ η $ICRE(X;t_1,t_2)$ φθίνουσα ως προς $t_1$.

Το θεώρημα που ακολουθεί αποδεικνύει ότι δεν υπάρχει καμία μη-αρνητική, συνεχής, τυχαία μεταβλητή $X$ με αύξουσα $ICRE$ ($IICRE$) στο [0,∞).

**Θεώρημα 6.7.1**

Έστω $X$ μη αρνητική, μη εκφυλισμένη τυχαία μεταβλητή. Τότε η $ICRE(X;t_1,t_2)$ δε μπορεί να είναι αύξουσα συνάρτηση ως προς $t_1$, για οποιοδήποτε σταθερό $t_2$.

**Απόδειξη:**

Με χρήση του Νόμου De l'Hopital έχουμε:

$$\lim_{t_1\to t_2} ICRE(X;t_1,t_2)$$

$$= \lim_{t_1\to t_2}\left\{\frac{-\int_{t_1}^{t_2}\bar{F}(x)\ln\bar{F}(x)\,dx + \ln[\bar{F}(t_1)-\bar{F}(t_2)]\int_{t_1}^{t_2}\bar{F}(x)dx}{\bar{F}(t_1)-\bar{F}(t_2)}\right\}$$

$$= \lim_{t_1\to t_2}\left\{\frac{\bar{F}(t_1)\ln\left(1-\frac{\bar{F}(t_2)}{\bar{F}(t_1)}\right)+\frac{f(t_1)}{\bar{F}(t_1)-\bar{F}(t_2)}\int_{t_1}^{t_2}\bar{F}(x)dx}{\bar{F}(t_1)-\bar{F}(t_2)}\right\} = -\infty.$$

Αντίστροφα, έστω ότι η $ICRE(X;t_1,t_2)$ είναι αύξουσα ως προς $t_1$, τότε $\forall t_1 \le t_2$,

$ICRE(X;t_1,t_2) \le ICRE(X;t_2,t_2) = -\infty$, άτοπο, καθώς εξ ορισμού ισχύει ότι:

$ICRE(X;t_1,t_2) \in \mathbb{R}, \forall (t_1,t_2) \in D$.

**Παρατήρηση:**

Στην ειδική περίπτωση που $ICRE(X;t_1,\infty) = DCRE(X;t_1)$, προφανώς οδηγούμαστε στις κλάσεις κατανομών της *DCRE*, *(DDCRE, IDCRE)* (Asadi and Zohrevand, 2007).

Στο Θεώρημα που ακολουθεί δίνεται άνω φράγμα για την $ICRE(X;t_2,t_2)$ όταν η κατανομή έχει την $DICRE$ ιδιότητα.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Θεώρημα 6.7.2**

Η τυχαία μεταβλητή $X$ έχει την $DICRE$ ιδιότητα αν και μόνο αν ικανοποιείται η παρακάτω ανισότητα $\forall (t_1, t_2) \in D$.

$$ICRE(X; t_1, t_2) \leq \mu(t_1, t_2) + \frac{(t_2 - t_1)\bar{F}(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)} - \log\left(\frac{\bar{F}(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)}\right)^{\frac{1}{h(t_1)}}. \qquad (6.23)$$

**Απόδειξη:**

Παίρνοντας της μερική παράγωγο της $ICRE(X; t_1, t_2)$ ως προς $t_1$ έχουμε:

$$\frac{\partial}{\partial t_1} ICRE(X; t_1, t_2) = -\frac{f(t_1)}{\left(\bar{F}(t_1) - \bar{F}(t_2)\right)^2} \int_{t_1}^{t_2} \bar{F}(x) \log \bar{F}(x) dx + \frac{\bar{F}(t_1) \log \bar{F}(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)}$$

$$-\frac{f(t_1)}{\left(\bar{F}(t_1) - \bar{F}(t_2)\right)^2} \left[ m(t_1, t_2) + (t_2 - t_1) \frac{\bar{F}(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)} \right]$$

$$+\left( \frac{\partial}{\partial t_1} m(t_1, t_2) - \frac{\bar{F}(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)} + \frac{(t_2 - t_1)\bar{F}(t_2) f(t_1)}{\left(\bar{F}(t_1) - \bar{F}(t_2)\right)^2} \right) \log[\bar{F}(t_1) - \bar{F}(t_2)].$$

