ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

ΚΑΤΑΝΟΜΗ PARETO ΚΑΙ ΧΑΡΑΚΤΗΡΙΣΜΟΙ ΑΥΤΗΣ

Δημήτριος Παν. Σπαθάρας

ΕΡΓΑΣΙΑ

Που υποβλήθηκε στο Τμήμα Στατιστικής

του Οικονομικού Πανεπιστημίου Αθηνών

ως μέρος των απαιτήσεων για την απόκτηση

Μεταπτυχιακού Διπλώματος

Συμπληρωματικής Ειδίκευσης στη Στατιστική

Μερικής Παρακολούθησης (Part-time)

Αθήνα

Νοέμβριος 2005

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

“ ΚΑΤΑΝΟΜΗ PARETO ΚΑΙ ΧΑΡΑΚΤΗΡΙΣΜΟΙ ΑΥΤΗΣ”

Δημήτριος Παν. Σπαθάρας

ΕΡΓΑΣΙΑ

Που υποβλήθηκε στο Τμήμα Στατιστικής

του Οικονομικού Πανεπιστημίου Αθηνών

ως μέρος των απαιτήσεων για την απόκτηση

Μεταπτυχιακού Διπλώματος

Συμπληρωματικής Ειδίκευσης στη Στατιστική

Μερικής Παρακολούθησης (Part-time)

Αθήνα
Ιούνιος 2005

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Εργασία που υποβλήθηκε ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Συμπληρωματικής Ειδίκευσης στη Στατιστική
Μερικής Παρακολούθησης (Part-time)

ΚΑΤΑΝΟΜΗ PARETO ΚΑΙ ΧΑΡΑΚΤΗΡΙΣΜΟΙ ΑΥΤΗΣ

Δημήτριος Παν. Σπαθάρας

Υπεύθυνο μέλος ΔΕΠ:
Κ. Δημάκη
Επίκουρη Καθηγήτρια

Ο Διευθυντής Μεταπτυχιακών Σπουδών

Μιχαήλ Ζαζάνης
Καθηγητής

Αφιερώνεται
στη μνήμη του πατέρα μου
που έφυγε τόσο νωρίς.

ΕΥΧΑΡΙΣΤΙΕΣ

Πρώτα απ' όλα αισθάνομαι την ανάγκη να εκφράσω τις θερμές και ειλικρινείς μου ευχαριστίες προς την επιβλέπουσα καθηγήτριά μου κ. Αικατερίνη Δημάκη, επίκουρο καθηγήτρια του Τμήματος Στατιστικής του Οικονομικού Πανεπιστημίου Αθηνών. Η ηθική της συμπαράσταση ήταν αδιάκοπη, οι επιστημονικές συμβουλές και υποδείξεις της ήταν πολύτιμες και η επιστημονική της καθοδήγηση ήταν συνεχής και ακούραστη καθ' όλη τη διάρκεια της εκπόνησης της παρούσας εργασίας.

Θα ήταν παράλειψη στο σημείο αυτό να μην εκφράσω τις ευχαριστίες μου προς όλους του διδάσκοντες καθηγητές του Μεταπτυχιακού Τμήματος Στατιστικής του Οικονομικού Πανεπιστημίου Αθηνών. Σε όλους αυτούς οφείλω την επιστημονική μου κατάρτιση κατά τις μεταπτυχιακές μου σπουδές. Θερμά ευχαριστώ επίσης τους υπεύθυνους του εργαστηρίου στατιστικής, τη γραμματεία και το προσωπικό της βιβλιοθήκης.

Τέλος πολλά ευχαριστώ οφείλω στη σύζυγό μου για την συμπαράσταση που μου παρείχε καθ' όλη τη διάρκεια των μεταπτυχιακών μου σπουδών και κατά το διάστημα εκπόνησης αυτής της εργασίας. Η κατανόηση και η υπομονή της ήταν καθοριστικό στοιχείο βοήθειας τις ατέλειωτες ώρες που χρειάστηκε να αφιερώσω. Ευχαριστώ επίσης όλους τους συναδέλφους και φίλους που με στήριξαν με τον ένα ή τον άλλο τρόπο κατά τη διάρκεια εκπόνησης αυτής της εργασίας.

ΒΙΟΓΡΑΦΙΚΟ ΣΗΜΕΙΩΜΑ

Αποφοίτησα το 1976 από το $2^{ο}$ εξατάξιο Γυμνάσιο Αρένων στη Λαμία όπου διαμένω μέχρι σήμερα. Στη συνέχεια φοίτησα στο Μαθηματικό τμήμα της σχολής Θετικών Επιστημών του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης και απέκτησα το πτυχίο του Μαθηματικού. Από το 1983 έως το 1994 εργάστηκα ως φροντιστής μαθηματικός και δίδαξα σε μαθητές λυκείου και σε υποψηφίους ανωτάτων και ανωτέρων εκπαιδευτικών ιδρυμάτων. Από το 1995 μέχρι σήμερα εργάζομαι σε δημόσια γυμνάσια και λύκεια της περιοχής μου. Έχω διδάξει επί σειρά ετών μαθηματικά σε μαθητές γυμνασίου και άλγεβρα, ευκλείδεια γεωμετρία, τριγωνομετρία, αναλυτική γεωμετρία, θεωρία αριθμών, γραμμική άλγεβρα, μιγαδικούς αριθμούς, ανάλυση, στατιστική και πιθανότητες σε μαθητές λυκείου.

Επιλέχτηκα και φοίτησα τα ακαδημαϊκά έτη 2002-2003 και 2003-2004 στο διετές Μεταπτυχιακό Πρόγραμμα Συμπληρωματικής Ειδίκευσης στη Στατιστική του Οικονομικού Πανεπιστημίου Αθηνών. Στα πλαίσια του προγράμματος αυτού εκπονώ την παρούσα εργασία. Τα ερευνητικά μου ενδιαφέροντα επικεντρώνονται κυρίως στη θεωρία Πιθανοτήτων και Κατανομών και γενικότερα στην επίλυση προβλημάτων στα οποία υπεισέρχεται ο παράγων της τυχαιότητας.

Δημήτρης Σπαθάρας

ABSTRACT

Dimitris Spatharas

Pareto Distribution and its Characterizations

June 2005

In order to present the personal income in an economy we use the distribution curve which is the result of the relation between the amount of the personal income and the percentage of the people possessing this income. Pareto distribution approaches the curve above quite efficiently. Its fatherhood is attributed to the Swiss professor of economics, Vilfredo Paretro, who at the end of the 19^{th} century formulated the aspect that the logarithm of the percentage of people possessing income greater than or equal to some value is a negative sloped linear function of the logarithm of this value.

In this project after the historical elements, we give the definition of the Pareto distribution and we present its functional relations with other known distributions. Then we give the definitions of the generalized Pareto distributions and we present their relations with each other. We estimate the moments and measures of the Pareto distribution and the alternative measures of dispersion and inequality which are of great importance to the practical applications of the specific distribution. We also mention the order statistics and apply this theory to the Pareto distribution. In the last part of this project we present the characterizations of the Pareto distribution. The practical significance of the characterizations is great because the relative theorems allow us, having a random sample of the population, to decide whether the sample comes from a population that follows Pareto distribution.

ΠΕΡΙΛΗΨΗ

Δημήτρης Σπαθάρας

Κατανομή Pareto και Χαρακτηρισμοί αυτής

Ιούνιος 2005

Για να παρουσιάσουμε τα προσωπικά εισοδήματα σε μια οικονομία χρησιμοποιούμε την καμπύλη κατανομής που είναι αποτέλεσμα της σχέσης μεταξύ του ύψους του προσωπικού εισοδήματος και του ποσοστού των ατόμων που κατέχουν αυτό το εισόδημα. Η κατανομή Pareto προσεγγίζει αρκετά ικανοποιητικά την παραπάνω καμπύλη. Η πατρότητά της οφείλεται στον Ελβετό καθηγητή των οικονομικών, Vilfredo Pareto, ο οποίος περί τα τέλη του $19^{ου}$ αιώνα διατύπωσε τη άποψη ότι ο λογάριθμος του ποσοστού των ατόμων που έχουν εισόδημα μεγαλύτερο ή ίσο από ένα ποσό είναι μια γραμμική συνάρτηση με αρνητική κλίση του λογαρίθμου του ποσού αυτού.

Στην παρούσα εργασία μετά τα ιστορικά στοιχεία, δίνουμε τον ορισμό της κατανομής Pareto και παρουσιάζουμε τις συναρτησιακές της σχέσεις με άλλες γνωστές κατανομές. Κατόπιν δίνουμε τους ορισμούς των γενικευμένων κατανομών Pareto και παρουσιάζουμε τις μεταξύ τους σχέσεις. Υπολογίζουμε τις ροπές και τα μέτρα της κατανομής Pareto καθώς επίσης και τα εναλλακτικά μέτρα διασποράς και ανισότητας που έχουν μεγάλη σημασία για τις πρακτικές εφαρμογές της συγκεκριμένης κατανομής. Επίσης κάνουμε αναφορά στα διατεταγμένα στατιστικά μιας συνεχούς κατανομής και εφαρμόζουμε την θεωρία αυτή στην κατανομή Pareto. Στο τελευταίο τμήμα της εργασίας αυτής παρουσιάζουμε τους χαρακτηρισμούς της κατανομής Pareto. Η πρακτική σημασία των χαρακτηρισμών είναι σημαντική διότι τα σχετικά θεωρήματα μας επιτρέπουν έχοντας ένα τυχαίο δείγμα του πληθυσμού να αποφανθούμε αν το δείγμα προέρχεται από πληθυσμό που ακολουθεί κατανομή Pareto.

ΚΑΤΑΛΟΓΟΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Σελίδα

Σελίδα

Κεφάλαιο 3° – Κατανομές διατεταγμένων στατιστικών.

Κεφάλαιο 4° – Χαρακτηρισμοί της κατανομής Pareto.

ΚΑΤΑΛΟΓΟΣ ΓΡΑΦΗΜΑΤΩΝ

ΚΕΦΑΛΑΙΟ 1^{Ο}

Εισαγωγή.

Στο κεφάλαιο αυτό μετά από μια σύντομη ιστορική αναφορά της κατανομής Pareto, δίνουμε τον ορισμό της και παρουσιάζουμε τις γραφικές της παραστάσεις. Στη συνέχεια παραθέτουμε τις σχέσεις της με άλλες γνωστές κατανομές και τέλος παρουσιάζουμε τις γενικευμένες μορφές αυτής καθώς επίσης και τις κατανομές στατιστικών από κατανομή Pareto.

1.1 Ιστορία της κατανομής Pareto.

Η κατανομή Pareto οφείλει την ονομασία της στον Ιταλικής καταγωγής Ελβετό καθηγητή των οικονομικών Vilfredo Pareto (1848-1923). Περί το τέλος του προηγούμενου αιώνα (1897), η μελέτη διαφόρων εμπειρικών κατανομών του εισοδήματος των νοικοκυριών διαφόρων χωρών, επέτρεψε στον οικονομολόγο Pareto να διατυπώσει ένα κανόνα, γνωστό στη βιβλιογραφία ως "Pareto's Law". Ο κανόνας αυτός έχει ως εξής:

Vilfredo Pareto

"Ο λογάριθμος του ποσοστού των ατόμων που έχουν εισόδημα μεγαλύτερο ή ίσο από κάποια τιμή x, είναι μια γραμμική συνάρτηση με αρνητική κλίση του λογαρίθμου της τιμής x."

Συμβολικά μπορούμε να γράψουμε

$$\log N = \log A - \alpha \log x \tag{1.1.1}$$

όπου N το ποσοστό των ατόμων του πληθυσμού που έχουν εισόδημα μεγαλύτερο ή ίσο από την τιμή x και α, A θετικές σταθερές.

Αν το εισόδημα X θεωρηθεί τυχαία μεταβλητή η οποία μπορεί να πάρει οποιαδήποτε τιμή στο διάστημα $[\theta,+\infty)$, όπου θ είναι το ελάχιστο των εισοδημάτων του πληθυσμού και $F(x)=P(X\leq x)$ είναι η συνάρτηση κατανομής της μεταβλητής αυτής, τότε για τη συνάρτηση επιβίωσης $\overline{F}(x)=1-F(x)=P(X>x)$ έχουμε

$$\overline{F}(x)=P(X>x)=Ax^{-\alpha}, \quad \alpha>0, \ A>0 \tag{1.1.2}$$

Πράγματι από τον τύπο (1.1.1) έχουμε

$$\begin{aligned}
\log[1-F(x)] = \log A - \alpha\log x \quad &\Leftrightarrow \quad \log[1-F(x)] = \log A + \log x^{-\alpha} \\
&\Leftrightarrow \quad \log[1-F(x)] = \log Ax^{-\alpha} \\
&\Leftrightarrow \quad 1-F(x) = Ax^{-\alpha} \\
&\Leftrightarrow \quad \overline{F}(x) = Ax^{-\alpha}
\end{aligned}$$

Δηλαδή αν θ είναι το ελάχιστο των εισοδημάτων του πληθυσμού τότε η σχετική συχνότητα των νοικοκυριών των οποίων το εισόδημα X είναι μεγαλύτερο από ένα οποιοδήποτε ποσό x, όπου $x\geq\theta$ δίνεται με ικανοποιητική προσέγγιση από τη σχέση (1.1.2)

Ο Pareto είχε την αίσθηση ότι ο νόμος αυτός ήταν γενικός και πάντοτε ανεξάρτητος από την φορολογία και τις κοινωνικές και πολιτικές συνθήκες μιας κοινωνίας. Μολονότι ο κανόνας του Pareto είναι εμπειρικός, έχει επαληθευτεί σε πολλές χώρες για τις ομάδες των ατόμων με υψηλό εισόδημα, αλλά με διαφορετικές τιμές της παραμέτρου α. Πριν από τους παγκόσμιους πολέμους η σταθερά αυτή ήταν περίπου 1,5 και μετά αυξήθηκε σε περισσότερο από 2,5. Στο Bresciani-Turroni (1939) διαβάζουμε την σημείωση:

> **"Οι πραγματικές τιμές της παραμέτρου *α* βρίσκονται σε συγκριτικά στενά όρια. Κυμαίνονται γύρω στο 1,5 και οι αποκλίσεις από αυτή την τιμή φαίνεται να είναι ένα πρόβλημα από ελλιπές στατιστικό υλικό παρά από την πραγματική αιτία."**

Μερικά χρόνια αργότερα ο Cramer (1971) σημειώνει:

"Οι τιμές της παραμέτρου α έχουν αυξηθεί από 1,6 έως 1,8 που ήταν τον 19ο αιώνα σε 1,9 έως 2,1 που είναι στις ανεπτυγμένες χώρες την παρούσα στιγμή."

Η μη σταθερότητα της παραμέτρου α έδωσε το έναυσμα ποικίλων συζητήσεων για το κατά πόσον η παραπάνω συνάρτηση κατανομής του εισοδήματος είναι ένας "νόμος" ή όχι. Ο ίδιος ο Pareto είχε πει:

"Αυτός ο νόμος είναι εμπειρικός, μπορεί να μη ισχύει πάντοτε, ή να μην ισχύει σε όλες τις κοινωνίες. Προς το παρών αφού τα στατιστικά δεδομένα που έχουμε δεν παρουσιάζουν εξαιρέσεις από το νόμο, μπορούμε ένεκα τούτου να δεχτούμε προσωρινά ότι ισχύει γενικά. Αλλά εξαιρέσεις μπορούν να βρεθούν και έχω πει ότι δεν θα εκπλαγώ και πολύ αν κάποια μέρα ανακαλυφθεί μια βάσιμη εξαίρεση."

Είναι φυσικό η απλότητα της συνάρτησης να μην προσεγγίζει ικανοποιητικά την καμπύλη κατανομής του εισοδήματος από την αρχή, δηλαδή για μικρά εισοδήματα των ατόμων. Διάφορες απόπειρες έγιναν για να προστεθούν μερικές συμπληρωματικές εκφράσεις για τα χαμηλότερα εισοδήματα. Οι προσπάθειες δεν φαίνεται να άξιζαν τον κόπο. Αναθεωρήσεις του νόμου έγιναν από αρκετούς πολύ γνωστούς οικονομολόγους τα τελευταία 80 χρόνια. Pigou (1932), Shirras (1935), Hayakawa (1951) και άλλοι ασχολήθηκαν με το θέμα αυτό. Περισσότερες προσπάθειες έγιναν πρόσφατα για να εξηγηθούν πολλά εμπειρικά φαινόμενα χρησιμοποιώντας την κατανομή Pareto ή κάποια παραπλήσια συγγενική έκφραση.

1.2 Ορισμός.

Μια τυχαία μεταβλητή Χ λέμε ότι ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha>0$ και $\theta>0$ και συμβολίζουμε $X\sim P(\theta,\alpha)$ αν η συνάρτηση πιθανότητας της μεταβλητής Χ είναι:

$$f(x)=\begin{cases}\alpha\theta^{\alpha}x^{-(\alpha+1)} & \text{αν } x\geq\theta,\quad \alpha>0,\ \theta>0\\ \\ 0 & \text{για κάθε άλλη τιμή του } x\end{cases} \tag{1.2.1}$$

Η συνάρτηση κατανομής της τυχαίας μεταβλητής Χ είναι

$$F(x)=1-\theta^{\alpha}x^{-\alpha},\quad \alpha>0,\ x\geq\theta>0 \tag{1.2.2}$$

Πράγματι από τον ορισμό προκύπτει

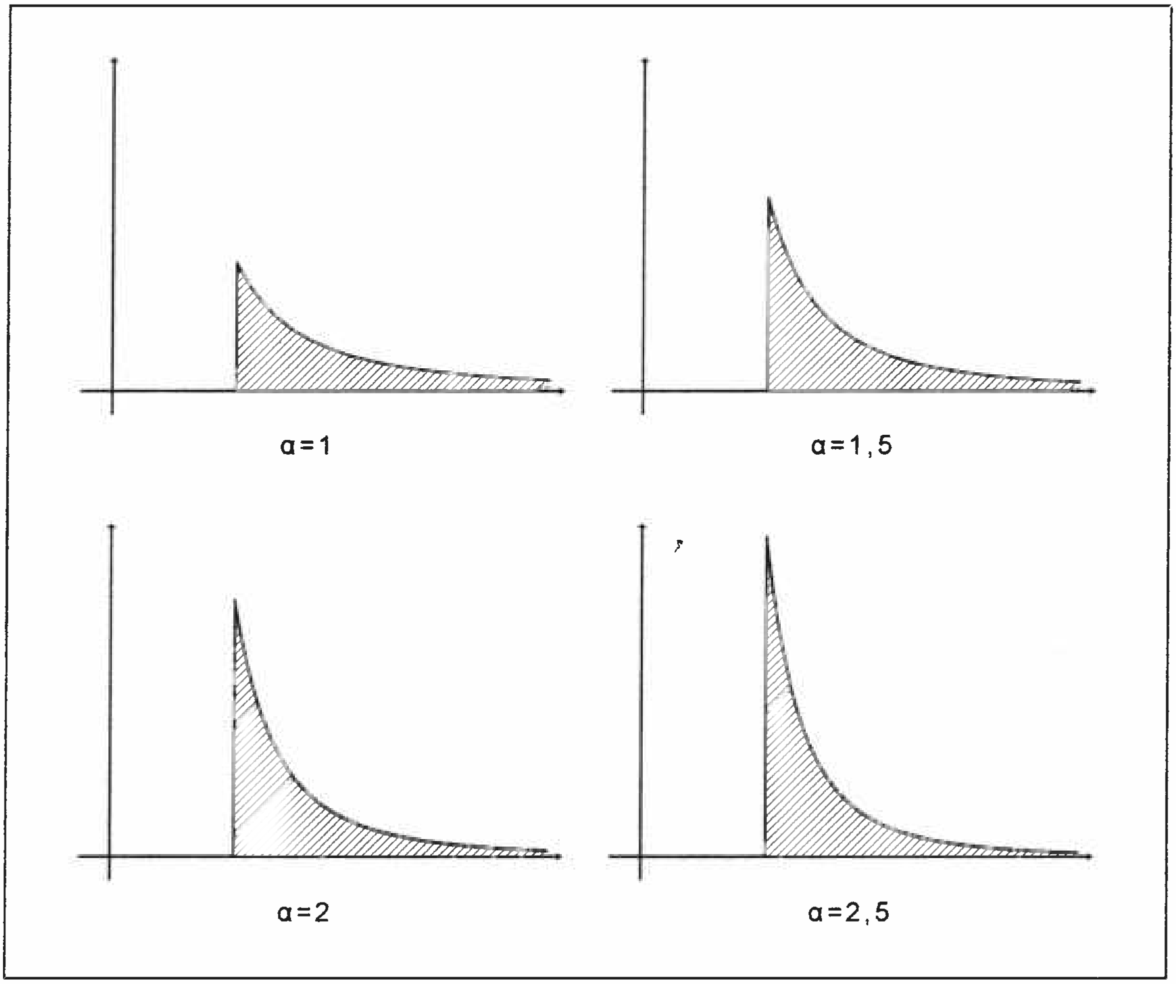
$$F(x)=\int_\theta^x f(t)dt=\int_\theta^x \alpha\theta^\alpha t^{-(\alpha+1)}dt$$

$$=-\theta^\alpha\int_\theta^x \frac{d}{dt}t^{-\alpha}dt$$

$$=-\theta^\alpha(x^{-\alpha}-\theta^{-\alpha})$$

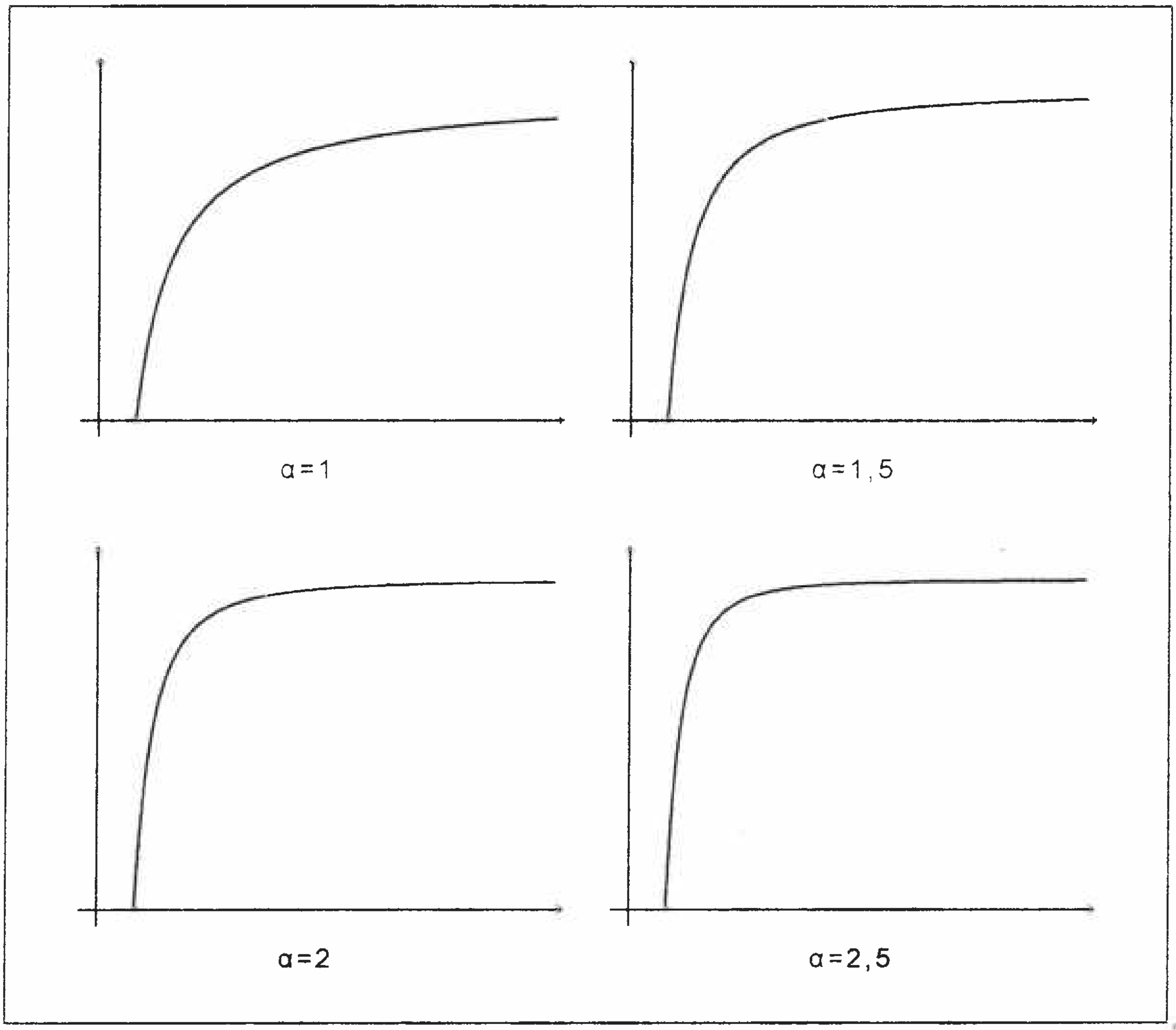
$$=1-\theta^\alpha x^{-\alpha} \qquad \alpha>0,\ x\geq\theta>0$$

1.3 Γραφικές παραστάσεις.



Σχήμα 1.3.1 Γραφικές παραστάσεις της συνάρτησης πιθανότητας μιας τυχαίας μεταβλητής που ακολουθεί κατανομή Pareto.

Το σχήμα (1.3.1) παρουσιάζει τις γραφικές παραστάσεις της συνάρτησης πιθανότητας μιας τυχαίας μεταβλητής Χ που ακολουθεί κατανομή Pareto για τις διάφορες τιμές της παραμέτρου α. Το σχήμα (1.3.2) παρουσιάζει τις αντίστοιχες γραφικές παραστάσεις της συνάρτησης κατανομής.



Σχήμα 1.3.2 Γραφικές παραστάσεις της συνάρτησης κατανομής μιας τυχαίας μεταβλητής που ακολουθεί κατανομή Pareto.

1.4 Συγγενείς κατανομές με την κατανομή Pareto.

Είναι σκόπιμο να παραθέσουμε εδώ τις συναρτησιακές συγγένειες μεταξύ μιας μεταβλητής από κατανομή Pareto και μεταβλητών από άλλες κατανομές. Τέτοιες συγγένειες μπορούν να αποδειχθούν μεταξύ μιας μεταβλητής που ακολουθεί κατανομή Pareto και μιας μεταβλητής που ακολουθεί εκθετική, λογιστική, Burr ή power function κατανομή. Οι συγγένειες αυτές είναι χρήσιμες στη συνέχεια για τη μελέτη της κατανομής Pareto και δίνονται από τα παρακάτω θεωρήματα.

ΘΕΩΡΗΜΑ 1.4.1

Έστω X μια τυχαία μεταβλητή που ακολουθεί εκθετική κατανομή με παράμετρο $\beta^{-1} > 0$. Σύμφωνα με τον τύπο (Α.3) η συνάρτηση πιθανότητας είναι

$$f(x) = \begin{cases} \dfrac{1}{\beta} e^{-\frac{x}{\beta}} & \text{αν } x > 0, \quad \beta > 0 \\ 0 & \text{αν } x \leq 0 \end{cases} \tag{1.4.1}$$

Τότε η μεταβλητή

$$Z = \theta e^{\frac{X}{\alpha\beta}}, \quad \alpha > 0, \ \theta > 0 \tag{1.4.2}$$

ακολουθεί κατανομή Pareto με $Z \sim P(\theta, \alpha)$. Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Z δίνεται από τον τύπο (1.2.1).

ΑΠΟΔΕΙΞΗ

Από το μετασχηματισμό (1.4.2) προκύπτει ότι

$$\begin{aligned} z = \theta e^{\frac{x}{\alpha\beta}} \quad &\Leftrightarrow \quad \frac{z}{\theta} = e^{\frac{x}{\alpha\beta}} \\ &\Leftrightarrow \quad \log\left(\frac{z}{\theta}\right) = \frac{x}{\alpha\beta} \\ &\Leftrightarrow \quad x = \alpha\beta \log\left(\frac{z}{\theta}\right) \end{aligned}$$

Επίσης

$$\begin{aligned} \frac{dx}{dz} &= \frac{d}{dz}\left[\alpha\beta\log\left(\frac{z}{\theta}\right)\right] \\ &= \frac{\alpha\beta}{z} > 0 \end{aligned}$$

και

$$\left|\frac{dx}{dz}\right| = \frac{\alpha\beta}{z}$$

Αν $x \geq 0$ τότε

$$\begin{aligned} x \geq 0 \quad &\Leftrightarrow \quad \alpha\beta\log\left(\frac{z}{\theta}\right) \geq 0 \\ &\Leftrightarrow \quad \log z \geq \log\theta \\ &\Leftrightarrow \quad z \geq \theta \end{aligned}$$

και

$$f_z(z) = f(x)\left|\frac{dx}{dz}\right|$$

$$= \frac{1}{\beta} e^{\frac{\alpha\beta\log\left(\frac{z}{\theta}\right)}{\beta}} \frac{\alpha\beta}{z}$$

$$= e^{\log\left(\frac{z}{\theta}\right)^{-\alpha}} \frac{\alpha}{z}$$

$$= \left(\frac{z}{\theta}\right)^{-\alpha} \frac{\alpha}{z}$$

$$= \alpha\theta^{\alpha} z^{-(\alpha+1)}$$

Αν $x < 0$ τότε

$$f_z(z) = 0$$

Τελικά για τη συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας μεταβλητής Ζ έχουμε

$$f_z(z) = \begin{cases} \alpha\theta^{\alpha} z^{-(\alpha+1)} & \text{αν } z \geq \theta \\ \\ 0 & \text{για κάθε άλλη τιμή του } z \end{cases}$$

δηλαδή $Z \sim P(\theta, \alpha)$

ΘΕΩΡΗΜΑ 1.4.2

Έστω Χ μια τυχαία μεταβλητή που ακολουθεί power function κατανομή με παραμέτρους $\alpha, \beta > 0$. Σύμφωνα με τον τύπο (Α.4) συνάρτηση πιθανότητας είναι

$$f(x) = \begin{cases} \alpha\beta^{-\alpha} x^{\alpha-1} & \text{αν } 0 \leq x \leq \beta \\ \\ 0 & \text{για κάθε άλλη τιμή του } x \end{cases} \tag{1.4.3}$$

Τότε η μεταβλητή

$$Z = \frac{1}{X} \tag{1.4.4}$$

ακολουθεί κατανομή Pareto με $Z \sim P\left(\beta^{-1}, \alpha\right)$. Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Ζ δίνεται από τον τύπο (1.2.1) με $\theta = \beta^{-1}$.

ΑΠΟΔΕΙΞΗ

Από το μετασχηματισμό (1.4.4) προκύπτει ότι

$$z=\frac{1}{x} \quad \Leftrightarrow \quad x=\frac{1}{z}$$

Επίσης

$$\frac{dx}{dz}=\frac{d}{dz}\left(\frac{1}{z}\right)$$

$$=-\frac{1}{z^2}<0$$

και

$$\left|\frac{dx}{dz}\right|=\frac{1}{z^2}$$

Αν $0 \le x \le \beta$ τότε

$$0 \le x \le \beta \quad \Leftrightarrow \quad 0 \le \frac{1}{z} \le \beta$$

$$\Leftrightarrow \quad z \ge \frac{1}{\beta}$$

και

$$f_z(z)=f(x)\left|\frac{dx}{dz}\right|$$

$$=\alpha\beta^{-\alpha}\left(\frac{1}{z}\right)^{\alpha-1}\frac{1}{z^2}$$

$$=\alpha\left(\beta^{-1}\right)^{\alpha} z^{-(\alpha+1)}$$

Αν x<0 τότε

$$f_z(z)=0$$

Τελικά για τη συνάρτηση πιθ**ανότητας $f_z(z)$ της τυχαίας μεταβ**λητής Ζ έχουμε

$$f_z(z)=\begin{cases} \alpha\left(\beta^{-1}\right)^{\alpha} z^{-(\alpha+1)} & \text{αν } z \ge \beta^{-1} \\ 0 & \text{για κάθε άλλη τιμή του } z \end{cases}$$

δηλαδή $Z \sim P\left(\beta^{-1},\alpha\right)$.

ΘΕΩΡΗΜΑ 1.4.3

Έστω X μια τυχαία μεταβλητή που ακολουθεί κατανομή Burr με παραμέτρους $\alpha > 0$ και $\beta > 1$. Σύμφωνα με τον τύπο (Α.5) η συνάρτηση πιθανότητας είναι

$$f(x) = \begin{cases} \dfrac{\alpha\beta x^{\beta-1}}{\left(1+x^{\beta}\right)^{\alpha+1}} & \text{αν } x \geq 0, \quad \alpha > 0,\ \beta > 1 \\ 0 & \text{για κάθε άλλη τιμή του x} \end{cases} \tag{1.4.5}$$

Τότε η μεταβλητή

$$Z = 1 + X^{\beta}, \quad \beta > 1 \tag{1.4.6}$$

ακολουθεί κατανομή Pareto με $Z \sim P(1,\alpha)$. Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Z δίνεται από τον τύπο (1.2.1) με $\theta = 1$.

ΑΠΟΔΕΙΞΗ

Από το μετασχηματισμό (1.4.6) προκύπτει ότι

$$z = 1 + x^{\beta} \iff x^{\beta} = z - 1$$

$$\iff x = (z-1)^{\frac{1}{\beta}}$$

Επίσης

$$\frac{dx}{dz} = \frac{d}{dz}\left[(z-1)^{\frac{1}{\beta}}\right]$$

$$= \frac{1}{\beta}(z-1)^{\frac{1}{\beta}-1}$$

$$= \frac{(z-1)^{\frac{1-\beta}{\beta}}}{\beta} > 0$$

και

$$\left|\frac{dx}{dz}\right| = \frac{(z-1)^{\frac{1-\beta}{\beta}}}{\beta}$$

Αν $x \geq 0$ τότε

$$x \geq 0 \iff (z-1)^{\frac{1}{\beta}} \geq 0$$

$$\iff z - 1 \geq 0$$

$$\iff z \geq 1$$

και

$$f_z(z)=f(x)\left|\frac{dx}{dz}\right|$$

$$=\frac{\alpha\beta\left[(z-1)^{\frac{1}{\beta}}\right]^{\beta-1}}{\left[1+\left[(z-1)^{\frac{1}{\beta}}\right]^{\beta}\right]^{\alpha+1}}\frac{(z-1)^{\frac{1-\beta}{\beta}}}{\beta}$$

$$=\frac{\alpha(z-1)^{\frac{\beta-1}{\beta}}(z-1)^{\frac{1-\beta}{\beta}}}{(1+z-1)^{\alpha+1}}$$

$$=\frac{\alpha}{z^{\alpha+1}}$$

$$=\alpha z^{-(\alpha+1)}$$

Αν x<0 τότε

$$f_z(z)=0$$

Τελικά για τη συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας μεταβλητής Z έχουμε

$$f_z(z)=\begin{cases}\alpha z^{-(\alpha+1)} & \text{αν } z\geq 1\\ \\ 0 & \text{για κάθε άλλη τιμή του } z\end{cases}$$

δηλαδή $Z\sim P(1,\alpha)$.

