15

ΤΜΗΜΑ ΔΗΜΟΣΙΑΣ ΔΙΟΙΚΗΣΗΣ
ΟΙΚΟΝΟΜΙΚΑ ΤΗΣ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΤΩΝ ΔΙΑΚΛΑΔΙΚΩΝ ΣΧΕΣΕΩΝ
ΠΑΝΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΟΙΝΩΝΙΚΩΝ & ΠΟΛΙΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Η σημασία της Εντροπίας στη Χρηματοοικονομική Επιστήμη

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Γεώργιος Ι. Τσώνης

Ηλεκτρολόγος Μηχανικός
& Μηχανικός Υπολογιστών - Ε.Μ.Π.

Επιβλέπων : Κώστας Συριόπουλος
Καθηγητής

Αθήνα, Ιούνιος 2008.

....

Γεώργιος Ι. Τσώνης
Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Περίληψη

Η εντροπία υπάρχει παντού, από την Φυσική μέχρι την καθημερινότητα του ανθρώπου, και δείχνει το μέτρο της αταξίας σ'ένα σύστημα. Ως έννοια θεμελιώνεται μέσα από τη θεωρία Πληροφορίας. Σ'αυτή την εργασία σκιαγραφείται η σημασία της στην ΧρηματοΟικονομική Επιστήμη. Πιο συγκεκριμένα, η λεγόμενη «αυξητική» εντροπία χρησιμοποιείται για να προβλεφθεί η μέγιστη απόδοση ενός χαρτοφυλακίου.

λέξεις κλειδιά: εντροπία, θεωρία πληροφορίας, χαρτοφυλάκιο, απόδοση, πρόβλεψη

Abstract

Entropy is everywhere, from Physics to the daily life, and indicates the level of disorder in a system. As a concept it derives from Information Theory. In this dissertation the importance of Entropy in Finance is highlighted. Specifically, incremental entropy is used in order to forecast the optimal return of a portfolio.

keywords: entropy, information theory, portfolio, return, forecast

Ευχαριστίες

Στο σημείο αυτό θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή κ. Κωνσταντίνο Συριόπουλο. Επίσης τον υποψήφιο διδάκτορα Διονύση Φίλιππα για τη βοήθεια και την αμέριστη συμπαράστασή του. Ακόμη και όλους τους συμφοιτητές μου, όπως τη Βιολέτα Παππά, το Γιάννη Γεωργόπουλο, τη Λίνα Μουστάκα, τον Παναγιώτη Μπατσινίλα, το Θέμη Ιωάννου, το Νίκο Λυβίτση, το Γιώργο Νίνο, τον Κλεάνθη Λιάμη, τη Σταυρούλα Μεϊμετέα, τον Παναγιώτη Μελαχροινό, τη Ματίνα Σφυρή, την Ηλιάνα Μύταλα, τη Βικτώρια Τρενεβά, τη Δήμητρα Αβδάλα, την Αγγελική Δεμερτζή, τη Δήμητρα Σκουρλή για το πολύ ωραίο κλίμα που είχαμε καθ'όλη τη διάρκεια του ΠΜΣ. Τέλος θα ήταν παράλειψη να μην αναφερθώ ιδιαίτερα στον Νίκο Πουγούνια για την σημαντική, πολυποίκιλη και ουσιαστική υποστήριξη που μου παρείχε.

Πίνακας Περιεχομένων

Πρόλογος

Η εργασία αυτή απευθύνεται κατά κύριο λόγο σε φοιτητές οικονομικών επιστημών αλλά και σε όσους ενδιαφέρονται για τα χρηματο-οικονομικά. Σκοπός της είναι να αποτελέσει μια γερή εισαγωγή στην υλοποίηση μιας νέας μεθόδου για την πρόβλεψη της μέγιστης απόδοσης ενός χαρτοφυλακίου.

Υπάρχει η αδήριτη ανάγκη για να βρεθεί ένας νέος τρόπος υπολογισμού της μέγιστης απόδοσης ενός χαρτοφυλακίου. Κι αυτό γιατί τα χρηματοοικονομικά παίζουν εξέχοντα ρόλο στο σύγχρονο παγκοσμιοποιημένο περιβάλλον. Αυτή η νέα μέθοδος σχετίζεται με την έννοια της **Εντροπίας**, η οποία έλκει την καταγωγή της από την Θεωρία Πληροφορίας.

Αρχικά παρουσιάζεται μια εισαγωγή στην έννοια της εντροπίας μέσα από το πρίσμα της Θεωρίας Πληροφορίας (Κεφάλαιο 1). Έπειτα αναλύεται η θεωρία χαρτοφυλακίου του Markowitz (Κεφάλαιο 2). Κατόπιν εξηγείται η νέα θεωρία του χαρτοφυλακίου, η οποία βασίζεται στην εντροπία (Κεφάλαιο 3). Τέλος παρουσιάζονται κάποιες ερευνητικές προεκτάσεις (Κεφάλαιο 4).

Κεφάλαιο 1

Εντροπία

Εισαγωγή

Στο κεφάλαιο αυτό αναφερόμαστε στην Εντροπία μέσα από τη Θεωρία Πληροφορίας. Πιο συγκεκριμένα στηριζόμαστε στο έργο του Shannon για να την ορίσουμε και στη συνέχεια αναλύουμε διάφορες εκφάνσεις της.

Θεωρία Πληροφορίας

Η θεωρία πληροφορίας είναι η επιστήμη η οποία ασχολείται με την έννοια της «πληροφορίας» ,τη μέτρηση της και τις εφαρμογές της. Η πρώτη προσπάθεια για να περιγραφεί με μαθηματικό τρόπο η μετάδοση των πληροφοριών έγινε από τον Claude E. Shannon, ο οποίος δημοσίευσε την εργασία του 'A mathematical theory of communication' to 1948 και γενικά θεωρείται ο ιδρυτής της θεωρίας πληροφοριών. Όμως ένας αριθμός επιστημόνων πριν από το Shannon προσπάθησαν να διατυπώσουν την αποδοτική χρήση των συστημάτων επικοινωνιών.

Το 1924 ο H.Nyquist δημοσίευσε μια εργασία στην οποία περιέγραψε τον τρόπο με τον οποίο μπορούσαν τα μηνύματα (ή χαρακτήρες για να χρησιμοποιήσουμε τις λέξεις του) να σταλούν με ένα τηλέγραφο με τη μέγιστη ταχύτητα , αλλά χωρίς παραμόρφωση. Εντούτοις, στην περιγραφή του αυτή δε χρησιμοποίησε τον όρο πληροφορία.

Ο R.V.L. Hartley το 1928 ήταν ο πρώτος που προσπάθησε να ορίσει ένα μέτρο πληροφορίας με τον παρακάτω τρόπο. Υποθέτουμε ότι για κάθε σύμβολο ενός μηνύματος έχουμε μια επιλογή από s δυνατότητες. Αλλά θεωρώντας τώρα μηνύματα l συμβόλων μπορούμε να πάρουμε s^l διακεκριμένα μηνύματα. Ο Hartley όρισε το ποσό πληροφορίας σαν το λογάριθμο του αριθμού των διακεκριμένων μηνυμάτων. Στην περίπτωση των μηνυμάτων μήκους l συνεπώς έχουμε

$$\mathrm{H_H}(s^l) = \log(s^l) = l * \log(s)$$

Για μηνύματα μήκους 1 βρίσκουμε

$$H_H(s^1) = \log(s)$$

Και συνεπώς

$$H_H(s^l)=l*H_H(s^1)$$

Αυτό αντιστοιχεί με τη διαίσθηση ότι ένα μήνυμα που αποτελείται από l σύμβολα, περιέχει l φορές περισσότερη πληροφορία από ένα μήνυμα που αποτελείται από ένα μόνο σύμβολο. Σε αυτό επίσης οφείλεται η εμφάνιση του λογαρίθμου στον ορισμό του Hartley. Μπορεί να δειχθεί άμεσα ότι η μόνη συνάρτηση που ικανοποιεί την εξίσωση

$$f\left(s^1\right)=l*f(s)$$

δίνεται από την $f(s)=\log(s)$

η οποία δίνει το μέτρο του Hartley για το ποσό πληροφορίας. Σημειώνουμε ότι ο λογάριθμος επίσης μας εξασφαλίζει ότι το ποσό πληροφορίας αυξάνει καθώς αυξάνει ο αριθμός των συμβόλων s, το οποίο συμφωνεί με τη διαίσθηση μας.

Η επιλογή της βάσης του λογαρίθμου είναι αυθαίρετη και είναι περισσότερο θέμα κανονικοποίησης. Αν χρησιμοποιήσουμε το φυσικό λογάριθμο, η μονάδα της πληροφορίας λέγεται nat (natural unit). Συνήθως χρησιμοποιείται το 2 σαν βάση. Το ποσό πληροφορίας εκφράζεται τότε σε bits (παράγεται από το binary digits). Στην περίπτωση που έχουμε να επιλέξουμε μία από δύο δυνατότητες το ποσό πληροφορίας που παίρνουμε όταν εμφανιστεί μία από τις δύο δυνατότητες είναι 1 bit. Είναι εύκολο να δούμε ότι η σχέση μεταξύ bit και nat δίνεται από την

$$1\,\text{nat}=1{,}44\,\text{bits}$$

Στην προσέγγιση του Hartley όπως δόθηκε παραπάνω δεν έγινε καμία ανοχή για το γεγονός ότι τα s σύμβολα μπορεί να έχουν άνισες πιθανότητες εμφάνισης ή ότι μπορεί να υπάρχει μια εξάρτηση μεταξύ των l διαδοχικών συμβόλων.

Το μεγαλύτερο επίτευγμα του Shannon είναι ότι επέκτεινε τις θεωρίες του Nyquist και του Hartley και οδήγησε στη θεμελίωση της θεωρίας πληροφοριών όπως είναι γνωστή σήμερα, συνδέοντας την πληροφορία με την αβεβαιότητα χρησιμοποιώντας την έννοια της πιθανότητας.

Με αναφορά στο μέτρο του Hartley, ο Shannon πρότεινε ότι αυτό μπορεί πραγματικά να ερμηνευτεί σαν το μέτρο για το ποσό πληροφορίας με την υπόθεση ότι όλα τα σύμβολα έχουν ίση πιθανότητα εμφάνισης. Για τη γενική περίπτωση ο

Shannon εισήγαγε ένα μέτρο πληροφορίας βασισμένο στην έννοια της πιθανότητας, το οποίο συμπεριλαμβάνει το μέτρο του Hartley σαν μια ειδική περίπτωση.

Εντροπία του Shannon

Όπως είδαμε πιο πάνω ο ορισμός της πληροφορίας του Hartley δεν λαμβάνει υπόψη τις διάφορες πιθανότητες εμφάνισης των συμβόλων ή ενδεχομένων. Πρώτος ο Shannon συνέδεσε την πληροφορία με την έννοια της πιθανότητας. Αυτή η σύνδεση είναι στην πραγματικότητα ένα πολύ λογικό γεγονός. Αν θεωρήσουμε ένα δειγματικό χώρο όπου όλα τα ενδεχόμενα έχουν την ίδια πιθανότητα εμφάνισης, υπάρχει μια μεγάλη αβεβαιότητα σχετικά με το ποιο από τα ενδεχόμενα θα εμφανιστεί. Δηλαδή, όταν εμφανιστεί ένα από αυτά τα ενδεχόμενα θα δώσει πολύ περισσότερη πληροφορία από ότι στις περιπτώσεις όπου ο δειγματικός χώρος είναι δομημένος με τέτοιον τρόπο ώστε ένα ενδεχόμενο να έχει μεγάλη πιθανότητα εμφάνισης. Η πληροφορία συνδέεται με την έννοια της πιθανότητας μέσω της αβεβαιότητας.