Οι Navarro and Ruiz (1996) όρισαν τις γενικευμένες συναρτήσεις κινδύνου (Generalized Failure Rate-GFR) οι οποίες δίνονται στις σχέσεις που ακολουθούν.

$$h_1(t_1, t_2) = \lim_{h \to 0^+} \left[ \frac{P(t_1 \leq X \leq t_1 + h | t_1 \leq X \leq t_2)}{h} \right] = \frac{f(t_1)}{\left(F(t_2) - F(t_1)\right)}, \qquad (6.24)$$

$$h_2(t_1, t_2) = \lim_{h \to 0^-} \left[ \frac{P(t_2 + h \leq X \leq t_2 | t_1 \leq X \leq t_2)}{h} \right] = \frac{f(t_2)}{\left(F(t_2) - F(t_1)\right)} \qquad (6.25)$$

Οι αντίστοιχες σχέσεις που τις συνδέουν με τους μέσους υπολειπόμενους χρόνους ζωής $m(t_1, t_2), m^*(t_1, t_2)$ είναι:

$$h_1(t_1, t_2) = \frac{1 + \frac{\partial}{\partial t_1} m(t_1, t_2)}{\mu(t_1, t_2)}, \qquad (6.26)$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$h_2(t_1,t_2) \ = \frac{1+\frac{\partial}{\partial t_2}\mu^*(t_1,t_2)}{\mu^*(t_1,t_2)}. \tag{6.27}$$

Έτσι, μετά από αντικατάσταση των εξισώσεων (6.26), (6.27) και μετά από απλοποίηση της εξίσωσης προκύπτει:

$$\frac{\partial}{\partial t_1} ICRE(X;t_1,t_2) = h_1(t_1,t_2)\Bigg[ICRE(X;t_1,t_2) - m(t_1,t_2) - \frac{(t_2-t_1)\bar{F}(t_2)}{\bar{F}(t_1)-\bar{F}(t_2)} + \log\left(\frac{\bar{F}(t_1)}{\bar{F}(t_1)-\bar{F}(t_2)}\right)^{\frac{1}{h(t_1)}}\Bigg].$$

**Πόρισμα 6.7.1**

Για την ειδική περίπτωση $t_2 \to \infty$ του παραπάνω θεωρήματος ισχύει ότι:

i. $DCRE(X;t)$ είναι φθίνουσα ως προς $t\,(DDCRE(X;t))$ αν και μόνο αν $DCRE(X;t) \le m(t), \forall t$.

ii. $DCRE(X;t)$ είναι αύξουσα ως προς $t\,(DDCRE(X;t))$ αν και μόνο αν $DCRE(X;t) \ge m(t), \forall t$.

**Απόδειξη:**

Χρησιμοποιώντας τη σχέση $h(t) = \frac{1+m'(t)}{m(t)}$ και απλοποιώντας την εξίσωση μπορούμε να γράψουμε:

$$\frac{\partial}{\partial t} DCRE(X;t) = h(t)[DCRE(X;t) - m(t)],$$

το οποίο αποδεικνύει το πόρισμα.

**Παράδειγμα:**

**Ομοιόμορφη κατανομή**

Έστω ότι η $X$ ακολουθεί την ομοιόμορφη κατανομή στο διάστημα $(0,b), b \ge 0$. Τότε,

$$ICRE(X;t_1,t_2) = \frac{2(t_1-b)^2\log\left(\frac{t_1-b}{t_1-t_2}\right)-2(t_2-b)^2\log\left(\frac{t_2-b}{t_1-t_2}\right)-(t_1-t_2)(t_1-2b+t_2)}{4(t_1-t_2)}. \qquad (6.28)$$

**Θεώρημα 6.7.3**

Έστω $X$ μη αρνητική, συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x)$ και συνάρτηση κατανομής $F(x)$. Τότε,

$$ICRE(X;t_1,t_2) \geq Ce^{H(X;t_1,t_2)}, \qquad (6.29)$$

όπου $C = exp\left\{\int_{p_1}^{p_2}\log(x|\log x|)dx\right\}$ και $p_i = \frac{\bar{F}(t_i)}{\bar{F}(t_1)-\bar{F}(t_2)}$, i = 1,2.