ΘΕΩΡΗΜΑ 1.4.4

Έστω Χ μια τυχαία μεταβλητή που ακολουθεί λογιστική κατανομή με παραμέτρους $\theta=1$ και $d=0$ δηλαδή $X\sim L(1,0)$. Σύμφωνα με τον τύπο (Α.6) του παραρτήματος η συνάρτηση πιθανότητας της Χ είναι

$$f(x)=\begin{cases}\dfrac{e^{-x}}{\left(1+e^{-x}\right)^2} & \text{αν } -\infty<x<+\infty\\ \\ 0 & \text{για κάθε άλλη τιμή του } x\end{cases} \tag{1.4.7}$$

Τότε η μεταβλητή

$$Z=1+e^{-X} \tag{1.4.8}$$

ακολουθεί κατανομή Pareto με $Z\sim P(1,1)$. Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Z δίνεται από τον τύπο (1.2.1) με $\theta=1$ και $\alpha=1$.

ΑΠΟΔΕΙΞΗ

Από το μετασχηματισμό (1.4.8) προκύπτει ότι

$$z = 1 + e^{-x} \Leftrightarrow e^{-x} = z - 1$$

$$\Leftrightarrow e^{x} = \frac{1}{z-1}$$

$$\Leftrightarrow x = \log\left(\frac{1}{z-1}\right)$$

$$\Leftrightarrow x = -\log(z-1)$$

Επίσης

$$\frac{dx}{dz} = \frac{d}{dz}\left[-\log(z-1)\right]$$

$$= -\frac{1}{z-1} < 0$$

και

$$\left|\frac{dx}{dz}\right| = \frac{1}{z-1}$$

Για κάθε x ισχύει

$$z > 1$$

και

$$f_z(z) = f(x)\left|\frac{dx}{dz}\right|$$

$$= \frac{e^{\log(z-1)}}{\left[1 + e^{\log(z-1)}\right]^2}\frac{1}{z-1}$$

$$= \frac{z-1}{(1+z-1)^2(z-1)}$$

$$= z^{-2}$$

Τελικά για τη συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας μεταβλητής Z έχουμε

$$f_z(z) = \begin{cases} z^{-2} & \text{αν } z \geq 1 \\ 0 & \text{για κάθε άλλη τιμή του } z \end{cases}$$

δηλαδή $Z \sim P(1,1)$.

1.5 Γενικευμένες κατανομές Pareto.

Οι κατανομές είναι περισσότερο πρόσφορο να περιγράφονται από τη συνάρτηση επιβίωσης $\bar{F}(x)=1-F(x)$. Η κατανομή Pareto έχει επικρατήσει με τον πρωτότυπο τύπο

$$\bar{F}(x)=P(X\geq x)=\left(\frac{x}{\theta}\right)^{-\alpha},\quad \alpha>0,\ x\geq\theta>0 \tag{1.5.1}$$

όπου $\bar{F}(x)$ είναι η πιθανότητα το εισόδημα να είναι μεγαλύτερο ή ίσο από την τιμή x και θ είναι κάποιο ελάχιστο εισόδημα. Σαν συνέπεια αυτού, η συνάρτηση κατανομής της τυχαίας μεταβλητής X του εισοδήματος, μπορεί να γραφτεί

$$F(x)=1-\left(\frac{x}{\theta}\right)^{-\alpha},\quad \alpha>0,\ x\geq\theta>0 \tag{1.5.2}$$

και η αντίστοιχη συνάρτηση πιθανότητας είναι

$$f(x)=\alpha\theta^{\alpha}x^{-(\alpha+1)},\quad \alpha>0,\ x\geq\theta>0 \tag{1.5.3}$$

Ο θετικός αριθμός θ λέγεται παράμετρος κλίμακας και ο θετικός αριθμός α λέγεται δείκτης ανισότητας (Pareto's index of inequality). Ο Arnold καλεί τη σχέση που δίνεται από τον τύπο (1.5.3) κατανομή Pareto πρώτου είδους. Αν μια τυχαία μεταβλητή X έχει την παραπάνω κατανομή τότε συμβολίζουμε $X\sim P(I)(\theta,\alpha)$.

Δύο άλλες μορφές της κατανομής αυτής διατυπώθηκαν από τον ίδιο τον Pareto ως εξής:

Η κατανομή Pareto δευτέρου είδους έχει συνάρτηση επιβίωσης

$$\bar{F}(x)=\left(1+\frac{x-d}{\theta}\right)^{-\alpha},\quad x>d,\ \theta>0,\ \alpha>0 \tag{1.5.4}$$

Η συνάρτηση κατανομής είναι

$$F(x)=1-\left(1+\frac{x-d}{\theta}\right)^{-\alpha},\quad x>d,\ \theta>0,\ \alpha>0 \tag{1.5.5}$$

και η συνάρτηση πιθανότητας είναι

$$f(x)=\alpha\theta^{-1}\left(1+\frac{x-d}{\theta}\right)^{-\alpha-1},\quad x>d,\ \theta>0,\ \alpha>0 \tag{1.5.6}$$

Η παράμετρος d λέγεται παράμετρος θέσης και στις περισσότερες εφαρμογές είναι μη αρνητικός αριθμός. Αν μια τυχαία μεταβλητή X έχει την παραπάνω κατανομή τότε συμβολίζουμε $X \sim P(II)(d,\theta,\alpha)$. Ο Lonax (1954) χρησιμοποίησε αυτή την κατανομή στην ανάλυση δεδομένων εμπορικής αποτυχίας.

Η κατανομή Pareto τρίτου είδους έχει συνάρτηση επιβίωσης

$$\bar{F}(x) = \left[1 + \left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1}, \quad x > d,\ \theta > 0,\ c > 0 \tag{1.5.7}$$

Η συνάρτηση κατανομής είναι

$$F(x) = 1 - \left[1 + \left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1}, \quad x > d,\ \theta > 0,\ c > 0 \tag{1.5.8}$$

και η συνάρτηση πιθανότητας είναι

$$f(x) = c^{-1}\theta^{-1}\left(\frac{x-d}{\theta}\right)^{\frac{1-c}{c}}\left[1 + \left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-2}, \quad x > d,\ \theta > 0,\ c > 0 \tag{1.5.9}$$

Ο θετικός αριθμός c λέγεται παράμετρος Gini. Αν μια τυχαία μεταβλητή X έχει την παραπάνω κατανομή τότε συμβολίζουμε $X \sim P(III)(d,\theta,c)$.

Μπορούμε να γενικεύσουμε ακόμη περισσότερο αν στην κατανομή Pareto τρίτου είδους εισάγουμε και την παράμετρο α οπότε καταλήγουμε στην κατανομή Pareto τετάρτου είδους.

Η κατανομή Pareto τετάρτου είδους έχει συνάρτηση επιβίωσης

$$\bar{F}(x) = \left[1 + \left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha}, \quad x > d,\ \theta > 0,\ c > 0,\ \alpha > 0 \tag{1.5.10}$$

Η συνάρτηση κατανομής είναι

$$F(x) = 1 - \left[1 + \left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha}, \quad x > d,\ \theta > 0,\ c > 0,\ \alpha > 0 \tag{1.5.11}$$

και η συνάρτηση πιθανότητας είναι

$$f(x) = \alpha c^{-1}\theta^{-1}\left(\frac{x-d}{\theta}\right)^{\frac{1-c}{c}}\left[1 + \left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha-1}, \quad x > d,\ \theta > 0,\ c > 0,\ \alpha > 0 \tag{1.5.12}$$

Αν μια τυχαία μεταβλητή X έχει την παραπάνω κατανομή τότε συμβολίζουμε $X \sim P(IV)(d,\theta,c,\alpha)$.

Εύκολα μπορούμε να συμπεράνουμε τις παρακάτω σχέσεις μεταξύ των γενικευμένων κατανομών Pareto.

$$P(I)(\theta,\alpha) \equiv P(IV)(\theta,\theta,1,\alpha) \tag{1.5.13}$$

$$P(II)(d,\theta,\alpha) \equiv P(IV)(d,\theta,1,\alpha) \tag{1.5.14}$$

$$P(III)(d,\theta,c) \equiv P(IV)(d,\theta,c,1) \tag{1.5.15}$$

1.6 Γενικευμένη Feller-Pareto κατανομή.

Ο Feller (1971) όρισε μια κατανομή Pareto με έναν διαφορετικό τρόπο κάνοντας χρήση της βήτα κατανομής ως ακολούθως.

Έστω η τυχαία μεταβλητή X που ακολουθεί βήτα κατανομή με παραμέτρους $c_1 > 0$ και $c_2 > 0$, δηλαδή $X \sim \beta(c_1, c_2)$. Τότε η τυχαία μεταβλητή

$$Y = X^{-1} - 1 \tag{1.6.1}$$

ακολουθεί αυτό που ο Feller ονομάζει κατανομή Pareto. Η συνάρτηση πιθανότητας $f_y(y)$ της τυχαίας μεταβλητής Y είναι.

$$f_y(y) = \frac{1}{B(c_1,c_2)} y^{c_2-1}(1+y)^{-c_1-c_2}, \quad y > 0 \tag{1.6.2}$$

όπου $B(c_1,c_2)$ η συνάρτηση βήτα όπως δίνεται από τον τύπο (Α.5) του παραρτήματος.

ΑΠΟΔΕΙΞΗ

Από το μετασχηματισμό (1.6.1) προκύπτει ότι

$$y = x^{-1} - 1 \;\Leftrightarrow\; x^{-1} = 1 + y$$

$$\Leftrightarrow\; x = (1+y)^{-1}$$

Επίσης

$$\frac{dx}{dy} = \frac{d}{dx}\left[(1+y)^{-1}\right]$$

$$= -(1+y)^{-2} < 0$$

και

$$\left|\frac{dx}{dy}\right| = (1+y)^{-2}$$

Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής X σύμφωνα με τον τύπο (A.10) του παραρτήματος είναι

$$f(x) = \frac{1}{B(c_1, c_2)} x^{c_1 - 1}(1-x)^{c_2 - 1}, \quad 0 < x < 1$$

όπου $B(c_1, c_2)$ η συνάρτηση βήτα όπως δίνεται από τον τύπο (A.9) του παραρτήματος.

Αν $0 < x < 1$ τότε

$$\begin{aligned} 0 < x < 1 &\Leftrightarrow 0 < (1+y)^{-1} < 1 \\ &\Leftrightarrow 1 + y > 1 \\ &\Leftrightarrow y > 0 \end{aligned}$$

Τελικά συνάρτηση πιθανότητας $f_y(y)$ της τυχαίας μεταβλητής Y είναι

$$\begin{aligned} f_y(y) &= f(x)\left|\frac{dx}{dy}\right| \\ &= \frac{1}{B(c_1, c_2)}\left[(1+y)^{-1}\right]^{c_1 - 1}\left[1-(1+y)^{-1}\right]^{c_2 - 1}(1+y)^{-2} \\ &= \frac{1}{B(c_1, c_2)}(1+y)^{1-c_1}\left(\frac{y}{1+y}\right)^{c_2 - 1}(1+y)^{-2} \\ &= \frac{1}{B(c_1, c_2)}(1+y)^{1-c_1} y^{c_2 - 1}(1+y)^{1-c_2}(1+y)^{-2} \\ &= \frac{1}{B(c_1, c_2)} y^{c_2 - 1}(1+y)^{-c_1 - c_2}, \quad y > 0 \end{aligned}$$

Ο Arnold πρότεινε μια άλλη πιο γενικευμένη έκδοση της κατανομής Pareto κάνοντας χρήση της βήτα κατανομής ως ακολούθως.

Έστω η τυχαία μεταβλητή X η οποία ακολουθεί βήτα κατανομή με παραμέτρους $c_1 > 0$ και $c_2 > 0$, δηλαδή $X \sim \beta(c_1, c_2)$. Τότε η τυχαία μεταβλητή

$$W = d + \theta\left(X^{-1} - 1\right)^{c}, \quad \theta > 0, \ c > 0 \tag{1.6.3}$$

ακολουθεί μια κατανομή που ονομάζεται Feller-Pareto κατανομή και συμβολίζεται με $W \sim FP(d, \theta, c, c_1, c_2)$. Η συνάρτηση πιθανότητας $f_w(w)$ της τυχαίας μεταβλητής W είναι

$$f_w(w)=\frac{c^{-1}\theta^{-1}}{B(c_1,c_2)}\left(\frac{w-d}{\theta}\right)^{\frac{c_2}{c}-1}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-c_1-c_2}, \quad w>d,\ \theta>0,\ c>0 \quad (1.6.4)$$

όπου $B(c_1,c_2)$ η συνάρτηση βήτα όπως δίνεται από τον τύπο (Α.9) του παραρτήματος.

<u>ΑΠΟΔΕΙΞΗ</u>

Από τους μετασχηματισμούς (1.6.1) και (1.6.3) προκύπτει ότι

$$w=d+\theta\left(x^{-1}-1\right)^{c} \iff w=d+\theta y^{c}$$

$$\iff \theta y^{c}=w-d$$

$$\iff y^{c}=\frac{w-d}{\theta}$$

$$\iff y=\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}$$

Αν $y>0$ τότε

$$y>0 \iff \left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}>0$$

$$\iff w>d$$

Επίσης

$$\frac{dy}{dw}=\frac{d}{dw}\left[\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]$$

$$=\frac{1}{c}\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}-1}\frac{d}{dw}\left(\frac{w-d}{\theta}\right)$$

$$=\frac{1}{c}\frac{1}{\theta}\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}-1}$$

$$=c^{-1}\theta^{-1}\left(\frac{w-d}{\theta}\right)^{\frac{1-c}{c}}>0 \qquad \text{αφού } w>d$$

και

$$\left|\frac{dy}{dw}\right| = c^{-1}\theta^{-1}\left(\frac{w-d}{\theta}\right)^{\frac{1-c}{c}}$$

Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Y δίνεται από τον τύπο (1.6.2) και είναι

$$f_y(y) = \frac{1}{B(c_1,c_2)} y^{c_2-1}(1+y)^{-c_1-c_2}, \quad y>0$$

Τελικά η συνάρτηση πιθανότητας $f_w(w)$ της τυχαίας μεταβλητής W είναι

$$f_w(w) = f_y(y)\left|\frac{dy}{dw}\right|$$

$$= \frac{1}{B(c_1,c_2)}\left[\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{c_2-1}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-c_1-c_2} c^{-1}\theta^{-1}\left(\frac{w-d}{\theta}\right)^{\frac{1-c}{c}}$$

$$= \frac{c^{-1}\theta^{-1}}{B(c_1,c_2)}\left(\frac{w-d}{\theta}\right)^{\frac{c_2-1}{c}+\frac{1-c}{c}}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-c_1-c_2}$$

$$= \frac{c^{-1}\theta^{-1}}{B(c_1,c_2)}\left(\frac{w-d}{\theta}\right)^{\frac{c_2}{c}-1}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-c_1-c_2}$$

Η κατανομή Pareto τετάρτου είδους όπως δίνεται από τους τύπους (1.5.10), (1.5.11) και (1.5.12) αποτελεί ειδική περίπτωση της κατανομής Feller-Pareto όπως δίνεται από τους τύπους (1.6.3) και (1.6.4) για τις τιμές των παραμέτρων $c_1 = \alpha > 0$ και $c_2 = 1$. Δηλαδή

$$P(IV)(d,\theta,c,\alpha) \equiv FP(d,\theta,c,\alpha,1) \qquad (1.6.6)$$

<u>ΑΠΟΔΕΙΞΗ</u>

Για $c_1 = \alpha > 0$ και $c_2 = 1$ σύμφωνα με τον τύπο (A.9) του παραρτήματος έχουμε

$$B(\alpha,1) = \alpha^{-1}$$

Από τον τύπο (1.6.4) για $c_1 = \alpha > 0$, $c_2 = 1$ και $B(\alpha,1) = \alpha^{-1}$ έχουμε

$$f_w(w)=\frac{c^{-1}\theta^{-1}}{\alpha^{-1}}\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}-1}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha-1}$$

$$=\alpha c^{-1}\theta^{-1}\left(\frac{w-d}{\theta}\right)^{\frac{1-c}{c}}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha-1}, \quad w>d,\ \theta>0,\ c>0,\ \alpha>0$$

οπότε λόγω του τύπου (1.5.12) συμπεραίνουμε ότι καταλήξαμε σε κατανομή Pareto τετάρτου είδους.

Η κατανομή Pareto τρίτου είδους όπως δίνεται από τους τύπους (1.5.7), (1.5.8) και (1.5.9) αποτελεί ειδική περίπτωση της κατανομής Feller-Pareto για $c_1=c_2=1$. Δηλαδή

$$P(III)(d,\theta,c)\equiv FP(d,\theta,c,1,1) \tag{1.6.7}$$

<u>ΑΠΟΔΕΙΞΗ</u>

Για $c_1=c_2=1$ από τον τύπο (A.9) του παραρτήματος έχουμε $B(1,1)=1$. Από τον τύπο (1.6.4) για $c_1=c_2=1$ και $B(1,1)=1$ έχουμε

$$f_w(w)=c^{-1}\theta^{-1}\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}-1}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1-1}$$

$$=c^{-1}\theta^{-1}\left(\frac{w-d}{\theta}\right)^{\frac{1-c}{c}}\left[1+\left(\frac{w-d}{\theta}\right)^{\frac{1}{c}}\right]^{-2}, \quad w>d,\ \theta>0,\ c>0$$

οπότε λόγω του τύπου (1.5.9) συμπεραίνουμε ότι καταλήξαμε σε κατανομή Pareto τρίτου είδους.

Συμπερασματικά μπορούμε να πούμε ότι η οικογένεια των κατανομών Feller-Pareto όπως δίνεται από τους τύπους (1.6.3) και (1.6.4) με πέντε παραμέτρους $FP(d,\theta,c,c_1,c_2)$ είναι μια γενικευμένη μορφή που μεταξύ των άλλων συμπεριλαμβάνει και όλες τις κατανομές Pareto πρώτου, δευτέρου, τρίτου και τετάρτου είδους όπως δίνονται από τους τύπους (1.5.1), (1.5.4), (1.5.7) και (1.5.10). Οι μεταξύ τους σχέσεις είναι

$$P(I)(\theta,\alpha) \equiv FP(\theta,\theta,1,\alpha,1) \quad (1.6.8)$$

$$P(II)(d,\theta,\alpha) \equiv FP(d,\theta,1,\alpha,1) \quad (1.6.9)$$

$$P(III)(d,\theta,c) \equiv FP(d,\theta,c,1,1) \quad (1.6.10)$$

$$P(IV)(d,\theta,c,\alpha) \equiv FP(d,\theta,c,\alpha,1) \quad (1.6.11)$$

1.7 Συγγενείς κατανομές με τις γενικευμένες κατανομές Pareto.

Παραθέσουμε εδώ συναρτησιακές συγγένειες μεταξύ μιας μεταβλητής από γενικευμένη κατανομή Pareto και μεταβλητών από άλλες γνωστές κατανομές.

ΘΕΩΡΗΜΑ 1.7.1

Έστω X μια τυχαία μεταβλητή που ακολουθεί κατανομή γάμμα με παραμέτρους $\alpha=\beta=1$ δηλαδή $X \sim \Gamma(1,1)$. Σύμφωνα με τον τύπο (A.8) του παραρτήματος η συνάρτηση πιθανότητας είναι

$$f(x)=e^{-x}, \quad x>0 \quad (\text{τυπική εκθετική κατανομή}) \quad (1.7.1)$$

Τότε η μεταβλητή

$$Z=d+\theta\left(e^{\frac{X}{\alpha}}-1\right), \quad z>d,\ \theta>0,\ \alpha>0 \quad (1.7.2)$$

ακολουθεί κατανομή Pareto δευτέρου είδους με $Z \sim P(II)(d,\theta,\alpha)$.

ΑΠΟΔΕΙΞΗ

Από το μετασχηματισμό (1.7.2) προκύπτει ότι

$$z=d+\theta\left(e^{\frac{x}{\alpha}}-1\right) \Leftrightarrow \theta\left(e^{\frac{x}{\alpha}}-1\right)=z-d$$

$$\Leftrightarrow e^{\frac{x}{\alpha}}-1=\frac{z-d}{\theta}$$

$$\Leftrightarrow e^{\frac{x}{\alpha}}=1+\frac{z-d}{\theta}$$

$$\Leftrightarrow e^{x}=\left(1+\frac{z-d}{\theta}\right)^{\alpha}$$

$$\Leftrightarrow x=\log\left(1+\frac{z-d}{\theta}\right)^{\alpha}$$

Ισχύει

$$
\begin{aligned}
x>0 \quad &\Leftrightarrow \quad \log\left(1+\frac{z-d}{\theta}\right)^{\alpha}>0 \\
&\Leftrightarrow \quad 1+\frac{z-d}{\theta}>1 \\
&\Leftrightarrow \quad \frac{z-d}{\theta}>0 \\
&\Leftrightarrow \quad z>d
\end{aligned}
$$

Επίσης

$$
\begin{aligned}
\frac{dx}{dz} &= \frac{d}{dx}\left[\log\left(1+\frac{z-d}{\theta}\right)^{\alpha}\right] \\
&= \alpha\left(1+\frac{z-d}{\theta}\right)^{-1}\frac{d}{dz}\left(1+\frac{z-d}{\theta}\right) \\
&= \alpha\theta^{-1}\left(1+\frac{z-d}{\theta}\right)^{-1}>0
\end{aligned}
$$

και

$$
\left|\frac{dx}{dz}\right| = \alpha\theta^{-1}\left(1+\frac{z-d}{\theta}\right)^{-1}
$$

Τελικά για τη συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας Z έχουμε

$$
\begin{aligned}
f_z(z) &= f(x)\left|\frac{dx}{dz}\right| \\
&= e^{\log\left(1+\frac{z-d}{\theta}\right)^{\alpha}}\alpha\theta^{-1}\left(1+\frac{z-d}{\theta}\right)^{-1} \\
&= \left(1+\frac{z-d}{\theta}\right)^{-\alpha}\alpha\theta^{-1}\left(1+\frac{z-d}{\theta}\right)^{-1} \\
&= \alpha\theta^{-1}\left(1+\frac{z-d}{\theta}\right)^{-\alpha-1}, \quad z>d,\ \theta>0,\ \alpha>0
\end{aligned}
$$

οπότε λόγω του τύπου (1.5.6) συμπεραίνουμε ότι η τυχαία μεταβλητή Z ακολουθεί κατανομή Pareto δευτέρου είδους με $Z \sim \mathbf{P(II)(d,\theta,\alpha)}$.

ΘΕΩΡΗΜΑ 1.7.2

Έστω X μια τυχαία μεταβλητή που ακολουθεί λογιστική κατανομή με παραμέτρους $\theta > 0$ και d δηλαδή $X \sim L(\theta, d)$. Σύμφωνα με τον τύπο (Α.6) του παραρτήματος η συνάρτηση πιθανότητας της X είναι

$$f(x) = \theta^{-1} e^{-\frac{x-d}{\theta}} \left(1 + e^{-\frac{x-d}{\theta}}\right)^{-2}, \quad -\infty < x < +\infty, \ \theta > 0 \tag{1.7.4}$$

Τότε η μεταβλητή

$$Z = e^{X} \tag{1.7.5}$$

ακολουθεί κατανομή Pareto τρίτου είδους με $Z \sim P(III)(0, e^{d}, \theta)$.

ΑΠΟΔΕΙΞΗ

Από το μετασχηματισμό (1.7.5) προκύπτει ότι

$$z = e^{x} \iff x = \log z, \quad z > 0$$

Ισχύει

$$\frac{dx}{dz} = \frac{d}{dz}(\log z)$$

$$= \frac{1}{z} > 0$$

και

$$\left|\frac{dx}{dz}\right| = \frac{1}{z}$$

Επίσης

$$-\infty < x < +\infty \iff z > 0$$

Η συνάρτηση πιθανότητας της X είναι

$$f(x) = \theta^{-1} e^{-\frac{x-d}{\theta}} \left(1 + e^{-\frac{x-d}{\theta}}\right)^{-2}, \quad -\infty < x < +\infty, \ \theta > 0$$

Τελικά για τη συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας Z έχουμε

$$f_z(z) = f(x)\left|\frac{dx}{dz}\right|$$

$$= \theta^{-1} e^{-\frac{\log z - d}{\theta}} \left(1 + e^{-\frac{\log z - d}{\theta}}\right)^{-2} \frac{1}{z}$$

$$= \theta^{-1} z^{-1} e^{-\frac{1}{\theta}(\log z - \log e^{d})} \left[1 + e^{-\frac{1}{\theta}(\log z - \log e^{d})}\right]^{-2}$$

$$= \theta^{-1} z^{-1} e^{-\frac{1}{\theta}\log\left(\frac{z}{e^{d}}\right)} \left[1 + e^{-\frac{1}{\theta}\log\left(\frac{z}{e^{d}}\right)}\right]^{-2}$$

$$= \theta^{-1} z^{-1} e^{\log\left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}}} \left[1 + e^{\log\left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}}}\right]^{-2}$$

$$= \theta^{-1} z^{-1} \left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}} \left[1 + \left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}}\right]^{-2}$$

$$= \theta^{-1} z^{-1} \frac{z}{e^{d}} \left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}-1} \left[1 + \left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}}\right]^{-2}$$

$$= \theta^{-1} e^{-d} \left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}-1} \left[1 + \left(\frac{z}{e^{d}}\right)^{-\frac{1}{\theta}}\right]^{-2}$$

οπότε λόγω του τύπου (1.5.9) συμπεραίνουμε ότι η τυχαία μεταβλητή Z ακολουθεί κατανομή Pareto τρίτου είδους με $Z \sim P(III)(0, e^{d}, \theta)$.

<u>1.8 Κατανομές Pareto ως μίξη άλλων κατανομών.</u>

Ο Harris (1968) έδειξε έναν άλλο τρόπο με τον οποίο η κατανομή Pareto μπορεί να γενικευτεί. Σκέφτηκε ένα μικτό σύστημα εκθετικών κατανομών το οποίο καταλήγει σε μια κατανομή Pareto.

ΘΕΩΡΗΜΑ 1.8.1

Έστω X **μια τυχαία μεταβλητή** με συνάρτηση κατανομής

$$F(x) = 1 - e^{-z\frac{x-d}{\theta}}, \quad x > d, \ \theta > 0 \tag{1.8.1}$$

όπου η παράμετρος z ακολουθεί γάμμα κατανομή με παραμέτρους $\alpha > 0$ και $\beta = 1$, δηλαδή $Z \sim \gamma(\alpha,1)$. **Τότε η τυχαία μεταβλητή** X ακολουθεί κατανομή Pareto δευτέρου είδους **με** $X \sim P(II)(d,\theta,\alpha)$.

ΑΠΟΔΕΙΞΗ

Για την παράμετρο z **από τον τύπο** (Α.8) του παραρτήματος έχουμε

$$f_z(z) = \frac{1}{\Gamma(\alpha)} z^{\alpha-1} e^{-z}, \quad z > 0, \ \alpha > 0$$

όπου $\Gamma(\alpha)$ **είναι η συνάρτηση** γάμμα όπως δίνεται από τον **τύπο (Α.7) του παραρτήματος**

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt$$

Η συνάρτηση κατανομής F(x) **της τυχαίας μεταβλητής** X που δίνεται από τον **τύπο** (1.8.1) **γράφεται**

$$F(x) = \int_0^{+\infty} \frac{1}{\Gamma(\alpha)} t^{\alpha-1} e^{-t} \left(1 - e^{-t\frac{x-d}{\theta}}\right) dt$$

$$= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} \left(t^{\alpha-1} e^{-t} - t^{\alpha-1} e^{-t} e^{-t\frac{x-d}{\theta}}\right) dt$$

$$= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1} e^{-t} dt - \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1} e^{-t-t\frac{x-d}{\theta}} dt$$

$$= \frac{1}{\Gamma(\alpha)} \Gamma(\alpha) - \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1} e^{-t-t\frac{x-d}{\theta}} dt$$

$$= 1 - \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1} e^{-t\left(1+\frac{x-d}{\theta}\right)} dt$$

Στο παραπάνω ολοκλήρωμα θέτουμε

$$t\left(1+\frac{x-d}{\theta}\right)=u$$

Ισχύουν

$$t=u\left(1+\frac{x-d}{\theta}\right)^{-1}$$

και

$$dt=\left(1+\frac{x-d}{\theta}\right)^{-1}du$$

Επίσης

$$\lim_{t\to 0}u=0$$

και

$$\lim_{t\to+\infty}u=+\infty \quad \text{αφού } x-d>0$$

Τελικά για τη συνάρτηση κατανομής F(x) της τυχαίας μεταβλητής X έχουμε

$$F(x)=1-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}\left[u\left(1+\frac{x-d}{\theta}\right)^{-1}\right]^{\alpha-1}e^{-u}\left(1+\frac{x-d}{\theta}\right)^{-1}du$$

$$=1-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}u^{\alpha-1}\left(1+\frac{x-d}{\theta}\right)^{-\alpha+1}e^{-u}\left(1+\frac{x-d}{\theta}\right)^{-1}du$$

$$=1-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}\left(1+\frac{x-d}{\theta}\right)^{-\alpha}u^{\alpha-1}e^{-u}du$$

$$=1-\left(1+\frac{x-d}{\theta}\right)^{-\alpha}\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}u^{\alpha-1}e^{-u}du$$

$$=1-\left(1+\frac{x-d}{\theta}\right)^{-\alpha}\frac{1}{\Gamma(\alpha)}\Gamma(\alpha)$$

$$=1-\left(1+\frac{x-d}{\theta}\right)^{-\alpha}, \quad x>d,\ \theta>0,\ \alpha>0$$

οπότε λόγω του τύπου (1.5.5) συμπεραίνουμε ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto δευτέρου είδους με $X\sim P(II)(d,\theta,\alpha)$.

ΘΕΩΡΗΜΑ 1.8.2

Έστω X μια τυχαία μεταβλητή με συνάρτηση κατανομής

$$F(x)=1-e^{-z\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}},\quad x>d,\ \theta>0,\ c>0 \tag{1.8.2}$$

όπου η παράμετρος z ακολουθεί **γάμμα κατανομή με παραμέτρους** $\alpha>0$ **και** $\beta=1$, **δηλαδή** $Z\sim\gamma(\alpha,1)$. **Τότε η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto τετάρτου είδους με** $X\sim P(IV)(d,\theta,c,\alpha)$

ΑΠΟΔΕΙΞΗ

Για την παράμετρο z από τον τύπο (A.4) του παραρτήματος έχουμε

$$f_z(z)=\frac{1}{\Gamma(\alpha)}z^{\alpha-1}e^{-z},\quad z>0,\ \alpha>0$$

όπου $\Gamma(\alpha)$ **είναι** η συνάρτηση γάμμα όπως δίνεται από τον τύπο (A.7) του παραρτήματος

$$\Gamma(\alpha)=\int_0^{+\infty}t^{\alpha-1}e^{-t}dt$$

Η συνάρτηση κατανομής F(x) **της** τυχαίας μεταβλητής X **που δίνεται από τον τύπο** (1.8.2) γράφεται

$$F(x)=\int_0^{+\infty}\frac{1}{\Gamma(\alpha)}t^{\alpha-1}e^{-t}\left[1-e^{-t\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}}\right]dt$$

$$=\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}\left[t^{\alpha-1}e^{-t}-t^{\alpha-1}e^{-t}e^{-t\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}}\right]dt$$

$$=\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}t^{\alpha-1}e^{-t}dt-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}t^{\alpha-1}e^{-t}e^{-t\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}}dt$$

$$=\frac{1}{\Gamma(\alpha)}\Gamma(\alpha)-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}t^{\alpha-1}e^{-t\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]}dt$$

$$=1-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}t^{\alpha-1}e^{-t\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]}dt$$

Στο παραπάνω ολοκλήρωμα θέτουμε

$$t\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]=u$$

Ισχύουν

$$t=u\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1}$$

και

$$dt=\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1}du$$

Επίσης

$$\lim_{t\to 0}u=0$$

και

$$\lim_{t\to+\infty}u=+\infty \quad \text{αφού } x-d>0$$

Τελικά για τη συνάρτηση κατανομής $F(x)$ της τυχαίας μεταβλητής X έχουμε

$$F(x)=1-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}\left[u\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1}\right]^{\alpha-1}e^{-u}\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1}du$$

$$=1-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}u^{\alpha-1}\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha+1}e^{-u}\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-1}du$$

$$=1-\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha}u^{\alpha-1}e^{-u}du$$

$$=1-\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha}\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}u^{\alpha-1}e^{-u}du$$

$$=1-\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-\alpha}$$

οπότε λόγω του τύπου (1.5.12) συμπεραίνουμε ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto τετάρτου είδους με $X \sim P(IV)(d,\theta,c,\alpha)$.

ΘΕΩΡΗΜΑ 1.8.3

Έστω Z_1 και Z_2 ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν γάμμα κατανομή με $Z_1 \sim \gamma(c_1,1)$ και $Z_2 \sim \gamma(c_2,1)$ όπου $c_1 > 0$, $c_2 > 0$. Τότε η τυχαία μεταβλητή

$$X = d + \theta\left(\frac{Z_2}{Z_1}\right)^c, \quad x > d,\ \theta > 0,\ c>0 \qquad (1.8.3)$$

ακολουθεί Feller-Pareto κατανομή με $X \sim FP(d,\theta,c,c_1,c_2)$

ΑΠΟΔΕΙΞΗ

Θα αποδείξουμε πρώτα ότι η συνάρτηση πιθανότητας της τυχαίας μεταβλητής $V = \frac{Z_2}{Z_1}$ είναι

$$f_v(v) = \frac{1}{B(c_1,c_2)} v^{c_2-1}(1+v)^{-c_1-c_2} \qquad (1.8.4)$$

όπου $B(c_1,c_2)$ είναι η συνάρτηση βήτα όπως δίνεται από τον τύπο (A.9) του παραρτήματος.

Αφού οι τυχαίες μεταβλητές Z_1 και Z_2 είναι ανεξάρτητες για την από κοινού συνάρτηση πιθανότητας ισχύει

$$f_{z_1z_2}(z_1,z_2) = f_{z_1}(z_1)\cdot f_{z_2}(z_2)$$

Όμως $Z_1 \sim \gamma(c_1,1)$ και $Z_2 \sim \gamma(c_2,1)$ οπότε σύμφωνα με τον τύπο (A.8) του παραρτήματος έχουμε

$$f_{z_1}(z_1) = \frac{1}{\Gamma(c_1)} z_1^{c_1-1} e^{-z_1}$$

και

$$f_{z_2}(z_2) = \frac{1}{\Gamma(c_2)} z_2^{c_2-1} e^{-z_2}$$

όπου $\Gamma(c_1)$ και $\Gamma(c_2)$ είναι η συνάρτηση γάμμα όπως δίνεται από τον τύπο (A.7) του παραρτήματος.