Πριν αναφέρουμε τις ιδιότητες που ικανοποιεί το μέτρο πληροφορίας του Shannon δίνουμε τον παρακάτω ορισμό

Ορισμός 1. Έστω Α ένα πιθανοθεωρητικό πείραμα με δειγματικό χώρο Χ και κατανομή πιθανότητας P, όπου $p(x_i)$ ή p_i είναι η πιθανότητα του ενδεχομένου $x_i \in X$. Τότε η εντροπία (μέσο ποσό πληροφορίας) δίνεται από την

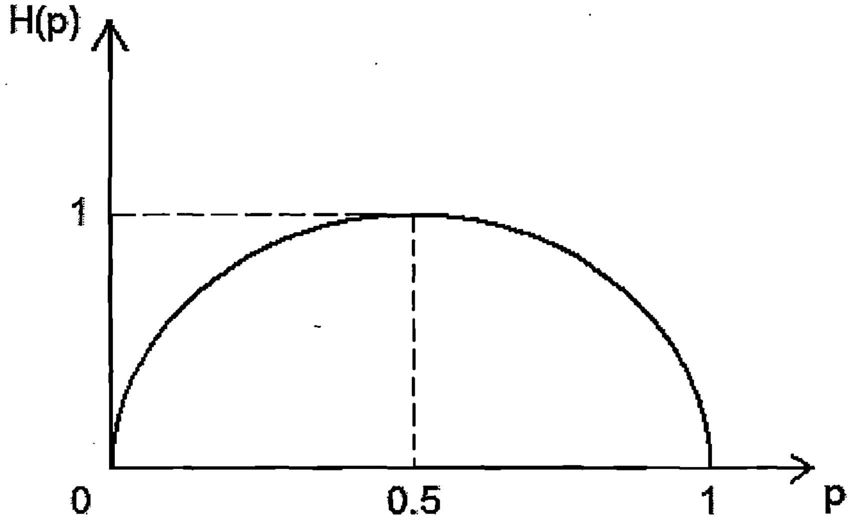
$$H(A) = -\sum_{i=1}^{n} p(x_i)\log p(x_i) = -\sum_{i=1}^{n} p_i \log p_i$$

Άλλοι συμβολισμοί για την εντροπία του Shannon είναι $H(X), H(P), H(p_1, p_2, \ldots, p_n)$. Στα παρακάτω θα χρησιμοποιήσουμε εναλλακτικά όλους αυτούς τους συμβολισμούς. Αυτό το μέτρο πληροφορίας παίρνει υπ'όψη τις n πιθανότητες $p_i, i = 1, \ldots, n$

Επειδή συνήθως επιλέγεται το 2 σαν βάση του λογαρίθμου, η μονάδα μέτρησης της πληροφορίας είναι το bit.

Στην περίπτωση δύο ενδεχομένων με πιθανότητες $p_1 = p$ και $p_2 = 1 - p$ βρίσκουμε

$$H(P) = -p\log p - (1-p)\log(1-p)$$



Στο πιο πάνω σχήμα φαίνεται η συμπεριφορά του H(P) σαν συνάρτηση του p. Μπορούμε να συμπεράνουμε ότι αν ένα ενδεχόμενο είναι βέβαιο, δηλαδή εμφανίζεται με πιθανότητα 1, το μέτρο πληροφορίας δίνει 0. Αυτό συμφωνεί με τη διαίσθηση ότι το βέβαιο ενδεχόμενο δε μας παρέχει πληροφορία. Το ίδιο ισχύει για p=0. Στην περίπτωση αυτή το άλλο ενδεχόμενο έχει πιθανότητα 1.

Όταν p=0.5 η H(P) προσεγγίζει τη μέγιστη τιμή της που είναι ίση με 1 bit. Για p=0.5 τα δύο ενδεχόμενα είναι ισοπίθανα και τότε έχουμε πλήρη αβεβαιότητα για το αποτέλεσμα. Η εμφάνιση του ενός από τα ενδεχόμενα μας παρέχει στην περίπτωση αυτή το μέγιστο ποσό πληροφορίας.

Ορίζουμε ακόμη plogp=0 όταν p=0 ($\lim_{x\to 0} x\log x = 0$)

Επιστρέφοντας στη γενική περίπτωση σημειώνουμε ότι η εντροπία ικανοποιεί τις παρακάτω τέσσερις απαιτήσεις:

I. Η H(P) είναι συνεχής ως προς p.
II. Η H(P) είναι συμμετρική. Δηλαδή η διάταξη των πιθανοτήτων $p_1,\ldots,p_n$ δεν έχει επίδραση στην τιμή της H(P).

III. Η H(P) είναι προσθετική. Αν X και Y είναι δύο δειγματικοί χώροι, όπου τα ενδεχόμενα του X είναι ανεξάρτητα με εκείνα του Y, τότε βρίσκουμε για την πληροφορία που σχετίζεται με τα από κοινού ενδεχόμενα (x_i, y_j)

$$H(p_1q_1,\dots,p_1q_m,\dots,p_nq_1,\dots p_nq_m) = H(p_1,\dots,p_n) + H(q_1,\dots,q_m)$$

IV. Η H(P) γίνεται μέγιστη όταν όλες οι πιθανότητες είναι ίσες. Αυτό αντιστοιχεί με την κατάσταση στην οποία υπάρχει μέγιστη αβεβαιότητα. Η H(P) είναι ελάχιστη αν ένα ενδεχόμενο έχει πιθανότητα ίση με 1.

Θα δώσουμε μια σύντομη εξήγηση των παραπάνω απαιτήσεων II, III και IV.

Το γεγονός ότι το μέτρο πληροφορίας του Shannon είναι συμμετρικό σημαίνει ότι η αλλαγή της σειράς στην αντικατάσταση των πιθανοτήτων δεν αλλάζει το ποσό της πληροφορίας. Μια συνέπεια αυτού είναι ότι διαφορετικοί δειγματικοί χώροι με κατανομές πιθανότητας που έχουν προκύψει από μεταθέσεις μιας κοινής κατανομής πιθανότητας θα δώσουν το ίδιο ποσό πληροφορίας.

Παράδειγμα 1. Θεωρούμε τα πειράματα A και B με τους παρακάτω δειγματικούς χώρους:

X={θα βρέχει αύριο, δε θα βρέχει αύριο}

Όπου P={0.8 , 0.2}

και

Y={ο Κώστας είναι νεότερος των 30 ετών, ο Κώστας είναι τουλάχιστον 30 ετών}

Όπου Q={0.2, 0.8}

Το ποσό της εντροπίας σε σχέση με το X είναι

H(A)=-(0.8 log0.8+0.2 log0.2)=0.72 bit

και σε σχέση με το Y είναι

H(B)=-(0.2 log0.2+0.8 log0.8)=0.72 bit

Συνεπώς H(A)=H(B)

Από το παραπάνω παράδειγμα μπορούμε να συμπεράνουμε ότι η εντροπία του Shannon δεν αναφέρεται στα περιεχόμενα της πληροφορίας. Οι πιθανότητες με τις οποίες εμφανίζονται τα ενδεχόμενα είναι αυτές που ενδιαφέρουν και όχι τα ενδεχόμενα αυτά καθαυτά.

Το ότι το μέτρο του Shannon ικανοποιεί την ιδιότητα που περιγράφεται στην III προκύπτει αμέσως γράφοντας αυτό αναλυτικά ως προς τις πιθανότητες. Η ιδιότητα της προσθετικότητας επεξηγείται καλύτερα με το παρακάτω παράδειγμα.

Θεωρούμε δύο ζάρια. Επειδή τα αποτελέσματα των δύο ζαριών είναι ανεξάρτητα το ένα με το άλλο δεν προκαλείται κάποια διαφορά όταν τα δύο ζάρια ρίχνονται ταυτόχρονα ή το ένα μετά το άλλο. Η πληροφορία που σχετίζεται με τα ζάρια όταν αυτά ρίχνονται ταυτόχρονα θα είναι η ίδια με τη διαδοχική πληροφορία που παίρνουμε ρίχνοντας πρώτα το ένα ζάρι και στη συνέχεια το άλλο.

Αν H(A) είναι το ποσό της πληροφορίας που σχετίζεται με τη ρίψη ενός ζαριού και H(B) το ποσό της πληροφορίας που σχετίζεται με τη ρίψη του άλλου ζαριού (σημειώνουμε ότι στην περίπτωση αυτή H(A)=H(B)) ενώ H(A,B) είναι το ποσό της πληροφορίας που σχετίζεται με τη ρίψη των δύο ζαριών ταυτόχρονα, τότε προκύπτει ότι

$$H(A,B)=H(A)+H(B)$$

Αυτό είναι ακριβώς που βεβαιώνει η προσθετική ιδιότητα.

Το ότι το ποσό της πληροφορίας θα μεγιστοποιείται στην περίπτωση των ίσων πιθανοτήτων είναι προφανές από το γεγονός ότι τότε η αβεβαιότητα είναι μέγιστη και η εμφάνιση ενός από τα ενδεχόμενα θα δώσει κατά συνέπεια τη μέγιστη πληροφορία.

Στο παρακάτω θεώρημα προσδιορίζεται το μέγιστο και το ελάχιστο ποσό της πληροφορίας.

Θεώρημα 1. Έστω $X=(x_1,\ldots x_n)$ είναι ο δειγματικός χώρος του πειράματος A, ενώ $P=(p_1,\ldots,p_n)$ είναι η αντίστοιχη κατανομή πιθανότητας. Τότε έχουμε ότι:

i. $H(P)\le \log n$ με ισότητα αν και μόνο αν

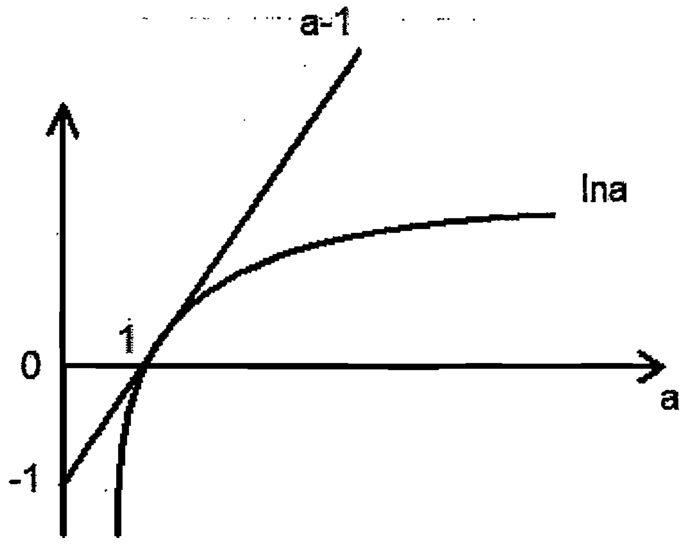
$$p_i = 1/n \text{ για όλα τα } i=1,\ldots n$$

ii. $H(P)\ge 0$ με ισότητα αν και μόνο αν υπάρχει ένα k τέτοιο ώστε $p_k=1$ ενώ για όλα τα άλλα $i\neq k, p_i=0$

Απόδειξη: (i) Κατά τη διάρκεια αυτής της απόδειξης θα χρησιμοποιήσουμε την παρακάτω ανισότητα

$$\ln a \le a-1$$

με την ακόλουθη γραφική παράσταση



Θεωρούμε τώρα τη διαφορά

$$H(P)-\log n=-\sum_{i=1}^{n}p_i\log p_i-\log n=-\sum_{i=1}^{n}p_i\{\log p_i+\log n\}=\sum_{i=1}^{n}p_i\log\left\{\frac{1}{p_i n}\right\}$$

Από τη δοθείσα ανισότητα προκύπτει

$$\log a=\frac{\ln a}{\ln 2}\le(a-1)\frac{\ln e}{\ln 2}=(a-1)\log e$$

Χρησιμοποιώντας αυτή την ανισότητα παίρνουμε

$$H(P)-\log n\le\sum_{i=1}^{n}p_i\left\{\frac{1}{p_i n}-1\right\}\log e=\left\{\sum_{i=1}^{n}\frac{1}{n}-\sum_{i=1}^{n}p_i\right\}\log e=\left\{n\frac{1}{n}-1\right\}\log e=0$$

Έτσι αποδείξαμε ότι $H(P)\le\log n$

$$H(P)\le\log n$$

με ισότητα αν και μόνο αν $1/(p_i n)=1$. Αυτό σημαίνει ότι $p_i=1/n$ για όλα τα i=1,...n

(ii) Επειδή p_i και $-\log p_i$ δεν μπορεί να είναι και τα δύο αρνητικά, το ποσό της πληροφορίας είναι πάντοτε θετικό ή ίσο με το μηδέν. Συνεπώς

$$H(P) \geq 0$$

Προφανώς η H(P) μπορεί να γίνει ίση με το μηδέν μόνο αν υπάρχει μια πιθανότητα στην P ίση με 1, ενώ όλες οι άλλες πιθανότητες θα είναι μηδέν.