**Απόδειξη:**

Έστω $P(X), Q(X)$ δύο οποιεσδήποτε θετικές συναρτήσεις. Τότε από την ιδιότητα Log-Sum προκύπτει:

$$\int P(X)\log\frac{P(X)}{Q(X)}dx \geq \left(\int P(X)dx\right)\log\left(\frac{\int P(X)dx}{\int Q(X)dx}\right).$$

Έτσι, αν θεωρήσουμε $P(X) = \frac{f(x)}{\bar{F}(t_1)-\bar{F}(t_2)}$ και $Q(X) = \frac{\bar{F}(x)}{\bar{F}(t_1)-\bar{F}(t_2)}\log\left|\frac{\bar{F}(x)}{\bar{F}(t_1)-\bar{F}(t_2)}\right|$ και ολοκληρώσουμε από $t_1$ ως $t_2$ προκύπτει το ζητούμενο.

Στο θεώρημα που ακολουθεί, βάσει της συνάρτησης $ICRE$, δίνεται ένα άνω φράγμα της διαφοράς δύο ανεξάρτητων, ισόνομα κατανεμημένων τυχαίων μεταβλητών.

**Θεώρημα 6.7.4**

Έστω *X, Y* δύο μη αρνητικές, συνεχείς τυχαίες μεταβλητές, οι οποίες είναι ανεξάρτητες και ακολουθούν την ίδια κατανομή.

Τότε,

$$E(|X-E(X)||t_1<X<t_2) \leq E(|X-Y||t_1<X<t_2, t_1<Y<t_2)$$



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

$$\leq \frac{2ICRE(X;t_1,t_2)}{\bar{F}(t_1)-\bar{F}(t_2)} - 2\frac{\log\left(\bar{F}(t_1)-\bar{F}(t_2)\right)}{\bar{F}(t_1)-\bar{F}(t_2)}\left[\mu(t_1,t_2)+(t_2-t_1)\frac{\bar{F}(t_2)}{\bar{F}(t_1)-\bar{F}(t_2)}\right].$$

(6.30)

**Απόδειξη:**

Για δύο ανεξάρτητες και ισόνομα κατανεμημένες τυχαίες μεταβλητές $X, Y$ έχουμε:

$$P\left((X>z)\cup(Y>z)|t_1<X<t_2, t_1<Y<t_2\right) = P\left((X>z)|t_1<X<t_2\right)$$

$$+P\left((Y\ >z)|t_1<X<t_2\right) - P\left((X>z)\cap(Y>z)|t_1<X<t_2, t_1<Y<t_2\right)$$

$$= 2\frac{\bar{F}(Z)}{\bar{F}(t_1)-\bar{F}(t_2)} - \left(\frac{\bar{F}(Z)}{\bar{F}(t_1)-\bar{F}(t_2)}\right)^2.$$

Επομένως,

$$2\frac{\bar{F}(z)}{\bar{F}(t_1)-\bar{F}(t_2)} - \left(\frac{\bar{F}(z)}{\bar{F}(t_1)-\bar{F}(t_2)}\right)^2 = P(\max(X,Y)>z|t_1<X<t_2, t_1<Y<t_2)$$

$$-P(\min(X,Y)>z\,|t_1<X<t_2, t_1<Y<t_2).$$

Ολοκληρώνοντας και τα δύο μέρη από $t_1$ έως $t_2$ έχουμε:

$$\frac{2}{\left(\bar{F}(t_1)-\bar{F}(t_2)\right)^2}\int_{t_1}^{t_2}\bar{F}(z)\left[\left(\bar{F}(t_1)-\bar{F}(t_2)\right)-\bar{F}(z)\right]dz = E(|X-Y||t_1<X,Y<t_2).$$

Γνωρίζουμε πως η ανισότητα $x(a-x)\leq x|\log x|$, ισχύει $\forall x\in(0,1)$ και $a\in(0,1)$.