Για την από κοινού συνάρτηση πιθανότητας $f_{z_1 z_2}(z_1, z_2)$ έχουμε

$$f_{z_1 z_2}(z_1, z_2) = f_{z_1}(z_1) \cdot f_{z_2}(z_2)$$

$$= \frac{1}{\Gamma(c_1)} z_1^{c_1 - 1} e^{-z_1} \frac{1}{\Gamma(c_2)} z_2^{c_2 - 1} e^{-z_2}$$

$$= \frac{1}{\Gamma(c_1)\Gamma(c_2)} z_1^{c_1 - 1} z_2^{c_2 - 1} e^{-z_1 - z_2}$$

Θέτουμε

$$U = Z_1 \text{ και } V = \frac{Z_2}{Z_1} \tag{1.8.5}$$

άρα έχουμε

$$z_1 = u \text{ και } z_2 = uv$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \frac{d(z_1, z_2)}{d(u, v)}$

$$J = \begin{vmatrix} \frac{dz_1}{du} & \frac{dz_1}{dv} \\ \frac{dz_2}{du} & \frac{dz_2}{dv} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

Για την από κοινού συνάρτηση πιθανότητας $f_{uv}(u, v)$ των τυχαίων μεταβλητών U και V έχουμε

$$f_{uv}(u, v) = f_{z_1 z_2}(z_1, z_2) \cdot |J|$$

$$= \frac{1}{\Gamma(c_1)\Gamma(c_2)} u^{c_1 - 1} (uv)^{c_2 - 1} e^{-u - uv} u$$

$$= \frac{1}{\Gamma(c_1)\Gamma(c_2)} u^{c_1 + c_2 - 1} v^{c_2 - 1} e^{-u(1+v)}$$

Για την περιθώρια συνάρτηση πιθανότητας της τυχαίας μεταβλητής V έχουμε

$$f_v(v) = \int_0^{+\infty} f_{uv}(u, v) du$$

$$= \int_0^{+\infty} \frac{1}{\Gamma(c_1)\Gamma(c_2)} u^{c_1 + c_2 - 1} v^{c_2 - 1} e^{-u(1+v)} du$$

$$= \frac{1}{\Gamma(c_1)\Gamma(c_2)} v^{c_2 - 1} \int_0^{+\infty} u^{c_1 + c_2 - 1} e^{-u(1+v)} du$$

Θέτουμε

$$u(1+v)=t$$

οπότε

$$u(1+v)=t \quad \Leftrightarrow \quad u=(1+v)^{-1}t$$

και

$$du=(1+v)^{-1}dt$$

Επίσης

$$\lim_{u\to 0} t=0$$

και

$$\lim_{u\to +\infty} t=+\infty$$

Η συνάρτηση πιθανότητας $f_v(v)$ της τυχαίας μεταβλητής V γράφεται

$$f_v(v)=\frac{1}{\Gamma(c_1)\Gamma(c_2)}v^{c_2-1}\int_0^{+\infty}\left[(1+v)^{-1}t\right]^{c_1+c_2-1}e^{-t}(1+v)^{-1}dt$$

$$=\frac{1}{\Gamma(c_1)\Gamma(c_2)}v^{c_2-1}\int_0^{+\infty}(1+v)^{-c_1-c_2+1}t^{c_1+c_2-1}e^{-t}(1+v)^{-1}dt$$

$$=\frac{1}{\Gamma(c_1)\Gamma(c_2)}v^{c_2-1}\int_0^{+\infty}(1+v)^{-c_1-c_2}t^{(c_1+c_2)-1}e^{-t}dt$$

$$=\frac{1}{\Gamma(c_1)\Gamma(c_2)}v^{c_2-1}(1+v)^{-c_1-c_2}\int_0^{+\infty}t^{(c_1+c_2)-1}e^{-t}dt$$

$$=\frac{\Gamma(c_1+c_2)}{\Gamma(c_1)\Gamma(c_2)}v^{c_2-1}(1+v)^{-c_1-c_2}$$

Από τις ιδιότητες της συνάρτησης βήτα όπως δίνονται στον τύπο (Α.9) του παραρτήματος έχουμε

$$B(c_1,c_2)=\frac{\Gamma(c_1)\Gamma(c_2)}{\Gamma(c_1+c_2)} \quad \Leftrightarrow \quad \frac{\Gamma(c_1+c_2)}{\Gamma(c_1)\Gamma(c_2)}=\frac{1}{B(c_1,c_2)}$$

Οπότε καταλήγουμε στον τύπο (1.8.4)

$$f_v(v)=\frac{1}{B(c_1,c_2)}v^{c_2-1}(1+v)^{-c_1-c_2}$$

Για την τυχαία μεταβλητή X από τον τύπο (1.8.3) και τον μετασχηματισμό (1.8.5) έχουμε

$$X = d + \theta\left(\frac{Z_2}{Z_1}\right)^c$$

$$= d + \theta V^c \quad , \quad x > d \, , \; \theta > 0 \, , \; c > 0$$

Οπότε

$$x = d + \theta v^c \quad \Leftrightarrow \quad \theta v^c = x - d$$

$$\Leftrightarrow \quad v^c = \frac{x-d}{\theta}$$

$$\Leftrightarrow \quad v = \left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}$$

και

$$\frac{dv}{dx} = \frac{d}{dx}\left[\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]$$

$$= \frac{1}{c}\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}-1}\frac{d}{dx}\left(\frac{x-d}{\theta}\right)$$

$$= c^{-1}\left(\frac{x-d}{\theta}\right)^{\frac{1-c}{c}}\frac{1}{\theta}$$

$$= c^{-1}\theta^{-1}\left(\frac{x-d}{\theta}\right)^{\frac{1-c}{c}}$$

Αφού είναι $x > d$ έχουμε

$$\left|\frac{dv}{dx}\right| = \left|c^{-1}\theta^{-1}\left(\frac{x-d}{\theta}\right)^{\frac{1-c}{c}}\right|$$

$$= c^{-1}\theta^{-1}\left(\frac{x-d}{\theta}\right)^{\frac{1-c}{c}}$$

Τελικά για τη συνάρτηση πιθανότητας της τυχαίας μεταβλητής X από τον τύπο (1.8.4) έχουμε

$$f(x)=f_v(v)\left|\frac{dv}{dx}\right|$$

$$=\frac{1}{B(c_1,c_2)}\left[\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{c_2-1}\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-c_1-c_2} c^{-1}\theta^{-1}\left(\frac{x-d}{\theta}\right)^{\frac{1-c}{c}}$$

$$=\frac{c^{-1}\theta^{-1}}{B(c_1,c_2)}\left(\frac{x-d}{\theta}\right)^{\frac{c_2-1}{c}+\frac{1-c}{c}}\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-c_1-c_2-1}$$

$$=\frac{c^{-1}\theta^{-1}}{B(c_1,c_2)}\left(\frac{x-d}{\theta}\right)^{\frac{c_2-c}{c}}\left[1+\left(\frac{x-d}{\theta}\right)^{\frac{1}{c}}\right]^{-c_1-c_2-1}$$

οπότε λόγω του τύπου (1.6.4) συμπεραίνουμε ότι η τυχαία μεταβλητή Χ **ακολουθεί** Feller-Pareto **κατανομή** με $X \sim FP(d,\theta,c,c_1,c_2)$.

1.9 Κατανομές στατιστικών από κατανομή Pareto.

Παραθέτουμε εδώ τις κατανομές κάποιων **στατιστικών από** κατανομή Pareto που είναι **απαραίτητες στη συνέχεια αυτής** της εργασίας.

ΘΕΩΡΗΜΑ 1.9.1

Έστω Χ μια τυχαία μεταβλητή που ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$

Τότε η τυχαία μεταβλητή

$$Y=cX,\quad c>0 \tag{1.9.1}$$

ακολουθεί κατανομή Pareto με $Y \sim P(c\theta,\alpha)$

ΑΠΟΔΕΙΞΗ

Έχουμε $X \sim P(\theta,\alpha)$ **άρα η συνάρτηση** κατανομής **δίνεται από τον τύπο** (1.2.2) **και** είναι

$$F(x)=1-\theta^{\alpha}x^{-\alpha},\quad \alpha>0,\ x\geq\theta>0$$

Από το μετασχηματισμό (1.9.1) **έχουμε**

$$y = cx \quad \Leftrightarrow \quad x = \frac{y}{c}, \quad c > 0$$

και

$$x \geq \theta > 0 \quad \Leftrightarrow \quad \frac{y}{c} \geq \theta > 0$$

$$\Leftrightarrow \quad y \geq c\theta > 0$$

Η συνάρτηση **κατανομής** $F_y(y)$ **της τυχαίας μεταβλητής** Y **είναι**

$$F_y(y) = F\left(\frac{y}{c}\right)$$

$$= 1 - \theta^{\alpha}\left(\frac{y}{c}\right)^{-\alpha}$$

$$= 1 - \theta^{\alpha} c^{\alpha} y^{-\alpha}$$

$$= 1 - (c\theta)^{\alpha} y^{-\alpha}, \quad \alpha > 0, \ y \geq c\theta > 0$$

Επομένως η τυχαία μεταβλητή Y ακολουθεί κατανομή Pareto με $Y \sim P(c\theta, \alpha)$ όπως **δίνεται από τον τύπο** (1.2.2)

<u>ΘΕΩΡΗΜΑ 1.9.2</u>

Έστω X μια τυχαία μεταβλητή που ακολουθεί κατανομή Pareto με $\mathbf{X \sim P(\theta, \alpha)}$

Τότε η τυχαία μεταβλητή

$$Y = X^{\beta}, \quad \beta > 0 \tag{1.9.2}$$

ακολουθεί κατανομή Pareto με $Y \sim P\left(\theta^{\beta}, \frac{\alpha}{\beta}\right)$

<u>ΑΠΟΔΕΙΞΗ</u>

Έχουμε $X \sim P(\theta, \alpha)$ άρα η συνάρτηση **κατανομής δίνεται από τον τύπο** (1.2.2) **και είναι**

$$\mathbf{F(x) = 1 - \theta^{\alpha} x^{-\alpha}}, \quad \alpha > 0, \ x \geq \theta > 0$$

Από το μετασχηματισμό (1.9.2) έχουμε

$$y = x^{\beta} \quad \Leftrightarrow \quad x = y^{\frac{1}{\beta}}, \quad \beta > 0$$

και

$$x \geq \theta > 0 \quad \Leftrightarrow \quad y^{\frac{1}{\beta}} \geq \theta > 0$$

$$\Leftrightarrow \quad y \geq \theta^{\beta} > 0$$

Η συνάρτηση κατανομής $F_y(y)$ της τυχαίας μεταβλητής Y είναι

$$F_y(y) = F\left(y^{\frac{1}{\beta}}\right)$$

$$= 1 - \theta^{\alpha}\left(y^{\frac{1}{\beta}}\right)^{-\alpha}$$

$$= 1 - \theta^{\alpha} y^{-\frac{\alpha}{\beta}}$$

$$= 1 - \left(\theta^{\beta}\right)^{\frac{\alpha}{\beta}} y^{-\frac{\alpha}{\beta}}, \quad \frac{\alpha}{\beta} > 0, \; y \geq \theta^{\beta} > 0$$

Επομένως η τυχαία μεταβλητή Y ακολουθεί κατανομή Pareto με $Y \sim P\left(\theta^{\beta}, \frac{\alpha}{\beta}\right)$ όπως δίνεται από τον τύπο (1.2.2)

<u>ΘΕΩΡΗΜΑ 1.9.3</u>

Έστω X_1 και X_2 ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κατανομή Pareto με $X_1 \sim P(\theta_1, \alpha_1)$ και $X_2 \sim P(\theta_2, \alpha_2)$. Τότε η τυχαία μεταβλητή

$$Z = \frac{X_1}{X_2} \tag{1.9.3}$$

έχει συνάρτηση πιθανότητας $f_z(z)$ που δίνεται από τον τύπο

$$f_z(z) = \frac{\alpha_2}{\alpha_1 + \alpha_2} \alpha_1 \left(\frac{\theta_1}{\theta_2}\right)^{\alpha_1} z^{-\alpha_1 - 1}, \quad z \geq \frac{\theta_1}{\theta_2} > 0 \tag{1.9.4}$$

<u>ΑΠΟΔΕΙΞΗ</u>

Έχουμε $X_1 \sim P(\theta_1, \alpha_1)$ και $X_2 \sim P(\theta_2, \alpha_2)$ άρα οι συναρτήσεις πιθανότητας $f_1(x_1)$ και $f_2(x_2)$ όπως δίνονται από τον τύπο (1.2.1) είναι

$$f_1(x_1) = \alpha_1 \theta_1^{\alpha_1} x^{-\alpha_1 - 1}, \quad \alpha_1 > 0, \; x_1 \geq \theta_1 > 0$$

και

$$f_2(x_2) = \alpha_2 \theta_2^{\alpha_2} x^{-\alpha_2 - 1}, \quad \alpha_2 > 0, \; x_2 \geq \theta_2 > 0$$

Επειδή οι τυχαίες μεταβλητές X_1 και X_2 **είναι ανεξάρτητες** για τη από κοινού συνάρτηση πιθανότητας $f_{x_1x_2}(x_1, x_2)$ έχουμε

$$f_{x_1x_2}(x_1, x_2) = f_1(x_1) f_2(x_2)$$

Θέτουμε

$$Z = \frac{X_1}{X_2}$$

και

$$W = X_2$$

άρα έχουμε

$$x_1 = zw, \quad x_1 \geq \theta_1$$

και

$$x_2 = w, \quad x_2 \geq \theta_2$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \dfrac{d(x_1, x_2)}{d(z, w)}$

$$J = \begin{vmatrix} \dfrac{dx_1}{dz} & \dfrac{dx_1}{dw} \\ \dfrac{dx_2}{dz} & \dfrac{dx_2}{dw} \end{vmatrix} = \begin{vmatrix} w & z \\ 0 & 1 \end{vmatrix} = w$$

Η από κοινού συνάρτηση πιθανότητας $f_{zw}(z, w)$ **των τυχαίων** μεταβλητών Z και W είναι

$$f_{zw}(z, w) = f_{x_1x_2}(x_1, x_2) \cdot |J|$$

$$= f_1(x_1) f_2(x_2) \cdot |J|$$

$$= \alpha_1 \theta_1^{\alpha_1} (zw)^{-\alpha_1 - 1} \alpha_2 \theta_2^{\alpha_2} w^{-\alpha_2 - 1} w$$

$$= \alpha_1 \alpha_2 \theta_1^{\alpha_1} \theta_2^{\alpha_2} z^{-\alpha_1 - 1} w^{-\alpha_1 - 1} w^{-\alpha_2 - 1} w$$

$$= \alpha_1 \alpha_2 \theta_1^{\alpha_1} \theta_2^{\alpha_2} z^{-\alpha_1 - 1} w^{-\alpha_1 - \alpha_2 - 1}$$

Θεωρούμε ότι

$$z \geq \frac{\theta_1}{\theta_2}$$

Ισχύει

$$\theta_2 \leq w < +\infty$$

Η περιθώρια συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας μεταβλητής Z είναι

$$f_z(z) = \int_{\theta_2}^{+\infty} f_{zw}(z, w)dw$$

$$= \int_{\theta_2}^{+\infty} \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}w^{-\alpha_1-\alpha_2-1}dw$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\int_{\theta_2}^{+\infty} w^{-\alpha_1-\alpha_2-1}dw$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\left[\frac{w^{-\alpha_1-\alpha_2}}{-\alpha_1-\alpha_2}\right]_{\theta_2}^{+\infty}$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\left(-\frac{\theta_2^{-\alpha_1-\alpha_2}}{-\alpha_1-\alpha_2}\right)$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\frac{\theta_2^{-\alpha_1-\alpha_2}}{\alpha_1+\alpha_2}$$

$$= \frac{\alpha_2}{\alpha_1+\alpha_2}\alpha_1\theta_1^{\alpha_1}\theta_2^{-\alpha_1}z^{-\alpha_1-1}$$

$$= \frac{\alpha_2}{\alpha_1+\alpha_2}\alpha_1\left(\frac{\theta_1}{\theta_2}\right)^{\alpha_1}z^{-\alpha_1-1}$$

ΘΕΩΡΗΜΑ 1.9.4

Έστω X_1 και X_2 ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κατανομή Pareto με $X_1 \sim P(\theta_1, \alpha_1)$ και $X_2 \sim P(\theta_2, \alpha_2)$. Τότε η τυχαία μεταβλητή

$$Z = X_1X_2 \tag{1.9.5}$$

έχει συνάρτηση πιθ**ανότητας $f_z(z)$ που δίνεται από τον τύπο**

$$f_z(z) = \frac{\alpha_1}{\alpha_1-\alpha_2}\alpha_2(\theta_1\theta_2)^{\alpha_2}z^{-\alpha_2-1} - \frac{\alpha_2}{\alpha_1-\alpha_2}\alpha_1(\theta_1\theta_2)^{\alpha_1}z^{-\alpha_1-1}, \quad z \geq \theta_1\theta_2 > 0 \tag{1.9.6}$$

ΑΠΟΔΕΙΞΗ

Έχουμε $X_1 \sim P(\theta_1,\alpha_1)$ και $X_2 \sim P(\theta_2,\alpha_2)$ **άρα** οι συναρτήσεις πιθανότητας $f_1(x_1)$ **και** $f_2(x_2)$ **όπως** δίνονται από τον **τύπο** (1.2.1) **είναι**

$$f_1(x_1)=\alpha_1\theta_1^{\alpha_1}x^{-\alpha_1-1}, \quad \alpha_1>0,\ x_1\ge\theta_1>0$$

και

$$f_2(x_2)=\alpha_2\theta_2^{\alpha_2}x^{-\alpha_2-1}, \quad \alpha_2>0,\ x_2\ge\theta_2>0$$

Επειδή οι τυχαίες μεταβλητές X_1 **και** X_2 **είναι ανεξάρτητες** για τη από κοινού συνάρτηση **πιθανότητας** $f_{x_1x_2}(x_1,x_2)$ **έχουμε**

$$f_{x_1x_2}(x_1,x_2)=f_1(x_1)f_2(x_2)$$

Θέτουμε

$$Z=X_1X_2$$

και

$$W=X_2$$

άρα έχουμε

$$x_1=\frac{z}{w}, \quad x_1\ge\theta_1$$

και

$$x_2=w, \quad x_2\ge\theta_2$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J=\dfrac{d(x_1,x_2)}{d(z,w)}$

$$J=\begin{vmatrix} \dfrac{dx_1}{dz} & \dfrac{dx_1}{dw} \\ \dfrac{dx_2}{dz} & \dfrac{dx_2}{dw} \end{vmatrix}=\begin{vmatrix} \dfrac{1}{w} & -\dfrac{z}{w^2} \\ 0 & 1 \end{vmatrix}=\frac{1}{w}$$

Ισχύει

$$x_1\ge\theta_1 \text{ και } x_2\ge\theta_2 \iff x_1x_2\ge\theta_1\theta_2$$
$$\iff z\ge\theta_1\theta_2$$

και

$$x_1\ge\theta_1 \text{ και } x_2\ge\theta_2 \iff \frac{z}{w}\ge\theta_1 \text{ και } w\ge\theta_2$$
$$\iff \frac{z}{\theta_1}\ge w\ge\theta_2$$

Η από κοινού συνάρτηση πιθανότητας $\mathbf{f_{zw}(z,w)}$ **των τυχαίων** μεταβλητών Ζ και W είναι

$$\mathbf{f_{zw}(z,w) = f_{x_1x_2}(x_1,x_2)\cdot|J|}$$

$$\mathbf{= f_1(x_1)f_2(x_2)\cdot|J|}$$

$$= \alpha_1\theta_1^{\alpha_1}\left(\frac{z}{w}\right)^{-\alpha_1-1}\alpha_2\theta_2^{\alpha_2}w^{-\alpha_2-1}\frac{1}{w}$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}w^{\alpha_1+1}w^{-\alpha_2-1}w^{-1}$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}w^{\alpha_1-\alpha_2-1}$$

Η περιθώρια συνάρτηση πιθανότητας $f_z(z)$ **της τυχαίας μεταβλητής** Ζ **είναι**

$$f_z(z) = \int_{\theta_2}^{\frac{z}{\theta_1}} f_{zw}(z,w)dw$$

$$= \int_{\theta_2}^{\frac{z}{\theta_1}} \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}w^{\alpha_1-\alpha_2-1}dw$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\int_{\theta_2}^{\frac{z}{\theta_1}} w^{\alpha_1-\alpha_2-1}dw$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\left[\frac{w^{\alpha_1-\alpha_2}}{\alpha_1-\alpha_2}\right]_{\theta_2}^{\frac{z}{\theta_1}}$$

$$= \alpha_1\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\left[\frac{\left(\theta_1^{-1}z\right)^{\alpha_1-\alpha_2}}{\alpha_1-\alpha_2}-\frac{\theta_2^{\alpha_1-\alpha_2}}{\alpha_1-\alpha_2}\right]$$

$$= \frac{\alpha_1}{\alpha_1-\alpha_2}\alpha_2\theta_1^{\alpha_1}\theta_2^{\alpha_2}z^{-\alpha_1-1}\theta_1^{-\alpha_1+\alpha_2}z^{\alpha_1-\alpha_2} - \frac{\alpha_2}{\alpha_1-\alpha_2}\alpha_1\theta_1^{\alpha_1}\theta_2^{\alpha_2}\theta_2^{\alpha_1-\alpha_2}z^{-\alpha_1-1}$$

$$= \frac{\alpha_1}{\alpha_1-\alpha_2}\alpha_2\theta_1^{\alpha_2}\theta_2^{\alpha_2}z^{-\alpha_1-1}z^{\alpha_1-\alpha_2} - \frac{\alpha_2}{\alpha_1-\alpha_2}\alpha_1\theta_1^{\alpha_1}\theta_2^{\alpha_1}z^{-\alpha_1-1}$$

$$= \frac{\alpha_1}{\alpha_1-\alpha_2}\alpha_2\left(\theta_1\theta_2\right)^{\alpha_2}z^{-\alpha_2-1} - \frac{\alpha_2}{\alpha_1-\alpha_2}\alpha_1\left(\theta_1\theta_2\right)^{\alpha_1}z^{-\alpha_1-1}$$

ΘΕΩΡΗΜΑ 1.9.5

Έστω $X_1, X_2, \ldots, X_m$ ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κατανομή Pareto με $X_j \sim P(1, \alpha_j)$, $j = 1,2,\ldots,m$. Τότε η τυχαία μεταβλητή

$$Z = X_1 X_2 \cdots X_m = \prod_{i=1}^{m} X_i \tag{1.9.7}$$

έχει συνάρτηση πιθανότητας $f_z(z)$ που δίνεται από τον τύπο

$$f_z(z) = \prod_{i=1}^{m} \alpha_i \sum_{j=1}^{m} \frac{z^{-\alpha_j - 1}}{\prod_{\substack{i=1 \\ i \neq j}}^{m} (\alpha_i - \alpha_j)} \tag{1.9.8}$$

ΑΠΟΔΕΙΞΗ

Από τον τύπο (A.11) του παραρτήματος έχουμε ότι αν $T_1, T_2, \ldots, T_m$ είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή με $T_j \sim \exp(\alpha_j)$, $j = 1,2,\ldots,m$ τότε η τυχαία μεταβλητή

$$V = T_1 + T_2 + \cdots + T_m = \sum_{j=1}^{m} T_j$$

έχει συνάρτηση κατανομής

$$H_v(v) = 1 - \prod_{i=1}^{m} \alpha_i \sum_{j=1}^{m} \frac{e^{-\alpha_j v}}{\alpha_j \prod_{\substack{i=1 \\ i \neq j}}^{m} (\alpha_i - \alpha_j)}$$

Θεωρούμε το μετασχηματισμό

$$T_j = \log X_j, \quad j = 1,2,\ldots,m$$

Τότε

$$\begin{aligned} V &= T_1 + T_2 + \cdots + T_m \\ &= \log X_1 + \log X_2 + \cdots + \log X_m \\ &= \log(X_1 X_2 \cdots X_m) \\ &= \log Z \end{aligned}$$

και

$$v = \log z$$

Η συνάρτηση κατανομής $F_z(z)$ της τυχαίας μεταβλητής Z είναι

$$
\begin{aligned}
F_z(z) &= H_v(\log z) \\
&= 1-\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{e^{-\alpha_j \log z}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)} \\
&= 1-\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{e^{\log z^{-\alpha_j}}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)} \\
&= 1-\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{z^{-\alpha_j}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)}
\end{aligned}
$$

Τελικά η συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας μεταβλητής Ζ είναι

$$
\begin{aligned}
f_z(z) &= \frac{d}{dz}F_z(z) \\
&= \frac{d}{dz}\left[1-\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{z^{-\alpha_j}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)}\right] \\
&= -\frac{d}{dz}\left[\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{z^{-\alpha_j}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)}\right] \\
&= -\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{\frac{d}{dz}z^{-\alpha_j}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)} \\
&= -\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{-\alpha_j z^{-\alpha_j-1}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)} \\
&= \prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{z^{-\alpha_j-1}}{\prod_{\substack{i=1\\ i\neq j}}^{m}\left(\alpha_i-\alpha_j\right)}
\end{aligned}
$$

ΘΕΩΡΗΜΑ 1.9.6

Έστω $X_1, X_2, \ldots, X_n$ **ανεξάρτητες** τυχαίες μεταβλητές που ακολουθούν **κατανομή Pareto με** $X_i \sim P(\theta,\alpha)$, $i=1,2,\ldots,n$. Τότε η τυχαία μεταβλητή

$$Z = X_1 X_2 \cdots X_n = \prod_{i=1}^{n} X_i \tag{1.9.9}$$

έχει συνάρτηση πιθανότητας $f_z(z)$ που δίνεται από τον τύπο

$$f_z(z) = \frac{\alpha^n \theta^{\alpha n}}{\Gamma(n)} z^{-\alpha-1} \left[\log\left(\frac{z}{\theta^n} \right) \right]^{n-1} \tag{1.9.10}$$

ΑΠΟΔΕΙΞΗ

Η συνάρτηση πιθανότητας των τυχαίων μεταβλητών $X_i \sim P(\theta,\alpha)$ είναι

$$f(x_i) = \alpha\theta^\alpha x_i^{-\alpha-1}, \quad \alpha > 0, \; x_i \geq \theta > 0$$

Θεωρούμε το μετασχηματισμό

$$Y_i = \log X_i$$

Τότε

$$Y_i = \log X_i \quad \Leftrightarrow \quad X_i = e^{Y_i}$$

και

$$x_i = e^{y_i}$$

Επίσης

$$\frac{dx_i}{dy_i} = e^{y_i} > 0$$

και

$$\left| \frac{dx_i}{dy_i} \right| = e^{y_i}$$

Η συνάρτηση πιθανότητας των **τυχαίων μεταβλητών** Y_i **είναι**

$$\begin{aligned} f(y_i) &= f(x_i) \left| \frac{dx_i}{dy_i} \right| \\ &= \alpha\theta^\alpha \left(e^{y_i}\right)^{-\alpha-1} e^{y_i} \\ &= \alpha e^{\alpha \log\theta} e^{-\alpha y_i} \\ &= \alpha e^{-\alpha(y_i - \log\theta)} \end{aligned}$$

Από την έκφραση της συνάρτησης πιθανότητας συμπεραίνουμε ότι οι τυχαίες μεταβλητές $T_i = Y_i - \log\theta$ ακολουθούν εκθετική κατανομή με παράμετρο $\alpha > 0$, δηλαδή $T_i \sim \exp(\alpha)$. Από τον τύπο (A.12) του παραρτήματος έχουμε ότι αν $T_1, T_2, \ldots, T_n$ είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή με $T_i \sim \exp(\alpha)$, $i = 1,2,\ldots,n$ τότε η τυχαία μεταβλητή

$$V = T_1 + T_2 + \cdots + T_n = \sum_{i=1}^{n} T_i$$

έχει συνάρτηση πιθανότητας

$$g_v(v) = \frac{\alpha^n}{\Gamma(n)} v^{n-1} e^{-\alpha v}$$

Όμως

$$T_i = Y_i - \log\theta$$

οπότε

$$\begin{aligned} V &= \sum_{i=1}^{n} T_i \\ &= \sum_{i=1}^{n} (Y_i - \log\theta) \\ &= \sum_{i=1}^{n} Y_i - n\log\theta \end{aligned}$$

Θεωρούμε το μετασχηματισμό

$$Z = X_1 X_2 \cdots X_n$$

οπότε

$$\begin{aligned} Z &= X_1 X_2 \cdots X_n \\ &= e^{Y_1} e^{Y_2} \cdots e^{Y_n} \\ &= e^{Y_1 + Y_2 + \cdots + Y_n} \\ &= e^{\sum_{i=1}^{n} Y_i - n\log\theta} e^{n\log\theta} \\ &= e^{\sum_{i=1}^{n} Y_i - n\log\theta} e^{\log\theta^n} \\ &= e^V \theta^n \end{aligned}$$

Για την τυχαία μεταβλητή V έχουμε

$$Z = e^{V}\theta^{n} \quad \Leftrightarrow \quad e^{V} = \frac{Z}{\theta^{n}}$$

$$\Leftrightarrow \quad V = \log\left(\frac{Z}{\theta^{n}}\right)$$

Είναι

$$v = \log\left(\frac{z}{\theta^{n}}\right)$$

και

$$\frac{dv}{dz} = \frac{d}{dz}\left[\log\left(\frac{z}{\theta^{n}}\right)\right]$$

$$= \frac{1}{z} > 0$$

Τελικά η συνάρτηση πιθανότητας $f_z(z)$ της τυχαίας μεταβλητής Z είναι

$$f_z(z) = g_v\left(\log\left(\frac{z}{\theta^{n}}\right)\right)\left|\frac{dv}{dz}\right|$$

$$= \frac{\alpha^{n}}{\Gamma(n)}\left[\log\left(\frac{z}{\theta^{n}}\right)\right]^{n-1} e^{-\alpha\log\left(\frac{z}{\theta^{n}}\right)}\frac{1}{z}$$

$$= \frac{\alpha^{n}}{\Gamma(n)}\left[\log\left(\frac{z}{\theta^{n}}\right)\right]^{n-1} e^{\log\left(\frac{z}{\theta^{n}}\right)^{-\alpha}}\frac{1}{z}$$

$$= \frac{\alpha^{n}}{\Gamma(n)}\left[\log\left(\frac{z}{\theta^{n}}\right)\right]^{n-1}\left(\frac{z}{\theta^{n}}\right)^{-\alpha}\frac{1}{z}$$

$$= \frac{\alpha^{n}}{\Gamma(n)}\left[\log\left(\frac{z}{\theta^{n}}\right)\right]^{n-1}\frac{z^{-\alpha}}{\theta^{-\alpha n}}\frac{1}{z}$$

$$= \frac{\alpha^{n}\theta^{\alpha n}}{\Gamma(n)} z^{-\alpha-1}\left[\log\left(\frac{z}{\theta^{n}}\right)\right]^{n-1}$$

ΚΕΦΑΛΑΙΟ 2^{O}

Ροπές και μέτρα της κατανομής Pareto.

Στο κεφάλαιο αυτό παρουσιάζουμε αρχικά τις ροπές γύρω από το 0 και τις κεντρικές ροπές της κατανομής Pareto. Στη συνέχεια παρουσιάζουμε τα μέτρα θέσης, διασποράς, σχετικής διασποράς, ασυμμετρίας και κύρτωσης. Επειδή τα μέτρα αυτά δεν ορίζονται για όλες τις τιμές των παραμέτρων της κατανομής, παρουσιάζουμε τα εναλλακτικά μέτρα θέσης, διασποράς και σχετικής διασποράς. Τέλος παρουσιάζουμε την καμπύλη Loreng και τα μέτρα ανισότητας που σχετίζονται με αυτή.

2.1 Ροπές r τάξης γύρω από το 0.

Έστω η τυχαία μεταβλητή X που ακολουθεί την κατανομή Pareto με παραμέτρους $\theta > 0$ και $\alpha > 0$. Η ροπή r τάξης μ'_r (rth moment) της X γύρω από το 0 ορίζεται μόνο αν $\alpha > r$ και είναι:

$$\mu'_r = E(X^r) = \frac{\alpha\theta^r}{\alpha - r}, \quad \alpha > r \tag{2.1.1}$$

Πράγματι με εφαρμογή του γνωστού ορισμού για τη ροπή r τάξης μ'_r της X γύρω από το 0 έχουμε:

$$\mu'_r = E(X^r)$$

$$= \int_{\theta}^{+\infty} x^r f(x)dx$$

$$= \int_{\theta}^{+\infty} x^r \alpha\theta^{\alpha} x^{-(\alpha+1)}dx$$

$$= \alpha\theta^{\alpha} \int_{\theta}^{+\infty} x^{r-(\alpha+1)}dx$$

Αν $0 < \alpha < r$ τότε:

$$\alpha\theta^{\alpha} \int_{\theta}^{+\infty} x^{r-(\alpha+1)}dx = \alpha\theta^{\alpha} \lim_{t\to+\infty} \int_{\theta}^{t} \frac{d}{dx}\left(\frac{x^{r-\alpha}}{r-\alpha}\right)dx$$

$$= \alpha\theta^{\alpha} \lim_{t\to+\infty}\left(\frac{t^{r-\alpha}}{r-\alpha} - \frac{\theta^{r-\alpha}}{r-\alpha}\right) = +\infty$$

Αν $\alpha = r$ τότε:

$$\alpha\theta^{\alpha} \int_{\theta}^{+\infty} x^{r-(\alpha+1)}dx = \alpha\theta^{\alpha} \int_{\theta}^{+\infty} x^{-1}dx$$

$$= \alpha\theta^{\alpha} \int_{\theta}^{+\infty} \frac{d}{dx}\log x dx$$

$$= \alpha\theta^{\alpha}\left[\lim_{t\to+\infty}(\log t) - \log\theta\right] = +\infty$$

Αν $\alpha > r$ τότε:

$$\alpha\theta^{\alpha} \int_{\theta}^{+\infty} x^{r-\alpha}dx = \alpha\theta^{\alpha} \lim_{t\to+\infty} \int_{\theta}^{t} \frac{d}{dx}\left(\frac{x^{r-\alpha}}{r-\alpha}\right)dx$$

$$= \alpha\theta^{\alpha} \lim_{t\to+\infty}\left(\frac{t^{r-\alpha}}{r-\alpha} - \frac{\theta^{r-\alpha}}{r-\alpha}\right)$$

$$= \alpha\theta^{\alpha} \frac{\theta^{r-\alpha}}{\alpha - r}$$

$$= \frac{\alpha\theta^{r}}{\alpha - r}$$

Επομένως η ροπή r τάξης μ'_r της X γύρω από το 0 ορίζεται μόνο αν $\alpha > r$ και δίνεται από τον τύπο (2.2.1)

2.2 Κεντρικές ροπές r τάξης.