Το μέγιστο ποσό πληροφορίας είναι συνεπώς ίσο με logn. Για να πάρουμε μια εικόνα για το ποσό της πληροφορίας που δίνει ένα σύστημα πληροφορίας θεωρούμε το παρακάτω παράδειγμα.

Παράδειγμα 2 . Μια εικόνα στην τηλεόραση αποτελείται από 576 γραμμές, κάθε μία από τις οποίες κατασκευάζεται από 720 στοιχεία. Συνεπώς μια απλή εικόνα στην οθόνη της τηλεόρασης αποτελείται συνολικά 414720 στοιχεία. Κάτω από την υπόθεση που κάνουμε για μια φωτεινή διαβάθμιση της εικόνας, όπου κάθε στοιχείο μπορεί να δείχνει ένα από τα δέκα διαστήματα της έντασης , υπάρχουν 10^{414720} δυνατές διαφορετικές εικόνες στην τηλεόραση. Αν κάθε μία από αυτές τις εικόνες έχει ίση πιθανότητα εμφάνισης, το ποσό της πληροφορίας που περιέχεται σε μια εικόνα είναι ίσο με

$$H(P) = \log n = \log(10^{411720}) \cong 1.4 * 10^{6} bits$$

Έχουμε δει ότι το ποσό της πληροφορίας δεν αλλάζει αν οι πιθανότητες αντικατασταθούν σε μια διαφορετική σειρά. Ας θεωρήσουμε τώρα τις παρακάτω δύο κατανομές πιθανότητας:

$$P = \{0.50, 0.25, 0.25\}$$

Και

$$Q = \{0.48, 0.32, 0.20\}$$

Υπολογίζοντας το αντίστοιχο ποσό της πληροφορίας για τις δύο περιπτώσεις βρίσκουμε

$$H(P) = H(Q) = 1.5 bits$$

Εμφανίζεται έτσι διαφορετικές κατανομές πιθανότητας να οδηγούν στο ίδιο ποσό πληροφορίας. Μερικά πειράματα μπορεί να έχουν διαφορετικές κατανομές πιθανότητας αλλά το ίδιο ποσό πληροφορίας.

Για την καλύτερη κατανόηση του μέτρου πληροφορίας του Shannon θεωρούμε το παρακάτω παράδειγμα.

Παράδειγμα 3. Έστω ο δειγματικός χώρος $X = \{x_1, x_2, x_3\}$ με αντίστοιχη κατανομή πιθανότητας $P = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\}$. Θεωρούμε το παιχνίδι των απαντήσεων με 'ναι' και 'όχι'. Είναι αυτονόητο να ρωτήσουμε πρώτα για το x_1 , αφού αυτό το αποτέλεσμα έχει τη μεγαλύτερη πιθανότητα. Αν η απάντηση είναι 'ναι', τότε βρήκαμε το αποτέλεσμα με μια προσπάθεια. Αν η απάντηση είναι 'όχι' , τότε το αποτέλεσμα είναι προφανώς το x_2 ή το x_3. Για να προσδιορίσουμε ποιο από τα δύο είναι κάνουμε άλλη μια ερώτηση και επομένως χρειαζόμαστε δύο ερωτήσεις συνολικά για να γνωρίζουμε το αποτέλεσμα.
Άρα κάνουμε είτε μία είτε δύο ερωτήσεις με ίσες πιθανότητες και συνεπώς ο μέσος όρος είναι 1.5 ερωτήσεις.

Αν υπολογίσουμε το ποσό της πληροφορίας σύμφωνα με τον Shannon βρίσκουμε

$$H(P) = -\left(\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4}\right) = 1.5 bits$$

Η ερμηνεία που δόθηκε προηγουμένως ισχύει κατά συνέπεια και για μη-ίσες πιθανότητες.

Από κοινού Εντροπία

Θεωρούμε το (Α, Β) σαν ένα πιθανοθεωρητικό πείραμα με ζεύγη γεγονότων (x_i, y_j) όπου $x_i \in X$ και $y_j \in Y$. Η πιθανότητα $r(x_i, y_j)$ που συμβολίζεται επίσης με r_{ij} ή $p(x_i, y_j)$ είναι η πιθανότητα ότι το πείραμα θα έχει σαν αποτέλεσμα το γεγονός (x_i, y_j) και ονομάζεται από κοινού πιθανότητα.

Με βάση το μέγεθος του δειγματικού χώρου (X, Y) μπορούμε να συμπεράνουμε ότι το πείραμα (A, B) έχει συνολικά nm δυνατά από κοινού ενδεχόμενα. Θα ορίσουμε παρακάτω το ποσό της εντροπίας όσον αφορά το πείραμα (A, B).

Υπάρχουν nm από κοινού ενεχόμενα (x_i, y_j) με πιθανότητες εμφάνισης $r(x_i, y_j)$ ή r_{ij}.

Υποθέτουμε τώρα ότι γράφουμε τα nm από κοινού ενδεχόμενα σαν τα ενδεχόμενα $z_1, z_2, \ldots, z_{nm}$ και τις αντίστοιχες πιθανότητες σαν $p(z_1), p(z_2), \ldots, p(z_{nm})$

Τότε έχουμε πάλι ένα δειγματικό χώρο μίας διάστασης και με τον ορισμό του περιθωρίου μέτρου πληροφορίας βρίσκουμε

$$H(Z) = -\sum_{k=1}^{nm} p(z_k)\log p(z_k)$$

Αλλά επειδή κάθε $p(z_k)$ θα είναι ίση με μια από τις πιθανότητες $r(x_i, y_j)$ το άθροισμα ως προς k θα δώσει το ίδιο αποτέλεσμα με το άθροισμα ως προς I και j. Με άλλα λόγια

$$H(Z) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log(r(x_i, y_j))$$

Αυτό οδηγεί στον παρακάτω ορισμό της από κοινού εντροπίας

Ορισμός 2. Θεωρούμε ένα πιθανοθεωρητικό πείραμα (A, B) με δειγματικό χώρο δύο διαστάσεων (X, Y), όπου r_{ij} ή $r(x_i, y_j)$ είναι η πιθανότητα του x_i και y_j, τότε η από κοινού εντροπία ορίζεται να είναι

$$H(A,B) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log(r(x_i, y_j)) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\log p_{ij}$$

Θα χρησιμοποιήσουμε εναλλακτικά τους συμβολισμούς

$$H(A,B), H(R), H(r_1, r_2, ..., r_{nm})$$

Δεσμευμένη Εντροπία

Μέχρι τώρα έχουμε δει ότι η περιθώρια εντροπία μπορεί να οριστεί με βάση τις περιθώριες πιθανότητες και ότι οι από κοινού πιθανότητες οδηγούν στην εισαγωγή της από κοινού εντροπίας. Θα δούμε τώρα πως μπορεί να οριστεί η δεσμευμένη εντροπία σε σχέση με τις δεσμευμένες πιθανότητες.

Θεωρούμε πάλι τα πιθανοθεωρητικά πειράματα Α και Β. Υποθέτουμε τώρα ότι ενδιαφερόμαστε για το ποσό της πληροφορίας όσον αφορά το Β κάτω από τη συνθήκη ότι έχει ήδη εμφανιστεί το ενδεχόμενο x_i . Έχουμε τότε τις πιθανότητες $q(y_j \mid x_i), j = 1, \ldots m,$ αντί για τις πιθανότητες $q(y_j), j = 1, \ldots, m,$ αλλά με το άθροισμά τους να εξακολουθεί να ισούται με 1.

Το ποσό της πληροφορίας όσον αφορά το Β δοθέντος του αποτελέσματος x_i, σε αναλογία με το περιθώριο μέτρο πληροφορίας μπορεί τότε να οριστεί σαν

$$H(B \mid x_i) = -\sum_{j=1}^{m} q(y_j \mid x_i) \log[q(y_j \mid x_i)]$$

Θεωρώντας τώρα κατά μέσο όρο όλες τις τιμές του x_i το μέσο ποσό της πληροφορίας του Β όταν δίνεται η πρόβλεψη του Α βρίσκεται:

$$\sum_{i=1}^{n} p(x_i) H(B \mid x_i) = -\sum_{i=1}^{n} \sum_{j=1}^{m} r(x_i, y_j) \log[q(y_j \mid x_i)]$$

Αυτή η ποσότητα συμβολίζει το δεσμευμένο ποσό της πληροφορίας $H(B \mid A)$. Αυτό οδηγεί στον παρακάτω ορισμό

Ορισμός 3. Το δεσμευμένο μέτρο πληροφορίας όσον αφορά το πείραμα B δοθέντος του A είναι ίσο με

$$H(B \mid A) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log[q(y_j \mid x_i)]$$

Με ανάλογο τρόπο μπορούμε να ορίσουμε το δεσμευμένο μέτρο πληροφορίας όσον αφορά το πείραμα A δοθέντος του B

$$H(A \mid B) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log[p(x_i \mid y_j)]$$

Αντί για τα $H(B \mid A)$ και $H(A \mid B)$ θα χρησιμοποιήσουμε επίσης τους συμβολισμούς σε σχέση με τους δειγματικούς χώρους αυτών $H(Y \mid X)$ και $H(X \mid Y)$.

Το παρακάτω θεώρημα δίνει την ελάχιστη και τη μέγιστη τιμή για το $H(B \mid A)$

Θεώρημα 2. Έστω $H(B \mid A)$ είναι το δεσμευμένο μέτρο πληροφορίας για το B δοθέντος του A. Τότε

i. $H(B \mid A) \geq 0$

ii. $H(B \mid A) \leq H(B)$ με ισότητα αν τα A και B είναι στοχαστικά ανεξάρτητα

Απόδειξη

i. Επειδή $q(y_j \mid x_i) \leq 1 \forall i, j$, προκύπτει ότι $\{-\log q(y_j \mid x_i)\} \geq 0$ και έτσι προκύπτει άμεσα από τον ορισμό ότι $H(B \mid A) \geq 0$

ii. $$H(B \mid A) - H(B) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log[q(y_j \mid x_i] + \sum_{j=1}^{m} q(y_j)\log q(y_j) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log\left[\frac{q(y_j)}{q(y_j \mid x_i)}\right]$$

Χρησιμοποιώντας την ανισότητα που αναφέραμε προηγουμένως, δηλαδή την $\ln a \leq a - 1$ προκύπτει

$$H(B \mid A) - H(B) \leq \sum_{i=1}^{n} \sum_{j=1}^{m} r(x_i, y_j) \left[\frac{q(y_j)}{q(y_j \mid x_i} - 1 \right] \log e$$

Το δεξί μέλος αυτής της ανισότητας μπορεί να γραφεί

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i) q(y_j \mid x_i) \frac{q(y_j)}{q(y_j \mid x_i)} \log e - \sum_{i=1}^{n} \sum_{j=1}^{m} r(x_i, y_j) \log e = \log e - \log e = 0$$

Συνεπώς $H(B \mid A) \leq H(B)$

Τα δύο ποσά της πληροφορίας είναι ίσα αν $q(y_j) = q(y_j \mid x_i)$ για όλα τα I και j , όπως συμβαίνει στην περίπτωση της στοχαστικής ανεξαρτησίας.