Επομένως,

$$\frac{2}{\left(\bar{F}(t_1)-\bar{F}(t_2)\right)^2}\int_{t_1}^{t_2}\bar{F}(z)\left[\left(\bar{F}(t_1)-\bar{F}(t_2)\right)-\bar{F}(z)\right]dz\ \leq \frac{2}{\left(\bar{F}(t_1)-\bar{F}(t_2)\right)^2}\int_{t_1}^{t_2}\bar{F}(z)|\log\bar{F}(z)|$$

Άρα,

$$E(|X-Y||t_1<X,Y<t_2) = \int_{t_1}^{t_2}E(|X-u||t_1<X<t_2)\,f(u)du$$

$$\geq \int_{t_1}^{t_2}|u-E(X)|\,f(u)du$$

$$= E(|X-E(X)||t_1<X<t_2).$$

**Πόρισμα 6.7.2**

Για τις ειδικές περιπτώσεις που $t_1 = 0$ και $t_2 = \infty$ από το παραπάνω θεώρημα προκύπτει:

$E(|X - E(X)|) \leq E(|X - Y|) \leq 2CRE(X)$ (Rao, 2005). (6.31)

**Παρατήρηση:**

Έστω $X, Y$ δύο μη αρνητικές, συνεχείς τυχαίες μεταβλητές, οι οποίες είναι ανεξάρτητες και ακολουθούν την ίδια κατανομή.

Τότε,

$$E(|X - E(X)||t_1 < X < t_2)$$

$$\leq E(|X - Y||t_1 < X < t_2, t_1 < Y < t_2)$$

$$\leq \frac{2ICPE(X; t_1, t_2)}{F(t_2) - F(t_1)} - 2\frac{\log\big(F(t_2) - F(t_1)\big)}{F(t_2) - F(t_1)}\left[\mu^*(t_1, t_2) + (t_2 - t_1)\frac{F(t_1)}{F(t_2) - F(t_1)}\right].$$

Για την ειδική περίπτωση: $t_1 = 0$ και $t_2 = \infty$ προκύπτει:

$$E(|X - E(X)|) \leq E(|X - Y|) \leq 2CPE(X). \qquad (6.32)$$

Οι Asadi and Zohrevand (2007) απέδειξαν, όπως είδαμε σε προηγούμενη ενότητα, ότι για μια απολύτως συνεχή κατανομή η $DCRE(X; t) = E(\mu(X)|X \geq t)$.

Στο θεώρημα που ακολουθεί παρουσιάζεται η επέκταση για την διπλά περικομμένη περίπτωση.

**Θεώρημα 6.7.5**

Έστω $F$ απολύτως συνεχής συνάρτηση κατανομής με $ICRE(X; t_1, t_2) < \infty$.

Τότε,

$$E(m(X)|t_1 \leq X \leq t_2) \geq ICRE(X; t_1, t_2). \qquad (6.33)$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**Απόδειξη:**

Γνωρίζουμε πως $m(X) = \frac{\int_x^\infty \bar{F}(t)dt}{\bar{F}(x)}$, επομένως,

$$E(m(X)|t_1 \leq X \leq t_2) = \int_{t_1}^{t_2} \left( \int_u^\infty \bar{F}(z)dz \right) \frac{f(u)}{\bar{F}(u)[\bar{F}(t_1) - \bar{F}(t_2)]} du$$

$$= \int_{t_1}^{t_2} \left( \int_{t_1}^{z} \frac{f(u)}{\bar{F}(u)} du \right) \frac{\bar{F}(z)}{[\bar{F}(t_1) - \bar{F}(t_2)]} dz$$

$$+ \int_{t_2}^{\infty} \left( \int_{t_1}^{t_2} \frac{f(u)}{\bar{F}(u)} du \right) \frac{\bar{F}(z)}{[\bar{F}(t_1) - \bar{F}(t_2)]} dz.$$