Γενικά οι κεντρικές ροπές r τάξης μ_r και οι ροπές r τάξης μ'_r γύρω από το 0 συνδέονται με τον αναδρομικό τύπο

$$\mu_r = \mu'_r - \binom{r}{1}\mu'_{r-1}\mu'_1 + \binom{r}{2}\mu'_{r-2}\mu_1'^{2} - \cdots + (-1)^r\binom{r}{r}\mu_1'^{r} \qquad (2.2.1)$$

Η κεντρική ροπή πρώτης είναι πάντοτε $\mu_1 = 0$. Ειδικά για την κατανομή Patero P(I)(θ,α) θα υπολογίσουμε τις κεντρικές ροπές δεύτερης, τρίτης και τέταρτης τάξης, δηλαδή μ_2, μ_3 και μ_4

Αν $\alpha > 2$ για την κεντρική ροπή δεύτερης τάξης μ_2 έχουμε

$$\mu_2 = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2 \qquad (2.2.2)$$

Πράγματι με εφαρμογή των τύπων (2.1.1) και (2.2.1) για $r = 2$ έχουμε

$$\mu_2 = \mu'_2 - \mu_1'^{2}$$

$$= \frac{\alpha\theta^2}{\alpha-2} - \left(\frac{\alpha\theta}{\alpha-1}\right)^2$$

$$= \frac{\alpha\theta^2}{\alpha-2} - \frac{\alpha^2\theta^2}{(\alpha-1)^2}$$

$$= \frac{(\alpha-1)^2\alpha\theta^2 - (\alpha-2)\alpha^2\theta^2}{(\alpha-1)^2(\alpha-2)}$$

$$= \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$$

Αν $\alpha > 3$ τότε η κεντρική ροπή τρίτης τάξης μ_3 είναι

$$\mu_3 = \frac{2\alpha(\alpha+1)\theta^3}{(\alpha-1)^3(\alpha-2)(\alpha-3)}, \quad \alpha > 3 \qquad (2.2.3)$$

Πράγματι με εφαρμογή των τύπων (2.1.1) και (2.2.1) για $r=3$ έχουμε

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$$

$$= \frac{\alpha\theta^3}{\alpha-3} - 3\frac{\alpha\theta^2}{\alpha-2}\frac{\alpha\theta}{\alpha-1} + 2\left(\frac{\alpha\theta}{\alpha-1}\right)^3$$

$$= \frac{\alpha\theta^3}{\alpha-3} - \frac{3\alpha^2\theta^3}{(\alpha-2)(\alpha-1)} + 2\frac{\alpha^3\theta^3}{(\alpha-1)^3}$$

$$= \frac{\alpha\theta^3(\alpha-1)^3(\alpha-2) - 3\alpha^2\theta^3(\alpha-1)^2(\alpha-3) + 2\alpha^3\theta^3(\alpha-2)(\alpha-3)}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= \alpha\theta^3\frac{(\alpha-1)^3(\alpha-2) - 3\alpha(\alpha-1)^2(\alpha-3) + 2\alpha^2(\alpha-2)(\alpha-3)}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= \alpha\theta^3\frac{(\alpha-1)^3(\alpha-2) - \alpha(\alpha-1)^2(\alpha-3) - 2\alpha(\alpha-1)^2(\alpha-3) + 2\alpha^2(\alpha-2)(\alpha-3)}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= \alpha\theta^3\frac{(\alpha-1)^2\left[(\alpha-1)(\alpha-2) - \alpha(\alpha-3)\right] - 2\alpha(\alpha-3)\left[(\alpha-1)^2 - \alpha(\alpha-2)\right]}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= \alpha\theta^3\frac{(\alpha-1)^2\left(\alpha^2-3\alpha+2-\alpha^2+3\alpha\right) - 2\alpha(\alpha-3)\left(\alpha^2-2\alpha+1-\alpha^2+2\alpha\right)}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= \alpha\theta^3\frac{2(\alpha-1)^2 - 2\alpha(\alpha-3)}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= 2\alpha\theta^3\frac{\alpha^2-2\alpha+1-\alpha^2+3\alpha}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= 2\alpha\theta^3\frac{\alpha+1}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

$$= \frac{2\alpha(\alpha+1)\theta^3}{(\alpha-1)^3(\alpha-2)(\alpha-3)}$$

Αν $\alpha>4$ τότε η κεντρική ροπή **τέταρτης τάξης μ_4 είναι**

$$\mu_4 = \frac{3\alpha(3\alpha^2+\alpha+2)\theta^4}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}, \quad \alpha>3 \qquad (2.2.4)$$

Πράγματι με εφαρμογή **των τύπων** (2.1.1) και (2.2.1) για $r=4$ έχουμε

$$\mu_4=\mu_4'-4\mu_3'\mu_1'+6\mu_2'\mu_1'^2-3\mu_1'^4$$

$$=\frac{\alpha\theta^4}{\alpha-4}-4\frac{\alpha\theta^3}{\alpha-3}\frac{\alpha\theta}{\alpha-1}+6\frac{\alpha\theta^2}{\alpha-2}\left(\frac{\alpha\theta}{\alpha-1}\right)^2-3\left(\frac{\alpha\theta}{\alpha-1}\right)^4$$

$$=\alpha\theta^4\frac{(\alpha-1)(\alpha-3)-4\alpha(\alpha-4)}{(\alpha-1)(\alpha-3)(\alpha-4)}+\alpha^3\theta^4\frac{6(\alpha-1)^2-3\alpha(\alpha-2)}{(\alpha-1)^4(\alpha-2)}$$

$$=\alpha\theta^4\frac{-3(\alpha^2-4\alpha-1)}{(\alpha-1)(\alpha-3)(\alpha-4)}+\alpha^3\theta^4\frac{3(\alpha^2-2\alpha+2)}{(\alpha-1)^2(\alpha-2)}$$

$$=3\alpha\theta^4\frac{-(\alpha^2-4\alpha-1)(\alpha-1)^3(\alpha-2)+\alpha^2(\alpha^2-2\alpha+2)(\alpha-3)(\alpha-4)}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=3\alpha\theta^4\frac{-\alpha(\alpha-4)(\alpha-1)^3(\alpha-2)+(\alpha-1)^3(\alpha-2)+\alpha^2(\alpha^2-2\alpha+2)(\alpha-3)(\alpha-4)}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=3\alpha\theta^4\frac{\alpha(\alpha-4)\left[-(\alpha-1)^3(\alpha-2)+\alpha(\alpha^2-2\alpha+2)(\alpha-3)\right]+(\alpha-1)^3(\alpha-2)}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=3\alpha\theta^4\frac{\alpha(\alpha-4)\left[-(\alpha-1)^3(\alpha-2)+\alpha^2(\alpha-2)(\alpha-3)+2\alpha(\alpha-3)\right]+(\alpha-1)^3(\alpha-2)}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=3\alpha\theta^4\frac{\alpha(\alpha-4)\left[(\alpha-2)\left[-(\alpha-1)^3+\alpha^2(\alpha-3)\right]+2\alpha(\alpha-3)\right]+(\alpha-1)^3(\alpha-2)}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=3\alpha\theta^4\frac{\alpha(\alpha-4)\left[(\alpha-2)(-3\alpha+1)+2\alpha^2-6\alpha\right]+(\alpha-1)^3(\alpha-2)}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=3\alpha\theta^4\frac{(\alpha^2-4\alpha)(-\alpha^2+\alpha-2)+(\alpha^3-3\alpha^2+3\alpha-1)(\alpha-2)}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=3\alpha\theta^4\frac{-\alpha^4+\alpha^3-2\alpha^2+4\alpha^3-4\alpha^2+8\alpha+\alpha^4-3\alpha^3+3\alpha^2-\alpha-2\alpha^3+6\alpha^2-6\alpha+2}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$=\frac{3\alpha(3\alpha^2+\alpha+2)\theta^4}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}$$

Με τον ίδιο τρόπο κάνοντας χρήση του αναδρομικού τύπου (2.2.1) μπορούμε να υπολογίσουμε κεντρικές ροπές οποιασδήποτε τάξης.

2.3 Μέτρα θέσης, διασποράς και σχετικής διασποράς.

Έστω η τυχαία μεταβλητή X που ακολουθεί την κατανομή Pareto με παραμέτρους $\theta > 0$ και $\alpha > 0$. Η αναμενόμενη τιμή (mean) της X ορίζεται αν $\alpha > 1$ και είναι:

$$\mu = E(X) = \frac{\alpha\theta}{\alpha - 1}, \quad \alpha > 1 \tag{2.3.1}$$

Αυτό προκύπτει άμεσα με εφαρμογή του τύπου (2.1.1) για $r = 1$.

Η διακύμανση (variance) της X ορίζεται αν $\alpha > 2$ και είναι:

$$\sigma^2 = Var(X) = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2 \tag{2.3.2}$$

Πράγματι ισχύει

$$\sigma^2 = Var(X) = E(X^2) - [E(X)]^2 = \mu_2$$

$$= \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$$

Ο συντελεστής μεταβλητότητας της X ορίζεται αν $\alpha > 2$ και είναι:

$$CV = \frac{1}{\sqrt{\alpha(\alpha-2)}}, \quad \alpha > 2 \tag{2.3.3}$$

διότι

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{Var(X)}}{E(X)}$$

$$= \sqrt{\frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}}\frac{\alpha-1}{\alpha\theta}$$

$$= \sqrt{\frac{\alpha}{(\alpha-2)}}\frac{\theta}{\alpha-1}\frac{\alpha-1}{\alpha\theta}$$

$$= \frac{1}{\sqrt{\alpha(\alpha-2)}}$$

2.4 Μέτρα ασυμμετρίας και κύρτωσης.

Έστω η τυχαία μεταβλητή X που ακολουθεί την κατανομή Pareto με παραμέτρους $\theta>0$ και $\alpha>0$. Ο συντελεστής ασυμμετρίας (skewness) της X ορίζεται αν $\alpha>3$ και είναι:

$$\alpha_3(X)=2\frac{\alpha+1}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}}, \quad \alpha>3 \qquad (2.4.1)$$

διότι

$$\alpha_3(X)=\frac{\mu_3}{\sigma^3}$$

$$=\frac{2\alpha(\alpha+1)\theta^3}{(\alpha-1)^3(\alpha-2)(\alpha-3)}\left(\sqrt{\operatorname{Var}(X)}\right)^{-3}$$

$$=\frac{2\alpha(\alpha+1)\theta^3}{(\alpha-1)^3(\alpha-2)(\alpha-3)}\left[\sqrt{\frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}}\right]^{-3}$$

$$=\frac{2\alpha(\alpha+1)\theta^3}{(\alpha-1)^3(\alpha-2)(\alpha-3)}\frac{(\alpha-2)(\alpha-1)^3}{\alpha\theta^3}\sqrt{\frac{\alpha-2}{\alpha}}$$

$$=2\frac{\alpha+1}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}}$$

Ο συντελεστής κυρτότητας (kurtosis) της τυχαίας μεταβλητής X ορίζεται αν $\alpha>4$ και είναι

$$\alpha_4(X)=\frac{3(\alpha-2)(3\alpha^2+\alpha+2)}{\alpha(\alpha-3)(\alpha-4)}, \quad \alpha>4 \qquad (2.4.2)$$

διότι

$$\alpha_4(X)=\frac{\mu_4}{\sigma^4}=\frac{3\alpha(3\alpha^2+\alpha+2)\theta^4}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}\left(\operatorname{Var}(X)\right)^{-2}$$

$$=\frac{3\alpha(3\alpha^2+\alpha+2)\theta^4}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}\left[\frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}\right]^{-2}$$

$$=\frac{3\alpha(3\alpha^2+\alpha+2)\theta^4}{(\alpha-1)^4(\alpha-2)(\alpha-3)(\alpha-4)}\frac{(\alpha-1)^4(\alpha-2)^2}{\alpha^2\theta^4}$$

$$=\frac{3(\alpha-2)(3\alpha^2+\alpha+2)}{\alpha(\alpha-3)(\alpha-4)}$$

2.5 Εναλλακτικά μέτρα θέσης της κατανομής Pareto.

Όπως φαίνεται παραπάνω ο αριθμητικός μέσος δεν ορίζεται για κάθε τιμή του α. Εναλλακτικά μέτρα θέσης είναι ο γεωμετρικός μέσος, ο αρμονικός μέσος και η διάμεσος τα οποία ορίζονται για κάθε θετικό αριθμό α. Αν η τυχαία μεταβλητή Χ ακολουθεί την κατανομή Pareto με παραμέτρους $\theta>0$ και $\alpha>0$ ο γεωμετρικός μέσος (geometric mean) της Χ είναι:

$$G=\theta e^{\frac{1}{\alpha}} \tag{2.5.1}$$

Πράγματι από τον ορισμό του γεωμετρικού μέσου έχουμε

$$G=\exp\left[\int_{\theta}^{+\infty}\log x f(x)dx\right]=\exp\left[\lim_{t\to\infty}\int_{\theta}^{t}\log x\alpha\theta^{\alpha}x^{-\alpha-1}dx\right]$$

$$=\exp\left[-\theta^{\alpha}\lim_{t\to+\infty}\int_{\theta}^{t}\log x\frac{d}{dx}x^{-\alpha}dx\right]$$

$$=\exp\left[-\theta^{\alpha}\lim_{t\to+\infty}\left(\left[x^{-\alpha}\log x\right]_{\theta}^{t}-\int_{\theta}^{t}x^{-\alpha-1}dx\right)\right]$$

$$=\exp\left[\log\theta+\theta^{\alpha}\lim_{t\to+\infty}\left[\int_{\theta}^{t}\frac{d}{dx}\left(-\frac{x^{-\alpha}}{\alpha}\right)dx\right]\right]$$

$$=\theta\cdot\exp\left[\theta^{\alpha}\lim_{t\to+\infty}\left(-\frac{t^{-\alpha}}{\alpha}+\frac{\theta^{-\alpha}}{\alpha}\right)\right]=\theta\cdot\exp\left(\frac{\theta^{\alpha}\theta^{-\alpha}}{\alpha}\right)=\theta e^{\frac{1}{\alpha}}$$

Ο αρμονικός μέσος (harmonic mean) της Χ είναι

$$H=\theta(1+\alpha^{-1}) \tag{2.5.2}$$

Πράγματι από τον ορισμό του αρμονικού μέσου έχουμε

$$H=[E(X^{-1})]^{-1}=\left[\int_{\theta}^{+\infty}x^{-1}f(x)dx\right]^{-1}$$

$$=\left(\int_{\theta}^{+\infty}x^{-1}\alpha\theta^{\alpha}x^{-\alpha-1}dx\right)^{-1}=\alpha^{-1}\theta^{-\alpha}\left(\int_{\theta}^{+\infty}x^{-\alpha-2}dx\right)^{-1}$$

$$=\alpha^{-1}\theta^{-\alpha}\left[\lim_{t\to+\infty}\left(\frac{t^{-\alpha-1}}{-\alpha-1}-\frac{\theta^{-\alpha-1}}{-\alpha-1}\right)\right]^{-1}$$

$$=\alpha^{-1}\theta^{-\alpha}\frac{\alpha+1}{\theta^{-\alpha-1}}=\theta\frac{\alpha+1}{\alpha}$$

$$=\theta(1+\alpha^{-1})$$

Η διάμεσος (median) της Χ είναι:

$$M = 2^{\frac{1}{\alpha}}\theta \qquad (2.5.3)$$

Πράγματι λύνοντας την εξίσωση $\int_M^{+\infty} f(x)dx = \frac{1}{2}$ ως προς Μ έχουμε

$$\int_M^{+\infty} f(x)dx = \frac{1}{2} \Leftrightarrow F(M) = \frac{1}{2}$$

$$\Leftrightarrow 1 - \theta^{\alpha} M^{-\alpha} = \frac{1}{2}$$

$$\Leftrightarrow \theta^{\alpha} M^{-\alpha} = \frac{1}{2}$$

$$\Leftrightarrow M = 2^{\frac{1}{\alpha}}\theta$$

Η ιδέα της διαμέσου της Χ επεκτείνεται με τον ορισμό του ποσοστιαίου σημείου τάξης p όπου $0 < p < 1$. Γενικά το ποσοστιαίο σημείο τάξης p με $0 < p < 1$ μιας τυχαίας μεταβλητής Χ είναι η ποσότητα X_p που αποτελεί λύση της εξίσωσης $F(X_p) = p$. Για την κατανομή Pareto έχουμε

$$X_p = \theta(1-p)^{-\frac{1}{\alpha}} \qquad (2.5.4)$$

Πράγματι λύνοντας την εξίσωση $F(X_p) = p$ ως προς X_p έχουμε

$$F(X_p) = p \Leftrightarrow 1 - \theta^{\alpha} X_p^{-\alpha} = p$$

$$\Leftrightarrow \theta^{\alpha} X_p^{-\alpha} = 1 - p$$

$$\Leftrightarrow X_p^{-\alpha} = \theta^{-\alpha}(1-p)$$

$$\Leftrightarrow X_p = \theta(1-p)^{-\frac{1}{\alpha}}$$

Η επικρατούσα τιμή (mode) της τυχαίας μεταβλητής Χ είναι

$$\text{mode}(X) = \theta \qquad (2.5.6)$$

Αυτό προκύπτει από το ότι η συνάρτηση πυκνότητας πιθανότητας $f(x)$ της τυχαίας μεταβλητής Χ είναι γνήσια φθίνουσα για $x \geq \theta$ αφού

$$\frac{df(x)}{dx} = \frac{d}{dx}\alpha\theta^{\alpha}x^{-\alpha-1}$$

$$= -\alpha(\alpha+1)\theta^{\alpha}x^{-\alpha-2} < 0$$

οπότε έχει μέγιστο για $x = \theta$

2.6 Εναλλακτικά μέτρα διασποράς της κατανομής Pareto.

Αν η τυχαία μεταβλητή X ακολουθεί την κατανομή Pareto με παραμέτρους $\theta > 0$ και $\alpha > 1$ τότε ο συντελεστής μέσης διαφοράς είναι

$$D = \frac{2\alpha\theta}{(\alpha-1)(2\alpha-1)}, \quad \alpha > 1 \tag{2.6.1}$$

Αυτό προκύπτει από τον ορισμό του συντελεστή μέσης διαφοράς μιας τυχαίας **μεταβλητής X που είναι**

$$D = \int_{R_Y}\int_{R_X} |x-y| dF(x)dF(y) \tag{2.6.2}$$

Στην περίπτωση της κατανομής Pareto με $\theta > 0$ και $\alpha > 1$ έχουμε

$$D = \int_\theta^{+\infty}\int_\theta^{+\infty} |x-y| \alpha\theta^\alpha x^{-\alpha-1} \alpha\theta^\alpha y^{-\alpha-1} dxdy$$

$$= \alpha^2\theta^{2\alpha}\int_\theta^{+\infty}\int_\theta^{+\infty} |x-y| x^{-\alpha-1} y^{-\alpha-1} dxdy$$

$$= 2\alpha^2\theta^{2\alpha}\int_\theta^{+\infty}\left[\int_\theta^x (x-y)x^{-\alpha-1}y^{-\alpha-1}dy\right]dx$$

$$= 2\alpha^2\theta^{2\alpha}\int_\theta^{+\infty}\left[\int_\theta^x x^{-\alpha}y^{-\alpha-1} - x^{-\alpha-1}y^{-\alpha}dy\right]dx$$

$$= 2\alpha^2\theta^{2\alpha}\int_\theta^{+\infty}\left[x^{-\alpha}\left[\frac{y^{-\alpha}}{-\alpha}\right]_\theta^x - x^{-\alpha-1}\left[\frac{y^{-\alpha+1}}{-\alpha+1}\right]_\theta^x\right]dx$$

$$= 2\alpha^2\theta^{2\alpha}\int_\theta^{+\infty}\left[x^{-\alpha}\left(\frac{\theta^{-\alpha}}{\alpha} - \frac{x^{-\alpha}}{\alpha}\right) - x^{-\alpha-1}\left(\frac{\theta^{-\alpha+1}}{\alpha-1} - \frac{x^{-\alpha+1}}{\alpha-1}\right)\right]dx$$

$$= 2\alpha^2\theta^{2\alpha}\int_\theta^{+\infty}\left(\frac{\theta^{-\alpha}x^{-\alpha}}{\alpha} - \frac{\theta^{-\alpha+1}x^{-\alpha-1}}{\alpha-1} + \frac{x^{-2\alpha}}{\alpha(\alpha-1)}\right)dx$$

$$= 2\alpha^2\theta^{2\alpha}\lim_{t\to\infty}\left[\frac{\theta^{-\alpha}x^{-\alpha+1}}{\alpha(-\alpha+1)} - \frac{\theta^{-\alpha+1}x^{-\alpha}}{-\alpha(\alpha-1)} + \frac{x^{-2\alpha+1}}{\alpha(\alpha-1)(-2\alpha+1)}\right]_\theta^t$$

$$= 2\alpha^2\theta^{2\alpha}\left[\frac{\theta^{-\alpha}\theta^{-\alpha+1}}{\alpha(\alpha-1)} - \frac{\theta^{-\alpha+1}\theta^{-\alpha}}{\alpha(\alpha-1)} + \frac{\theta^{-2\alpha+1}}{\alpha(\alpha-1)(2\alpha-1)}\right]$$

$$= \frac{2\alpha\theta}{(\alpha-1)(2\alpha-1)}$$

Όπως έχουμε αποδείξει η διακύμανση της **κατανομής Pareto δεν ορίζεται για** $\alpha \le 2$. **Λόγω αυτού ο Muniruzzaman** (1957) **εισήγαγε ένα νέο μέτρο διασποράς το**

οποίο ονομάζει γεωμετρική τυπική απόκλιση. Γενικά για μια τυχαία μεταβλητή X η γεωμετρική τυπική απόκλιση δίνεται από τον τύπο

$$\lambda = \exp\left[\sqrt{E\left[(\log X - \log G)^2\right]}\right] \tag{2.6.3}$$

όπου G είναι ο γεωμετρικός μέσος της X. Αν η τυχαία μεταβλητή X ακολουθεί την κατανομή Pareto με παραμέτρους $\theta > 0$ και $\alpha > 0$ τότε η γεωμετρική τυπική απόκλιση είναι

$$\lambda = e^{\frac{1}{\alpha}} \tag{2.6.4}$$

Για να αποδείξουμε αυτό χρησιμοποιούμε τις ιδιότητες της μέσης τιμής και έχουμε

$$\begin{aligned} E\left[(\log X - \log G)^2\right] &= E\left[\log^2 X - 2\log X \cdot \log G + \log^2 G\right] \\ &= E\left[\log^2 X\right] - 2E[\log X \cdot \log G] + E\left[\log^2 G\right] \\ &= E\left[\log^2 X\right] - 2\log G \cdot E[\log X] + \log^2 G \\ &= E\left[\log^2 X\right] - 2\log G \cdot \log G + \log^2 G \\ &= E\left[\log^2 X\right] - \log^2 G \end{aligned}$$

Όμως

$$\begin{aligned} E\left[\log^2 X\right] &= \int_\theta^{+\infty} \log^2 x f(x)dx = \int_\theta^{+\infty} \log^2 x \alpha\theta^\alpha x^{-\alpha-1}dx \\ &= -\theta^\alpha \int_\theta^{+\infty} \log^2 x \frac{d}{dx} x^{-\alpha} dx \\ &= -\theta^\alpha \left(\lim_{t\to\infty} \left[x^{-\alpha}\log^2 x\right]_\theta^t - \int_\theta^{+\infty} x^{-\alpha} \frac{d}{dx} \log^2 x dx \right) \\ &= -\theta^\alpha \lim_{t\to\infty}\left(t^{-\alpha}\log^2 t - \theta^{-\alpha}\log^2\theta\right) + 2\theta^\alpha \int_\theta^{+\infty} x^{-\alpha-1}\log x dx \\ &= -\theta^\alpha \left(\lim_{t\to\infty} \frac{2\log t}{\alpha t^\alpha} - \theta^{-\alpha}\log^2\theta \right) + 2\frac{1}{\alpha}\int_\theta^{+\infty} \log x \alpha\theta^\alpha x^{-\alpha-1}dx \\ &= -\theta^\alpha \left(\lim_{t\to\infty} \frac{2}{\alpha^2 t^\alpha} - \theta^{-\alpha}\log^2\theta \right) + 2\frac{1}{\alpha}\int_\theta^{+\infty} \log x f(x)dx \\ &= -\theta^\alpha\left(0 - \theta^{-\alpha}\log^2\theta\right) + 2\frac{1}{\alpha}\log G = \log^2\theta + 2\frac{1}{\alpha}\log G \end{aligned}$$

Λαμβάνοντας υπόψη ότι $G=\theta e^{\frac{1}{\alpha}}$ λόγω του τύπου (2.5.1) έχουμε

$$E\left[(\log X-\log G)^2\right]=\log^2\theta+2\frac{1}{\alpha}\log G-\log^2 G$$

$$=\log^2\theta+2\frac{1}{\alpha}\log\theta e^{\frac{1}{\alpha}}-\log^2\theta e^{\frac{1}{\alpha}}$$

$$=\log^2\theta+2\frac{1}{\alpha}\left(\log\theta+\frac{1}{\alpha}\right)-\left(\log\theta+\frac{1}{\alpha}\right)^2$$

$$=\log^2\theta+2\frac{1}{\alpha}\log\theta+\frac{2}{\alpha^2}-\left(\log^2\theta+2\frac{1}{\alpha}\log\theta+\frac{1}{\alpha^2}\right)$$

$$=\log^2\theta+2\frac{1}{\alpha}\log\theta+\frac{2}{\alpha^2}-\log^2\theta-2\frac{1}{\alpha}\log\theta-\frac{1}{\alpha^2}$$

$$=\frac{1}{\alpha^2}$$

Τελικά με εφαρμογή του τύπου (2.6.3) έχουμε

$$\lambda=\exp\left[\sqrt{E\left[(\log X-\log G)^2\right]}\right]$$

$$=\exp\left(\sqrt{\frac{1}{\alpha^2}}\right)$$

$$=e^{\frac{1}{\alpha}}$$

Το τετράγωνο της παραπάνω γεωμετρικής τυπικής απόκλισης λέγεται γεωμετρική διασπορά.

2.7 Συντελεστές σχετικής διασποράς ή ανισότητας της κατανομής Pareto.

Θα εισάγουμε τώρα δύο άλλους συντελεστές διασποράς οι οποίοι είναι ανεξάρτητοι από τις μονάδες μέτρησης της μεταβλητής. Είναι δηλαδή καθαροί αριθμοί και μας διευκολύνουν στη σύγκριση των διασπορών διαφορετικών πληθυσμών που έχουν διαφορετικές μονάδες μέτρησης.

Ο πρώτος είναι ο συντελεστής μεταβλητότητας Karl Pearson's και δίνεται από τον τύπο

$$V = 100\frac{\sigma}{\mu_1'} \tag{2.7.1}$$

Αν η τυχαία μεταβλητή X ακολουθεί την κατανομή Pareto με παραμέτρους $\theta > 0$ και $\alpha > 2$ τότε ο συντελεστής μεταβλητότητας Karl Pearson's είναι

$$V = \frac{100}{\alpha}\sqrt{\frac{\alpha}{\alpha-2}}, \quad \alpha > 2 \tag{2.7.2}$$

Πράγματι από τον τύπο (2.3.2) για $\alpha > 2$ έχουμε

$$\sigma = \sqrt{\frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}}$$

$$= \frac{\theta}{\alpha-1}\sqrt{\frac{\alpha}{\alpha-2}}$$

και από τον τύπο (2.2.1) για $r = 1$ έχουμε

$$\mu_1' = \frac{\alpha\theta}{\alpha-1}$$

οπότε ο τύπος (2.1.7) γράφεται

$$V = 100\frac{\sigma}{\mu_1'}$$

$$= 100\frac{\frac{\theta}{\alpha-1}\sqrt{\frac{\alpha}{(\alpha-2)}}}{\frac{\alpha\theta}{\alpha-1}}$$

$$= 100\frac{\theta(\alpha-1)}{\alpha\theta(\alpha-1)}\sqrt{\frac{\alpha}{(\alpha-2)}}$$

$$= \frac{100}{\alpha}\sqrt{\frac{\alpha}{(\alpha-2)}}$$

Ο δεύτερος σχετικός συντελεστής μεταβλητότητας λέγεται συντελεστής συγκέντρωσης Gini's και δίνεται από τον τύπο

$$G^* = \frac{D}{2\mu_1'} \tag{2.7.3}$$

όπου D είναι συντελεστής μέσης διαφοράς. Αν η τυχαία μεταβλητή X ακολουθεί την κατανομή Pareto με παραμέτρους $\theta > 0$ και $\alpha > 1$ τότε ο συντελεστής συγκέντρωσης Gini's δίνεται από τον τύπο

$$G^* = (2\alpha - 1)^{-1}, \quad \alpha > 1 \tag{2.7.4}$$

Πράγματι από τον τύπο (2.6.1) για $\alpha > 1$ έχουμε

$$D = \frac{2\alpha\theta}{(\alpha-1)(2\alpha-1)}$$

και από τον τύπο (2.2.1) για $r = 1$ έχουμε

$$\mu_1' = \frac{\alpha\theta}{\alpha - 1}$$

οπότε ο τύπος (2.1.3) γράφεται

$$G^* = \frac{D}{2\mu_1'}$$

$$= \frac{\frac{2\alpha\theta}{(\alpha-1)(2\alpha-1)}}{\frac{2\alpha\theta}{\alpha-1}}$$

$$= (2\alpha - 1)^{-1}$$

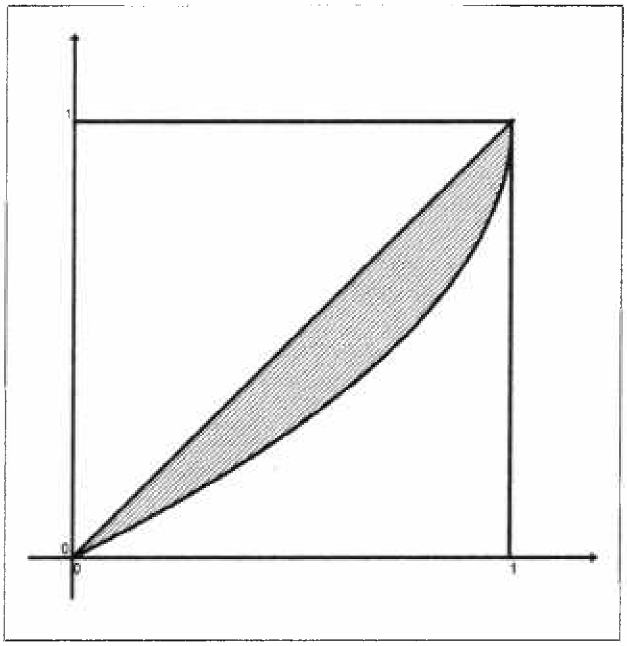
2.8 Η καμπύλη Lorenz ως δείκτης ανισότητας κατανομής Pareto.

Δεν υπάρχει αμφιβολία ότι, όσον αφορά τα οικονομικά η καμπύλη Lorenz και τα μέτρα ανισότητας που άμεσα ή έμμεσα σχετίζονται με αυτή βρίσκονται στο κέντρο του ενδιαφέροντος. Πρόκειται για μια συνάρτηση L(u) ορισμένη στο διάστημα $[0,1]$, τέτοια ώστε για καθορισμένο u το L(u) να παριστάνει το ποσοστό του συνολικού εισοδήματος του πληθυσμού για τα 100u% φτωχότερα άτομα του πληθυσμού. Ο Gastwirth το 1971 έδωσε τον ορισμό της καμπύλης Lorenz ως εξής

Έστω η τυχαία μεταβλητή X με συνάρτηση κατανομής $F(x)$ η οποία έχει πεπερασμένο μέσο και $F^{-1}(t)$, $0 \le t \le 1$ η αντίστροφη συνάρτηση της $F(x)$. Η καμπύλη Lorenz ορίζεται από τον τύπο

$$L(u) = \frac{\int_0^u F^{-1}(t)dt}{\int_0^1 F^{-1}(t)dt}, \quad 0 \le u \le 1 \tag{2.8.1}$$

Η συνάρτηση $L(u)$ είναι συνεχής στο $[0,1]$ με $L(0)=0$ και $L(1)=1$. Παριστάνει μια καμπύλη σχήματος τόξου που βρίσκεται κάτω από την εξισωτική γραμμή που ορίζεται από τη συνάρτηση $S(u)=u$, $0 \le u \le 1$. Όσο μεγαλύτερη είναι η επιφάνεια μεταξύ του παραπάνω τόξου και της εξισωτικής γραμμής τόσο μεγαλύτερη είναι η τιμή της ανισότητας. Η καμπύλη Lorenz και η επιφάνεια που αναφέρουμε παραπάνω φαίνονται στο διπλανό σχήμα (2.8.1).



Σχήμα 2.8.1 Καμπύλη Lorenz

Αν η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta, \alpha)$ τότε η καμπύλη Lorenz δίνεται από τον τύπο

$$L(u) = 1 - (1-u)^{\frac{\alpha-1}{\alpha}}, \quad 0 \le u \le 1 \tag{2.8.2}$$

<u>ΑΠΟΔΕΙΞΗ</u>

Η συνάρτηση κατανομής της τυχαίας μεταβλητής X από τον τύπο (1.2.2) είναι

$$F(x) = 1 - \theta^{\alpha} x^{-\alpha}, \quad \alpha > 0, \; x \ge \theta > 0$$

Θέτουμε

$$t = F(x), \quad 0 \le t \le 1$$

και έχουμε

$$\begin{aligned} t = F(x) \quad &\Leftrightarrow \quad t = 1 - \theta^{\alpha} x^{-\alpha} \\ &\Leftrightarrow \quad \theta^{\alpha} x^{-\alpha} = 1 - t \\ &\Leftrightarrow \quad x^{-\alpha} = \frac{1-t}{\theta^{\alpha}} \\ &\Leftrightarrow \quad x^{\alpha} = \theta^{\alpha} (1-t)^{-1} \\ &\Leftrightarrow \quad x = \theta (1-t)^{-\frac{1}{\alpha}} \end{aligned}$$

Επομένως

$$F^{-1}(t) = \theta(1-t)^{-\frac{1}{\alpha}}, \quad \alpha > 0, \ 0 \le t \le 1$$

Για την καμπύλη Lorenz εφαρμόζουμε τον τύπο (2.8.1) και έχουμε

$$L(u) = \frac{\int_0^u F^{-1}(t)dt}{\int_0^1 F^{-1}(t)dt}$$

$$= \frac{\int_0^u \theta(1-t)^{-\frac{1}{\alpha}} dt}{\int_0^1 \theta(1-t)^{-\frac{1}{\alpha}} dt}$$

$$= \frac{\theta \frac{\alpha}{\alpha-1} \int_0^u d\left[(1-t)^{\frac{\alpha-1}{\alpha}}\right]}{\theta \frac{\alpha}{\alpha-1} \int_0^1 d\left[(1-t)^{\frac{\alpha-1}{\alpha}}\right]}$$

$$= \frac{(1-u)^{\frac{\alpha-1}{\alpha}} - 1}{-1}$$

$$= 1 - (1-u)^{\frac{\alpha-1}{\alpha}}, \quad 0 \le u \le 1$$

Το διπλάσιο του εμβαδού του χωρίου που περικλείεται μεταξύ της καμπύλης Lorenz και της εξισωτικής γραμμής είναι ο δείκτης Gini όπως δίνεται από τον τύπο (2.7.4).

ΑΠΟΔΕΙΞΗ

Το εμβαδό του χωρίου που περικλείεται μεταξύ της καμπύλης Lorenz και της εξισωτικής γραμμής είναι

$$E = \int_0^1 [u - L(u)]du$$

Άρα έχουμε

$$G^* = 2\int_0^1 [u - L(u)]du$$

$$= 2\int_0^1 udu - 2\int_0^1 \left[1 - (1-u)^{\frac{\alpha-1}{\alpha}}\right]du$$

$$= 2\left[\frac{u^2}{2}\right]_0^1 - 2\int_0^1 du + 2\int_0^1 (1-u)^{\frac{\alpha-1}{\alpha}} du$$

$$= 1 - 2 - \frac{2\alpha}{2\alpha - 1}\int_0^1 d\left[(1-u)^{\frac{2\alpha-1}{\alpha}}\right] =$$

$$= -1 - \frac{2\alpha}{2\alpha-1}\left[(1-u)^{\frac{2\alpha-1}{\alpha}}\right]_0^1$$

$$= -1 + \frac{2\alpha}{2\alpha-1}$$

$$= \frac{-2\alpha + 1 + 2\alpha}{2\alpha - 1}$$

$$= (2\alpha - 1)^{-1}$$

Ένας άλλος δείκτης ανισότητας σχετικός με την καμπύλη Lorenz είναι ο δείκτης Pietra. Ο δείκτης αυτός ορίζεται ως η μέγιστη κάθετη απόκλιση ανάμεσα στην καμπύλη Lorenz και στην εξισωτική γραμμή. Στην περίπτωση που η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta, \alpha)$ τότε ο δείκτης Pietra είναι

$$\text{Pietra} = \frac{(\alpha-1)^{\alpha-1}}{\alpha^{\alpha}}, \quad \alpha > 1 \qquad (2.8.3)$$

ΑΠΟΔΕΙΞΗ

Η συνάρτηση που εκφράζει την κάθετη απόκλιση ανάμεσα στην καμπύλη Lorenz και στην εξισωτική γραμμή είναι

$$H(u) = u - L(u)$$

$$= u - 1 + (1-u)^{\frac{\alpha-1}{\alpha}}, \quad \alpha > 1,\ 0 \le u \le 1$$

Η παράγωγος της συνάρτησης αυτής είναι

$$\frac{d}{du}H(u) = \frac{d}{du}\left[u - 1 + (1-u)^{\frac{\alpha-1}{\alpha}}\right]$$

$$= 1 - \frac{\alpha-1}{\alpha}(1-u)^{-\frac{1}{\alpha}} \quad , \quad \alpha > 0,\ 0 \le u \le 1$$

Η τιμή του u για την οποία παρουσιάζει μέγιστο η συνάρτηση H(u) είναι η λύση της εξίσωσης

$$\mathbf{H(u)} = 0 \iff 1 - \frac{\alpha-1}{\alpha}(1-u)^{-\frac{1}{\alpha}} = 0$$

$$\iff \frac{\alpha-1}{\alpha}(1-u)^{-\frac{1}{\alpha}} = 1$$

$$\iff (1-u)^{\frac{1}{\alpha}} = \frac{\alpha-1}{\alpha}$$

$$\iff u = 1 - \frac{(\alpha-1)^{\alpha}}{\alpha^{\alpha}}$$

Το μέγιστο που παρουσιάζει η συνάρτηση H(u) δηλαδή ο δείκτης Pietra είναι

$$H\left[1 - \frac{(\alpha-1)^{\alpha}}{\alpha^{\alpha}}\right] = 1 - \frac{(\alpha-1)^{\alpha}}{\alpha^{\alpha}} - 1 + \left[1 - 1 + \frac{(\alpha-1)^{\alpha}}{\alpha^{\alpha}}\right]^{\frac{\alpha-1}{\alpha}}$$

$$= -\frac{(\alpha-1)^{\alpha}}{\alpha^{\alpha}} + \frac{(\alpha-1)^{\alpha-1}}{\alpha^{\alpha-1}}$$

$$= \frac{(\alpha-1)^{\alpha-1}}{\alpha^{\alpha-1}}\left(1 - \frac{\alpha-1}{\alpha}\right)$$

$$= \frac{(\alpha-1)^{\alpha-1}}{\alpha^{\alpha}}$$

ΚΕΦΑΛΑΙΟ 3$^{\text{Ο}}$

Κατανομές διατεταγμένων στατιστικών.