Το συμπέρασμα που μπορούμε να βγάλουμε από το θεώρημα αυτό είναι ότι το δεσμευμένο ποσό της πληροφορίας είναι πάντοτε μικρότερο ή ίσο από το περιθώριο ποσό της πληροφορίας. Με άλλα λόγια η πληροφορία σχετικά με το A γενικά θα οδηγεί σε μείωση της αβεβαιότητας, κάτι που συμφωνεί με τη διαίσθηση.

Όπως φαίνεται στο παρακάτω θεώρημα υπάρχει μια άμεση σχέση μεταξύ του περιθώριου , του δεσμευμένου και του από κοινού μέτρου της πληροφορίας

Θεώρημα 3. Για όλα τα πειράματα A και B ισχύει

$$H(A, B) = H(A) + H(B \mid A) = H(B) + H(A \mid B)$$

Απόδειξη. Έχουμε

$$H(A, B) = -\sum_{i=1}^{n} \sum_{j=1}^{m} r(x_i, y_j) \log[p(x_i) q(y_j \mid x_i)] =$$

$$= -\sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log p(x_i) - \sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log\left[q(y_j \mid x_i)\right] = H(A) + H(B \mid A)$$

Με ανάλογο τρόπο αποδεικνύεται ότι

$$H(A,B) = H(B) + H(A \mid B)$$

Το θεώρημα αυτό μας λέει ότι το από κοινού ποσό της πληροφορίας είναι το άθροισμα του ποσού της πληροφορίας όσον αφορά το Α και του δεσμευμένου ποσού της πληροφορίας του Β δοθέντος του Α.

Με βάση τα θεωρήματα 2 και 3 μπορούμε επιπλέον να συμπεράνουμε ότι

$$H(A,B) = H(A) + H(B \mid A) \le H(A) + H(B)$$

Με ισότητα αν τα Α, Β είναι ανεξάρτητα.
Μπορούμε συνεπώς να υποθέσουμε ότι το από κοινού ποσό της πληροφορίας είναι μέγιστο αν τα δύο πιθανοθεωρητικά πειράματα είναι ανεξάρτητα και μειώνεται καθώς αυξάνεται η εξάρτηση. Στην περίπτωση της απόλυτης εξάρτησης το αποτέλεσμα του Β είναι γνωστό αν το αποτέλεσμα του Α είναι γνωστό έτσι ώστε H(B|A)=0. Στην περίπτωση αυτή είναι H(A,B)=H(A).

Αμοιβαία Εντροπία

Θα δώσουμε τώρα έναν ακόμη ορισμό, αυτόν της αμοιβαίας εντροπίας.

Ορισμός 4. Η αμοιβαία εντροπία των Α και Β ορίζεται να είναι

$$I(A;B) = H(B) - H(B \mid A) = \sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log\left[\frac{r(x_i, y_j)}{p(x_i, q(y_j))}\right]$$

Το $I(A;B)$ μπορεί να ερμηνευτεί σαν ένα μέτρο εξάρτησης μεταξύ των B και A. Όταν τα A και B είναι ανεξάρτητα, τότε το $I(A;B)$ είναι ελάχιστο και ισούται με το μηδέν. Αν το B είναι πλήρως εξαρτώμενο από το A, τότε H(B|A)=0 και το $I(A;B)$ παίρνει τη μέγιστή τιμή του που είναι H(B).

Θα δείξουμε τώρα το παρακάτω θεώρημα.

Θεώρημα 4. Για όλα τα πειράματα A και B ισχύει

i. $I(A;B) = H(A) - H(A \mid B)$

ii. $I(A;B) = H(A) + H(B) - H(A,B)$

Απόδειξη.

i. $I(A;B) = H(B) - H(B \mid A) = \sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log\left[\frac{r(x_i, y_j)}{p(x_i, q(y_j))}\right] =$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log\left[\frac{r(x_i, y_j)}{q(y_j)}\right] - \sum_{i=1}^{n}\sum_{j=1}^{m} r(x_i, y_j)\log p(x_i) =$$

$$= -H(A \mid B) - \sum_{i=1}^{n} p(x_i)\log p(x_i) = H(A) - H(A \mid B)$$

ii. Από το θεώρημα 3 γνωρίζουμε ότι

$H(A,B) = H(B) + H(A \mid B)$ και μέσω του i μπορούμε να βρούμε ότι

$$H(A,B) = H(B) + H(A) - I(A;B) \Rightarrow I(A;B) = H(A) + H(B) - H(A,B)$$

Επίσης προκύπτει άμεσα ότι το $I(A;B)$ είναι συμμετρικό, δηλαδή για όλα τα A και B ισχύει

$$I(A;B) = I(B;A)$$

Μπορούμε να συμπεράνουμε ότι ισχύουν οι παρακάτω σχέσεις για τα μέτρα πληροφορίας

$$I(A;B) = H(A) \cap H(B)$$
$$H(A,B) = H(A) \cup H(B)$$

$$H(A \mid B) \leq H(A)$$
$$H(B \mid A) \leq H(B)$$

$$I(A;B) \leq H(B)$$
$$I(A;B) \leq H(A)$$

$$I(A;B) = H(A) - H(A \mid B) = H(B) - H(B \mid A)$$

$$H(A,B) = H(A \mid B) + I(A;B) + H(B \mid A) = H(B) + H(A \mid B) = H(A) + H(B \mid A)$$

$$H(A,B) \leq H(A) + H(B)$$

Αξιωματική θεμελίωση

Προηγουμένως εισάγαμε το μέτρο πληροφορίας του Shannon και είδαμε μερικές ιδιότητες αυτού του μέτρου πληροφορίας. Αυτές οι ιδιότητες ήταν σε αντιστοιχία με τις ιδιότητες που θα περίμενε κάποιος για ένα μέτρο πληροφορίας. Στην ομοιόμορφη περίπτωση το μέτρο πληροφορίας του Shannon ισούται με το μέτρο πληροφορίας του Hartley, το οποίο είναι ο λογάριθμος του αριθμού των μηνυμάτων. το μέτρο πληροφορίας του Shannon που βασίζεται στις πιθανότητες μπορεί να προκύψει άμεσα από την ομοιόμορφη περίπτωση και το μέτρο του Hartley.

Υποθέτουμε ότι ένα μέτρο πληροφορίας πρέπει να ικανοποιεί τις παρακάτω τρεις απαιτήσεις.

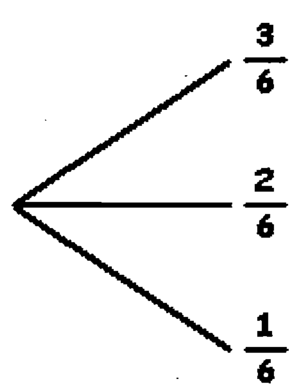
i. Αν τα αποτελέσματα χωρίζονται σε ομάδες, τότε όλες οι τιμές του H για τις διάφορες ομάδες, πολλαπλασιαζόμενες με τα στατιστικά τους βάρη, πρέπει να οδηγούν στη συνολική τιμή του H.

ii. Το H πρέπει να είναι συνεχής συνάρτηση ως προς p_i.

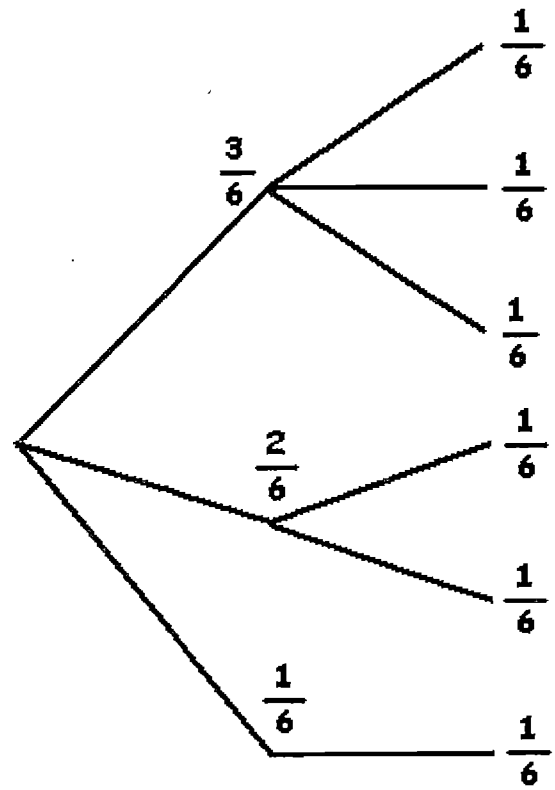
iii. Αν όλα τα p_i είναι ίσα, δηλαδή για όλα τα i ισχύει $p_i = 1/n$, τότε H θα αυξάνει μονότονα σαν συνάρτηση του n. Αυτό σημαίνει ότι η αβεβαιότητα θα αυξάνει για έναν αυξανόμενο αριθμό ίσων πιθανοτήτων.

Για n ισοπίθανα αποτελέσματα η H πρέπει να ικανοποιεί την $H = \log n$ σύμφωνα με τον Hartley και την απαίτηση iii.
Για άνισες πιθανότητες θεωρούμε την παρακάτω περίπτωση.Υποθέτουμε ότι οι πιθανότητες είναι $\frac{3}{6}, \frac{2}{6}, \frac{1}{6}$

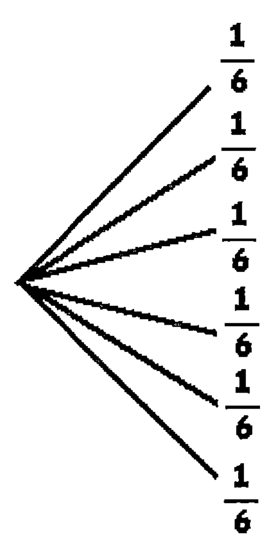
Τα πιο κάτω σχήματα δείχνουν το δέντρο απόφασης για το οποίο πρέπει να υπολογιστεί η H.



Σχήμα 1



Σχήμα 2



Σχήμα 3

Η τιμή του Η όσον αφορά το δέντρο απόφασης στο σχήμα (3) πρέπει να είναι ίση με $\mathrm{H}^{c} = \log 6$. Εντούτοις επειδή τα δύο δέντρα απόφασης (2) και (3) είναι ουσιαστικά ταυτόσημα, η τιμή του Η όσον αφορά το δέντρο απόφασης στο σχήμα (2) πρέπει επίσης να είναι ίση με $\mathrm{H}^{b} = \log 6$.

Με βάση την απαίτηση i αυτή η τιμή πρέπει να είναι ίση με τη αβεβαιότητα αναφορικά με την επιλογή μεταξύ των κλάδων που συμβολίζονται με $\frac{3}{6}, \frac{2}{6}$ και $\frac{1}{6}$ στο σχήμα (2) (δηλαδή την H^{a} για την οποία ψάχνουμε) συν τις αβεβαιότητες αναφορικά με τους υποκλάδους πολλαπλασιαζόμενες με τα αντίστοιχα υποβάρη τους.

Έτσι προκύπτει ότι

$$H^{a} + \frac{3}{6}\log 3 + \frac{2}{6}\log 2 + \frac{1}{6}\log 1 = \log 6$$

Και $H^{a} = -\left(\frac{1}{2}\log\frac{1}{2} + \frac{1}{3}\log\frac{1}{3} + \frac{1}{6}\log\frac{1}{6}\right)$

Πιο γενικά προκύπτει ότι

$$H(A) = -\sum_{i=1}^{n} p_i \log p_i$$

Οι Chaudy και Mcleod (1960) έδωσαν το παρακάτω θεώρημα το οποίο προσδιορίζει μοναδικά το μέτρο πληροφορίας του Shannon.