Χρησιμοποιώντας τη σχέση $\int_{t_1}^{t_2} h(u)du = -\log \frac{\bar{F}(t_2)}{\bar{F}(t_1)}$ προκύπτει ότι:

$$E(m(X)|t_1 \leq X \leq t_2) = -\int_{t_1}^{t_2} \frac{\bar{F}(z)}{[\bar{F}(t_1) - \bar{F}(t_2)]} \log \frac{\bar{F}(z)}{\bar{F}(t_1)} dz - \log \frac{\bar{F}(t_2)}{\bar{F}(t_1)} \int_{t_2}^{\infty} \frac{\bar{F}(z)}{[\bar{F}(t_1) - \bar{F}(t_2)]} dz$$

$$= -\int_{t_1}^{t_2} \frac{\bar{F}(z)}{[\bar{F}(t_1) - \bar{F}(t_2)]} \log \bar{F}(z)\, dz + \log \bar{F}(t_1) \int_{t_1}^{\infty} \frac{\bar{F}(z)}{[\bar{F}(t_1) - \bar{F}(t_2)]} dz$$

$$+ \log \bar{F}(t_2) \int_{t_2}^{\infty} \frac{\bar{F}(z)}{[\bar{F}(t_1) - \bar{F}(t_2)]} dz.$$

Επομένως, το ζητούμενο αποτέλεσμα θα ικανοποιηθεί χρησιμοποιώντας τις ανισότητες $-\log \bar{F}(t_1) \leq -\log \bar{F}(t_2)$ και $\log[\bar{F}(t_1) - \bar{F}(t_2)] \leq \log \bar{F}(t_1)$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Κεφάλαιο 7ο

## Σύνοψη

Στο πρώτο κεφάλαιο παρατέθηκαν Αναλογιστικά ζητήματα στα οποία βρίσκουν εφαρμογή τα μέτρα αβεβαιότητας εντροπίας.

Κατόπιν στο δεύτερο παρουσιάστηκε η εφαρμογή του μέτρου του Shannon και η εφαρμογή σε Μοντέλα Απώλειας υπό επίδραση πληθωρισμού. Είδαμε ότι το φαινόμενο του πληθωρισμού συνεπάγεται και αύξηση αβεβαιότητας.

Στο επόμενο παρουσιάστηκε το μέτρο Kullback–Leibler, απεδείχθη η ισοδύναμη σχέση του με την Αρχή Μέγιστης Εντροπίας και παρουσιάστηκε εφαρμογή αυτού σε προβλήματα υποθέσεων θνησιμότητας.

Στο τέταρτο κεφάλαιο εξετάστηκε το μέτρο Υπολειπόμενης εντροπίας και παρουσιάστηκε η εφαρμογή του σε Μοντέλα Απώλειας με επίδραση αφαιρετέου ποσού.

Παρατηρήθηκαν τα εξής:

- Η εντροπία κόστους ανά πληρωμή δεν επηρεάζει την εκθετική κατανομή. Ισοδύναμα, σε σταθερό ρίσκο η αβεβαιότητα δεν εξαρτάται από την απαλλαγή.
- Από τη συσχέτιση των μέτρων συνάρτησης κινδύνου και μέσου υπολειπόμενου χρόνου ζωής με την υπολειπόμενη εντροπία κόστους ανά πληρωμή, είδαμε πως αν η αβεβαιότητα αυξάνει (μειώνεται) ως προς την απαλλαγή τότε αυξάνεται (μειώνεται ) και το μέσο υπολειπόμενο κόστος.
- Για δείγματα ζημιών που παρουσιάζουν αύξουσα αβεβαιότητα κόστους ανά πληρωμή ως προς κάποιο ποσό απαλλαγής, μπορεί να προσδιοριστεί μονοσήμαντα η κατανομή του δείγματος.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Στο κεφάλαιο πέντε εξετάστηκε το μέτρο Παρελθοντικής Εντροπίας και η εφαρμογή του σε Μοντέλα Απώλειας υπό επίδραση ιδίας κράτησης. Αναφέρθηκε πως για δείγμα ζημιών με φθίνουσα αντίστροφη συνάρτηση κινδύνου, η αβεβαιότητα κόστους ανά πληρωμή είναι αύξουσα ως προς το όριο ιδίας κράτησης.

Κατόπιν μελετήθηκε η εφαρμογή των μέτρων Υπολειπόμενης και Παρελθοντικής Εντροπίας σε Μοντέλα Απώλειας και στην Ανάλυση Επιβίωσης.