Στο κεφάλαιο αυτό θα εισάγουμε τη έννοια των διατεταγμένων στατιστικών μιας συνεχούς κατανομής και θα προσδιορίσουμε την κατανομή και άλλα χαρακτηριστικά αυτών. Στη συνέχεια θα εφαρμόσουμε την θεωρία των διατεταγμένων στατιστικών στην κατανομή Pareto.

3.1 Η έννοια του στοιχείου πιθανότητας.

Έστω η συνεχής τυχαία μεταβλητή X με συνάρτηση πυκνότητας πιθανότητας f(x) και συνάρτηση κατανομής F(x). Η πιθανότητα ώστε η τυχαία μεταβλητή X να παίρνει τιμές στο διάστημα $(x, x+dx]$ είναι

$$P(x < X \leq x + dx) = \int_{x}^{x+dx} f(u)du$$

Επειδή το παραπάνω ολοκλήρωμα προσεγγιστικά είναι ίσο με το διαφορικό f(x)dx, έχουμε

$$P(x < X \leq x + dx) = f(x)dx \qquad (3.1.1)$$

Η ποσότητα f(x)dx η οποία δίνεται από τον τύπο (3.1.1) λέγεται στοιχείο πιθανότητας της συνεχούς τυχαίας μεταβλητής X.

Η πιθανότητα ώστε η τυχαία μεταβλητή X να παίρνει τιμές στο διάστημα $(x+dx, y]$ με $x < y$ είναι

$$P(x+dx < X \leq y) = \int_{x+dx}^{y} f(u)du$$

Επειδή το παραπάνω ολοκλήρωμα προσεγγιστικά είναι ίσο με

$$\int_{x}^{y} f(u)du = F(y) - F(x)$$

έχουμε ότι

$$P(x+dx < X \leq y) = F(y) - F(x) \tag{3.1.2}$$

Η πιθανότητα ώστε η τυχαία μεταβλητή X να παίρνει τιμές στο διάστημα $(x+dx, +\infty)$ είναι

$$P(x+dx < X) = \int_{x+dx}^{+\infty} f(u)du$$

Επειδή το παραπάνω ολοκλήρωμα προσεγγιστικά είναι ίσο με

$$\int_{x}^{+\infty} f(u)du = 1 - \int_{-\infty}^{x} f(u)du = 1 - F(x)$$

έχουμε ότι

$$P(x+dx < X) = 1 - F(x) \tag{3.1.3}$$

Η έννοια του στοιχείου πιθανότητας επεκτείνεται και στη n-διάστατο συνεχή τυχαία μεταβλητή δηλαδή στην περίπτωση του n-διάστατου τυχαίου διανύσματος $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$. Αν $f(x_1, x_2, \ldots, x_n)$ είναι η από κοινού συνάρτηση πυκνότητας πιθανότητας του n-διάστατου τυχαίου διανύσματος $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ τότε η πιθανότητα ώστε οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ να παίρνουν συγχρόνως τιμές στα διαστήματα $(x_1, x_1 + dx_1]$, $(x_2, x_2 + dx_2]$, ... ,$(x_n, x_n + dx_n]$ αντιστοίχως είναι

$$P(x_1 < X_1 \leq x_1 + dx_1, \ldots, x_n < X_n \leq x_n + dx_n) = \int_{x_1}^{x_1+dx_1} \ldots \int_{x_n}^{x_n+dx_n} f(u_1, \ldots, u_n)du_1 \cdots du_n$$

Επειδή το παραπάνω ολοκλήρωμα προσεγγιστικά είναι ίσο με $f(x_1, x_2, \ldots, x_n)dx_1 dx_2 \ldots dx_n$ έχουμε ότι

$$P(x_1 < X_1 \leq x_1 + dx_1, \ldots, x_n < X_n \leq x_n + dx_n) = f(x_1, x_2, \ldots, x_n)dx_1 dx_2 \cdots dx_n \tag{3.1.4}$$

Η ποσότητα $f(x_1,x_2,\ldots,x_n)dx_1dx_2\cdots dx_n$ η οποία δίνεται από τον τύπο (3.1.4) λέγεται στοιχείο πιθανότητας της n-διάστατης συνεχούς τυχαίας μεταβλητής $\underset{\sim}{X}=(X_1,X_2,\ldots,X_n)$.

3.2 Ορισμός διατεταγμένων στατιστικών συνεχούς κατανομής.

Έστω $x_1,x_2,\ldots,x_n$ οι αντίστοιχες τιμές τυχαίου δείγματος μεγέθους n $X_1,X_2,\ldots,X_n$ ανεξάρτητων τυχαίων μεταβλητών που κάθε μια έχει συνεχή κατανομή με συνάρτηση πυκνότητας πιθανότητας f(x) και συνάρτηση κατανομής F(x). Αν οι τιμές $x_1,x_2,\ldots,x_n$ διαταχθούν κατά αύξουσα σειρά και στη συνέχεια συμβολιστούν με $x_{j:n}$ όπου $j=1,2,\ldots,n$ τότε έχουμε $x_{1:n}<x_{2:n}<\ldots<x_{n:n}$. Στην προηγούμενη διάταξη αποκλείεται η ισότητα δύο ή περισσοτέρων τιμών εφόσον η κατανομή είναι συνεχής και η πιθανότητα δύο ή περισσότερες τιμές να συμπίπτουν είναι 0. Οι τυχαίες μεταβλητές $X_{1:n},X_{2:n},\ldots,X_{n:n}$ που έχουν τιμές $x_{1:n},x_{2:n},\ldots,x_{n:n}$ αντίστοιχα, λέγονται διατεταγμένα στατιστικά του τυχαίου δείγματος $X_1,X_2,\ldots,X_n$ μεγέθους n του πληθυσμού. Γενικά η τυχαία μεταβλητή $X_{r:n}$ όπου $r=1,2,\ldots,n$ λέγεται διατεταγμένο στατιστικό τάξεως r. Είναι αυτονόητο ότι ισχύει $X_{1:n}<X_{2:n}<\cdots<X_{n:n}$.

Η τυχαία μεταβλητή $R=X_{n:n}-X_{1:n}$ λέγεται εύρος του τυχαίου δείγματος.

Η τυχαία μεταβλητή $\frac{1}{2}(X_{1:n}+X_{n:n})$ λέγεται μεσαίο σημείο του τυχαίου δείγματος.

Τέλος η τυχαία μεταβλητή

$$\tilde{X}=\begin{cases} X_{\frac{n+1}{2}:n} & \text{αν n περιττός} \\ \frac{1}{2}\left(X_{\frac{n}{2}:n}+X_{\frac{n}{2}+1:n}\right) & \text{αν n άρτιος} \end{cases}$$

λέγεται διάμεσος του τυχαίου δείγματος.

3.3 Συνάρτηση πιθανότητας διατεταγμένων στατιστικών.

Συμβολίζουμε με $f_r(x)$ όπου $r = 1,2,\ldots,n$ τη συνάρτηση πιθανότητας του διατεταγμένου στατιστικού $X_{r:n}$ τάξεως r. Μπορούμε να υπολογίσουμε τη συνάρτηση αυτή.

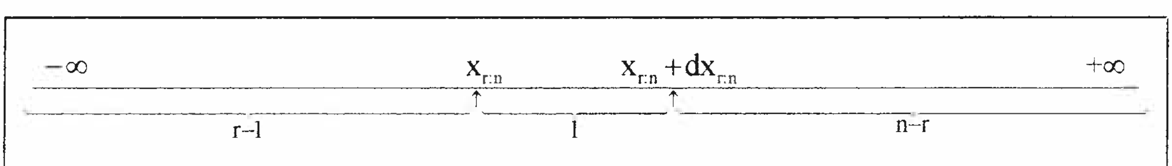
ΘΕΩΡΗΜΑ 3.3.1

Έστω $X_1, X_2, \ldots, X_n$ τυχαίο δείγμα του πληθυσμού με συνεχή κατανομή που έχει συνάρτηση πιθανότητας f(x) και συνάρτηση κατανομής F(x). Τότε η συνάρτηση πιθανότητας του διατεταγμένου στατιστικού $X_{r:n}$ τάξης r όπου $r = 1,2,\ldots,n$ είναι

$$f_r(x) = r\binom{n}{r} f(x) F^{r-1}(x)\left[1 - F(x)\right]^{n-r} \tag{3.3.1}$$

ΑΠΟΔΕΙΞΗ

Αν $x_{r:n}$ είναι η τιμή του διατεταγμένου στατιστικού $X_{r:n}$, τότε αυτή βρίσκεται στο διάστημα $(x_{r:n}, x_{r:n} + dx_{r:n}]$. Οι υπόλοιπες $n-1$ στο πλήθος τιμές των υπολοίπων διατεταγμένων $X_{j:n}$ με $j = 1,2,\ldots,r-1,r+1,\ldots,n$ βρίσκονται στα διαστήματα $(-\infty, x_{r:n}]$ και $(x_{r:n} + dx_{r:n}, +\infty)$. Πιο συγκεκριμένα στο διάστημα $(-\infty, x_{r:n}]$ βρίσκονται $r-1$ πλήθος τιμών και στο διάστημα $(x_{r:n} + dx_{r:n}, +\infty)$ βρίσκονται $n-r$ πλήθος τιμών. Αυτό φαίνεται στο σχήμα (3.3.1) όπου οι αριθμοί $r-1, 1, n-r$ παριστάνουν το πλήθος των τιμών στα αντίστοιχα διαστήματα.



Σχήμα 3.3.1 Πλήθος τιμών των διατεταγμένων στατιστικών στα διαστήματα.

Για τη λήψη των τιμών $x_1, x_2, \ldots, x_n$ των τυχαίων μεταβλητών $X_1, X_2, \ldots, X_n$ αντιστοίχως, το πείραμα εκτελέστηκε προφανώς n ανεξάρτητες φορές που αντιστοιχούν στις τυχαίες μεταβλητές του δείγματος αυτού. Μια συγκεκριμένη τυχαία μεταβλητή X_t με $t = 1,2,\ldots,n$ λόγω της τυχαιότητας μπορεί να πάρει οποιαδήποτε τιμή μικρότερη της $x_{r:n}$ με πιθανότητα

$$p_1 = P(X_t \le x_{r:n})$$

$$= \int_{-\infty}^{x_{r:n}} f(u)du$$

$$= F(x_{r:n})$$

μπορεί να πάρει τη τιμή $x_{r:n}$ με πιθανότητα

$$p_2 = P(x_{r:n} < X_t \le x_{r:n} + dx_{r:n})$$

$$= \int_{x_{r:n}}^{x_{r:n}+dx_{r:n}} f(u)du$$

$$= f(x_{r:n})dx_{r:n}$$

και τέλος μπορεί να πάρει οποιαδήποτε τιμή μεγαλύτερη της $x_{r:n}$ με πιθανότητα

$$p_3 = P(x_{r:n} + dx_{r:n} < X_t)$$

$$= \int_{x_{r:n}+dx_{r:n}}^{+\infty} f(u)du$$

$$= 1 - F(x_{r:n})$$

Σε κάθε εκτέλεση του πειράματος υπάρχουν τρία δυνατά αποτελέσματα με σταθερές πιθανότητες p_1, p_2, p_3 αντίστοιχα όπου

$$p_1 + p_2 + p_3 = 1$$

Επομένως η πιθανότητα $P(x_{r:n} < X_{r:n} \le x_{r:n} + dx_{r:n})$ μπορεί να βρεθεί από την πολυωνυμική κατανομή ως εξής

$$P(x_{r:n} < X_{r:n} \le x_{r:n} + dx_{r:n}) = \frac{n!}{(r-1)!1!(n-r)!} p_1^{r-1} p_2^1 p_3^{n-r}$$

Από τον τύπο (3.1.1) ισχύει

$$P(x_{r:n} < X_{r:n} \le x_{r:n} + dx_{r:n}) = f_r(x_{r:n})dx_{r:n}$$

οπότε για τη συνάρτηση πιθανότητας $f_r(x_{r:n})$ έχουμε.

$$f_r(x_{r:n})dx_{r:n} = P(x_{r:n} < X_{r:n} \le x_{r:n} + dx_{r:n})$$

$$= \frac{n!}{(r-1)!1!(n-r)!} p_1^{r-1} p_2^1 p_3^{n-r}$$

$$= \frac{n!}{(r-1)!(n-r)!} F^{r-1}(x_{r:n}) f(x_{r:n}) dx_{r:n} \left[1 - F(x_{r:n})\right]^{n-r}$$

$$= r\frac{n!}{r(r-1)!(n-r)!} f(x_{r:n}) F^{r-1}(x_{r:n}) \left[1 - F(x_{r:n})\right]^{n-r} dx_{r:n}$$

$$= r\frac{n!}{r!(n-r)!} f(x_{r:n}) F^{r-1}(x_{r:n}) \left[1 - F(x_{r:n})\right]^{n-r} dx_{r:n}$$

$$= r\binom{n}{r} f(x_{r:n}) F^{r-1}(x_{r:n}) \left[1 - F(x_{r:n})\right]^{n-r} dx_{r:n}$$

και επειδή όταν τα στοιχεία πιθανότητας είναι ίσα και οι κατανομές αυτών είναι ίσες, έχουμε

$$f_r(x_{r:n}) = r\binom{n}{r} f(x_{r:n}) F^{r-1}(x_{r:n}) \left[1 - F(x_{r:n})\right]^{n-r}$$

Τέλος αν στην παραπάνω συνάρτηση στη θέση του $x_{r:n}$ θέσουμε το x, προκύπτει η συνάρτηση πιθανότητας $f_r(x)$ του διατεταγμένου στατιστικού $X_{r:n}$ όπως δίνεται από τον τύπο (3.3.1)

$$f_r(x) = r\binom{n}{r} f(x) F^{r-1}(x) \left[1 - F(x)\right]^{n-r}$$

Με εφαρμογή του τύπου (3.3.1) για $r = 1$ προκύπτει η συνάρτηση πιθανότητας του μικρότερου διατεταγμένου στατιστικού $X_{1:n}$

$$f_1(x) = nf(x)\left[1 - F(x)\right]^{n-1} \qquad (3.3.2)$$

Επίσης με εφαρμογή του ίδιου τύπου (3.3.1) για $r = n$ προκύπτει η συνάρτηση πιθανότητας του μεγαλύτερου διατεταγμένου στατιστικού $X_{n:n}$

$$f_n(x) = nf(x)F^{n-1}(x) \qquad (3.3.3)$$

3.4 Συνάρτηση κατανομής διατεταγμένων στατιστικών.

Συμβολίζουμε με $F_r(x)$ όπου $r = 1,2,\ldots,n$ τη συνάρτηση κατανομής του διατεταγμένου στατιστικού $X_{r:n}$ τάξεως r. Μπορούμε να υπολογίσουμε τη συνάρτηση αυτή.

ΘΕΩΡΗΜΑ 3.4.1

Έστω $X_1, X_2, \ldots, X_n$ τυχαίο δείγμα του πληθυσμού με συνεχή κατανομή που έχει συνάρτηση πιθανότητας f(x) και συνάρτηση κατανομής F(x). Τότε η συνάρτηση κατανομής $F_r(x)$ του διατεταγμένου στατιστικού $X_{r:n}$ τάξης r όπου $r = 1,2,\ldots,n$ είναι

$$F_r(x) = r\binom{n}{r}\int_0^{F(x)} t^{r-1}(1-t)^{n-r}dt \tag{3.4.1}$$

ΑΠΟΔΕΙΞΗ

Η συνάρτηση κατανομής $F_r(x)$ του διατεταγμένου στατιστικού $X_{r:n}$ τάξης r είναι

$$F_r(x) = \int_{-\infty}^{x} f_r(u)du$$

και λόγω του τύπου (3.3.1) η παραπάνω ισότητα γράφεται

$$F_r(x) = \int_{-\infty}^{x} r\binom{n}{r} f(u)F^{r-1}(u)[1-F(u)]^{n-r}\,du$$

$$= r\binom{n}{r}\int_{-\infty}^{x} f(u)F^{r-1}(u)[1-F(u)]^{n-r}\,du$$

Θέτουμε

$$F(u) = t$$

οπότε

$$dF(u) = dt \Leftrightarrow f(u)du = dt$$

και

$$\lim_{u\to-\infty} F(u) = 0$$

$$\lim_{u\to x} F(u) = F(x)$$

Τότε προκύπτει η συνάρτηση κατανομής $F_r(x)$ όπως δίνεται από τον τύπο (3.4.1)

$$F_r(x) = r\binom{n}{r}\int_0^{F(x)} t^{r-1}(1-t)^{n-r}dt$$

Με εφαρμογή του τύπου (3.4.1) για $r = 1$ προκύπτει

$$F_1(x) = 1\binom{n}{1}\int_0^{F(x)} (1-t)^{n-1}dt$$

$$= n\int_0^{F(x)} (1-t)^{n-1}dt$$

$$= -\int_0^{F(x)} d[(1-t)^n]$$

$$= 1-[1-F(x)]^n$$

Άρα η συνάρτηση κατανομής του μικρότερου διατεταγμένου στατιστικού $X_{1:n}$ είναι

$$F_1(x) = 1-[1-F(x)]^n \tag{3.4.2}$$

Με εφαρμογή του τύπου (3.4.1) για $r = n$ προκύπτει

$$F_n(x) = n\binom{n}{n}\int_0^{F(x)} t^{n-1}dt$$

$$= \int_0^{F(x)} d(t^n)$$

$$= F^n(x)$$

Άρα η συνάρτηση κατανομής του μεγαλύτερου διατεταγμένου στατιστικού $X_{n:n}$ είναι

$$F_n(x) = F^n(x) \tag{3.4.3}$$

3.5 Η από κοινού κατανομή των διατεταγμένων στατιστικών.

Συμβολίζουμε με $f_{ij}(x_i, x_j)$ την από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{i:n} < X_{j:n}$ όπου $1 \le i < j \le n$, με $f_{r_1 r_2 \dots r_k}(x_1, x_2, \dots x_k)$ την από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{r_1:n} < X_{r_2:n} < \cdots < X_{r_k:n}$ όπου $1 \le k \le n$, $1 \le r_1 < r_2 < \cdots < r_k \le n$ και με

$f_{r_1 r_2 \dots r_n}(x_1, x_2, \dots, x_n)$ την από κοινού συνάρτηση πιθανότητας όλων των διατεταγμένων στατιστικών $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$

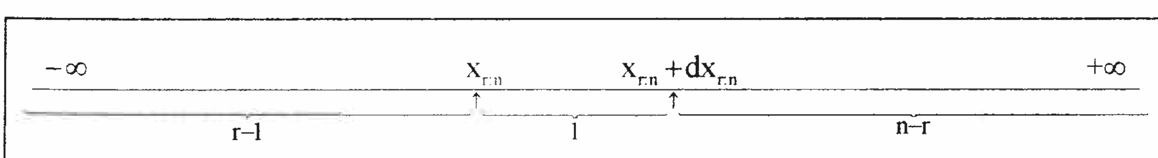
ΘΕΩΡΗΜΑ 3.5.1

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα του πληθυσμού με συνεχή κατανομή που έχει συνάρτηση πιθανότητας f(x) και συνάρτηση κατανομής F(x). Τότε η από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{i:n} < X_{j:n}$ με $1 \le i < j \le n$ είναι

$$f_{ij}(x_i, x_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F^{i-1}(x_i)[F(x_j) - F(x_i)]^{j-i-1}[1 - F(x_j)]^{n-j} f(x_i) f(x_j) \tag{3.5.1}$$

ΑΠΟΔΕΙΞΗ

Αν $x_{i:n} < x_{j:n}$ είναι οι αντίστοιχες τιμές των διατεταγμένων στατιστικών $X_{i:n} < X_{j:n}$ με $1 \le i < j \le n$ τότε αυτές βρίσκονται στα διαστήματα $(x_{i:n}, x_{i:n} + dx_{i:n}]$ και $(x_{j:n}, x_{j:n} + dx_{j:n}]$ αντίστοιχα. Οι υπόλοιπες $n-2$ στο πλήθος τιμές των υπολοίπων διατεταγμένων στατιστικών $X_{k:n}$ με $k = 1, 2, \dots, i-1, i+1, \dots, j-1, j+1, \dots, n$ βρίσκονται στα διαστήματα $(-\infty, x_{i:n}]$, $(x_{i:x} + dx_{i:x}, x_{j:n}]$ και $(x_{j:n} + dx_{j:n}, +\infty)$. Πιο συγκεκριμένα στο διάστημα $(-\infty, x_{i:n}]$ βρίσκονται $i-1$ πλήθος τιμών, στο διάστημα $(x_{i:x} + dx_{i:x}, x_{j:n}]$ βρίσκονται $j-i-1$ πλήθος τιμών και στο διάστημα $(x_{j:n} + dx_{j:n}, +\infty)$ βρίσκονται $n-j$ πλήθος τιμών. Αυτό φαίνεται στο παρακάτω σχήμα (3.5.1) όπου οι αριθμοί $i-1, 1, j-i-1, 1, n-j$ παριστάνουν το πλήθος των τιμών στα αντίστοιχα διαστήματα.



Σχήμα 3.5.1 Πλήθος τιμών των διατεταγμένων στατιστικών στα διαστήματα.

Για τη λήψη των τιμών $x_1, x_2, \dots, x_n$ των τυχαίων μεταβλητών $X_1, X_2, \dots, X_n$ αντιστοίχως, το πείραμα εκτελέστηκε προφανώς n ανεξάρτητες φορές που αντιστοιχούν στις τυχαίες μεταβλητές του δείγματος αυτού. Μια συγκεκριμένη τυχαία μεταβλητή X_t με $t = 1,2,\ldots,n$ λόγω της τυχαιότητας μπορεί να πάρει οποιαδήποτε τιμή μικρότερη της $x_{i:n}$ με πιθανότητα

$$p_1 = P(X_t \le x_{i:n})$$

$$= \int_{-\infty}^{x_{i:n}} f(u)du$$

$$= F(x_{i:n})$$

μπορεί να πάρει να πάρει τη τιμή $x_{i:n}$ με πιθανότητα

$$p_2 = P(x_{i:n} < X_t \le x_{i:n} + dx_{i:n})$$

$$= \int_{x_{i:n}}^{x_{i:n}+dx_{i:n}} f(u)du$$

$$= f(x_{i:n})dx_{i:n}$$

μπορεί πάρει οποιαδήποτε τιμή μεταξύ των $x_{i:n}$ και $x_{j:n}$ με πιθανότητα

$$p_3 = P(x_{i:n} + dx_{i:n} < X_t \le x_{j:n})$$

$$= \int_{x_{i:n}+dx_{i:n}}^{x_{j:n}} f(u)du$$

$$= \int_{x_{i:n}}^{x_{j:n}} f(u)du$$

$$= F(x_{j:n}) - F(x_{i:n})$$

μπορεί να πάρει τη τιμή $x_{j:n}$ με πιθανότητα

$$p_4 = P(x_{j:n} < X_t \le x_{j:n} + dx_{j:n})$$

$$= \int_{x_{j:n}}^{x_{j:n}+dx_{j:n}} f(u)du$$

$$= f(x_{j:n})dx_{j:n}$$

και τέλος μπορεί να πάρει οποιαδήποτε τιμή μεγαλύτερη της $x_{j:n}$ με πιθανότητα

$$p_5 = P(x_{j:n} + dx_{j:n} < X_t)$$

$$= \int_{x_{j:n}+dx_{j:n}}^{+\infty} f(u)du$$

$$= 1 - F(x_{j:n})$$

Σε κάθε εκτέλεση του πειράματος υπάρχουν πέντε δυνατά αποτελέσματα με σταθερές πιθανότητες $p_1, p_2, \ldots, p_5$ αντίστοιχα όπου

$$p_1 + p_2 + p_3 + p_4 + p_5 = 1$$

Επομένως η πιθανότητα $P(x_{i:n} < X_{i:n} \le x_{i:n} + dx_{i:n},\ x_{j:n} < X_{j:n} \le x_{j:n} + dx_{j:n})$ μπορεί να βρεθεί από την πολυωνυμική κατανομή ως εξής

$$p = P(x_{i:n} < X_{i:n} \le x_{i:n} + dx_{i:n},\ x_{j:n} < X_{j:n} \le x_{j:n} + dx_{j:n})$$

$$= \frac{n!}{(i-1)!1!(j-i-1)!1!(n-j)!} p_1^{i-1} p_2^1 p_3^{j-i-1} p_4^1 p_5^{n-j}$$

Από τον τύπο (3.1.4) ισχύει

$$P(x_{i:n} < X_{i:n} \le x_{i:n} + dx_{i:n},\ x_{j:n} < X_{j:n} \le x_{j:n} + dx_{j:n}) = f_{ij}(x_{i:n}, x_{j:n}) dx_{i:n} dx_{j:n}$$

οπότε για την από κοινού συνάρτηση πιθανότητας $f_{ij}(x_{i:n}, x_{j:n})$ έχουμε.

$$p = f_{ij}(x_{i:n}, x_{j:n}) dx_{i:n} dx_{j:n}$$

$$= P(x_{i:n} < X_{i:n} \le x_{i:n} + dx_{i:n},\ x_{j:n} < X_{j:n} \le x_{j:n} + dx_{j:n})$$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} p_1^{i-1} p_2 p_3^{j-i-1} p_4 p_5^{n-j}$$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F^{i-1}(x_{i:n})[F(x_{j:n}) - F(x_{i:n})]^{j-i-1}[1 - F(x_{j:n})]^{n-j} f(x_{i:n}) f(x_{j:n}) dx_{i:n} dx_{j:n}$$

και επειδή όταν τα στοιχεία πιθανότητας είναι ίσα και οι κατανομές αυτών είναι ίσες, έχουμε

$$f_{ij}(x_{i:n}, x_{j:n}) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F^{i-1}(x_{i:n})[F(x_{j:n}) - F(x_{i:n})]^{j-i-1}[1 - F(x_{j:n})]^{n-j} f(x_{i:n}) f(x_{j:n})$$

Τέλος αν στην παραπάνω συνάρτηση στη θέση των $x_{i:n}, x_{j:n}$ θέσουμε τα x_i, x_j αντίστοιχα προκύπτει η από κοινού συνάρτηση πιθανότητας $f_{ij}(x_i, x_j)$ των διατεταγμένων στατιστικών $X_{i:n} < X_{j:n}$ όπως αυτή δίνεται από τον τύπο (3.5.1)

$$f_{ij}(x_i, x_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F^{i-1}(x_i)[F(x_j) - F(x_i)]^{j-i-1}[1 - F(x_j)]^{n-j} f(x_i) f(x_j)$$

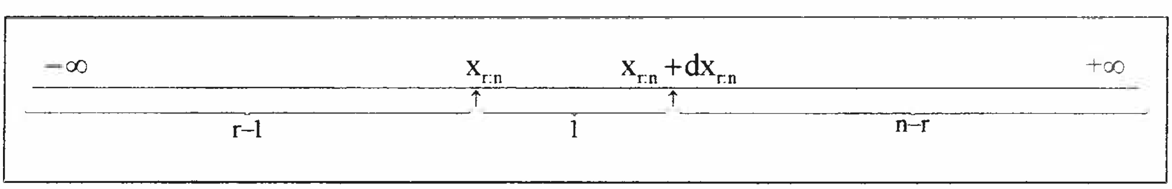
ΘΕΩΡΗΜΑ 3.5.2

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα του πληθυσμού με συνεχή κατανομή που έχει συνάρτηση πιθανότητας f(x) και συνάρτηση κατανομής F(x). Τότε η κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{r_1:n} < X_{r_2:n} < \cdots < X_{r_k:n}$ όπου $1 \le r_1 < r_2 < \cdots < r_k \le n$ και $r_0 = 0$, $F(x_0) = 0$, $r_{k+1} = n+1$, $F(x_{k+1}) = 1$ είναι

$$f_{r_1 r_2 \dots r_k}(x_1, x_2, \dots x_k) = n! \prod_{i=1}^{k} f(x_i) \prod_{j=0}^{k} \frac{\left[F(x_{j+1}) - F(x_j)\right]^{r_{j+1}-r_j-1}}{(r_{j+1} - r_j - 1)!} \qquad (3.5.2)$$

ΑΠΟΔΕΙΞΗ

Αν $x_{r_1:n} < x_{r_2:n} < \cdots < x_{r_k:n}$ είναι οι αντίστοιχες τιμές των διατεταγμένων στατιστικών $X_{r_1:n} < X_{r_2:n} < \cdots < X_{r_k:n}$ με $1 \le r_1 < r_2 < \cdots < r_k \le n$, τότε αυτές βρίσκονται η κάθε μια αντίστοιχα σε καθένα από στα διαστήματα της μορφής $\left(x_{r_i:n}, x_{r_i:n} + dx_{r_i:n}\right]$ με $i = 1, 2, \dots, k$. Από τις υπόλοιπες $n - (r_1 + r_2 + \cdots + r_k)$ στο πλήθος τιμές των υπολοίπων διατεταγμένων στατιστικών, $r_1 - 1$ πλήθος από αυτές βρίσκεται στο διάστημα $\left(-\infty, x_{r_1:n}\right]$, $r_j - r_{j-1} - 1$ πλήθος από αυτές με $j = 2, 3, \dots, k$ βρίσκονται στα διαστήματα της μορφής $\left(x_{r_{j-1}:n} + dx_{r_{j-1}:n}, x_{r_j:n}\right]$ αντίστοιχα και $n - r_k$ πλήθος από αυτές βρίσκονται στο διάστημα $\left(x_{r_k} + dx_{r_k}, +\infty\right)$. Αυτό φαίνεται στο σχήμα (3.5.2) όπου οι αριθμοί $r_1 - 1, 1, r_2 - r_1 - 1, 1, \dots, 1, r_k - r_{k-1} - 1, 1, n - r_k$ παριστάνουν το πλήθος των τιμών στα αντίστοιχα διαστήματα.



Σχήμα 3.5.2 Πλήθος τιμών των διατεταγμένων στατιστικών στα διαστήματα.