Θεώρημα 5. Θεωρούμε μια συνάρτηση $\mathrm{f(A) = f(P) = f(p_1, \ldots} p_n)$ και μια συνάρτηση $g(.)$, οι οποίες ικανοποιούν τις παρακάτω ιδιότητες:

i. $f(P)=\sum_{i=1}^{n} g(p_i)$

ii. η $f(.)$είναι συνεχής στο διάστημα $[0,1]$

iii. η $f(P)$ είναι προσθετική, δηλαδή

$$f(p_1 q_1,\ldots,p_n q_m)=f(p_1,\ldots,p_n)+f(q_1,\ldots,q_m)$$

iv. $f(\frac{1}{2},\frac{1}{2})=1$

Τότε θα ισχύει

$$f(P)=H(P)=-\sum_{i=1}^{n} p_i \log p_i$$

Μπορεί να προκύψει από το θεώρημα αυτό ότι στην πραγματικότητα είναι η προσθετική ιδιότητα αυτή που προσδιορίζει μοναδικά την εντροπία του Shannon.

Σύνοψη

Στο κεφάλαιο αυτό έγινε μια αναφορά στη Θεωρία Πληροφορίας. Μετά την σύντομη ιστορική αναδρομή καταλήξαμε στον Shannon, ο οποίος όρισε την **Εντροπία**

$$H(A)=-\sum_{i=1}^{n} p_i \log p_i$$

Όπου p_i είναι η πιθανότητα του κάθε ενδεχομένου στο δειγματικό χώρο.

Στη συνέχεια ορίσαμε την από κοινού εντροπία.

$$\mathrm{H(A,B)}=-\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} \log p_{ij}$$

Όπου p_{ij} είναι η πιθανότητα να συμβούν δύο ενδεχόμενα i και j στο δειγματικό χώρο δύο διαστάσεων, όπου το ένα ανήκει στη μία διάσταση και το άλλο στην άλλη.

Έπειτα αναφερθήκαμε στη δεσμευμένη εντροπία.

$$H(A \mid B) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log[p(x_i \mid y_j)]$$

Δηλαδή η πιθανότητα να συμβεί το A με δεδομένο ότι συμβαίνει το B.

Τέλος επισημάναμε την αμοιβαία εντροπία.

$$I(A;B) = H(A) - H(A \mid B)$$

Με βάση τα παραπάνω χρήσιμα εργαλεία θα παρουσιάσουμε μια μέθοδο εύρεσης των καλύτερων αναλογιών επένδυσης σε ένα χαρτοφυλάκιο.

Κεφάλαιο 2
Θεωρία Χαρτοφυλακίου

Εισαγωγή

Ως χαρτοφυλάκιο ορίζουμε ένα σύνολο από περιουσιακά στοιχεία (μετοχές, ομόλογα, καταθέσεις στην τράπεζα). Στο κεφάλαιο αυτό αναφερόμαστε στη θεωρία του Markowitz για τη μέγιστη απόδοση ενός χαρτοφυλακίου και επιπλέον εισάγουμε μια νέα θεωρία που βασίζεται στην εντροπία.
Αντικαθιστώντας τον αριθμητικό μέσο των αποδόσεων και την τυπική απόκλιση που υιοθετήθηκε από το M.H. Markowitz με το γεωμετρικό μέσο των αποδόσεων σαν ένα κριτήριο εκτίμησης ενός χαρτοφυλακίου λαμβάνουμε την αυξητική εντροπία: μία από τις γενικευμένες εντροπίες. Αυτή δείχνει την αυξανόμενη ταχύτητα του κεφαλαίου και είναι ένα πιο αντικειμενικό και εξετάσιμο μέτρο. Το διαφορετικό από τη θεωρία του Markowitz είναι ότι η νέα θεωρία δίνει έμφαση στο ότι για δεδομένη πιθανότητα των κερδών υπάρχει ένα αντικειμενικά ωφέλιμο χαρτοφυλάκιο, το οποίο μπορεί να εξετασθεί είτε παίζοντας νομίσματα είτε με εξομοίωση σε υπολογιστή.
Παρακάτω εξετάζουμε κάποιους απλούς και εφαρμόσιμους τύπους για μεγιστοποίηση των ποσοστών επένδυσης και κάποια δείγματα που δείχνουν τη μεγιστοποίηση ενός χαρτοφυλακίου με πολλά στοιχεία από υπολογιστή. Επίσης εξετάζουμε ένα μέτρο για την αξία της πληροφορίας βασισμένο στη νέα θεωρία του χαρτοφυλακίου, αναλύουμε ομοιότητες και διαφορές μεταξύ αυτού του μέτρου και του μέτρου του K.J. Arrow και εξετάζουμε πως να μεγιστοποιηθούν οι προβλέψεις με το νέο μέτρο για την αξία της πληροφορίας σαν κριτήριο.

Πολλά χρόνια πριν είχε παρουσιαστεί μια γενικευμένη θεωρία πληροφορίας και αποδείχτηκε ότι το μέτρο της γενικευμένης πληροφορίας είχε επίσης τη δική του κωδική σημασία (Lu,1992,1993, 1994). Αργότερα, βρέθηκε ότι η νέα γενικευμένη εντροπία, όπως π.χ. η αυξητική εντροπία, μπορούσε να χρησιμοποιηθεί για τη μεγιστοποίηση χαρτοφυλακίου. Η νέα θεωρία του χαρτοφυλακίου βασίζεται σε μερικές έννοιες της θεωρίας του Markowitz (1959,1991), αλλά έχει και σημαντικές διαφορές. Δίνει έμφαση στο ότι υπάρχει ένα αντικειμενικό κριτήριο το οποίο είναι η αυξανόμενη ταχύτητα του κεφαλαίου. Σε αυτό το κριτήριο, για δεδομένη πιθανότητα πρόβλεψης των αποδόσεων, μπορούμε να πάρουμε τα μέγιστα ποσοστά των επενδύσεων σε διαφορετικά περιουσιακά στοιχεία. Συνδυάζοντας τη νέα θεωρία του χαρτοφυλακίου και τη γενική θεωρία πληροφορίας, μπορούμε να προσεγγίσουμε ένα σαφές μέτρο της αξίας της πληροφορίας, το οποίο σχεδιάζει την αύξηση της ταχύτητας του αυξανόμενου κεφαλαίου αφού η πληροφορία παρέχεται ενώ κάποιος παίρνει πάντοτε αποφάσεις με τον πιο ωφέλιμο τρόπο. Η θεωρία του χαρτοφυλακίου εφαρμόζεται για τις επενδύσεις σε αγορές αποθέματος και μέλλοντος. Η πρακτική δείχνει ότι η θεωρία έχει σαφή υπεροχή.

Μεγιστοποίηση Χαρτοφυλακίου

-Συζητώντας για το ρίσκο όταν παίζεις νομίσματα

Πρώτα από όλα, ασχολούμαστε με το θέμα των χαρτοφυλακίων από ένα απλό παράδειγμα.

Παράδειγμα 1. Υποθέτουμε ότι υπάρχει μια επένδυση ή ένα στοίχημα του οποίου η απόδοση είναι μη συγκεκριμένη και αποφασίζεται παίζοντας ένα νόμισμα. Αν εμφανιστεί η πλευρά Α θα κερδίσεις 200% του στοιχήματός σου και αν εμφανιστεί η πλευρά Β θα χάσεις 100%. Έχεις μόνο 100 $ και δεν μπορείς να δανειστείς αν χάσεις. Το ερώτημα είναι, για πόσες φορές πονταρίσματος, ποια είναι η καλύτερη αναλογία του κεφαλαίου σου που μπορεί να έχει το στοίχημα έτσι ώστε το κεφάλαιο σου να αυξάνει γρηγορότερα, ή να το πούμε αλλιώς πως θα γίνεις εκατομμυριούχος πιο γρήγορα.

Το να ποντάρεις όλα τα λεφτά σου κάθε φορά δεν είναι συμφέρον γιατί αν εμφανιστεί μια φορά η πλευρά Β θα χάσεις όλα τα λεφτά σου. Από την άλλη, το να ποντάρεις 10% των χρημάτων σου δεν είναι τόσο άσχημο μια και δε θα τα χάσεις κατευθείαν. Μετά από 2 φορές κατά μέσο όρο η απόδοσή σου θα είναι (1-1*0.1)(1+2*0.1)=1.08 φορές. Αλλά γιατί, με αυτόν τον τρόπο η ταχύτητα αύξησης του κεφαλαίου είναι τόσο αργή; Μπορούμε να βρούμε την καλύτερη αναλογία επένδυσης;

Σε μια πιο γενική περίπτωση, ο αριθμός των στοιχείων που πρέπει να επενδυθούν είναι μεγαλύτερος από ένα και οι μελλοντικές αποδόσεις δεν είναι συγκριμένες. Πώς μπορούμε να αποφασίσουμε την καλύτερη αναλογία αποδόσεων με διαφορετικά στοιχεία, για δεδομένες από κοινού κατανομές πιθανοτήτων των μελλοντικών αποδόσεων;

Η θεωρία του χαρτοφυλακίου που εισήχθη από το Markowitz (1959) έχει σπουδαία επιτεύγματα. Για αυτό το λόγο, ο M.H. Markowitz, ο W.F.Shape και ο M. Miller κέρδισαν το βραβείο Nobel το 1990. Η θεωρία μας λέει ότι όσο μεγαλύτερα είναι τα αναμενόμενα κέρδη τόσο καλύτερη είναι η απόδοση. Επίσης όσο μικρότερη είναι η τυπική απόκλιση των κερδών, το οποίο σημαίνει ρίσκο, τόσο το καλύτερο. Μπορούμε να μειώσουμε την τυπική απόκλιση των κερδών ή του κινδύνου συνδυάζοντας anti-covariant αξίες μαζί. Πώς μπορούμε να εκτιμήσουμε τη χρησιμότητα ενός χαρτοφυλακίου με δεδομένα τα αναμενόμενα κέρδη και την τυπική απόκλιση; Δεν υπάρχει κάποιο αντικειμενικό κριτήριο, ή να το πούμε διαφορετικά το κριτήριο αλλάζει από πρόσωπο σε πρόσωπο.

Το πιο αποδοτικό χαρτοφυλάκιο που προτάθηκε από το Markowitz είναι αυτό που έχει το μικρότερο κίνδυνο για ένα δεδομένο επίπεδο αναμενόμενων κερδών ή το μεγαλύτερο αναμενόμενο κέρδος για ένα δεδομένο επίπεδο κινδύνου. Ωστόσο, δεν είναι ένα χαρτοφυλάκιο που πετυχαίνει τη μεγαλύτερη αύξηση του κεφαλαίου. Έτσι, χρειαζόμαστε νέα μαθηματικά μοντέλα.

Κεφάλαιο 3

Μαθηματικό Μοντέλο

Εισαγωγή

Στο κεφάλαιο αυτό προσπαθούμε να εξάγουμε ένα μαθηματικό μοντέλο για την μεγιστοποίηση των κερδών ενός χαρτοφυλακίου χρησιμοποιώντας την έννοια της εντροπίας που αναφέρθηκε στο πρώτο κεφάλαιο.