Τέλος, εξετάστηκαν μέτρα Αθροιστικών Υπολειπόμενων εντροπιών και οι σχέσεις που τις συνδέουν με τη συνάρτηση κινδύνου και το μέσο υπολειπόμενο χρόνο, και αναφέρθηκαν οικογένειες κατανομών που προκύπτουν από τη μονοτονία των μέτρων.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

## Παράρτημα Κώδικα R

*Σχήμα 2: Γραφική απεικόνιση της εντροπίας μιας Bernoulli κατανομής*

```
library(lattice)

distribution(x)<-function(x) {

-x*log(x)-(1-x)*log(1-x)}

X<-seq(0,1,0.00001)

xyplot(distribution(X)~X,type="l", col = "blue", cex.lab = 3.5, cex.axis = 3.5, lwd=2,main="Entropy of Bernoulli Process",ylab="H(X)", xlab="p(x)" )
```

*Σχήμα 2: Γραφική απεικόνιση ιστογράμματος και πυκνότητας ζημιών παραδείγματος (Ενότητα 4.5)*

```
hist(LOSSES,50,freq=FALSE,main="CLAIMS FOR THE YEAR 2018")

lines(density.default(LOSSES))
```

*Εκτίμηση κατανομής του δείγματος ζημιών*

```
a<-
read.table("C:/Users/ALEXANDRATIPAPIDI/Desktop/zimies.txt
"); a

attach(a)

library(moments)

library(MDM)

library(MASS)

library(survival)

library(fitdistrplus)
```



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

```
library(logspline)

library(actuar)

summary(LOSSES)

summary(fitg)

plot(fitg)

plot(fitg, demp = TRUE)

plot(fitg, histo = FALSE, demp = TRUE)

cdfcomp(fitg, addlegend=FALSE)

denscomp(fitg, addlegend=FALSE)

ppcomp(fitg, addlegend=FALSE)

qqcomp(fitg, addlegend=FALSE)

fitg <- fitdist(LOSSES, "weibull")

fitg
```

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

# Ελληνική Βιβλιογραφία

Σαχλάς Π. Α., (2010). Θέματα Στατιστικής Θεωρίας Πληροφοριών και Αναλογισμού, Παν. Πειραιώς, τμ. Στατιστικής και Ασφαλιστικής Επιστήμης.

Λιάβας Π. Α., (2012). Στοιχεία Θεωρίας Πληροφορίας, Σημειώσεις Μαθήματος. Πολυτεχνείο Κρήτης, τμ. ΗΜΜΥ.

# Ξένη Βιβλιογραφία

Abramowitz, M. and Stegun, I.A. (1972). Handbook of Mathematical Functions. National Bureau of Standards, Applied Mathematics Series, No. 55.

Asadi, M., Ebrahimi, N., Hamedani, G., Soofi E.S. (2004). Maximum dynamic entropy models. Journal of Applied Probability 41:379-390.

Asadi, M. and Zohrevand, Y. (2007). On the Dynamic Cumulative Residual Entropy. Journal of Statistical Planning and Inference 137, 1931-1941.

Asadi, M. and Ebrahimi, N. (2000). Residual entropy and its characterizations in terms of hazard function and mean residual life function. Statistics and Probability Letters 49, 263-269.

Asha, G. and Rejeesh, C.J. (2015). Characterizations using past entropy measures, Metron 73, 119-134.

Asok K. Nanda and Prasanta P. (2006). Some properties on past entropy and their applications, Metrika 64, 47-61.

Belunze, F., Navarro. J., Ruiz J.M. and Del Aguila, Y. (2004). Some Results on Residual Entropy Function. Metrika 59, 147-161.

Bernardo, J.M. and Rueda, R. (2002). Bayesian Hypothesis Testing: A Reference Approach. International Statistical Review 70(3):351-370.

Block, H.W., Borges, W.S., Savits T.H. (1985). Age-dependent minimal repair. Journal of Applied Probability 22, 370-385.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Brockett, P.L. (1991). Information Theoretic Approach to Actuarial Science: A Unification and Extension of Relevant Theory and Applications. Transactions of Society of Actuaries 1991, 43, 73-114.

Brockett, P.L. and Cox, S.H. (1984). Statistical Adjustment of Mortality Tables to Reflect Known Information. TSA, 36, 63-71.