Για τη λήψη των τιμών $x_1, x_2, \dots, x_n$ των τυχαίων μεταβλητών $X_1, X_2, \dots, X_n$ αντιστοίχως, το πείραμα εκτελέστηκε προφανώς n ανεξάρτητες φορές που αντιστοιχούν στις τυχαίες μεταβλητές του δείγματος αυτού. Μια συγκεκριμένη

τυχαία μεταβλητή X_t με $t=1,2,\ldots,n$ λόγω της τυχαιότητας μπορεί να πάρει οποιαδήποτε τιμή στο διάστημα $\left(-\infty, x_{r_1:n}\right]$ **με πιθανότητα**

$$p_1 = P(X_t \le x_{r_1:n})$$

$$= \int_{-\infty}^{x_{r_1:n}} f(u)du$$

$$= F(x_{r_1:n})$$

μπορεί να πάρει οποιαδήποτε **από τις τιμές στα διαστήματα της μορφής** $\left(x_{r_i:n}, x_{r_i:n} + dx_{r_i:n}\right]$ όπου $i = 1,2,\ldots,k$ **με πιθανότητες αντίστοιχα**

$$p_{2i} = P(x_{r_i:n} < X_t \le x_{r_i:n} + dx_{r_i:n})$$

$$= \int_{x_{r_i:n}}^{x_{r_i:n}+dx_{r_i:n}} f(u)du$$

$$= f(x_{r_i:n})dx_{r_i:n}$$

μπορεί να **πάρει** οποιαδήποτε από τις **τιμές στα** διαστήματα της μορφής $\left(x_{r_{j-1}:n} + dx_{r_{j-1}:n}, x_{r_j:n}\right]$ όπου $j = 2,3,\ldots,k$ **με πιθανότητες αντίστοιχα**

$$p_{3j} = P(x_{r_{j-1}:n} + dx_{r_{j-1}:n} < X_t \le x_{r_j:n})$$

$$= \int_{x_{r_{j-1}:n}+dx_{r_{j-1}:n}}^{x_{r_j:n}} f(u)du$$

$$= \int_{x_{r_{j-1}:n}}^{x_{r_j:n}} f(u)du$$

$$= F(x_{r_j:n}) - F(x_{r_{j-1}:n})$$

και τέλος μπορεί να πάρει οποιαδήποτε τιμή στο διάστημα $\left(x_{r_k} + dx_{r_k}, +\infty\right)$ με πιθανότητα

$$p_4 = P(x_{r_k:n} + dx_{r_k:n} < X_t)$$

$$= \int_{x_{r_k:n}+dx_{r_k:n}}^{+\infty} f(u)du$$

$$= 1 - F(x_{r_k:n})$$

Σε κάθε εκτέλεση του πειράματος υπάρχουν 2κ+1 δυνατά αποτε**λέσματα με** σταθερές **πιθανότητες αντίστοιχα** p_1, p_{2i}, p_{3j}, p_4 **όπου** $i = 1, 2, \ldots, k$ και $j = 2, 3 \ldots, k$. Για τις πιθανότητες **αυτές ισχύει**

$$p_1 + \sum_{i=1}^{k} p_{2i} + \sum_{j=2}^{k} p_{3j} + p_4 = 1$$

Επομένως η πιθανότητα $P(x_{r_1:n} < X_{r_1:n} \le x_{r_1:n} + dx_{r_1:n}, \ldots, x_{r_k:n} < X_{r_k:n} \le x_{r_k:n} + dx_{r_k:n})$ **μπορεί να βρεθεί από την πολυωνυμική κατανομή ως εξής**

$$p = P(x_{r_1:n} < X_{r_1:n} \le x_{r_1:n} + dx_{r_1:n}, \ldots, x_{r_k:n} < X_{r_k:n} \le x_{r_k:n} + dx_{r_k:n})$$

$$= n! \frac{1}{(r_1 - 1)!} p_1^{r_1 - 1} \frac{1}{(n - r_k)!} p_4^{n - r_k} \prod_{i=1}^{k} \frac{p_{2i}^{1}}{1!} \prod_{j=2}^{k} \frac{p_{3j}^{r_j - r_{j-1} - 1}}{(r_j - r_{j-1} - 1)!}$$

$$= n! \frac{p_1^{r_1 - 1} p_4^{n - r_k}}{(r_1 - 1)!(n - r_k)!} p_1^{r_1 - 1} p_4^{n - r_k} \prod_{i=1}^{k} p_{2i} \prod_{j=2}^{k} \frac{p_{3j}^{r_j - r_{j-1} - 1}}{(r_j - r_{j-1} - 1)!}$$

Από τον τύπο (3.1.4) ισχύει

$$P(x_{r_1:n} < X_{r_1:n} \le x_{r_1:n} + dx_{r_1:n}, \ldots, x_{r_k:n} < X_{r_k:n} \le x_{r_k:n} + dx_{r_k:n}) = f_{r_1 r_2 \ldots r_k}(x_{r_1:n}, x_{r_2:n}, \ldots x_{r_k:n}) dx_{r_1:n} dx_{r_2:n} \cdots dx_{r_k:n}$$

οπότε για την από κοινού **συνάρτηση πιθανότητας** $f_{r_1 r_2 \ldots r_k}(x_{r_1:n}, x_{r_2:n}, \ldots x_{r_k:n})$ έχουμε.

$$p = f_{r_1 r_2 \ldots r_k}(x_{r_1:n}, x_{r_2:n}, \ldots x_{r_k:n}) dx_{r_1:n} dx_{r_2:n} \cdots dx_{r_k:n}$$

$$= P(x_{r_1:n} < X_{r_1:n} \le x_{r_1:n} + dx_{r_1:n}, \ldots, x_{r_k:n} < X_{r_k:n} \le x_{r_k:n} + dx_{r_k:n})$$

$$= n! \frac{p_1^{r_1 - 1} p_4^{n - r_k}}{(r_1 - 1)!(n - r_k)!} \prod_{i=1}^{k} p_{2i} \prod_{j=2}^{k} \frac{p_{3j}^{r_j - r_{j-1} - 1}}{(r_j - r_{j-1} - 1)!}$$

$$= n! \frac{F^{r_1 - 1}(x_{r_1:n}) \left[1 - F(x_{r_k:n})\right]^{n - r_k}}{(r_1 - 1)!(n - r_k)!} \prod_{i=1}^{k} f(x_{r_i:n}) dx_{r_i:n} \prod_{j=2}^{k} \frac{\left[F(x_{r_j:n}) - F(x_{r_{j-1}:n})\right]^{r_j - r_{j-1} - 1}}{(r_j - r_{j-1} - 1)!}$$

και επειδή όταν τα στοιχεία πιθανότητας **είναι ίσα και οι κατανομές** αυτών είναι **ίσες, έχουμε**

$$p = f_{r_1 r_2 \dots r_k}(x_{r_1:n}, x_{r_2:n}, \dots x_{r_k:n})$$

$$= n! \frac{F^{r_1 - 1}(x_{r_1:n})\left[1 - F(x_{r_k:n})\right]^{n-r_k}}{(r_1 - 1)!(n - r_k)!} \prod_{i=1}^{k} f(x_{r_i:n}) \prod_{j=2}^{k} \frac{\left[F(x_{r_j:n}) - F(x_{r_{j-1}:n})\right]^{r_j - r_{j-1} - 1}}{(r_j - r_{j-1} - 1)!}$$

Αν στη συνάρτηση αυτή στη θέση των $x_{r_1:n}, x_{r_2:n}, \dots x_{r_k:n}$ θέσουμε τα $x_1, x_2, \dots x_k$ αντίστοιχα, έχουμε

$$p = f_{r_1 r_2 \dots r_k}(x_1, x_2, \dots x_k)$$

$$= n! \frac{F^{r_1 - 1}(x_1)[1 - F(x_k)]^{n-r_k}}{(r_1 - 1)!(n - r_k)!} \prod_{i=1}^{k} f(x_i) \prod_{j=2}^{k} \frac{\left[F(x_j) - F(x_{j-1})\right]^{r_j - r_{j-1} - 1}}{(r_j - r_{j-1} - 1)!}$$

Τέλος αν θεωρήσουμε ότι $r_0 = 0$, $F(x_0) = 0$ και $r_{k+1} = n + 1$, $F(x_{k+1}) = 1$ τότε η παραπάνω μορφή του τύπου απλουστεύεται και προκύπτει η από κοινού συνάρτηση πιθανότητας $f_{r_1 r_2 \dots r_k}(x_1, x_2, \dots x_k)$ των διατεταγμένων στατιστικών $X_{r_1:n} < X_{r_2:n} < \dots < X_{r_k:n}$ όπως αυτή δίνεται από τον τύπο (3.5.2)

$$f_{r_1 r_2 \dots r_k}(x_1, x_2, \dots x_k) = n! \prod_{i=1}^{k} f(x_i) \prod_{j=1}^{k+1} \frac{\left[F(x_j) - F(x_{j-1})\right]^{r_j - r_{j-1} - 1}}{(r_j - r_{j-1} - 1)!}$$

ΘΕΩΡΗΜΑ 3.5.3

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα του πληθυσμού με συνεχή κατανομή που έχει συνάρτηση πιθανότητας f(x) και συνάρτηση κατανομής F(x). Τότε η από κοινού συνάρτηση πιθανότητας όλων των διατεταγμένων στατιστικών $X_{1:n} < X_{2:n} < \dots < X_{n:n}$ είναι

$$f_{r_1 r_2 \dots r_n}(x_1, x_2, \dots, x_n) = n! \prod_{j=1}^{n} f(x_j) \qquad (3.5.3)$$

ΑΠΟΔΕΙΞΗ

Αν $x_{1:n}, x_{2:n}, \dots, x_{n:n}$ είναι οι αντίστοιχες τιμές των διατεταγμένων στατιστικών $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ τότε αυτές βρίσκονται από μια αντίστοιχα στα διαστήματα $(x_{1:n}, x_{1:n} + dx_{1:n}]$, $(x_{2:n}, x_{2:n} + dx_{2:n}], \dots, (x_{n:n}, x_{n:n} + dx_{n:n}]$. Για τη λήψη των τιμών

$x_1, x_2, \ldots, x_n$ **των τυχαίων μεταβλητών** $X_1, X_2, \ldots, X_n$ **αντιστοίχως,** το πείραμα **εκτελέστηκε προφανώς n ανεξάρτητες φορές που αντιστοιχούν στις** τυχαίες **μεταβλητές του δείγματος αυτού. Μια συγκεκριμένη τυχαία μεταβλητή** X_t **με** $t = 1,2,\ldots,n$ **λόγω της τυχαιότητας μπορεί να πάρει οποιαδήποτε από** τις **τιμές** $x_{1:n}, x_{2:n}, \ldots, x_{n:n}$ που βρίσκονται στα διαστήματα $(x_{1:n}, x_{1:n} + dx_{1:n}]$, $(x_{2:n}, x_{2:n} + dx_{2:n}], \ldots, (x_{n:n}, x_{n:n} + dx_{n:n}]$ με πιθανότητες αντίστοιχα

$$p_1 = P(x_{1:n} < X_t \le x_{1:n} + dx_{1:n}) = \int_{x_{1:n}}^{x_{1:n}+dx_{1:n}} f(u)du = f(x_{1:n})dx_{1:n}$$

$$p_2 = P(x_{2:n} < X_t \le x_{2:n} + dx_{2:n}) = \int_{x_{2:n}}^{x_{2:n}+dx_{2:n}} f(u)du = f(x_{2:n})dx_{2:n}$$

$$p_3 = P(x_{3:n} < X_t \le x_{3:n} + dx_{3:n}) = \int_{x_{3:n}}^{x_{3:n}+dx_{3:n}} f(u)du = f(x_{3:n})dx_{3:n}$$

$$\vdots$$

$$\vdots$$

$$p_n = P(x_{n:n} < X_t \le x_{n:n} + dx_{n:n}) = \int_{x_{n:n}}^{x_{n:n}+dx_{n:n}} f(u)du = f(x_{n:n})dx_{n:n}$$

Σε κάθε εκτέλεση υπάρχουν n δυνατά αποτελέσματα με σταθερές πιθανότητες $p_1, p_2, \ldots, p_n$ αντίστοιχα όπου

$$p_1 + p_2 + \ldots + p_n = 1$$

Επομένως η πιθανότητα $P(x_{1:n} < X_{1:n} \le x_{1:n} + dx_{1:n}, \ldots, x_{n:n} < X_{n:n} \le x_{n:n} + dx_{n:n})$ **μπορεί** να βρεθεί από την πολυωνυμική κατανομή ως εξής

$$p = P(x_{1:n} < X_{1:n} \le x_{1:n} + dx_{1:n}, \ldots, x_{n:n} < X_{n:n} \le x_{n:n} + dx_{n:n})$$

$$= \frac{n!}{1!1!\cdots 1!} p_1^1 p_2^1 \cdots p_n^1$$

$$= n!\, p_1 p_2 \cdots p_n$$

Από τον τύπο (3.1.4) ισχύει

$$P(x_{1:n} < X_{1:n} \le x_{1:n} + dx_{1:n}, \ldots, x_{n:n} < X_{n:n} \le x_{n:n} + dx_{n:n}) = f_{r_1 r_2 \ldots r_n}(x_{1:n}, x_{2:n}, \ldots, x_{n:n}) dx_{1:n} dx_{2:n} \cdots dx_{n:n}$$

οπότε για την από κοινού συνάρτηση πιθανότητας $f_{r_1 r_2 \dots r_n}(x_{1:n}, x_{2:n}, \dots, x_{n:n})$ έχουμε.

$$p = f_{r_1 r_2 \dots r_n}(x_{1:n}, x_{2:n}, \dots, x_{n:n}) dx_{1:n} dx_{2:n} \cdots dx_{n:n}$$

$$= P(x_{1:n} < X_{1:n} \leq x_{1:n} + dx_{1:n}, \dots, x_{n:n} < X_{n:n} \leq x_{n:n} + dx_{n:n})$$

$$= n! p_1 p_2 \cdots p_n$$

$$= n! f(x_{1:n}) dx_{1:n} \cdots f(x_{n:n}) dx_{n:n}$$

$$= n! \prod_{j=1}^{n} f(x_{j:n}) dx_{j:n}$$

και επειδή όταν τα στοιχεία πιθανότητας είναι ίσα και οι κατανομές αυτών είναι ίσες, έχουμε

$$f_{r_1 r_2 \dots r_n}(x_{1:n}, x_{2:n}, \dots, x_{n:n}) = n! \prod_{j=1}^{n} f(x_{j:n})$$

Τέλος αν στη παραπάνω συνάρτηση στη θέση των $x_{1:n}, x_{2:n}, \dots, x_{n:n}$ θέσουμε τα $x_1, x_2, \dots, x_n$ αντίστοιχα, προκύπτει η από κοινού συνάρτηση πιθανότητας $f_{r_1 r_2 \dots r_n}(x_1, x_2, \dots, x_n)$ όλων των διατεταγμένων στατιστικών $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ όπως αυτή δίνεται από τον τύπο (3.5.3)

$$f_{r_1 r_2 \dots r_n}(x_1, x_2, \dots, x_n) = n! \prod_{j=1}^{n} f(x_j)$$

3.6 Ροπές k τάξης γύρω από το 0 διατεταγμένων στατιστικών.

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα του πληθυσμού με συνεχή κατανομή που έχει συνάρτηση πιθανότητας f(x) και συνάρτηση κατανομής F(x). Συμβολίζουμε με $\mu'_{r:n}(k)$ την ροπή k τάξης γύρω από το 0 του διατεταγμένου στατιστικού $X_{r:n}$ τάξης r για την οποία έχουμε

$$\mu'_{r:n}(k) = E(X^k_{r:n})$$

$$= \int_{-\infty}^{+\infty} x^k f_r(x) dx$$

και λόγω του τύπου (3.3.1) έχουμε

$$\mu'_{r:n}(k) = r\binom{n}{r}\int_{-\infty}^{+\infty} x^k f(x)F^{r-1}(x)\left[1-F(x)\right]^{n-r}dx$$

Θέτουμε

$$u = F(x)$$

οπότε

$$du = dF(x) = f(x)dx$$

$$x = F^{-1}(u)$$

$$\lim_{x\to-\infty} F(x) = 0$$

$$\lim_{x\to+\infty} F(x) = 1$$

Η ροπή $\mu'_{r:n}(k)$ **k τάξης γύρω από το 0 γράφεται**

$$\mu'_{r:n}(k) = r\binom{n}{r}\int_0^1 \left[F^{-1}(u)\right]^k u^{r-1}(1-u)^{n-r}du \qquad (3.6.1)$$

Ειδικότερα αν θέσουμε $k=1$ στον τύπο (3.6.1) έχουμε τη ροπή $\mu'_{r:n}(1)$ $1^{ης}$ **τάξης γύρω από το 0 του διατεταγμένου στατιστικού** $X_{r:n}$ τάξης r

$$\mu'_{r:n}(1) = r\binom{n}{r}\int_0^1 F^{-1}(u)u^{r-1}(1-u)^{n-r}du \qquad (3.6.2)$$

Συμβολίζουμε με $\mu'_{r,s:n}(k,m)$ την από κοινού ή αλλιώς παραγοντική ροπή τάξης **(k,m) γύρω από το 0 των διατεταγμένων στατιστικών $X_{r:n} < X_{s:n}$** η οποία δίνεται **από τον τύπο**

$$\mu'_{r,s:n}(k,m) = E\left(X_{r:n}^k X_{s:n}^m\right)$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_r^k, x_s^m f_{rs}(x_r, x_s)dx_r dx_s$$

Αν λάβουμε υπόψη μας **τον τύπο** (3.5.1) **έχουμε**

$$\mu'_{r,s:n}(k,m) = \rho\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_r^k, x_s^m F^{r-1}(x_r)[F(x_s)-F(x_r)]^{s-r-1}[1-F(x_s)]^{n-s}f(x_r)f(x_s)dx_r dx_s$$

όπου

$$\rho = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}$$

Θέτουμε

$$u = F(x_r)$$

και

$$v = F(x_s)$$

οπότε για $r < s$ έχουμε

$$du = dF(x_r) = f(x_r)dx_r$$

$$dv = dF(x_s) = f(x_s)dx_s$$

$$x_r = F^{-1}(u)$$

$$x_s = F^{-1}(v)$$

$$\lim_{x_r \to -\infty} F(x_r) = 0$$

$$\lim_{x_r \to +\infty} F(x_r) = F(x_s) = v$$

$$\lim_{x_s \to -\infty} F(x_s) = 0$$

$$\lim_{x_s \to +\infty} F(x_s) = 1$$

Η παραγοντική ροπή $\mu'_{r,s:n}(k,m)$ τάξης (k,m) γύρω από το 0 γράφεται

$$\mu'_{r,s:n}(k,m) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}\int_0^1\int_0^v \left[F^{-1}(u)\right]^k \left[F^{-1}(v)\right]^m u^{r-1}(v-u)^{s-r-1}(1-v)^{n-s}dudv \tag{3.6.3}$$

3.7 Διακύμανση και συνδιακύμανση διατεταγμένων στατιστικών.

Έστω $X_1, X_2, \ldots, X_n$ τυχαίο δείγμα του πληθυσμού με συνεχή κατανομή. Συμβολίζουμε με $\sigma^2_{r:n}$ τη διακύμανση του διατεταγμένου στατιστικού $X_{r:n}$ και έχουμε

$$\sigma^2_{r:n} = \mu'_{r:n}(2) - [\mu'_{r:n}(1)]^2 \tag{3.7.1}$$

Συμβολίζουμε με $\sigma_{r,s:n}$ τη συνδιακύμανση των διατεταγμένων στατιστικών $X_{r:n} < X_{s:n}$ και έχουμε

$$\sigma_{r,s:n} = \mu'_{r,s:n}(1,1) - \mu'_{r:n}(1)\mu'_{s:n}(1) \tag{3.7.2}$$

Ισχύουν

$$\sigma_{r,s:n} = \sigma_{s,r:n} \text{ και } \sigma_{r,r:n} = \sigma^2_{r:n}$$

3.8 Αναδρομικές σχέσεις μεταξύ των ροπών διατεταγμένων στατιστικών.

Για οποιαδήποτε τυχαία μεταβλητή με συνεχή κατανομή, μεταξύ των ροπών τάξης k γύρω από το 0 των διατεταγμένων στατιστικών ισχύουν οι παρακάτω αναδρομικές σχέσεις.

ΣΧΕΣΗ 3.8.1

$$i\cdot\mu'_{i+1:n}(k)+(n-i)\mu'_{i:n}(k)=n\cdot\mu'_{i:n-1}(k),\quad i=1,2,\ldots,n-1 \tag{3.8.1}$$

ΑΠΟΔΕΙΞΗ

Από τον **τύπο (3.6.1)** έχουμε

$$A_{\mu\acute{\varepsilon}\lambda o\varsigma}=i\cdot\mu'_{i+1:n}(k)+(n-i)\mu'_{i:n}(k)$$

$$=i(i+1)\binom{n}{i+1}\int_0^1\left[F^{-1}(u)\right]^k u^i(1-u)^{n-i-1}du+(n-i)i\binom{n}{i}\int_0^1\left[F^{-1}(u)\right]^k u^{i-1}(1-u)^{n-i}du$$

$$=\frac{i(i+1)n!}{(i+1)!(n-i-1)!}\int_0^1\left[F^{-1}(u)\right]^k u^i(1-u)^{n-i-1}du+\frac{i(n-i)n!}{i!(n-i)!}\int_0^1\left[F^{-1}(u)\right]^k u^{i-1}(1-u)^{n-i}du$$

$$=n\cdot i\cdot\binom{n-1}{i}\int_0^1\left[F^{-1}(u)\right]^k u^i(1-u)^{n-i-1}du+n\cdot i\cdot\binom{n-1}{i}\int_0^1\left[F^{-1}(u)\right]^k u^{i-1}(1-u)^{n-i}du$$

$$=n\cdot i\cdot\binom{n-1}{i}\int_0^1\left[\left[F^{-1}(u)\right]u^i(1-u)^{n-i-1}+\left[F^{-1}(u)\right]^k u^{i-1}(1-u)^{n-i}\right]du$$

$$=n\cdot i\cdot\binom{n-1}{i}\int_0^1\left[F^{-1}(u)\right]^k u^{i-1}(1-u)^{n-i-1}(u+1-u)du$$

$$=n\cdot i\cdot\binom{n-1}{i}\int_0^1\left[F^{-1}(u)\right]^k u^{i-1}(1-u)^{n-i-1}du$$

$$=n\cdot\mu'_{i:n-1}(k)$$

ΣΧΕΣΗ 3.8.2

Αν ν άρτιος τότε

$$\mu'_{\frac{n}{2}+1:n}(k)+\mu'_{\frac{n}{2}:n}(k)=2\mu'_{\frac{n}{2}:n-1}(k) \tag{3.8.2}$$

ΑΠΟΔΕΙΞΗ

Από τον τύπο (3.8.1) για $i=\frac{n}{2}$ έχουμε

$$\frac{n}{2}\mu'_{\frac{n}{2}+1:n}(k)+(n-\frac{n}{2})\mu'_{\frac{n}{2}:n}(k)=n\cdot\mu'_{\frac{n}{2}:n-1}(k) \Leftrightarrow \frac{n}{2}\mu'_{\frac{n}{2}+1:n}(k)+\frac{n}{2}\mu'_{\frac{n}{2}:n}(k)=n\cdot\mu'_{\frac{n}{2}:n-1}(k)$$

$$\Leftrightarrow \mu'_{\frac{n}{2}+1:n}(k)+\mu'_{\frac{n}{2}:n}(k)=2\mu'_{\frac{n}{2}:n-1}(k)$$

ΣΧΕΣΗ 3.8.3

$$\sum_{i=1}^{n}\mu'_{i:n}(k)=nE\left[X^k\right] \tag{3.8.3}$$

ΑΠΟΔΕΙΞΗ

Από τον τύπο (3.6.1) έχουμε

$$\sum_{i=1}^{n}\mu'_{i:n}(k)=\mu'_{1:n}(k)+\mu'_{2:n}(k)+\cdots+\mu'_{n:n}(k)$$

$$=1\binom{n}{1}\int_0^1\left[F^{-1}(u)\right]^k u^{1-1}(1-u)^{n-1}du+\cdots+n\binom{n}{n}\int_0^1\left[F^{-1}(u)\right]^k u^{n-1}(1-u)^{n-n}du$$

$$=\int_0^1\left[F^{-1}(u)\right]^k\left[1\binom{n}{1}u^{1-1}(1-u)^{n-1}+2\binom{n}{2}u^{2-1}(1-u)^{n-2}+\cdots+n\binom{n}{n}u^{n-1}(1-u)^{n-n}\right]du$$

$$=\int_0^1\left[F^{-1}(u)\right]^k n\left[\binom{n-1}{0}u^{1-1}(1-u)^{n-1}+\binom{n-1}{1}u^{2-1}(1-u)^{n-2}+\cdots+\binom{n-1}{n-1}u^{n-1}(1-u)^{n-n}\right]du$$

$$=n\int_0^1\left[F^{-1}(u)\right]^k(u+1-u)^{n-1}du$$

$$=n\int_0^1\left[F^{-1}(u)\right]^k du$$

Θέτουμε

$$F^{-1}(u) = t$$

οπότε

$$u = F(t)$$

$$du = dF(t) = f(t)dt$$

$$\lim_{u \to 0} F^{-1}(u) = -\infty$$

$$\lim_{u \to 1} F^{-1}(u) = +\infty$$

και το παραπάνω ολοκλήρωμα γίνεται

$$\sum_{i=1}^{n} \mu'_{i:n}(k) = n\int_0^1 \left[F^{-1}(u)\right]^k du$$

$$= n\int_{-\infty}^{+\infty} t^k f(t)dt$$

$$= nE\left[X^k\right]$$

3.9 Διατεταγμένα στατιστικά κατανομής Pareto.

Στην ενότητα αυτή θα παρουσιάσουμε μερικά συμπεράσματα για την κατανομή Pareto τα οποία προκύπτουν κάνοντας χρήση της παραπάνω θεωρίας των διατεταγμένων στατιστικών.

ΘΕΩΡΗΜΑ 3.9.1

Έστω η τυχαία μεταβλητή X που ακολουθεί κατανομή Pareto με παραμέτρους $\theta = 1$ και $\alpha > 0$, δηλαδή $X \sim P(1,\alpha)$. Τότε το διατεταγμένο στατιστικό $X_{k:n}$ τάξης κ μπορεί να γραφεί

$$X_{k:n} = \prod_{i=1}^{k} Y_i^{\frac{1}{n-i+1}} \tag{3.9.1}$$

για κάθε ακέραιο κ με $1 \le k \le n$ όπου οι τυχαίες μεταβλητές Y_i με $i = 1,2,\ldots,k$ είναι ανεξάρτητες και έχουν ακριβώς με τη ίδια συνάρτηση κατανομής με την X, δηλαδή $Y_i \sim P(1,\alpha)$.

<u>ΑΠΟΔΕΙΞΗ</u>

Σύμφωνα με τον τύπο (3.5.3) η από κοινού συνάρτηση πιθανότητας όλων των **διατεταγμένων στατιστικών** $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ είναι

$$f_{r_1 r_2 \ldots r_n}(x_1, x_2, \ldots, x_n) = n! \prod_{j=1}^{n} f(x_j)$$

Όμως για την κατανομή Pareto $P(I)(1,\alpha)$ **έχουμε**

$$f(x) = \alpha x^{-\alpha-1}$$

οπότε

$$f_{r_1 r_2 \ldots r_n}(x_1, x_2, \ldots, x_n) = n! \prod_{j=1}^{n} \alpha^n x_j^{-\alpha-1} = n! \alpha^n \prod_{j=1}^{n} x_j^{-\alpha-1}$$

$$= n! \alpha^n \left(\prod_{j=1}^{n} x_j \right)^{-\alpha-1}$$

Θεωρούμε το μετασχηματισμό

$$Y_1 = X_{1:n}^n \qquad \text{άρα } x_1 = y_1^{\frac{1}{n}} = \prod_{i=1}^{1} y_i^{\frac{1}{n-i+1}}$$

$$Y_2 = \left(\frac{X_{2:n}}{X_{1:n}} \right)^{n-1} \qquad \text{άρα } x_2 = x_1 y_2^{\frac{1}{n-1}} = y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} = \prod_{i=1}^{2} y_i^{\frac{1}{n-i+1}}$$

$$Y_3 = \left(\frac{X_{3:n}}{X_{2:n}} \right)^{n-2} \qquad \text{άρα } x_3 = x_2 y_3^{\frac{1}{n-2}} = y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} y_3^{\frac{1}{n-2}} = \prod_{i=1}^{3} y_i^{\frac{1}{n-i+1}}$$

$$\vdots$$

$$Y_k = \left(\frac{X_{k:n}}{X_{k-1:n}} \right)^{n-k+1} \qquad \text{άρα } x_k = x_{k-1} y_k^{\frac{1}{n-k+1}} = y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} \cdots y_k^{\frac{1}{n-k+1}} = \prod_{i=1}^{k} y_i^{\frac{1}{n-i+1}}$$

$$\vdots$$

$$Y_n = \frac{X_{n:n}}{X_{n-1:n}} \qquad \text{άρα } x_n = x_{n-1} y_n = y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} y_3^{\frac{1}{n-2}} \ldots y_n = \prod_{i=1}^{n} y_i^{\frac{1}{n-i+1}}$$

(3.9.2)

οπότε

$$f_{r_1 r_2 \ldots r_n}(y_1, y_2, \ldots, y_n) = n!\,\alpha^n \left(\prod_{j=1}^{n} \left(\prod_{i=1}^{j} y_i^{\frac{1}{n-i+1}} \right) \right)^{-\alpha-1} |J|$$

όπου $J = \dfrac{d(x_1, x_2, \ldots, x_n)}{d(y_1, y_2, \ldots, y_n)}$ είναι η Ιακωβιανή ορίζουσα για τη οποία έχουμε

$$J = \begin{vmatrix} \dfrac{dx_1}{dy_1} & \dfrac{dx_1}{dy_2} & \cdots & \dfrac{dx_1}{dy_n} \\ \dfrac{dx_2}{dy_1} & \dfrac{dx_2}{dy_2} & \cdots & \dfrac{dx_2}{dy_n} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{dx_n}{dy_1} & \dfrac{dx_n}{dy_2} & \cdots & \dfrac{dx_n}{dy_n} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{1}{n} y_1^{\frac{1}{n}-1} & 0 & \cdots & 0 \\ \dfrac{dx_2}{dy_1} & \dfrac{1}{n-1} y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}-1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{dx_n}{dy_1} & \dfrac{dx_n}{dy_2} & \cdots & y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} \cdots y_{n-1}^{\frac{1}{2}} \cdot 1 \end{vmatrix}$$

$$= \frac{1}{n} y_1^{\frac{1}{n}-1} \frac{1}{n-1} y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}-1} \frac{1}{n-2} y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} y_3^{\frac{1}{n-2}-1} \cdots y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} y_3^{\frac{1}{n-2}} \cdots y_{n-1}^{\frac{1}{2}} \cdot 1$$

$$= \frac{1}{n} \frac{1}{n-1} \cdots \frac{1}{2} 1 \cdot y_1^{\frac{1}{n}-1+\frac{n-1}{n}} y_2^{\frac{1}{n-1}-1+\frac{n-2}{n-1}} y_3^{\frac{1}{n-2}-1+\frac{n-3}{n-2}} \cdots y_{n-1}^{\frac{1}{2}-1+\frac{1}{2}}$$

$$= \frac{1}{1 \cdot 2 \cdots (n-1)n} y_1^0 y_2^0 \cdots y_n^0$$

$$= \frac{1}{n!}$$

Η συνάρτηση $f_{r_1 r_2 \dots r_n}(y_1, y_2, \dots, y_n)$ γράφεται

$$f_{r_1 r_2 \dots r_n}(y_1, y_2, \dots, y_n) = n! \frac{1}{n!} \alpha^n \left(\prod_{j=1}^{n} \left(\prod_{i=1}^{j} y_i^{\frac{1}{n-i+1}} \right) \right)^{-\alpha-1}$$

$$= \alpha^n \left[\left(y_1^{\frac{1}{n}} \right) \cdot \left(y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} \right) \cdot \left(y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} y_3^{\frac{1}{n-2}} \right) \cdots \left(y_1^{\frac{1}{n}} y_2^{\frac{1}{n-1}} y_3^{\frac{1}{n-2}} \dots y_n \right) \right]^{-\alpha-1}$$

$$= \alpha^n \left(y_1^{\frac{n}{n}} y_2^{\frac{n-1}{n-1}} y_3^{\frac{n-2}{n-2}} \cdots y_n^{\frac{1}{1}} \right)^{-\alpha-1}$$

$$= \alpha^n (y_1 y_2 y_3 \cdots y_n)^{-\alpha-1}$$

$$= \prod_{j=1}^{n} \alpha y_j^{-\alpha-1}$$

$$= \prod_{j=1}^{n} f(y_j) \tag{3.9.3}$$

Από τις παραπάνω **σχέσεις μετασχηματισμού** (3.9.2) **συμπεραίνουμε** ότι

$$X_{k:n} = \prod_{i=1}^{k} Y_i^{\frac{1}{n-i+1}}, \quad k = 1, 2, \dots, n$$

Από τη σχέση (3.9.3) συμπεραίνουμε **ότι οι τυχαίες μεταβλητές** $Y_1, Y_2, \dots, Y_n$ είναι **ανεξάρτητες** και έχουν την **ίδια συνάρτηση κατανομής με την** X.

ΘΕΩΡΗΜΑ 3.9.2

Έστω **η τυχαία μεταβλητή** X που ακολουθεί κατανομή **Pareto με παραμέτρους** $\theta = 1$ **και** $\alpha > 0$ δηλαδή $X \sim P(1, \alpha)$. Τότε οι λόγοι

$$R_{k:n} = \frac{X_{k:n}}{X_{k-1:n}}, \quad k = 1, 2, \dots, n \text{ και } X_{0:n} = 1 \tag{3.9.4}$$

είναι ανεξάρτητες μεταβλητές και ακολουθούν κατανομή Pareto με

$$R_{k:n} \sim P(1, (n-k+1)\alpha)$$

Επίσης οι τυχαίες μεταβλητές

$$d_{k:n} = (R_{k:n})^{n-k+1}, \quad k = 1, 2, \dots, n \tag{3.9.5}$$

είναι ανεξάρτητες μεταβλητές και ακολουθούν κατανομή Pareto με

$$d_{k:n} \sim P(1, \alpha).$$

ΑΠΟΔΕΙΞΗ

Συγκρίνοντας τους τύπους (3.9.2) και (3.9.4) παρατηρούμε ότι

$$R_{k:n} = Y_k^{\frac{1}{n-k+1}}, \quad k = 1,2,\ldots,n$$

άρα

$$r_j = y_j^{\frac{1}{n-j+1}} \Leftrightarrow y_j = r_j^{n-j+1}$$

και

$$\frac{dy_j}{dr_j} = (n-j+1)r_j^{n-j} > 0$$

Για τη συνάρτηση πιθανότητας $f_{r_1 r_2 \ldots r_n}(r_1, r_2, \ldots, r_n)$ από τον τύπο (3.9.3) έχουμε

$$f_{r_1 r_2 \ldots r_n}(r_1, r_2, \ldots, r_n) = \prod_{j=1}^{n} \alpha\left(r_j^{n-j+1}\right)^{-\alpha-1} \left|\frac{dy_j}{dr_j}\right|$$

$$= \prod_{j=1}^{n} \alpha(n-j+1) r_j^{n-j} r_j^{-(n-j+1)(\alpha+1)}$$

$$= \prod_{j=1}^{n} \alpha(n-j+1) r_j^{n-j-(n-j+1)\alpha-n+j-1}$$

$$= \prod_{j=1}^{n} \alpha(n-j+1) r_j^{-(n-j+1)\alpha-1}$$

$$= \prod_{j=1}^{n} f_j(r_j)$$

όπου

$$f_j(r_j) = \alpha(n-j+1) r_j^{-(n-j+1)\alpha-1}$$

Από τα παραπάνω συμπεραίνουμε ότι οι $R_{k:n}$ με $k = 1,2,\ldots,n$ είναι ανεξάρτητες μεταβλητές Pareto με

$$R_{k:n} \sim P\left(1,(n-k+1)\alpha\right)$$

Για να αποδείξουμε ότι οι $d_{k:n} = \left(R_{k:n}\right)^{n-k+1}$ με $k = 1,2,\ldots,n$ είναι ανεξάρτητες μεταβλητές Pareto με $d_{k:n} \sim P(1,\alpha)$ συγκρίνουμε τους τύπους (3.9.2) και (3.9.5) παρατηρούμε ότι $d_{k:n} = Y_k$ με $k = 1,2,\ldots,n$. Η από κοινού συνάρτηση πιθανότητας των Y_k δίνεται από τον τύπο (3.9.3) και είναι

$$f_{r_1 r_2 \ldots r_n}(y_1, y_2, \ldots, y_n) = \prod_{j=1}^{n} f(y_j)$$

από όπου προκύπτει ότι οι τυχαίες μεταβλητές Y_k άρα και οι $d_{k:n}$ είναι ανεξάρτητες μεταβλητές Pareto με

$$d_{k:n} \sim P(1,\alpha).$$

<u>ΘΕΩΡΗΜΑ 3.9.3</u>

Έστω η τυχαία μεταβλητή Χ που ακολουθεί κατανομή Pareto με παραμέτρους $\theta > 0$ και $\alpha > 0$ δηλαδή $X \sim P(\theta,\alpha)$. Τότε το διατεταγμένο στατιστικό $X_{k:n}$ τάξης k μπορεί να γραφεί

$$X_{k:n} = \theta \prod_{j=1}^{k} R_{j:n} \tag{3.9.6}$$

για κάθε ακέραιο κ με $1 \le k \le n$ όπου οι τυχαίες μεταβλητές $R_{j:n}$ με $j = 1,2,\ldots,n$ είναι ανεξάρτητες και ακολουθούν κατανομή Pareto.

<u>ΑΠΟΔΕΙΞΗ</u>

Θα κάνουμε χρήση του θεωρήματος Renyi για τα διατεταγμένα στατιστικά της εκθετικής κατανομής που αναφέρεται στον τύπο (Α13) του παραρτήματος. Σύμφωνα με αυτό αν η τυχαία μεταβλητή Ζ ακολουθεί εκθετική κατανομή τότε το διατεταγμένο στατιστικό $Z_{k:n}$ τάξης κ μπορεί να γραφεί

$$Z_{k:n} = \sum_{j=1}^{k} \frac{d_j}{n-j+1}$$

όπου

$$d_j = (n-j+1)\left(Z_{j:n} - Z_{j-1:n}\right), \quad j = 1,2,\ldots,n \text{ και } Z_{0:n} = 0$$

είναι ανεξάρτητες μεταβλητές που ακολουθούν εκθετική κατανομή.

Θεωρούμε το μετασχηματισμό

$$X = \theta \cdot \exp\left(\frac{Z}{\alpha\beta}\right), \quad \theta > 0,\ \alpha > 0,\ \beta > 0$$

Λόγω του θεωρήματος (1.4.1) η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta, \alpha)$.