Αυξητική Εντροπία σε χαρτοφυλάκια

Ας θεωρήσουμε ότι οι τιμές των N στοιχείων σε ένα χαρτοφυλάκιο σχηματίζουν ένα N-διάστατο διάνυσμα και η τιμή του k-οστού αξιόγραφου έχει n_k πιθανές αξίες, k=1,2,...,N. Ως εκ τούτου υπάρχουν $W = n_1 \times n_2 \times \ldots n_N$ πιθανά διανύσματα τιμών. Ας υποθέσουμε ότι το i-οστό διάνυσμα τιμών είναι $x_i = (x_{i1}, x_{i2}, \ldots, x_{ik}), i = 1,2,\ldots,W$ και το τρέχον διάνυσμα τιμών είναι $x_0 = (x_{01}, x_{02}, \ldots, x_{0k})$. Όταν ισχύει το x_i διάνυσμα τιμών η απόδοση από το k-οστό αξιόγραφό είναι r_{ik}. Τότε η απόδοση από το χαρτοφυλάκιο είναι

$$r_i = \sum_{k=0}^{N} q_k r_{ik} \quad (3.1)$$

Όπου $q_k, k = 1,2,\ldots,N$ είναι η αναλογία του επενδεδυμένου κεφαλαίου στο k-οστό αξιόγραφο, q_0 είναι η αναλογία του χρήματος που κρατιέται από τον επενδυτή, $r_{ik} = x_{ik}/x_{0k}, r_{ok} = r_o = (1 + \text{int}\,erest - rate)$. Θεωρώντας τις χρεώσεις περιλαμβανομένου και του φόρου τότε

$$r_{ik} = \left(1 - d\left|q_k - q_k^{'}\right|\right) x_{ik}/x_0 \quad (3.2)$$

όπου d είναι οι χρεώσεις για ανταλλαγή του κάθε €,

$q_k^{'}$ είναι η προηγούμενη αναλογία των επενδύσεων στο k-οστό αξιόγραφο

Θεωρούμε ότι ένα διάνυσμα τιμών r_i συμβαίνει N_i φορές σε N πειράματα επενδύσεων. Τότε η απόδοση του κεφαλαίου μετά N πειράματα επενδύσεων είναι

$$\prod_{i=1}^{N} r_i^{N_i} \quad (3.3)$$

Το μέσο κέρδος για κάθε περίοδο είναι

$$\prod_{i=1}^{N} r_i^{N_i/N} \quad (3.4)$$

Το πιο πάνω μαθηματικό μοντέλο πρώτα προτάθηκε από τους Henry A.L.atane και το Donald L. Tuttle(1967) και ονομαζόταν μοντέλο μεγιστοποίησης πλούτου (wealth-maximizing model).

Παίρνοντας τη λογαριθμική αξία έχουμε

$$H = \sum_{i=1}^{N} P(x_i)\log r_i = \sum_{i=1}^{N} P(x_i)\log \sum_{k=o}^{N} q_k r_{ik} \quad (3.5)$$

Ονομάζουμε αυξητική εντροπία το H, που είναι μία από τις γενικευμένες εντροπίες και έχει το ίδιο μετρικό σύστημα με την πληροφορία. Προτείνεται η βάση του λογαρίθμου να είναι το 2. Τότε το H σημαίνει τις φορές του διπλού κεφαλαίου.

Σύγκριση μεγιστοποίησης χαρτοφυλακίου - κώδικα επικοινωνίας

Σε σύγκριση με τις γενικευμένες εντροπίες, η αυξητική εντροπία στερείται ενός αρνητικού συμβόλου. Αν θεωρήσουμε $\bar{r_i} = 1/r_i$, τότε $\bar{r_i}$ σημαίνει πόσα δολλάρια σαν κεφάλαιο χρειάζονται τώρα για να πάρουμε ένα δολλάριο μετά την επένδυση. Έτσι έχουμε μια αυξητική εντροπία με αρνητικό σύμβολο:

$$H = -\sum_{i=1}^{N} P(x_i)\log \bar{r_i} \quad (4.1)$$

Από το πρώτο κεφάλαιο ξέρουμε ότι στη θεωρία πληροφορίας ο τύπος του μέσου μήκους κώδικα είναι

$$\bar{c} = \sum_{i=1}^{N} P(x_i)c_i = \sum_{i=1}^{N} P(x_i)\log m_i \qquad (4.2)$$

όπου c_i είναι το μήκος της κωδικοποίησης για το ψηφίο x_i, m_i είναι ο αριθμός των πιθανών κωδικών με μήκος c_i. Για παράδειγμα, τα δυαδικά ψηφία 0 και 1 χρησιμοποιούνται σαν κώδικες, η βάση του λογαρίθμου είναι 2. Για $c_i = 2$ υπάρχουν 4 πιθανοί κώδικες: 00, 01, 10, 11. Μεγιστοποίηση της κωδικοποίησης σημαίνει να αλλάζεις τους κανόνες κωδικοποίησης έτσι ώστε το $\bar{c}$ να προσεγγίζει το ελάχιστό του. Η θεωρία της πληροφορίας του Shannon (1948) μας λέει ότι καθώς το m_i απειρίζεται προσεγγίζει το $1/P(x_i)$ για κάθε i, το $\bar{c}$ απειριζόμενο προσεγγίζει την πιο μικρή του τιμή που είναι απλώς η εντροπία του Shannon:

$$H = -\sum_{i=1}^{N} P(x_i)\log P(x_i) \qquad (4.3)$$

Συγκρίνοντας την (3.5) με την (4.1) μπορούμε να δούμε ότι το r_i αντιστοιχίζεται με το m_i και ένα διάνυσμα (q_k) των αναλογιών επένδυσης αντιστοιχεί σε έναν κανόνα κωδικοποίησης. Για να μεγιστοποιηθεί ένα χαρτοφυλάκιο πρέπει να αλλάξουμε το (q_k) έτσι ώστε το H να φτάσει το μέγιστό του αντίθετα από το ελάχιστό του.

Αν ένα κανάλι επικοινωνίας μεταδίδει ένα κώδικα ανά μονάδα χρόνου και αυτοί οι κώδικες δε σχετίζονται ο ένας με τον άλλο, τότε το H σημαίνει την ταχύτητα της μεταδιδόμενης πληροφορίας. Ομοίως, σε όρους χαρτοφυλακίων, το H σημαίνει την αυξητική ταχύτητα του κεφαλαίου. Αν θεωρήσουμε $r_g = 2^H$, τότε το r_g δηλώνει το γεωμετρικό μέσο των αποδόσεων (geometric mean return, GMR). Αν ένα χαρτοφυλάκιο το οποίο έχει αυξητική πολλαπλότητα r_g κάθε χρονιά έχει τιμή M μετά T χρόνια, τότε

$$r_g^T = M\,,\quad T = \log M / \log r_g = (\log M)/H \quad (4.4)$$

Συγκρίνοντας με τον τύπο της Φυσικής: **χρόνος = απόσταση / χρόνος** μπορούμε να δούμε πιο ξεκάθαρα ότι το H έχει τη σημασία της ταχύτητας.

Βελτιστοποίηση των αναλογιών επένδυσης

Στα επόμενα θεωρούμε ότι συμβολίζουμε το $P(x_i)$ με P_i, το $(P_1, P_2, \ldots, P_W)$ με (P_i), το $(q_0, q_1, q_2, \ldots)$ από **q**=(q_k) και ούτω κάθε εξής. Αλλάζουμε το **q** έτσι ώστε το H να προσεγγίζει το μέγιστό του. Με τη μέθοδο των πολλαπλασιαστών Lagrangre ξέρουμε ότι το H φτάνει το μέγιστό του όταν

$$\sum_{i=1}^{W} \frac{P_i \Delta_{ik}}{r_0 + \sum_k \Delta_{ik} q_k} = 0, k = 1,2,\ldots,N \quad (5.1)$$

όπου $\Delta_{ik} = r_{ik} - r_0$ σημαίνει υπερβάλλουσες αποδόσεις. Το q^* βρίσκεται λύνοντας το πιο πάνω σύστημα N εξισώσεων. Καθώς το q^* είναι συνάρτηση των P_i και r_{ik} μπορούμε να γράψουμε **q***=**q**$^*((P_i),(r_{ik}))$. Όταν το r_{ik} δίνεται, το r_i είναι συνάρτηση του q^*. Έτσι μπορούμε να γράψουμε $r_i^* = r_i(q^*)$

Ας θεωρήσουμε το θέμα της βελτιστοποίησης της αναλογίας επένδυσης που αναλύσαμε πιο πριν. Σε αυτή την περίπτωση $N = 1, i = 1,2$ και ο τύπος (5.1) γίνεται

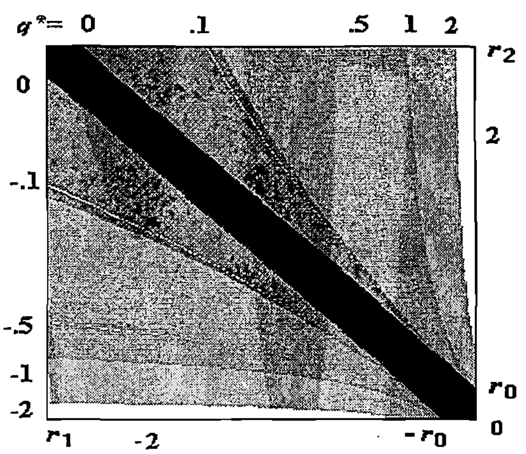
$$\frac{P_1 \Delta_1}{r_0 + \Delta_1 q} + \frac{P_2}{r_0 + \Delta_2 q} = 0 \quad (5.2)$$

όπου P_1, P_2 είναι οι πιθανότητες εμφάνισης της Α και Β πλευράς αντίστοιχα,

q είναι η αναλογία επένδυσης, $\Delta_1 = r_1 - r_0, \Delta_2 = r_2 - r_0$ είναι το ποσό αύξησης και μείωσης από το επιτόκιο αντίστοιχα. Από την (5.2) μπορούμε να λάβουμε τη βέλτιστη αναλογία επένδυσης

$$q^* = -\frac{P_1\Delta_1 + P_2\Delta_2}{\Delta_1\Delta_2} r_0 \quad (5.3)$$

Το πιο κάτω σχήμα δείχνει τη σχέση μεταξύ του q^* και των r_o, r_1, r_2



Ας σημειώσουμε ότι ο αριθμητής είναι απλώς τα αναμενόμενα πρόσθετα κέρδη.

Αν στον πιο πάνω τύπο θέσουμε $P_1 = P_2 = \frac{1}{2}, \Delta_1 = -1, \Delta_2 = 2, r_0 = 1$ βρίσκουμε $q^* = 0.25$. Λέγεται ότι η μέγιστη αναλογία επένδυσης είναι 25%. Η απόδοση θα είναι (1-1*0.25)(1+2*0.25)=1.125 φορές κατά μέσο όρο μετά κάθε 2 ρίψεις. Αν τα Δ_1, Δ_2 είναι και τα δύο θετικά ή αρνητικά , ο πιο πάνω τύπος δεν είναι κατάλληλος. Σε αυτές τις περιπτώσεις το q^* θα έπρεπε να είναι ίσο με το 1 ή το -1 αν η υπέρβαση δεν είναι επιτρεπτή.

Από την (5.3) μπορούμε να λάβουμε ένα ενδιαφέρον συμπέρασμα. Αν θεωρήσουμε $P_1 = P_2 = 1/2, \Delta_1 = -1, r_0 = 1$ τότε παίρνουμε

$$q^* = \frac{-1 + \Delta_2}{2\Delta_2} < \frac{1}{2} \quad (5.4)$$

Το οποίο συμβαίνει όταν κέρδη και απώλειες είναι ίσες. Αν οι απώλειες μπορεί να είναι μέχρι 100% τότε μην ποντάρεις περισσότερο από 50% του κεφαλαίου σου ασχέτως του πόσο υψηλό μπορεί να είναι το πιθανό κέρδος. Το συμπέρασμα είναι πολύ σημαντικό για μελλοντικές επενδύσεις. Πολλοί νεοεισερχόμενοι στις ανερχόμενες αγορές χάνουν όλα τα χρήματά τους πολύ γρήγορα επειδή οι αναλογίες επένδυσης δεν είναι καλά ελεγχόμενες και γενικά πολύ υψηλές.

Από την (5.3) $q^* \rangle 1$ και $q^* < 0$ είναι πιθανά. Το $q^* \rangle 1$ σημαίνει ότι θα ήταν καλύτερο να διακινδύνευες 100% ή $q^* = 100\%$, αν ο δανεισμός ή η υπέρβαση του πιστωτικού ορίου είναι επιτρεπτή, του κεφαλαίου σου. Το $q^* \langle 0$ σημαίνει ότι θα ήταν καλύτερο να μη διακινδύνευες τίποτα ή $|q^*| = 100\%$ υπέρβαση πωλήσεων αν αυτό είναι επιτρεπτό.

Για στοιχεία με πιθανότητα πολύ-αποδόσεων, μπορούμε να χρησιμοποιήσουμε τον υπολογιστή να πάρουμε τα μέγιστα ποσοστά στην (3.5) και μπορούμε να χρησιμοποιήσουμε προσεγγιστικούς τύπους.