Cover, T.M. and Thomas, J.A. (1991). Elements of Information Theory. Wiley, New York.

Cox, D.R. (1972). Regression Models and Life-Tables. Journal of the Royal Statistical Society, Series B, 34(2), 187-220.

Cox, S.H. (1991). Bounds on expected values of insurance payments and option prices. TSA 43, 231-260.

Di Crescenzo, A. (2000). Some results on the proportional reversed hazards model. Statistics and Probability Letters, 50(4), 313-321.

Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. Journal of Applied Probability 39, 434-440.

Di Crescenzo, A. and Longobardi, M. (2009). On cumulative entropies. Journal of Statistical Planning and Inference 139, 4072-4087.

Ebrahimi, N. (1996). How to measure uncertainty in the residual life distributions. Sankhya 58:48-57.

Ebrahimi, N. and Kirmani, S.N.U.A. (1996). Some results on ordering of survival functions through uncertainty. Statistics and Probability Letters 29, 167-176.

Ebrahimi, N. and Pellerey, F. (1995). New partial ordering of survival functions based on notion of uncertainty. Journal of Applied Probability 32(1), 202-211.

Ebrahimi, N., Kirmani, S.N.U.A. and Soofi E.S. (2007). Multivariate dynamic information. Journal of Multivariate Analysis 98:328-349.

Gupta, R.C., Gupta, R.D. (2007). Proportional reversed hazard rate model and its applications. Journal of Statistical Planning and Inference 137:3525-3536.

Hartley, R.V.L. (1928). Transmission of information. Bell System Technical Journal 7, 535-563.

Jeffreys, H. (1948). Theory of probability. Oxford University Press.

Khorashadizadeh, M., Roknabadi, A.H.R. and Borzadaran, G.R.M. (2013). Doubly truncated interval) cumulative residual and past entropy. Statistics and Probability Letters 83, 1464-1471.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Kluggman, A.S., Panjer, H.H., Willmot. E.G. (2008). Loss models, from data to decisions, 3rd edition. Wiley, New York.

Kullback, S.P. (1959). Information Theory and Statistics. Wiley, New York.

Lai, C.D. and Xie, M. (2006). Stochastic Ageing and Dependence for Reliability. Springer, New York.

Liu, J. (2007). Information Theoretic Content and Probability. Ph.D. Thesis, University of Florida, USA.

Nair, C., Prabhakar, B., Shah, D. (2006). On entropy for mixtures of discrete and continuous variables. Computing Research Repository (CoRR), abs/cs/0607075.

Nair, K.R.M. and Rajesh, G. (1998). Characterization of Probability distributions using the residual entropy function. Journal of Indian Statistical Association 36:157-166.

Navarro, J. and Ruiz, J.M. (1996). Failure-rate functions for doubly-truncated random variables. IEEE Transactions on reliability 45(4): 685-690.

Rao, M. (2005). More on a new concept of entropy and information. Journal of Theoretical Probability 18, 967-981.

Rao, M., Chen, Y., Vemuri, B.C. and Wang, F. (2004). Cumulative residual entropy: A New Measure of Information. IEEE Transaction on Information Theory 50, 1220-1228.

Sankaran, P.G., Gleeja, C.L. (2008). Proportional reversed hazard and frailty models. Metrika 68:333-342.

Shannon, C.E. (1948). A mathematical theory of communication. Bell System Technical Journal 27:379-423.

Schlas, A. and Papaioannou, T. (2012). Residual and Past Entropy in Actuarial Science and Survival Models. Methodology and Computing in Applied Probability 16, 79-99 (2014).

Soofi, E.S. (1994). Capturing the Intangible Concept of Information. Journal of the American Statistical Association, Vol. 89, No. 428, 1244-1254.

Vonta F. (1996). Efficient estimation in a nonproportional hazards model in survival analysis. Scandinavian Journal of Statistics 23:49-62.

Vonta F., Karagrigoriou A. (2010). Generalized measures of divergence in survival analysis and reliability. Journal of Applied Probability 47(1):216-234.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Zografos, K. (2008). On some entropy and divergence type measures of variability and dependence for mixed continuous and discrete variables. Journal of Statistical Planning and Inference 138, 3899-3914.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