Από τη σχέση μετασχηματισμού έχουμε

$$X = \theta \cdot \exp\left(\frac{Z}{\alpha\beta}\right) \Leftrightarrow \frac{X}{\theta} = \exp\left(\frac{Z}{\alpha\beta}\right)$$

$$\Leftrightarrow \log\left(\frac{X}{\theta}\right) = \frac{Z}{\alpha\beta}$$

$$\Leftrightarrow Z = \log\left(\frac{X}{\theta}\right)^{\alpha\beta}$$

οπότε

$$Z_{k:n} = \sum_{j=1}^{k} \frac{d_j}{n-j+1} \Leftrightarrow Z_{k:n} = \sum_{j=1}^{k} \frac{(n-j+1)\left(Z_{j:n} - Z_{j-1:n}\right)}{n-j+1}$$

$$\Leftrightarrow Z_{k:n} = \sum_{j=1}^{k} \left(Z_{j:n} - Z_{j-1:n}\right)$$

$$\Leftrightarrow \log\left(\frac{X_{k:n}}{\theta}\right)^{\alpha\beta} = \sum_{j=1}^{k}\left[\log\left(\frac{X_{j:n}}{\theta}\right)^{\alpha\beta} - \log\left(\frac{X_{j-1:n}}{\theta}\right)^{\alpha\beta}\right]$$

$$\Leftrightarrow \alpha\beta \log\left(\frac{X_{k:n}}{\theta}\right) = \alpha\beta \sum_{j=1}^{k} \log\left(\frac{X_{j:n}}{X_{j-1:n}}\right)$$

$$\Leftrightarrow \log\left(\frac{X_{k:n}}{\theta}\right) = \log \prod_{j=1}^{k} \frac{X_{j:n}}{X_{j-1:n}}$$

$$\Leftrightarrow X_{k:n} = \theta \prod_{j=1}^{k} \frac{X_{j:n}}{X_{j-1:n}}$$

Αν θέσουμε

$$R_{k:n} = \frac{X_{k:n}}{X_{k-1:n}}, \quad k = 1, 2, \ldots, n \text{ και } X_{0:n} = 1$$

τότε

$$X_{k:n} = \theta \prod_{j=1}^{k} R_{j:n}$$

όπου σύμφωνα με το θεώρημα (3.9.2) οι λόγοι $R_{k:n}$ είναι ανεξάρτητες μεταβλητές Pareto με $R_{k:n} \sim P\left(1, (n-r+1)\alpha\right)$ αφού $X \sim P(\theta, \alpha)$.

3.10 Ροπές διατεταγμένων στατιστικών κατανομής Pareto.

Πρώτος ο Malik υπολόγισε τις ροπές των διατεταγμένων στατιστικών από Pareto κατανομές. Για να βρει τις ροπές k τάξης γύρω από το 0 χρησιμοποίησε τη χαρακτηριστική συνάρτηση. Για να βρει τις από κοινού ροπές χρησιμοποίησε απευθείας ολοκλήρωση της από κοινού συνάρτησης πυκνότητας πιθανότητας των διατεταγμένων στατιστικών. Η μέθοδος είναι εύκολη για τις από κοινού ροπές δύο διατεταγμένων στατιστικών. Οι υπολογισμοί γίνονται πολύπλοκοι όταν έχουμε να υπολογίσουμε από κοινού ροπές περισσοτέρων των δύο διατεταγμένων στατιστικών. Μια ενδιαφέρουσα ιδέα είναι αυτή που πρότεινε ο Kabe η οποία απλοποιεί τη διαδικασία της παραπάνω απευθείας μεθόδου. Η απλοποίηση βασίζεται στη χρήση της θεωρίας Dirichlet's για την πολλαπλή ολοκλήρωση και το μετασχηματισμό που χρειάζεται για να υπολογιστεί το πολλαπλό ολοκλήρωμα. Η μέθοδος είναι απλούστερη από την απευθείας ολοκλήρωση που πρότεινε ο Malik, αλλά ακόμη χρειάζεται πολλαπλή ολοκλήρωση όπως και υπολογισμό της Ιακωβιανής ορίζουσας n διάστασης. Ο Houang χρησιμοποίησε το παραπάνω θεώρημα (3.9.3) για να λύσει το πρόβλημα του υπολογισμού των ροπών των διατεταγμένων στατιστικών. Στη συνέχεια θα παρουσιάσουμε τη μέθοδό του.

Το θεώρημα (3.9.3) μας επιτρέπει να εκφράσουμε το διατεταγμένο στατιστικό $X_{k:n}$ από κατανομή Pareto P(1,α) ως γινόμενο από ανεξάρτητες Pareto μεταβλητές

$$R_{j:n} \sim P\left(1,(n-j+1)\alpha\right), \quad j=1,2,\ldots,n$$

που σύμφωνα με τον τύπο (2.3.1) έχουν μέση τιμή

$$E\left[R_{j:n}\right]=\frac{(n-j+1)\alpha}{(n-j+1)\alpha-1}, \quad \alpha>\frac{1}{n-j+1}$$

Λαμβάνοντας υπόψιν το θεώρημα (1.9.2), οι **ανεξάρτητες τυχαίες** μεταβλητές $R_{j:n}^{r}$ όπου $r>0$ ακολουθούν κατανομή Pareto με $P\left(1,\frac{(n-j+1)\alpha}{r}\right)$ και σύμφωνα με τον τύπο (2.3.1) έχουν μέση τιμή

$$E\left[R_{j:n}^{r}\right]=\frac{\frac{(n-j+1)\alpha}{r}}{\frac{(n-j+1)\alpha}{r}-1}, \quad \alpha>\frac{r}{n-j+1}$$

Το γεγονός ότι η μέση τιμή του γινομένου των ανεξάρτητων τυχαίων μεταβλητών ισούται με το γινόμενο των επιμέρους μέσων τιμών των μεταβλητών αυτών, σε συνδυασμό με το θεώρημα (3.9.3) και τις ιδιότητες της συνάρτησης γάμμα μας επιτρέπει να υπολογίσουμε τις ροπές $\mu'_{k:n}(r)$ τάξης r γύρω από το 0 και τις από κοινού ή παραγοντικές ροπές $\mu'_{k,m:n}(r,s)$ τάξης (r,s) γύρω από το 0 των διατεταγμένων στατιστικών.

Η ροπή $\mu'_{k:n}(r)$ τάξης r γύρω από το 0 του διατεταγμένου στατιστικού $X_{k:n}$ τάξης k είναι

$$\mu'_{k:n}(r) = \theta^r \frac{n!}{(n-k)!} \frac{\Gamma\left(n-k+1-\frac{r}{\alpha}\right)}{\Gamma\left(n+1-\frac{r}{\alpha}\right)}, \quad \alpha > \frac{r}{n-k+1} \tag{3.10.1}$$

όπου Γ είναι η γνωστή συνάρτηση γάμμα όπως δίνεται από τον τύπο (Α.7)

<u>ΑΠΟΔΕΙΞΗ</u>

Έχουμε

$$\begin{aligned}
\mu'_{k:n}(r) &= E\left[X^r_{k:n}\right] \\
&= E\left[\left(\theta\prod_{j=1}^{k} R_{j:n}\right)^r\right] \\
&= \theta^r E\left[\prod_{j=1}^{k} R^r_{j:n}\right] \\
&= \theta^r \prod_{j=1}^{k} E\left[R^r_{j:n}\right] \\
&= \theta^r \prod_{j=1}^{k} \frac{(n-j+1)\alpha r^{-1}}{(n-j+1)\alpha r^{-1}-1} \\
&= \theta^r \left[\frac{n\alpha r^{-1}}{n\alpha r^{-1}-1}\cdot\frac{(n-1)\alpha r^{-1}}{(n-1)\alpha r^{-1}-1}\cdot\ldots\cdot\frac{(n-k+1)\alpha r^{-1}}{(n-k+1)\alpha r^{-1}-1}\right] \\
&= \theta^r n(n-1)\cdot\ldots\cdot(n-k+1)\left(\frac{\alpha r^{-1}}{n\alpha r^{-1}-1}\cdot\frac{\alpha r^{-1}}{(n-1)\alpha r^{-1}-1}\cdot\ldots\cdot\frac{\alpha r^{-1}}{(n-k+1)\alpha r^{-1}-1}\right) \\
&= \theta^r \frac{n!}{(n-k)!}\left(\frac{1}{n-r\alpha^{-1}}\cdot\frac{1}{n-1-r\alpha^{-1}}\cdot\ldots\cdot\frac{1}{n-k+1-r\alpha^{-1}}\right)
\end{aligned}$$

Από τις ιδιότητες της συνάρτησης γάμμα όπως αυτές παρουσιάζονται στον τύπο (Α.7) του παραρτήματος έχουμε

$$\left(n-r\alpha^{-1}\right)\left(n-1-r\alpha^{-1}\right)\cdots\left(n-k+1-r\alpha^{-1}\right)=\frac{\Gamma\left(n+1-r\alpha^{-1}\right)}{\Gamma\left(n-k+1-r\alpha^{-1}\right)}$$

Τελικά η ροπή $\mu'_{k:n}(r)$ τάξης r γύρω από το 0 του διατεταγμένου στατιστικού $X_{k:n}$ τάξης k γράφεται

$$\mu'_{k:n}(r)=\theta^r\frac{n!}{(n-k)!}\left(\frac{1}{n-r\alpha^{-1}}\cdot\frac{1}{n-1-r\alpha^{-1}}\cdot\ldots\cdot\frac{1}{n-k+1-r\alpha^{-1}}\right)$$

$$=\theta^r\frac{n!}{(n-k)!}\frac{1}{\left(n-r\alpha^{-1}\right)\left(n-1-r\alpha^{-1}\right)\cdots\left(n-k+1-r\alpha^{-1}\right)}$$

$$=\theta^r\frac{n!}{(n-k)!}\frac{\Gamma\left(n-k+1-r\alpha^{-1}\right)}{\Gamma\left(n+1-r\alpha^{-1}\right)}$$

$$=\theta^r\frac{n!}{(n-k)!}\frac{\Gamma\left(n-k+1-\frac{r}{\alpha}\right)}{\Gamma\left(n+1-\frac{r}{\alpha}\right)}$$

Με εφαρμογή του τύπου (3.10.1) για $r=1$ έχουμε τη ροπή $\mu'_{k:n}(1)$ πρώτης τάξης γύρω από το 0, δηλαδή τη μέση τιμή του διατεταγμένου στατιστικού $X_{k:n}$ τάξης k που είναι

$$\mu'_{k:n}(1)=\theta\frac{n!}{(n-k)!}\frac{\Gamma\left(n-k+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{1}{\alpha}\right)},\quad \alpha>\frac{1}{n-k+1} \tag{3.10.2}$$

Η διακύμανση $\sigma^2_{k:n}$ του διατεταγμένου στατιστικού $X_{k:n}$ τάξης είναι

$$\sigma^2_{k:n}=\mu'_{r:n}(2)-\left[\mu'_{r:n}(1)\right]^2$$

$$\theta^2\frac{n!}{(n-k)!}\frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)}-\left[\theta\frac{n!}{(n-k)!}\frac{\Gamma\left(n-k+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{1}{\alpha}\right)}\right]^2,\ \alpha>\frac{2}{n-k+1}$$

(3.10.3)

Η από κοινού ή παραγοντική ροπή $\mu'_{k,m:n}(r,s)$ γύρω από το 0 των διατεταγμένων στατιστικών $X_{k:n} < X_{m:n}$ με $\alpha > \max\left(\frac{r+s}{n-k+1}, \frac{s}{n-m+1}\right)$ είναι

$$\mu'_{k,m:n}(r,s) = \theta^{r+s} \frac{n!}{(n-m)!} \frac{\Gamma\left(n-k+1-\frac{r+s}{\alpha}\right)\Gamma\left(n-m+1-\frac{s}{\alpha}\right)}{\Gamma\left(n+1-\frac{r+s}{\alpha}\right)\Gamma\left(n-k+1-\frac{s}{\alpha}\right)} \tag{3.10.4}$$

<u>ΑΠΟΔΕΙΞΗ</u>

Έχουμε

$$\mu'_{k,m:n}(r,s) = E\left[X_{k:n}^r X_{m:n}^s\right]$$

$$= E\left[\left(\theta\prod_{j=1}^{k} R_{j:n}\right)^r \left(\theta\prod_{j=1}^{m} R_{j:n}\right)^s\right]$$

$$= E\left[\theta^r \prod_{j=1}^{k} R_{j:n}^r \cdot \theta^s \prod_{j=1}^{m} R_{j:n}^s\right] =$$

$$= E\left[\theta^{r+s} \prod_{j=1}^{k} R_{j:n}^r \cdot \prod_{j=1}^{m} R_{j:n}^s\right]$$

$$= \theta^{r+s} E\left[\prod_{j=1}^{k} R_{j:n}^r \cdot \prod_{j=1}^{m} R_{j:n}^s\right]$$

$$= \theta^{r+s} E\left[\prod_{j=1}^{k} R_{j:n}^r \cdot \prod_{j=1}^{k} R_{j:n}^s \cdot \prod_{j=k+1}^{m} R_{j:n}^s\right]$$

$$= \theta^{r+s} E\left[\prod_{j=1}^{k} R_{j:n}^{r+s} \cdot \prod_{j=k+1}^{m} R_{j:n}^s\right]$$

$$= \theta^{r+s} \prod_{j=1}^{k} E\left[R_{j:n}^{r+s}\right] \cdot \prod_{j=k+1}^{m} E\left[R_{j:n}^s\right]$$

$$= \theta^{r+s} \prod_{j=1}^{k} \frac{\frac{(n-j+1)\alpha}{r+s}}{\frac{(n-j+1)\alpha}{r+s} - 1} \cdot \prod_{j=k+1}^{m} \frac{\frac{(n-j+1)\alpha}{s}}{\frac{(n-j+1)\alpha}{s} - 1}$$

Για τα παραπάνω γινόμενα έχουμε

$$\prod_{j=1}^{k}\frac{\frac{(n-j+1)\alpha}{r+s}}{\frac{(n-j+1)\alpha}{r+s}-1}=\frac{\frac{n\alpha}{r+s}}{\frac{n\alpha}{r+s}-1}\,\frac{\frac{(n-1)\alpha}{r+s}}{\frac{(n-1)\alpha}{r+s}-1}\cdots\frac{\frac{(n-k+1)\alpha}{r+s}}{\frac{(n-k+1)\alpha}{r+s}-1}$$

$$=n(n-1)\cdots(n-k+1)\left[\frac{\frac{\alpha}{r+s}}{\frac{n\alpha}{r+s}-1}\,\frac{\frac{\alpha}{r+s}}{\frac{(n-1)\alpha}{r+s}-1}\cdots\frac{\frac{\alpha}{r+s}}{\frac{(n-k+1)\alpha}{r+s}-1}\right]$$

$$=\frac{n!}{(n-k)!}\left(\frac{1}{n-\frac{r+s}{\alpha}}\,\frac{1}{n-1\frac{r+s}{\alpha}}\cdots\frac{1}{n-k+1-\frac{r+s}{\alpha}}\right)$$

$$=\frac{n!}{(n-k)!}\,\frac{\Gamma\left(n-k+1-\frac{r+s}{\alpha}\right)}{\Gamma\left(n+1-\frac{r+s}{\alpha}\right)}$$

και

$$\prod_{j=k+1}^{m}\frac{\frac{(n-j+1)\alpha}{s}}{\frac{(n-j+1)\alpha}{s}-1}=\frac{\frac{(n-k)\alpha}{s}}{\frac{(n-k)\alpha}{s}-1}\,\frac{\frac{(n-k-1)\alpha}{s}}{\frac{(n-k-1)\alpha}{s}-1}\cdots\frac{\frac{(n-m+1)\alpha}{s}}{\frac{(n-m+1)\alpha}{s}-1}$$

$$=(n-k)(n-k-1)\cdots(n-m+1)\left[\frac{\frac{\alpha}{s}}{\frac{(n-k)\alpha}{s}-1}\,\frac{\frac{\alpha}{s}}{\frac{(n-k-1)\alpha}{s}-1}\cdots\frac{\frac{\alpha}{s}}{\frac{(n-m+1)\alpha}{s}-1}\right]$$

$$=\frac{(n-k)!}{(n-m)!}\left(\frac{1}{n-k-\frac{s}{\alpha}}\,\frac{1}{n-k-1\frac{s}{\alpha}}\cdots\frac{1}{n-m+1-\frac{s}{\alpha}}\right)$$

$$=\frac{(n-k)!}{(n-m)!}\,\frac{\Gamma\left(n-m+1-\frac{s}{\alpha}\right)}{\Gamma\left(n-k+1-\frac{s}{\alpha}\right)}$$

Τελικά **η από** κοινού ή παραγοντική ροπή $\mu'_{k,m:n}(r,s)$ γύρω από το 0 των **διατεταγμένων στατιστικών** $X_{k:n} < X_{m:n}$ γράφεται

$$\mu'_{k,m:n}(r,s) = \theta^{r+s} \frac{n!}{(n-k)!} \frac{\Gamma\left(n-k+1-\frac{r+s}{\alpha}\right)}{\Gamma\left(n+1-\frac{r+s}{\alpha}\right)} \frac{(n-k)!}{(n-m)!} \frac{\Gamma\left(n-m+1-\frac{s}{\alpha}\right)}{\Gamma\left(n-k+1-\frac{s}{\alpha}\right)} a$$

$$= \theta^{r+s} \frac{n!}{(n-m)!} \frac{\Gamma\left(n-k+1-\frac{r+s}{\alpha}\right)\Gamma\left(n-m+1-\frac{s}{\alpha}\right)}{\Gamma\left(n+1-\frac{r+s}{\alpha}\right)\Gamma\left(n-k+1-\frac{s}{\alpha}\right)}$$

Η συνδιακύμανση $\sigma_{k,m:n}$ **τη** διατεταγμένων στατιστικών $X_{k:n} < X_{m:n}$ με $1 \le k < m \le n$ και $\alpha > \max\left(\frac{2}{n-k+1}, \frac{1}{n-m+1}\right)$ δίνεται από τον τύπο

$$\sigma_{k,mn} = \frac{\theta^2 n!}{(n-m)!} \frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)\Gamma\left(n-m+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)\Gamma\left(n-k+1-\frac{1}{\alpha}\right)} - \frac{\theta^2 (n!)^2}{(n-k)!(n-m)!} \frac{\Gamma\left(n-k+1-\frac{1}{\alpha}\right)\Gamma\left(n-m+1-\frac{1}{\alpha}\right)}{\left[\Gamma\left(n+1-\frac{1}{\alpha}\right)\right]^2} \quad (3.10.5)$$

<u>ΑΠΟΔΕΙΞΗ</u>

Από τους τύπους (3.7.2), (3.10.2) και (3.10.4) έχουμε

$$\sigma_{k,mn} = \mu'_{k,mn}(1,1) - \mu'_{kn}(1)\mu'_{mn}(1)$$

$$= \frac{\theta^2 n!}{(n-m)!} \frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)\Gamma\left(n-m+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)\Gamma\left(n-k+1-\frac{1}{\alpha}\right)} - \frac{\theta n!}{(n-k)!} \frac{\Gamma\left(n-k+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{1}{\alpha}\right)} \frac{\theta n!}{(n-m)!} \frac{\Gamma\left(n-m+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{1}{\alpha}\right)}$$

$$= \frac{\theta^2 n!}{(n-m)!} \frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)\Gamma\left(n-m+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)\Gamma\left(n-k+1-\frac{1}{\alpha}\right)} - \frac{\theta^2 (n!)^2}{(n-k)!(n-m)!} \frac{\Gamma\left(n-k+1-\frac{1}{\alpha}\right)\Gamma\left(n-m+1-\frac{1}{\alpha}\right)}{\left[\Gamma\left(n+1-\frac{1}{\alpha}\right)\right]^2}$$

ΘΕΩΡΗΜΑ 3.10.1

Έστω η τυχαία μεταβλητή X που ακολουθεί **κατανομή Pareto με παραμέτρους** $\theta > 0$ και $\alpha > 1$. Τότε για το διατεταγμένο στατιστικό $X_{1:n}$ **πρώτης τάξης ισχύει**

$$E\left[X_{1:n}^{n}\right] = E[X], \quad \alpha > 1 \tag{3.10.6}$$

ΑΠΟΔΕΙΞΗ

Έχουμε

$$E\left[X_{1:n}^{n}\right] = \theta \frac{n!}{(n-1)!} \frac{\Gamma(n - n\alpha^{-1})}{\Gamma(n+1-n\alpha^{-1})}$$

$$= \theta \frac{n!}{(n-1)!} \frac{\Gamma(n - n\alpha^{-1})}{(n - n\alpha^{-1})\Gamma(n - n\alpha^{-1})}$$

$$= \theta \frac{n!}{(n-1)!} \frac{1}{n(1-\alpha^{-1})}$$

$$= \theta \frac{1}{1 - \frac{1}{\alpha}}$$

$$= \frac{\alpha\theta}{\alpha - 1}$$

$$= E[X]$$

3.11 Σχέσεις μεταξύ ροπών διατεταγμένων στατιστικών κατανομής Pareto.

Έστω η τυχαία μεταβλητή X που ακολουθεί κατανομή Pareto **με** $X \sim P(\theta, \alpha)$. Ο **Malik πρότεινε τις παρακάτω αναδρομικές σχέσεις που ισχύουν μεταξύ των ροπών των διατεταγμένων στατιστικών της τυχαίας μεταβλητής** X.

ΣΧΕΣΗ 3.11.1

$$\mu'_{k:n}(r) = \frac{n-k+1}{n-k+1-\frac{r}{\alpha}} \mu'_{k-1:n}(r), \quad \alpha > \frac{r}{n-r+2} \tag{3.11.1}$$

ΑΠΟΔΕΙΞΗ

Από τις ιδιότητες της συνάρτησης γάμμα όπως δίνονται από τον τύπο (Α.7) του **παραρτήματος έχουμε**

$$\Gamma\left(n-k+2-\frac{r}{\alpha}\right)=\left(n-k+1-\frac{r}{\alpha}\right)\Gamma\left(n-k+1-\frac{r}{\alpha}\right)$$

$$\Leftrightarrow \Gamma\left(n-k+1-\frac{r}{\alpha}\right)=\frac{\Gamma\left(n-k+2-\frac{r}{\alpha}\right)}{\left(n-k+1-\frac{r}{\alpha}\right)}$$

οπότε από τον **τύπο (3.10.1) έχουμε**

$$\mu'_{k:n}(r)=\theta^r\frac{n!}{(n-k)!}\frac{\Gamma\left(n-k+1-\frac{r}{\alpha}\right)}{\Gamma\left(n+1-\frac{r}{\alpha}\right)}$$

$$=\theta^r\frac{n!(n-k+1)}{(n-k+1)!}\frac{\Gamma\left(n-k+2-\frac{r}{\alpha}\right)}{\left(n-k+1-\frac{r}{\alpha}\right)\Gamma\left(n+1-\frac{r}{\alpha}\right)}$$

$$=\frac{(n-k+1)}{\left(n-k+1-\frac{r}{\alpha}\right)}\theta^r\frac{n!}{(n-k+1)!}\frac{\Gamma\left(n-k+2-\frac{r}{\alpha}\right)}{\Gamma\left(n+1-\frac{r}{\alpha}\right)}$$

$$=\frac{(n-k+1)}{\left(n-k+1-\frac{r}{\alpha}\right)}\mu'_{k-1:n}(r)$$

ΣΧΕΣΗ 3.11.2

$$\mu'_{k,m:n}(1,1)=\frac{n-m+1}{\left(n-m+1-\frac{1}{\alpha}\right)}\mu'_{k,m-1:n}(1,1),\quad \alpha>\max\left(\frac{2}{n-k+1},\frac{1}{n-m+1}\right)$$

(3.11.2)

ΑΠΟΔΕΙΞΗ

Από τις ιδιότητες της συνάρτησης γάμμα όπως δίνονται από τον τύπο (Α.7) του παραρτήματος έχουμε

$$\Gamma\left(n-m+2-\frac{r}{\alpha}\right)=\left(n-m+1-\frac{r}{\alpha}\right)\Gamma\left(n-m+1-\frac{r}{\alpha}\right)$$

$$\Leftrightarrow\quad \Gamma\left(n-m+1-\frac{r}{\alpha}\right)=\frac{\Gamma\left(n-m+2-\frac{r}{\alpha}\right)}{\left(n-m+1-\frac{r}{\alpha}\right)}$$

οπότε από τον τύπο (3.10.4) έχουμε

$$\mu'_{k,m:n}(1,1)=\theta^2\frac{n!}{(n-m)!}\frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)\Gamma\left(n-m+1-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)\Gamma\left(n-k+1-\frac{1}{\alpha}\right)}$$

$$=\theta^2\frac{n!}{\frac{(n-m+1)!}{n-m+1}}\frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)\frac{\Gamma\left(n-m+2-\frac{1}{\alpha}\right)}{n-m+1-\frac{1}{\alpha}}}{\Gamma\left(n+1-\frac{2}{\alpha}\right)\Gamma\left(n-k+1-\frac{1}{\alpha}\right)}$$

$$=\theta^2\frac{n!(n-m+1)}{(n-m+1)!}\frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)\Gamma\left(n-m+2-\frac{1}{\alpha}\right)}{\left(n-m+1-\frac{1}{\alpha}\right)\Gamma\left(n+1-\frac{2}{\alpha}\right)\Gamma\left(n-k+1-\frac{1}{\alpha}\right)}$$

$$=\frac{n-m+1}{n-m+1-\frac{1}{\alpha}}\theta^2\frac{n!}{(n-m+1)!}\frac{\Gamma\left(n-k+1-\frac{2}{\alpha}\right)\Gamma\left(n-m+2-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)\Gamma\left(n-k+1-\frac{1}{\alpha}\right)}$$

$$=\frac{n-m+1}{n-m+1-\frac{1}{\alpha}}\mu'_{k,m-1:n}(1,1)$$

Η σχέση που ακολουθεί είναι μια γενικευμένη μορφή της (3.11.2)

<u>ΣΧΕΣΗ 3.11.3</u>

$$\mu'_{k,m:n}(1,1)=\prod_{w=1}^{v}\frac{n-k+w}{n-m+w-\dfrac{1}{\alpha}}\mu'_{k,m-v:n}(1,1),\quad \alpha>\max\left(\frac{2}{n-k+1},\frac{1}{n-m+1}\right)$$

(3.11.3)

<u>ΑΠΟΔΕΙΞΗ</u>

Εφαρμόζοντας διαδοχικά τον τύπο (3.11.2) έχουμε

$$\mu'_{k,m:n}(1,1)=\frac{n-m+1}{n-m+1-\dfrac{1}{\alpha}}\mu'_{k,m-1:n}(1,1)$$

$$=\frac{n-m+1}{n-m+1-\dfrac{1}{\alpha}}\cdot\frac{n-m+2}{n-m+2-\dfrac{1}{\alpha}}\mu'_{k,m-2:n}(1,1)$$

$$=\frac{(n-m+1)(n-m+2)\cdots(n-m+v)}{\left(n-m+1-\dfrac{1}{\alpha}\right)\left(n-m+2-\dfrac{1}{\alpha}\right)\cdots\left(n-m+v-\dfrac{1}{\alpha}\right)}\mu'_{k,m-v:n}(1,1)$$

$$=\prod_{w=1}^{v}\frac{n-m+w}{n-m+w-\dfrac{1}{\alpha}}\mu'_{k,m-v:n}(1,1)$$

ΚΕΦΑΛΑΙΟ 4^{O}

Χαρακτηρισμοί της κατανομής Pareto.

Στο κεφάλαιο αυτό θα αναφερθούμε σε χαρακτηρισμούς (characterizations) της κατανομής Pareto. Με τον όρο αυτό εννοούμε τις ικανές και αναγκαίες συνθήκες που όταν ισχύουν για την κατανομή μιας τυχαίας μεταβλητής, προκύπτει ότι αυτή είναι κατανομή Pareto. Στη βιβλιογραφία υπάρχουν πάρα πολλοί χαρακτηρισμοί για την κατανομή Pareto οι οποίοι χρονολογούνται από το 1925, αλλά σημαντική έρευνα έγινε από το 1960 και μετά. Δεν υπάρχει μια συστηματική καταγραφή τους, και δεν είναι δυνατόν να συμπεριληφθούν όλοι στα πλαίσια αυτής της εργασίας. Θα αναφερθούμε σε ορισμένους μόνο χαρακτηρισμούς με βάση τις υπό συνθήκη αναμενόμενες τιμές, τις πιθανότητες των υπό συνθήκη κατανομών, τις συναρτήσεις των διατεταγμένων στατιστικών και τις ροπές των διατεταγμένων στατιστικών.

4.1 Χαρακτηρισμοί με βάση τις υπό συνθήκη αναμενόμενες τιμές.

Ο Shanbhag πρότεινε ένα χαρακτηρισμό για την εκθετική κατανομή με βάση τις αναμενόμενες τιμές των υπό συνθήκη κατανομών. Στο ίδιο πνεύμα μπορούμε να διατυπώσουμε το παρακάτω θεώρημα για την κατανομή Pareto.

ΘΕΩΡΗΜΑ 4.1.1

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [\theta, +\infty)$, $\theta > 0$ και $E[X] < +\infty$. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(\theta, \alpha)$ είναι

$$E[X \mid X > y] = \frac{y}{\theta} E[X], \quad y \geq \theta, \ \alpha > 0 \tag{4.1.1}$$

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta, \alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha\theta^{\alpha}x^{-\alpha-1} \quad \text{και} \quad F(x) = 1 - \theta^{\alpha}x^{-\alpha}, \quad x \geq \theta > 0, \ \alpha > 0$$

Το πρώτο μέλος του τύπου (4.1.1) γράφεται

$$\begin{aligned}
E[X \mid X > y] &= \int_{y}^{+\infty} x d\left[\frac{F(x) - F(y)}{1 - F(y)}\right] \\
&= \frac{1}{1 - F(y)} \int_{y}^{+\infty} x dF(x) \\
&= \frac{1}{\theta^{\alpha}y^{-\alpha}} \int_{y}^{+\infty} x d\left(1 - \theta^{\alpha}x^{-\alpha}\right) \\
&= \theta^{-\alpha}y^{\alpha} \int_{y}^{+\infty} x\alpha\theta^{\alpha}x^{-\alpha-1} dx \\
&= \alpha y^{\alpha}\left[\frac{x^{-\alpha+1}}{-\alpha+1}\right]_{y}^{+\infty} \\
&= \frac{\alpha y}{\alpha - 1}
\end{aligned}$$

Το δεύτερο μέλος του τύπου (4.1.1) **λόγω του τύπου** (2.3.1) γράφεται

$$\frac{y}{\theta}E[X]=\frac{y}{\theta}\frac{\alpha\theta}{\alpha-1}=\frac{\alpha y}{\alpha-1}$$

Επομένως ισχύει ο τύπος (4.1.1)

$$E[X\,|\,X>y]=\frac{y}{\theta}E[X]$$

Αντιστρόφως

Έστω ότι για την τυχαία μεταβλητή X ισχύει ο τύπος (4.1.1). Τότε

$$E[X\,|\,X>y]=\frac{y}{\theta}E[X] \Leftrightarrow \int_y^{+\infty} x d\left[\frac{F(x)-F(y)}{1-F(y)}\right]=\frac{y}{\theta}E[X]$$

$$\Leftrightarrow \frac{1}{1-F(y)}\int_y^{+\infty} x dF(x)=\frac{y}{\theta}E[X]$$

$$\Leftrightarrow \int_y^{+\infty} x dF(x)=\frac{E[X]}{\theta}y[1-F(y)]$$

$$\Leftrightarrow \int_y^{+\infty} x dF(x)=\frac{E[X]}{\theta}y\bar{F}(y)$$

Παραγωγίζουμε **και τα δύο μέλη της παραπάνω ισότητας και έχουμε**

$$\frac{d}{dy}\int_y^{+\infty} x dF(x)=\frac{d}{dy}\left[\frac{E[X]}{\theta}y\bar{F}(y)\right] \Leftrightarrow -y\frac{d}{dy}F(y)=\frac{E[X]}{\theta}\left[\bar{F}(y)+y\frac{d}{dy}\bar{F}(y)\right]$$

$$\Leftrightarrow y\frac{d}{dy}[1-F(y)]=\frac{E[X]}{\theta}\bar{F}(y)+\frac{E[X]}{\theta}y\frac{d}{dy}\bar{F}(y)$$

$$\Leftrightarrow y\frac{d}{dy}\bar{F}(y)=\frac{E[X]}{\theta}\bar{F}(y)+\frac{E[X]}{\theta}y\frac{d}{dy}\bar{F}(y)$$

$$\Leftrightarrow \left[1-\frac{E[X]}{\theta}\right]y\frac{d}{dy}\bar{F}(y)=\frac{E[X]}{\theta}\bar{F}(y)$$

$$\Leftrightarrow [\theta-E[X]]y\frac{d}{dy}\bar{F}(y)=E[X]\bar{F}(y)$$

$$\Leftrightarrow \frac{1}{\bar{F}(y)}\frac{d}{dy}\bar{F}(y)=\frac{E[X]}{[\theta-E[X]]}\frac{1}{y}$$

Λύνουμε τη διαφορική εξίσωση που προέκυψε

$$\frac{1}{\overline{F}(y)}\frac{d}{dy}\overline{F}(y)=\frac{E[X]}{[\theta-E[X]]}\frac{1}{y} \Leftrightarrow \frac{d}{dy}\left[\log\overline{F}(y)\right]=\frac{d}{dy}\left[\frac{E[X]}{[\theta-E[X]]}\log y\right]$$

$$\Leftrightarrow \log\overline{F}(y)=\log y^{\frac{E[X]}{[\theta-E[X]]}}+\log c$$

$$\Leftrightarrow \log\overline{F}(y)=\log c y^{\frac{E[X]}{[\theta-E[X]]}}$$

$$\Leftrightarrow \overline{F}(y)=c y^{-\frac{E[X]}{[E[X]-\theta]}}$$

$$\Leftrightarrow F(y)=1-c y^{-\frac{E[X]}{[E[X]-\theta]}}$$

Θέτουμε

$$\frac{E[X]}{[E[X]-\theta]}=\alpha$$

οπότε

$$F(y)=1-cy^{-\alpha}$$

Ισχύει

$$\lim_{y\to+\infty}F(y)=1 \Leftrightarrow \lim_{y\to+\infty}\left(1-cy^{-\alpha}\right)=1$$

$$\Leftrightarrow \lim_{y\to+\infty}\left(y^{-\alpha}\right)=0$$

από όπου συμπεραίνουμε ότι $\alpha>0$.

Όμως

$$F(\theta)=0 \Leftrightarrow 1-c\theta^{-\alpha}=0$$

$$\Leftrightarrow c\theta^{-\alpha}=1$$

$$\Leftrightarrow c=\theta^{\alpha}$$

Τελικά είναι

$$F(x)=1-\theta^{\alpha}x^{-\alpha},\quad x\ge\theta>0,\ \alpha>0$$

που σημαίνει ότι **η τυχαία μεταβλητή Χ ακολουθεί κατανομή Pareto με** $X\sim P(\theta,\alpha)$.

Μια διαφορετική εκδοχή του παραπάνω χαρακτηρισμού δίνεται από **το παρακάτω θεώρημα**.