Επίδραση της συνδιακύμανσης στο γεωμετρικό μέσο των αποδόσεων

Η επίδραση της συνδιακύμανσης των αποδόσεων του αξιόγραφου στην αξία της επένδυσης ενός χαρτοφυλακίου έχει αναλυθεί πολύ καλά από το Markowitz.(1959, 1991). Τώρα την εξηγούμε από τη νέα θεωρία από την οποία μπορούμε να λάβουμε παρόμοια συμπεράσματα για το πώς να μειώσουμε τον κίνδυνο και να βελτιώσουμε την αυξανόμενη ταχύτητα του κεφαλαίου μεγιστοποιώντας τις αναλογίες επένδυσης.

Παράδειγμα 2. Αν θεωρήσουμε ότι οι υπερβάλλουσες αποδόσεις μιας μελλοντικής επένδυσης αναπαρίστανται από τις δύο όψεις ενός ζευγαριού νομισμάτων. Κερδίζεις 300% αν και τα 2 νομίσματα δείξουν την όψη Α, χάνεις 200% αν και τα 2 νομίσματα δείξουν την όψη Β και κερδίζεις 100% αν τα 2 νομίσματα δείξουν διαφορετικές όψεις. Ας υποθέσουμε ότι υπάρχουν 2 προθεσμιακές συναλλαγές των οποίων οι αποδόσεις καθορίζονται όπως πιο πάνω, αλλά πιθανόν 1 ή 2 νομίσματα είναι πιο συχνά χρησιμοποιούμενα. Υπολόγισε το Γεωμετρικό Μέσο των Κερδών όπως επενδύεις όλο το κεφάλαιο (50% για κάθε ένα) και επένδυσε στις μέγιστες αναλογίες.

Είναι προφανές ότι όσο πιο μεγάλος είναι ο συντελεστής συνδιακύμανσης τόσο πιο άσχημο θα είναι το αποτέλεσμα για το χαρτοφυλάκιο.

Σύνοψη

Στο κεφάλαιο αυτό παρουσιάστηκε ένα μοντέλο για τη μεγιστοποίηση των αναλογιών επένδυσης ενός χαρτοφυλακίου. Αυτό έγινε συγκρίνοντας τη μεγιστοποίηση του χαρτοφυλακίου με αυτή ενός κώδικα επικοινωνίας.
Με βάση τον ορισμό της εντροπίας η αυξητική ταχύτητα του κεφαλαίου δίνεται από τη σχέση

$$H = -\sum_{i=1}^{N} P(x_i)\log \overline{r_i}$$

Όπου το $P(x_i)$ είναι η πιθανότητα τα στοιχεία του χαρτοφυλακίου να έχουν το x_i διάνυσμα τιμών. Το $\overline{r_i}$ δείχνει το ποσό των χρημάτων που πρέπει να έχουμε ως κεφάλαιο για να υπάρξει κέρδος 1 €.
Με τη μέθοδο πολλαπλασιαστών Lagrange βρίσκουμε ότι η αυξητική εντροπία φτάνει τη μέγιστή τιμή της όταν

$$\sum_{i=1}^{W} \frac{P(x_i)}{r_0 + \sum_k \Delta_{ik} q_k} = 0, k = 1,2,\ldots,N$$

Όπου $q_k, k = 1,2,\ldots,N$ είναι η αναλογία του επενδεδυμένου κεφαλαίου στο k-οστό στοιχείο, N είναι τα στοιχεία του χαρτοφυλακίου και $\Delta_{ik} = r_{ik} - r_0$ είναι οι υπερβάλλουσες αποδόσεις.
Ένα πρακτικό συμπέρασμα είναι ότι αν οι απώλειες από την επένδυση είναι περοσσότερες από 100% του κεφαλαίου τότε δεν πρέπει να ποντάρουμε περισσότερο από το 50% του κεφαλαίου.

Κεφάλαιο 4
Ερευνητικές προεκτάσεις

Εισαγωγή

Στο κεφάλαιο αυτό προσπαθούμε να κάνουμε μια σύγκριση της νέας θεωρίας του χαρτοφυλακίου που αναλύσαμε στο προηγούμενο κεφάλαιο με τις ήδη υπάρχουσες θεωρίες. Επίσης γίνεται μια εφαρμογή της νέας μεθόδου στην εκτίμηση των προβλέψεων.

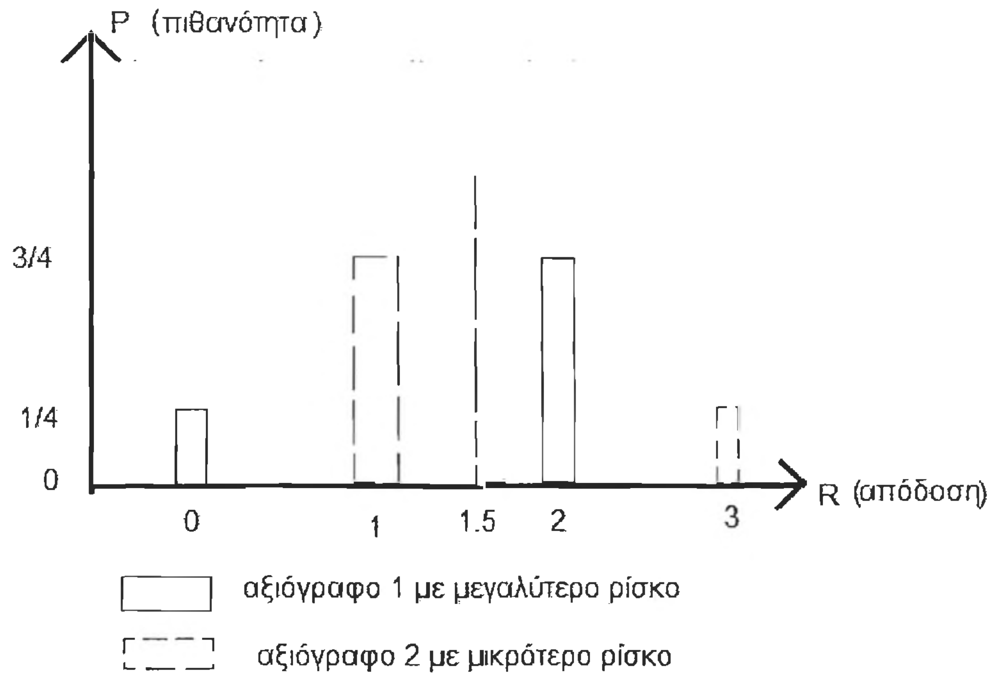
Σύγκριση με τη θεωρία του Markowitz

Η νέα θεωρία υποστηρίζει τα συμπεράσματα του Markowitz για τη μείωση του κόστους επένδυσης με το πιο αποτελεσματικό χαρτοφυλάκιο. Τα διαφορετικά είναι

1) Η νέα θεωρία αναφέρεται στο γεωμετρικό μέσο των κερδών σαν αντικειμενικό κριτήριο για μεγιστοποίηση του χαρτοφυλακίου και παρέχει μέγιστες αναλογίες επένδυσης για την πιο γρήγορη αύξηση του κεφαλαίου σε πιο μεγάλο χρονικό διάστημα.

2) Η νέα θεωρία χρησιμοποιεί τις πιθανότητες κέρδους και απωλειών ή τη συνάρτηση χρησιμότητας $u = u(q)$ αντί της αναμενόμενης τιμής και της τυπικής απόκλισης των αποδόσεων σαν περιγραφή της αξίας επένδυσης ενός αξιόγραφου ή ενός χαρτοφυλακίου.

Παράδειγμα 3. Οι τρέχουσες τιμές δύο αξιόγραφων 1 και 2 είναι 1 δολάριο και οι δύο. Η πιθανή τιμή του Α στο μέλλον είναι 0 και 2 με πιθανότητα $1/4$ και $3/4$. Η τιμή του αξιόγραφου Β στο μέλλον έχει την ίδια αναμενόμενη τιμή (1.5) και τυπική απόκλιση (0.886) αλλά αντίστροφη πιθανότητα κατανομής. Η ανάλυση των αξιών επένδυσης των 2 αξιόγραφων δείχνεται στο πιο κάτω σχήμα, στο οποίο το επιτόκιο των τραπεζών αμελείται.



Πίνακας 3

	Αναμενόμενη τιμή	Τυπική απόκλιση	Μέσο Σύνθετο επιτόκιο όταν το ποντάρισμα είναι 100%	Αναλογία Μεγιστοποίησης	Μέσο Σύνθετο επιτόκιο μετά τη μεγιστοποίηση
Αξιόγραφο 1	0.5	0.886	-100%	50	15%
Αξιόγραφο 2	0.5	0.886	32%	>=100	>=32%

Η μέγιστη αναλογία $q^{*}\rangle = 100\%$ σημαίνει ότι αν η υπέρβαση του πιστωτικού ορίου είναι επιτρεπτή θα ήταν καλύτερο να το υπερβείς για να αγοράσεις και όσο πιο πολύ το υπερβείς τόσο το καλύτερο. Στη θεωρία του Markowitz, η A και η B πλευρά έχουν την ίδια αξία επένδυσης. Αλλά στη νέα θεωρία η Β είναι πιο καλή από την Α. Για επενδύσεις με πιθανότητα απωλειών σε μεγάλη αναλογία, όπως όψιον, προθεσμιακές συναλλαγές, δάνεια και ασφαλιστική κάλυψη το πιο πάνω ελάττωμα στη θεωρία του Markowitz είναι πιο ξεκάθαρο.

Σύγκριση με τον τύπο του Arrow

Ο Arrow όρισε την αξία της πληροφορίας ή τη χρησιμότητα της πληροφορίας σαν αύξηση της χρησιμότητας αφού παρέχεται η πληροφορία. Η συνάρτηση χρησιμότητας είναι

$$U = \sum_i P_i U\,(a_i r_i) = \sum_i P_i \log(a_i r_i) = \sum_i P_i \log a_i + \sum_i P_i \log r_i \qquad (9.1)$$

Όπου r_i συμβολίζει την i-οστή απόδοση μιας επένδυσης ή ενός πονταρίσματος.

P_i είναι η ανταποκρινόμενη πιθανότητα

a_i είναι η αναλογία του κεφαλαίου επενδυμένη στην i- οστή απόδοση αντί για το αξιόγραφο

$U(a_i r_i)$ είναι η χρησιμότητα που αποκτάται από τον επενδυτή όταν συμβαίνει η i-οστή απόδοση

Η λογαριθμική συνάρτηση χρησιμοποιείται γιατί ο Arrow τη θεώρησε σαν συνάρτηση χρησιμότητας χωρίς να υπολογίζει το γεωμετρικό μέσο των αποδόσεων.

Κάτω από το όριο $\sum_i a_i = 1$, το U πλησιάζει το μέγιστο του καθώς $(a_i) = (P_i)$. Το μέγιστο είναι

$$U^{*} = \sum P_i \log P_i + \sum_i P_i \log r_i \qquad (9.2)$$

Αφού η πληροφορία παρέχεται, ο επενδυτής γνωρίζει ποια απόδοση τελικά θα συμβεί και για αυτό επενδύει όλο το κεφάλαιό του σε αυτή την απόδοσή του. Τότε

$$U^{**} = \sum_i P_i \log r_i \qquad (9.3)$$

Η αξία της πληροφορίας είναι η εντροπία του Shannon, δηλαδή

$$V = U^{**} - U^{*} = -\sum_i P_i \log P_i = H(X) \quad (9.4)$$

Ο ορισμός του Arthur για την αξία της πληροφορίας βασίζεται στη θεωρία του Arrow:

1) για δεδομένες προβλέψεις των αποδόσεων υπάρχει μια αντικειμενική μέγιστη αναλογία επένδυσης

2) η αξία της πληροφορίας είναι ίση στην αύξηση της χρησιμότητας αφού παρέχεται η πληροφορία. Ωστόσο το μοντέλο επένδυσης που αναπτύσσεται σε αυτό το φυλλάδιο είναι πολύ διαφορετικό. Από την άποψη του συγγραφέα το μοντέλο επένδυσης που αναπτύσσεται από τον Arthur είναι πολύ περίεργο. Μπορούμε να επενδύσουμε σε ένα αξιόγραφο ή ένα αντικείμενο: αλλά πως μπορούμε να επενδύσουμε σε κάποια απόδοση ενός αξιόγραφου ή ενός αντικειμένου; Η εξίσωση (9.1) απαιτεί το κάθε r_i να είναι θετικό, το οποίο σημαίνει ότι επενδύσεις με πιθανές απώλειες δεν είναι επιτρεπτές καθώς ο λογάριθμος ενός αρνητικού αριθμού δεν ορίζεται. Είναι αυτή η επένδυση ή το ποντάρισμα κοινό όπως το βλέπουμε στην καθημερινή ζωή; Αλλά είναι ακόμη δύσκολο να καταλάβουμε ότι όταν $r_i \langle 1$ καταλήγουμε σε αρνητική χρησιμότητα. Για αυτούς τους λόγους η εφαρμογή του τύπου του Arrow για την αξία της πληροφορίας είναι δύσκολη.

Η αξία της πληροφορίας βασισμένη στην αυξητική εντροπία

Στη γενικευμένη θεωρία της πληροφορίας η πιθανότητα διακρίνεται στην υποκειμενική πιθανότητα, η οποία υποκειμενικά προβλέπεται από κάποιον, αντικειμενική πιθανότητα, στην οποία ένα γεγονός πραγματικά συμβαίνει και στη λογική πιθανότητα. Η λογική πιθανότητα μιας πρόβλεψης ή μιας πρότασης λέγεται επίσης διάστημα εμπιστοσύνης και παίρνει πραγματικές τιμές στο διάστημα $[0,1]$

Ας θεωρήσουμε ότι Χ είναι μια τυχαία μεταβλητή παίρνοντας μια τιμή από μια σειρά γεγονότων $A=\{X_1,X_2,\ldots\}$ και Υ μια τυχαία μεταβλητή παίρνοντας τιμή από μια σειρά προτάσεων $B=\{y_1,y_2,\ldots\}$. Η Υ μεταδίδει πληροφορίες για τη Χ. Τα γεγονότα τα οποία κάνουν μια πρόταση y_j να είναι αληθείς σχηματίζουν μια σειρά A_j, η οποία είναι μια συγκεντρωτική σειρά γιατί η λογική πιθανότητα των y_j σαν συνάρτηση των A_j παίρνει μια τιμή από $[0,1]$. η πληροφορία για το x_i και η μέση πληροφορία για το Χ μεταδιδόμενο από το y_j είναι

$$I(x_i;y_j)=\log\frac{Q(x_i|A_j)}{Q(x_i)} \qquad (8.1)$$

$$I(X;y_j)=\sum_i P(x_i|y_j)\log\frac{Q(x_i|A_j)}{Q(x_i)} \qquad (8.2)$$

Όπου $P(x_i|y_j)$ είναι η υπό συνθήκη πιθανότητα, $Q(x_i)$ είναι η εκ των προτέρων υποκειμενική πιθανότητα και $Q(x_i|A_j)$ είναι η εκ των υστέρων υποκειμενική πιθανότητα που συνάγεται από το y_j. Γίνεται από το παρατηρούμενο στοιχείο $Z=z'$ και η μέση πληροφορία είναι

$$I(X;y_j(z'))=\sum_i P(x_i|z')\log\frac{Q(x_i|A_j)}{Q(x_i)} \qquad (8.3)$$

Δεδομένου ενός διανύσματος πιθανοτήτων κατανομή *q*=(Q_i)=Q(*X*) και ενός πίνακα αποδόσεων (r_{ik}) υποθέτουμε *P(X)*= Q(*X*) για να αποκτήσουμε το διάνυσμα μέγιστων αναλογιών επένδυσης $\boldsymbol{q}^*=\boldsymbol{q}^*((Q_i),(r_{ik}))$ από την (3.5) . Με τις αλλαγές πρόβλεψης από (Q_i) σε $(Q_{ij})=Q(X/A_j)$, το $\mathbf{q}^*$ γίνεται $\boldsymbol{q}^{**}=\boldsymbol{q}^{**}((Q_{ij}),(r_{ik}))$. Ορίζουμε την αυξητική εντροπία μετά την πληροφορία που παρέχεται σαν αξία πληροφορίας

$$V(X;y_j(z')) = \sum_i P(x_i|z')\log\frac{r_i(q^{**})}{r_i(q^*)} \qquad (8.4)$$

Μπορεί να δειχθεί ότι $V(X;y_j(z'))$ και $I(X;y_j(z'))$ έχουν παρόμοια μαθηματική κατασκευή. Η αξία της πληροφορίας του y_j όταν συμβαίνει το x_i είναι

$$v(x_i;y_j(z')) = \log\frac{r_i(q^{**})}{r_i(q^*)} \qquad (8.5)$$

Και η πληροφορία και η αξία της πληροφορίας είναι σχετικές και εξαρτώνται από το πόσο ένας λήπτης πληροφορίας καταλαβαίνει μια πρόβλεψη. Η σχετικότητα της αξίας πληροφορίας οφείλεται στον πίνακα αποδόσεων (r_{ik}) είναι διαφορετική για διαφορετικούς επενδυτές, με διαφορετικά όρια επενδυτικών εργαλείων. Η ασυμμετρία της πληροφορίας και της αξίας της πληροφορίας πηγάζει από τους ίδιους λόγους.

Εφαρμογή σε εκτιμήσεις προβλέψεων

Ο τύπος (8.4) μπορεί να χρησιμοποιηθεί για να μεγιστοποιήσει προβλέψεις. Ας θεωρήσουμε $P_i = P(x_i/z')$. Με δεδομένο το (P_i) ψάχνουμε την πιθανότητα πρόβλεψης $Q(X/y_j = y^*)$ η οποία μεγιστοποιεί την $V(X;y_j(z'))$. Τότε y^*είναι η μέγιστη πρόβλεψη. Δεν είναι πολύ δύσκολο να αποδείξουμε ότι όταν $(Q_{ij}) = (P_i)$ ή να το πούμε διαφορετικά η πρόβλεψη είναι ακριβής τα $I(X; y_j(z'))$ και $V(X; y_j(z'))$ φτάνουν τα μέγιστά τους την ίδια στιγμή. Αν η σειρά προβλέψεων B είναι περιορισμένη έτσι ώστε να μην υπάρχει καμία πρόβλεψη που να ισχύει $(Q_{ij}) = (P_i)$. Σε αυτήν την περίπτωση η πρόβλεψη η οποία φέρνει το (Q_{ij}) πιο κοντά στο (P_i) είναι η μέγιστη πρόβλεψη.

Παράδειγμα 4 Ένα περιεχόμενο αποθέματος X παίρνει μια τιμή από το set $A = [100,110,120,130,\ldots]$. Μια σειρά από προβλέψεις για το X είναι $B = \{y_j =$"X είναι x_j"$\mid j = 1,2,\ldots\}$. Η λογική πιθανότητα του y_j είναι $Q(A_j / x_i)$. Για παράδειγμα είναι συνάρτηση με x_j στο κέντρο της και σ σαν τυπική της απόκλιση

$$Q(A_j \mid x_i) = \exp\left[-\frac{(x_i - x_j)^2}{2\sigma^2}\right] \qquad (10.1)$$

Το τρέχον περιεχόμενο αποθέματος είναι x_0. Η προηγούμενη πιθανότητα πρόβλεψης είναι $Q(X)$ όπως μια συνάρτηση κανονικής κατανομής με x_0 στο κέντρο. Ακόμα ένα γεγονός συμβαίνει στην αντικειμενική πιθανότητα $P(X \mid z')$. Υπολόγισε την αξία της πληροφορίας πρόβλεψης $y_j =$"X είναι x_j" και τη μέγιστη πρόβλεψη y^*.

Λύση. Με το γενικευμένο τύπο του Bayes έχουμε

$$Q(x_i \mid A_j) = Q(x_i) Q(A_j \mid x_i) / Q(A_j) \qquad (10.2)$$

στην οποία το $Q(A_j)$ σημαίνει τη λογική πιθανότητα του predicate $y_j(.)$ και

$$Q(A_j) = \sum_i Q(x_i) Q(A_j \mid x_i) \qquad (10.3)$$

Από την αυξητική εντροπία

$$H_0 = \sum_i Q(x_i) \log r_i = \sum_i Q(x_i) \log[1 + q(x_i - x_0)/x_0] \quad (10.4)$$

Λαμβάνουμε τη μέγιστη αναλογία q^*.
Από την αυξητική εντροπία

$$H_1 = \sum_i Q(x_i \mid A_j) \log r_i = \sum_i Q(x_i \mid A_j) \log[1 + q(x_i - x_0)/x_0] \quad (10.5)$$

λαμβάνουμε τη μέγιστη αναλογία q^{**}. Η αξία της πληροφορίας του y_j είναι ως εκ τούτου

$$V(X; y_j(z')) = \sum_i P(x_i \mid z') \log \frac{x_0 + q^{**}(x_i - x_0)}{x_0 + q^*(x_i - x_0)}. \qquad (10.6)$$

Για κάθε y_j στο B μπορούμε να έχουμε το αντίστοιχο q^{**}. Το y_j το οποίο κάνει την αξία της πληροφορίας $V(X; y_j(z'))$ ή την εκ των υστέρων αυξητική εντροπία

$$H = \sum_i P(x_i \mid z') \log[1 + q^{**}(x_i - x_0)/x_0] \qquad (10.7)$$

να φτάνει το μέγιστό της είναι η μέγιστη πρόβλεψη y^*.

Επίλογος

Στην εργασία αυτή είδαμε αρχικά την έννοια της Εντροπίας μέσα από το πρίσμα της Θεωρίας Πληροφορίας. Στη συνέχεια αναλύσαμε ένα μοντέλο για την εύρεση των μέγιστων αναλογιών επένδυσης ενός χαρτοφυλακίου. Τελικώς βρέθηκε ότι οι μέγιστες αναλογίες επένδυσης δίνονται από την επίλυση του παρακάτω συστήματος εξισώσεων

$$\sum_{i=1}^{W} \frac{P(x_i)}{r_0 + \sum_k \Delta_{ik} q_k} = 0, k = 1,2,\ldots,N$$

Ένα ενδιαφέρον συμπέρασμα που βγαίνει είναι ότι αν οι απώλειες από την επένδυση είναι περοσσότερες από 100% του κεφαλαίου τότε δεν πρέπει να ποντάρουμε περισσότερο από το 50% του κεφαλαίου.

Το επόμενο βήμα είναι να επιλεγούν τα κατάλληλα δεδομένα και να πραγματοποιηθεί μια σειρά εξομοιώσεων (simulation). Αυτό είναι βέβαια ένα ανοιχτό ερευνητικό πεδίο με πολλές προεκτάσεις. ■

Βιβλιογραφία - Πηγές

[1]

Arrow, K.J. : The Economics of Information

[2]

Latane, h.A. and Tuttle D.L. : Criterion for portfolio building, The Journal of Finance

[3]

LU, Chen-Guant : Applying a generalized information theory to assessment and optimization of forecasts

[4]

Markowitz, M. H. : Foundation of Portfolio Theory, The Journal of Finance

[5]

Shannon, C. E. : A mathematical theory of communication

[6]

Δημήτρης Φιλιππάκος : Διερεύνηση κινδύνου σε κύρια χρηματοπιστωτικά προϊόντα

[7]

Ανδριανού Ελένη: Αποτίμηση Κινδύνου, Πιστωτικός Κίνδυνος

[8]

Jing Chen: Information, entropy and evolutionary finance

[9]

Cyprian J. Bruck : The Approximate Entropy of Various Financial Asset Prices

[10]

Andreia Dionísio, Rui Menezes and Diana A. Mendes : Uncertainty analysis in financial markets: can entropy be a solution?