ΘΕΩΡΗΜΑ 4.1.2

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [\theta, +\infty)$, $\theta > 0$ και $E[X] < +\infty$. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(\theta, \alpha)$ είναι

$$E[X \mid X > y] = \frac{\alpha y}{\alpha - 1}, \quad y \geq \theta, \ \alpha > 0 \tag{4.1.2}$$

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta, \alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha \theta^{\alpha} x^{-\alpha-1} \quad \text{και} \quad F(x) = 1 - \theta^{\alpha} x^{-\alpha}, \quad x \geq \theta > 0, \ \alpha > 0$$

Το πρώτο μέλος του τύπου (4.1.2) γράφεται

$$E[X \mid X > y] = \int_y^{+\infty} x d\left[\frac{F(x) - F(y)}{1 - F(y)}\right]$$

$$= \frac{1}{1 - F(y)} \int_y^{+\infty} x dF(x)$$

$$= \frac{1}{\theta^{\alpha} y^{-\alpha}} \int_y^{+\infty} x d\left(1 - \theta^{\alpha} x^{-\alpha}\right)$$

$$= \theta^{-\alpha} y^{\alpha} \int_y^{+\infty} x \alpha \theta^{\alpha} x^{-\alpha-1} dx$$

$$= \alpha y^{\alpha} \int_y^{+\infty} x^{-\alpha} dx$$

$$= \alpha y^{\alpha} \left[\frac{x^{-\alpha+1}}{-\alpha + 1}\right]_y^{+\infty}$$

$$= \alpha y^{\alpha} \frac{y^{-\alpha+1}}{\alpha - 1}$$

$$= \frac{\alpha y}{\alpha - 1}$$

Αντιστρόφως

Έστω ότι για την τυχαία μεταβλητή X ισχύει ο τύπος (4.1.2). Τότε

$$E[X \mid X>y]=\frac{\alpha y}{\alpha-1} \Leftrightarrow \int_{y}^{+\infty} x d\left[\frac{F(x)-F(y)}{1-F(y)}\right]=\frac{\alpha y}{\alpha-1}$$

$$\Leftrightarrow \frac{1}{1-F(y)} \int_{y}^{+\infty} x dF(x)=\frac{\alpha y}{\alpha-1}$$

$$\Leftrightarrow \frac{1}{\bar{F}(y)} \int_{y}^{+\infty} x dF(x)=\frac{\alpha y}{\alpha-1}$$

$$\Leftrightarrow \int_{y}^{+\infty} x dF(x)=\frac{\alpha}{\alpha-1} y \bar{F}(y)$$

Παραγωγίζουμε και τα δύο μέλη της παραπάνω ισότητας και έχουμε

$$\frac{d}{dy} \int_{y}^{+\infty} x F(x) dx=\frac{\alpha}{\alpha-1} \frac{d}{dy}[y \bar{F}(y)] \Leftrightarrow -y \frac{d}{dy} F(y)=\frac{\alpha}{\alpha-1}\left[\bar{F}(y)+y \frac{d}{dy} \bar{F}(y)\right]$$

$$\Leftrightarrow y \frac{d}{dy}[1-F(y)]=\frac{\alpha}{\alpha-1} \bar{F}(y)+\frac{\alpha}{\alpha-1} y \frac{d}{dy} \bar{F}(y)$$

$$\Leftrightarrow y \frac{d}{dy} \bar{F}(y)=\frac{\alpha}{\alpha-1} \bar{F}(y)+\frac{\alpha}{\alpha-1} y \frac{d}{dy} \bar{F}(y)$$

$$\Leftrightarrow\left(\frac{\alpha}{\alpha-1}-1\right) y \frac{d}{dy} \bar{F}(y)=-\frac{\alpha}{\alpha-1} \bar{F}(y)$$

$$\Leftrightarrow \frac{\alpha-\alpha+1}{\alpha-1} y \frac{d}{dy} \bar{F}(y)=-\frac{\alpha}{\alpha-1} \bar{F}(y)$$

$$\Leftrightarrow \frac{1}{\alpha-1} y \frac{d}{dy} \bar{F}(y)=-\frac{\alpha}{\alpha-1} \bar{F}(y)$$

$$\Leftrightarrow y \frac{d}{dy} \bar{F}(y)=-\alpha \bar{F}(y)$$

$$\Leftrightarrow y d\bar{F}(y)=-\alpha \bar{F}(y) dy$$

$$\Leftrightarrow \frac{1}{\bar{F}(y)} d\bar{F}(y)=-\frac{\alpha}{y} dy$$

Λύνουμε τη διαφορική εξίσωση που προέκυψε

$$\frac{1}{\overline{F}(y)} d\overline{F}(y) = -\frac{\alpha}{y} dy \quad \Leftrightarrow \quad \frac{d}{dx}\left[\log \overline{F}(x)\right] = \frac{d}{dx}\left(-\alpha \log y\right)$$

$$\Leftrightarrow \quad \log \overline{F}(x) = -\alpha \log y + \log c$$

$$\Leftrightarrow \quad \log \overline{F}(x) = \log c y^{-\alpha}$$

$$\Leftrightarrow \quad \overline{F}(x) = c y^{-\alpha}$$

$$\Leftrightarrow \quad F(x) = 1 - c y^{-\alpha}$$

Όμως

$$F(\theta) = 0 \quad \Leftrightarrow \quad 1 - c\theta^{-\alpha} = 0$$

$$\Leftrightarrow \quad c\theta^{-\alpha} = 1$$

$$\Leftrightarrow \quad c = \theta^{\alpha}$$

Τελικά είναι

$$F(x) = 1 - \theta^{\alpha} x^{-\alpha}, \quad x \ge \theta > 0, \ \alpha > 0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$.

Ένας άλλος χαρακτηρισμός με βάση την αναμενόμενη τιμή της υπό συνθήκη κατανομής του λογαρίθμου της τυχαίας μεταβλητής δίνεται από το παρακάτω θεώρημα.

ΘΕΩΡΗΜΑ 4.1.3

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [\theta, +\infty)$, $\theta > 0$ και $E[\log X] < +\infty$. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(\theta,\alpha)$ είναι

$$E\left[\log X \mid X > y\right] = \log y + \frac{1}{\alpha}, \quad y \ge 1, \ \alpha > 0 \qquad (4.1.3)$$

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x)=\alpha\theta^{\alpha}x^{-\alpha-1} \quad \text{και} \quad F(x)=1-\theta^{\alpha}x^{-\alpha}, \quad x\ge\theta>0,\ \alpha>0$$

Το πρώτο μέλος του τύπου (4.1.3) γράφεται

$$E[\log X \mid X>y]=\int_y^{+\infty}\log x\, d\left[\frac{F(x)-F(y)}{1-F(y)}\right]$$

$$=\frac{1}{1-F(y)}\int_y^{+\infty}\log x\, dF(x)$$

$$=\frac{1}{\theta^{\alpha}y^{-\alpha}}\int_y^{+\infty}\log x\, d\left(1-\theta^{\alpha}x^{-\alpha}\right)$$

$$=\theta^{-\alpha}y^{\alpha}\left[-\int_y^{+\infty}\theta^{\alpha}\log x\, d\left(x^{-\alpha}\right)\right]$$

$$=y^{\alpha}\left[-\left[x^{-\alpha}\log x\right]_y^{+\infty}+\int_y^{+\infty}x^{-\alpha}d\log x\right]$$

$$=y^{\alpha}\left(y^{-\alpha}\log y+\int_y^{+\infty}x^{-\alpha}\frac{1}{x}dx\right) \quad \left(\lim_{x\to+\infty}\frac{\log x}{x^{\alpha}}=0,\ \alpha>0\right)$$

$$=\log y+y^{\alpha}\int_y^{+\infty}x^{-\alpha-1}dx=\log y-\frac{1}{\alpha}y^{\alpha}\int_y^{+\infty}dx^{-\alpha}$$

$$=\log y-\frac{1}{\alpha}y^{\alpha}\left[x^{-\alpha}\right]_y^{+\infty}=\log y+\frac{1}{\alpha}y^{\alpha}y^{-\alpha}$$

$$=\log y+\frac{1}{\alpha}$$

<u>Αντιστρόφως</u>

Έστω ότι για την τυχαία μεταβλητή X ισχύει ο τύπος (4.1.3). Τότε

$$E[\log X \mid X>y]=\log y+\frac{1}{\alpha} \iff \int_y^{+\infty}\log x\, d\left[\frac{F(x)-F(y)}{1-F(y)}\right]=\log y+\frac{1}{\alpha}$$

$$\iff \frac{1}{1-F(y)}\int_y^{+\infty}\log x\, dF(x)=\log y+\frac{1}{\alpha}$$

$$\iff \int_y^{+\infty}\log x\, dF(x)=\left(\log y+\frac{1}{\alpha}\right)\overline{F}(y)$$

Παραγωγίζουμε και τα δύο μέλη της παραπάνω ισότητας και έχουμε

$$\frac{d}{dy}\int_y^{+\infty}\log x dF(x) = \frac{d}{dy}\left[\left(\log y + \frac{1}{\alpha}\right)\overline{F}(y)\right]$$

$$\Leftrightarrow \quad -\log y \frac{d}{dy}F(y) = \frac{1}{y}\overline{F}(y) + \left(\log y + \frac{1}{\alpha}\right)\frac{d}{dy}\overline{F}(y)$$

$$\Leftrightarrow \quad \log y \frac{d}{dy}\overline{F}(y) = \frac{1}{y}\overline{F}(y) + \log y \frac{d}{dy}\overline{F}(y) + \frac{1}{\alpha}\frac{d}{dy}\overline{F}(y)$$

$$\Leftrightarrow \quad \frac{1}{\alpha}\frac{d}{dy}\overline{F}(y) = -\frac{1}{y}\overline{F}(y)$$

$$\Leftrightarrow \quad \frac{1}{\overline{F}(y)}\frac{d}{dy}\overline{F}(y) = -\frac{\alpha}{y}$$

Λύνουμε τη διαφορική εξίσωση που προέκυψε

$$\frac{1}{\overline{F}(y)}\frac{d}{dy}\overline{F}(y) = -\frac{\alpha}{y} \quad \Leftrightarrow \quad \frac{d}{dy}\left[\log\overline{F}(y)\right] = \frac{d}{dy}(-\alpha\log y)$$

$$\Leftrightarrow \quad \log\overline{F}(y) = \log y^{-\alpha} + \log c$$

$$\Leftrightarrow \quad \log\overline{F}(y) = \log c y^{-\alpha}$$

$$\Leftrightarrow \quad \overline{F}(y) = c y^{-\alpha}$$

$$\Leftrightarrow \quad F(y) = 1 - c y^{-\alpha}$$

Όμως

$$F(\theta) = 0 \quad \Leftrightarrow \quad 1 - c\theta^{-\alpha} = 0$$

$$\Leftrightarrow \quad c\theta^{-\alpha} = 1$$

$$\Leftrightarrow \quad c = \theta^{\alpha}$$

Τελικά είναι

$$F(x) = 1 - \theta^{\alpha}x^{-\alpha}, \quad x \geq \theta > 0, \ \alpha > 0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$.

Το 1975 ο Dallas πρότεινε ένα χαρακτηρισμό με βάση την ροπή r τάξης της δεσμευμένης κατανομής. Ακολουθεί το σχετικό θεώρημα.

ΘΕΩΡΗΜΑ 4.1.4

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [\theta, +\infty)$, $\theta > 0$ και $E\left[X^r\right] < +\infty$. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(\theta, \alpha)$ είναι

$$E\left[X^r \mid X > y\right] = E\left[\left(\frac{y}{\theta}X\right)^r\right], \quad y \geq \theta, \ \alpha > r, \ r = 1, 2, \ldots \tag{4.1.4}$$

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta, \alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha\theta^{\alpha}x^{-\alpha-1} \quad \text{και} \quad F(x) = 1 - \theta^{\alpha}x^{-\alpha}, \quad x \geq \theta > 0, \ \alpha > 0$$

Το πρώτο μέλος του τύπου (4.1.4) γράφεται

$$E\left[X^r \mid X > y\right] = \int_y^{+\infty} x^r d\left[\frac{F(x) - F(y)}{1 - F(y)}\right]$$

$$= \frac{1}{1 - F(y)} \int_y^{+\infty} x^r dF(x)$$

$$= \frac{1}{\theta^{\alpha}y^{-\alpha}} \int_y^{+\infty} x^r d\left(1 - \theta^{\alpha}x^{-\alpha}\right)$$

$$= \theta^{-\alpha}y^{\alpha} \int_y^{+\infty} x^r \alpha\theta^{\alpha}x^{-\alpha-1}dx$$

$$= \alpha y^{\alpha} \int_y^{+\infty} x^{r-\alpha-1}dx$$

$$= \alpha y^{\alpha}\left[\frac{x^{r-\alpha}}{r-\alpha}\right]_y^{+\infty}$$

$$= \alpha y^{\alpha}\frac{y^{r-\alpha}}{\alpha - r}$$

$$= \frac{\alpha y^r}{\alpha - 1}$$

Το δεύτερο μέλος του τύπου (4.1.4) λόγω του τύπου (2.2.1) γράφεται

$$E\left[\left(\frac{y}{\theta}X\right)^r\right]=E\left[\frac{y^r}{\theta^r}X^r\right]=\frac{y^r}{\theta^r}E\left[X^r\right]$$

$$=\frac{y^r}{\theta^r}\frac{\alpha\theta^r}{\alpha-r}=\frac{\alpha y^r}{\alpha-r}$$

Επομένως ισχύει ο τύπος (4.1.4)

$$E\left[X^r \mid X>y\right]=E\left[\left(\frac{y}{\theta}X\right)^r\right]$$

<u>Αντιστρόφως</u>

Έστω ότι για την τυχαία μεταβλητή X ισχύει ο τύπος (4.1.4). Τότε

$$E\left[X^r \mid X>y\right]=E\left[\left(\frac{y}{\theta}X\right)^r\right] \Leftrightarrow \int_y^{+\infty} x^r d\left[\frac{F(x)-F(y)}{1-F(y)}\right]=E\left[\frac{y^r}{\theta^r}X^r\right]$$

$$\Leftrightarrow \frac{1}{1-F(y)}\int_y^{+\infty} x^r dF(x)=\frac{y^r}{\theta^r}E\left[X^r\right]$$

$$\Leftrightarrow \int_y^{+\infty} x^r dF(x)=\frac{E\left[X^r\right]}{\theta^r}y^r\overline{F}(y)$$

Παραγωγίζουμε και τα δύο μέλη της παραπάνω ισότητας και έχουμε

$$\frac{d}{dy}\int_y^{+\infty} x^r dF(x)=\frac{d}{dy}\left[\frac{E\left[X^r\right]}{\theta^r}y^r\overline{F}(y)\right]$$

$$\Leftrightarrow -y^r\frac{d}{dy}F(y)=\frac{E\left[X^r\right]}{\theta^r}\left[ry^{r-1}\overline{F}(y)+y^r\frac{d}{dy}\overline{F}(y)\right]$$

$$\Leftrightarrow y^r\frac{d}{dy}\overline{F}(y)=r\frac{E\left[X^r\right]}{\theta^r}y^{r-1}\overline{F}(y)+\frac{E\left[X^r\right]}{\theta^r}y^r\frac{d}{dy}\overline{F}(y)$$

$$\Leftrightarrow \left[1-\frac{E\left[X^r\right]}{\theta^r}\right]y^r\frac{d}{dy}\overline{F}(y)=r\frac{E\left[X^r\right]}{\theta^r}y^{r-1}\overline{F}(y)$$

$$\Leftrightarrow \left[\theta^r-E\left[X^r\right]\right]y^r\frac{d}{dy}\overline{F}(y)=rE\left[X^r\right]y^{r-1}\overline{F}(y)$$

$$\Leftrightarrow \frac{1}{\overline{F}(y)}\frac{d}{dy}\overline{F}(y)=\frac{E\left[X^r\right]}{\left[\theta^r-E\left[X^r\right]\right]}\frac{1}{y}$$

Λύνουμε τη διαφορική εξίσωση που **προέκυψε**

$$\frac{1}{\overline{F}(y)}\frac{d}{dy}\overline{F}(y)=\frac{E\left[X^r\right]}{\left[\theta^r-E\left[X^r\right]\right]}\frac{1}{y} \Leftrightarrow \frac{d}{dy}\left[\log\overline{F}(y)\right]=\frac{d}{dy}\left[\frac{E\left[X^r\right]}{\left[\theta^r-E\left[X^r\right]\right]}\log y\right]$$

$$\Leftrightarrow \text{og}\,\overline{F}(y)=\frac{E\left[X^r\right]}{\left[\theta^r-E\left[X^r\right]\right]}\log y+\log c$$

$$\Leftrightarrow \log\overline{F}(y)=\log y^{\frac{E\left[X^r\right]}{\left[\theta^r-E\left[X^r\right]\right]}}+\log c$$

$$\Leftrightarrow \log\overline{F}(y)=\log cy^{\frac{E\left[X^r\right]}{\left[\theta^r-E\left[X^r\right]\right]}}$$

$$\Leftrightarrow \overline{F}(y)=cy^{\frac{-E\left[X^r\right]}{\left[E\left[X^r\right]-\theta^r\right]}}$$

$$\Leftrightarrow F(y)=1-cy^{\frac{-E\left[X^r\right]}{\left[E\left[X^r\right]-\theta^r\right]}}$$

Θέτουμε

$$\frac{E\left[X^r\right]}{\left[E\left[X^r\right]-\theta^r\right]}=\alpha$$

οπότε

$$F(y)=1-cy^{-\alpha}$$

Ισχύει

$$\lim_{y\to+\infty}F(y)=1 \Leftrightarrow \lim_{y\to+\infty}\left(1-cy^{-\alpha}\right)=1$$

$$\Leftrightarrow \lim_{y\to+\infty}\left(y^{-\alpha}\right)=0$$

από όπου συμπεραίνουμε ότι $\alpha>0$.

Όμως

$$F(\theta)=0 \Leftrightarrow 1-c\theta^{-\alpha}=0$$

$$\Leftrightarrow c\theta^{-\alpha}=1$$

$$\Leftrightarrow c=\theta^{\alpha}$$

Τελικά είναι

$$F(x)=1-\theta^{\alpha}x^{-\alpha},\quad x\geq\theta>0,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή Χ ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$.

Το 1967 ο Ferguson πρότεινε τον ακόλουθο χαρακτηρισμό με βάση την αναμενόμενη τιμή της δεσμευμένης κατανομής του λόγου των διατεταγμένων στατιστικών.

ΘΕΩΡΗΜΑ 4.1.5

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [1,+\infty)$ και $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n της κατανομής αυτής. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(1,\alpha)$ είναι

$$E\left[\frac{X_{m+1:n}}{X_{m:n}} \mid X_{m:n} = x\right] = c, \quad c > 0, \ 1 \le m \le n-1 \tag{4.1.5}$$

ΑΠΟΔΕΙΞΗ

Ευθέως

Η από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{m:n} < X_{m+1:n}$ από τον τύπο (3.5.1) είναι

$$f_{m,m+1}(x_m, x_{m+1}) = \frac{n!}{(m-1)!(n-m-1)!} F^{m-1}(x_m)\left[1 - F(x_{m+1})\right]^{n-m-1} f(x_m) f(x_{m+1})$$

Θεωρούμε το μετασχηματισμό

$$U = X_{m:n} \text{ και } V = \frac{X_{m+1:n}}{X_{m:n}} \tag{4.1.6}$$

οπότε

$$X_{m:n} = U \text{ και } X_{m+1:n} = UV$$

και

$$x_m = u \text{ και } x_{m+1} = uv, \quad u \ge 1, \ v \ge 1$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \dfrac{d(x_m, x_{m+1})}{d(u,v)}$

$$J = \begin{vmatrix} \dfrac{dx_m}{du} & \dfrac{dx_m}{dv} \\ \dfrac{dx_{m+1}}{du} & \dfrac{dx_{m+1}}{dv} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

Η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V είναι

$$f_{uv}(u,v) = f_{m,m+1}(x_m, x_{m+1})|J|$$

$$= \frac{n!}{(m-1)!(n-m-1)!} F^{m-1}(u)[1-F(uv)]^{n-m-1} f(u)f(uv)u \qquad (4.1.7)$$

Η συνάρτηση πιθανότητας $f_u(u)$ της τυχαίας μεταβλητής $U = X_{m:n}$ από τον τύπο (3.3.1) είναι

$$f_u(u) = m\binom{n}{m} F^{m-1}(u)[1-F(u)]^{n-m} f(u)$$

$$= m\frac{n!}{m!(n-m)!} F^{m-1}(u)[1-F(u)]^{n-m} f(u)$$

$$= \frac{n!}{(m-1)!(n-m)!} F^{m-1}(u)[1-F(u)]^{n-m} f(u) \qquad (4.1.8)$$

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(1,\alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha x^{-\alpha-1} \quad \text{και} \quad F(x) = 1 - x^{-\alpha}, \quad x \geq 1,\ \alpha > 0$$

Στην περίπτωση αυτή η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V γράφεται

$$f_{uv}(u,v) = \frac{n!}{(m-1)!(n-m-1)!}\left(1-u^{-\alpha}\right)^{m-1}\left[(uv)^{-\alpha}\right]^{n-m-1} \alpha u^{-\alpha-1}\alpha(uv)^{-\alpha-1}u$$

$$= \frac{n!}{(m-1)!(n-m-1)!}\alpha^2\left(1-u^{-\alpha}\right)^{m-1} u^{-\alpha(n-m-1)}v^{-\alpha(n-m-1)}u^{-2\alpha-1}v^{-\alpha-1}$$

$$= \frac{n!}{(m-1)!(n-m-1)!}\alpha^2\left(1-u^{-\alpha}\right)^{m-1} u^{-\alpha(n-m+1)-1}v^{-\alpha(n-m)-1}$$

και η συνάρτηση πιθανότητας $f_u(u)$ της τυχαίας μεταβλητής U γράφεται

$$f_u(u) = \frac{n!}{(m-1)!(n-m)!}\left(1-u^{-\alpha}\right)^{m-1}\left(u^{-\alpha}\right)^{n-m}\alpha u^{-\alpha-1}$$

$$= \frac{n!}{(m-1)!(n-m)!}\alpha\left(1-u^{-\alpha}\right)^{m-1} u^{-\alpha(n-m)}u^{-\alpha-1}$$

$$= \frac{n!}{(m-1)!(n-m)!}\alpha\left(1-u^{-\alpha}\right)^{m-1} u^{-\alpha(n-m+1)-1}$$

Το πρώτο μέλος του τύπου (4.1.5) γράφεται

$$E\left[\frac{X_{m+1:n}}{X_{m:n}} \mid X_{m:n}=x\right]=E[V \mid U=u]$$

$$=\int_1^{+\infty} v f_{v|u}(v \mid u)du$$

$$=\int_1^{+\infty} v \frac{f_{uv}(uv)}{f_u(u)}dv$$

Για το παραπάνω ολοκλήρωμα έχουμε

$$\int_1^{+\infty} v \frac{f_{uv}(uv)}{f_u(u)}dv=\int_1^{+\infty} v \frac{\frac{n!}{(m-1)!(n-m-1)!}\alpha^2\left(1-u^{-\alpha}\right)^{m-1}u^{-\alpha(n-m+1)-1}v^{-\alpha(n-m)-1}}{\frac{n!}{(m-1)!(n-m)!}\alpha\left(1-u^{-\alpha}\right)^{m-1}u^{-\alpha(n-m+1)-1}}dv$$

$$=\alpha\frac{(n-m)!}{(n-m-1)!}\int_1^{+\infty} v^{-\alpha(n-m)}dv$$

$$=\alpha(n-m)\left[\frac{v^{-\alpha(n-m)+1}}{-\alpha(n-m)+1}\right]_1^{+\infty}$$

$$=\frac{\alpha(n-m)}{\alpha(n-m)+1}$$

Τελικά αν θέσουμε

$$c=\frac{\alpha(n-m)}{\alpha(n-m)+1}$$

είναι $c>0$ και

$$E\left[\frac{X_{m+1:n}}{X_{m:n}} \mid X_{m:n}=x\right]=c$$

<u>Αντιστρόφως</u>

Έστω ότι ισχύει η σχέση (4.1.5). Τότε λόγω του τύπου (4.1.6) έχουμε

$$E\left[\frac{X_{m+1:n}}{X_{m:n}} \mid X_{m:n}=x\right]=c \iff E[V \mid U=u]=c$$

$$\iff \int_1^{+\infty} v f_{v|u}(v \mid u)du=c$$

$$\iff \int_1^{+\infty} v \frac{f_{uv}(uv)}{f_u(u)}dv=c$$

Από τους τύπους (4.1.7) και (4.1.8) το **παραπάνω ολοκλήρωμα γράφεται**

$$\int_1^{+\infty} v\frac{f_{uv}(uv)}{f_u(u)}dv=\int_1^{+\infty} v\frac{\frac{n!}{(m-1)!(n-m-1)!}F^{m-1}(u)\left[1-F(uv)\right]^{n-m-1}f(u)f(uv)u}{\frac{n!}{(m-1)!(n-m)!}F^{m-1}(u)\left[1-F(u)\right]^{n-m}f(u)}dv$$

$$=\frac{(n-m)!}{(n-m-1)!}\int_1^{+\infty}\frac{\left[1-F(uv)\right]^{n-m-1}}{\left[1-F(u)\right]^{n-m-1}}\frac{f(uv)u}{1-F(u)}vdv$$

$$=(n-m)\int_1^{+\infty}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m-1}\frac{f(uv)u}{1-F(u)}vdv$$

Άρα

$$E\left[\frac{X_{m+1:n}}{X_{m:n}}\mid X_{m:n}=x\right]=c\quad\Leftrightarrow\quad(n-m)\int_1^{+\infty}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m-1}\frac{f(uv)u}{1-F(u)}vdv=c$$

Από την παραπάνω ισότητα συμπεραίνουμε **ότι η συνάρτηση**

$$\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m-1}\frac{f(uv)u}{1-F(u)}$$

δεν εξαρτάται από τη μεταβλητή u, δηλαδή είναι μόνο συνάρτηση του v και έστω **ότι είναι H(v). Τότε έχουμε**

$$\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m-1}\frac{f(uv)u}{1-F(u)}=H(v)\quad\Leftrightarrow\quad-\frac{1}{n-m}\frac{d}{dv}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m}=H(v)$$

$$\Leftrightarrow\quad\frac{d}{dv}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m}=-(n-m)H(v)$$

$$\Leftrightarrow\quad\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m}=-\int(n-m)H(v)dv+c_1$$

$$\Leftrightarrow\quad\frac{1-F(uv)}{1-F(u)}=\left[-\int(n-m)H(v)dv+c_1\right]^{\frac{1}{n-m}}$$

που σημαίνει ότι η **συνάρτηση** $\frac{1-F(uv)}{1-F(u)}$ **είναι ανεξ**άρτητη του u δηλαδή είναι συνάρτηση μόνο του v και έστω ότι είναι G(v). Τότε έχουμε

$$\frac{1-F(uv)}{1-F(u)} = G(v) \quad \Leftrightarrow \quad \frac{\overline{F}(uv)}{\overline{F}(u)} = G(v)$$

$$\Leftrightarrow \quad \overline{F}(uv) = \overline{F}(u)G(v)$$

Με χρήση του τύπου (Α.16) όπως δίνεται στο παράρτημα **λύνουμε τη** παραπάνω **συναρτησιακή εξίσωση** και έχουμε

$$\overline{F}(x) = kx^{c_2} \text{ και } G(x) = x^{c_2}, \quad x \geq 1 \text{ αφού } u \geq 1 \text{ και } v \geq 1$$

άρα

$$F(x) = 1 - kx^{c_2}, \quad x \geq 1$$

Ισχύει

$$\lim_{x\to+\infty} F(x) = 1 \Leftrightarrow \lim_{x\to+\infty}\left(1 - kx^{c_2}\right) = 1$$

$$\Leftrightarrow \lim_{x\to+\infty}\left(x^{c_2}\right) = 0$$

από όπου συμπεραίνουμε ότι $c_2 < 0$ και έστω ότι είναι $c_2 = -\alpha$ με $\alpha > 0$. Τότε

$$F(x) = 1 - kx^{-\alpha}$$

και

$$F(1) = 0 \quad \Leftrightarrow \quad 1 - k = 0$$

$$\Leftrightarrow \quad k = 1$$

Τελικά έχουμε

$$F(x) = 1 - x^{-\alpha}, \quad x \geq 1, \ \alpha > 0$$

που σημαίνει **ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με** $X \sim P(1,\alpha)$.

4.2 Χαρακτηρισμοί με βάση τις πιθανότητες των υπό συνθήκη κατανομών.

Οι χαρακτηρισμοί της εκθετικής κατανομής που βασίζονται στις πιθανότητες των υπό συνθήκη κατανομών είναι από τους πρώτους που προτάθηκαν. Αν τους μετασχηματίσουμε κατάλληλα τότε προκύπτουν χαρακτηρισμοί για την κατανομή Pareto. Σχετικά είναι τα θεωρήματα (4.2.1) και (4.2.2) που ακολουθούν.

ΘΕΩΡΗΜΑ 4.2.1

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [\theta,+\infty)$, $\theta > 0$. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(\theta,\alpha)$ είναι

$$P(X > y \mid X > z) = P\left(\frac{z}{\theta} X > y\right), \quad y \geq z > 0,\ \theta > 0 \qquad (4.2.1)$$

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha\theta^{\alpha}x^{-\alpha-1} \quad \text{και} \quad F(x) = 1-\theta^{\alpha}x^{-\alpha}, \quad x \geq \theta > 0,\ \alpha > 0$$

Το πρώτο μέλος του τύπου (4.2.1) γράφεται

$$\begin{aligned} P(X > y \mid X > z) &= \frac{P(X > y)}{P(X > z)} \\ &= \frac{\overline{F}(y)}{\overline{F}(z)} \\ &= \frac{\theta^{\alpha}y^{-\alpha}}{\theta^{\alpha}z^{-\alpha}} \\ &= z^{\alpha}y^{-\alpha} \end{aligned}$$

Το δεύτερο μέλος του τύπου (4.2.1) γράφεται

$$\begin{aligned} P\left(\frac{z}{\theta}X > y\right) &= P\left(X > \frac{\theta y}{z}\right) \\ &= \overline{F}\left(\frac{\theta y}{z}\right) \\ &= \theta^{\alpha}\left(\frac{\theta y}{z}\right)^{-\alpha} \\ &= \theta^{\alpha}\frac{\theta^{-\alpha}y^{-\alpha}}{z^{-\alpha}} \\ &= z^{\alpha}y^{-\alpha} \end{aligned}$$

Επομένως ισχύει ο τύπος (4.2.1)

$$P(X > y \mid X > z) = P\left(\frac{z}{\theta}X > y\right), \quad y \geq z > 0,\ \theta > 0$$

<u>Αντιστρόφως</u>

Έστω ότι για την **τυχαία μεταβλητή** Χ ισχύει **ο τύπος** (4.2.1). **Τότε**

$$P(X > y \mid X > z) = P\left(\frac{z}{\theta}X > y\right) \Leftrightarrow \frac{P(X > y)}{P(X > z)} = P\left(X > \frac{\theta y}{z}\right)$$

$$\Leftrightarrow \frac{\overline{F}(y)}{\overline{F}(z)} = \overline{F}\left(\frac{\theta y}{z}\right)$$

$$\Leftrightarrow \overline{F}(y) = \overline{F}(z)\overline{F}\left(\frac{\theta y}{z}\right)$$

Θέτουμε

$$y = zt$$

Ισχύει

$$t \geq 1 \quad \text{αφού} \quad y \geq z > 0$$

και έχουμε

$$\overline{F}(zt) = \overline{F}(z)\overline{F}(\theta t)$$

Με χρήση του τύπου (Α.15) **όπως δίνεται** στο παράρτημα λύνουμε τη παραπάνω **συναρτησιακή εξίσωση και έχουμε**

$$\overline{F}(t) = kt^{c} \quad \text{και} \quad \overline{F}(\theta t) = t^{c}, \quad t \geq 1$$

Θέτουμε

$$x = \theta t$$

Ισχύει

$$x \geq \theta \quad \text{αφού} \quad t \geq 1$$

και έχουμε

$$\overline{F}(x) = \left(\frac{x}{\theta}\right)^{c} \Leftrightarrow \overline{F}(x) = \left(\frac{x}{\theta}\right)^{c}$$

$$\Leftrightarrow 1 - F(x) = \theta^{-c}x^{c}$$

$$\Leftrightarrow F(x) = 1 - \theta^{-c}x^{c}$$

Επίσης

$$\lim_{x\to+\infty} F(x)=1 \Leftrightarrow \lim_{x\to+\infty}\left(1-\theta^{c}x^{-c}\right)=1$$

$$\Leftrightarrow \lim_{x\to+\infty}\left(x^{c}\right)=0$$

από όπου συμπεραίνουμε ότι $c<0$ και έστω ότι είναι $c=-\alpha$ με $\alpha>0$.

Τελικά έχουμε

$$F(x)=1-\theta^{\alpha}x^{-\alpha}, \quad x\geq\theta>0,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$.

ΘΕΩΡΗΜΑ 4.2.2

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x\in[\theta,+\infty)$, $\theta>0$. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X\sim P(\theta,\alpha)$ είναι

$$P\left(\frac{X}{\theta}>yz \mid \frac{X}{\theta}>y\right)=P\left(\frac{X}{\theta}>z\right), \quad y\geq1,\ z\geq1,\ \theta>0 \tag{4.2.2}$$

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x)=\alpha\theta^{\alpha}x^{-\alpha-1} \quad \text{και} \quad F(x)=1-\theta^{\alpha}x^{-\alpha}, \quad x\geq\theta>0,\ \alpha>0$$

Το πρώτο μέλος του τύπου (4.2.2) γράφεται

$$\begin{aligned} P\left(\frac{X}{\theta}>yz \mid \frac{X}{\theta}>y\right) &= P\left(X>\theta yz \mid X>\theta y\right) \\ &= \frac{P(X>\theta yz)}{P(X>\theta y)} \\ &= \frac{\overline{F}(\theta yz)}{\overline{F}(\theta y)} \\ &= \frac{\theta^{\alpha}(\theta yz)^{-\alpha}}{\theta^{\alpha}(\theta y)^{-\alpha}} \\ &= z^{-\alpha} \end{aligned}$$

Το δεύτερο μέλος του τύπου (4.2.2) γράφεται

$$P\left(\frac{X}{\theta} > z\right) = P(X > \theta z)$$

$$= \bar{F}(\theta z)$$

$$= \theta^{\alpha}(\theta z)^{-\alpha}$$

$$= \theta^{\alpha}\theta^{-\alpha}z^{-\alpha}$$

$$= z^{-\alpha}$$

Επομένως ισχύει ο τύπος (4.2.2)

$$P\left(\frac{X}{\theta} > yz \mid \frac{X}{\theta} > y\right) = P\left(\frac{X}{\theta} > z\right), \quad y \geq 1,\ z \geq 1,\ \theta > 0$$

<u>Αντιστρόφως</u>

Έστω ότι για την τυχαία μεταβλητή X ισχύει ο τύπος (4.2.2). Τότε

$$P\left(\frac{X}{\theta} > yz \mid \frac{X}{\theta} > y\right) = P\left(\frac{X}{\theta} > z\right) \Leftrightarrow P(X > \theta yz \mid X > \theta y) = P(X > \theta z)$$

$$\Leftrightarrow \frac{P(X > \theta yz)}{P(X > \theta y)} = P(X > \theta z)$$

$$\Leftrightarrow \frac{\bar{F}(\theta yz)}{\bar{F}(\theta y)} = \bar{F}(\theta y)$$

$$\Leftrightarrow \bar{F}(\theta yz) = \bar{F}(\theta z)\bar{F}(\theta y)$$

Με χρήση του τύπου (Α.15) όπως δίνεται στο παράρτημα λύνουμε τη παραπάνω συναρτησιακή εξίσωση και έχουμε

$$\bar{F}(\theta t) = t^{c}$$

Ισχύει

$t \geq 1$ αφού $y \geq 1$ και $z \geq 1$

Θέτουμε

$$x = \theta t$$

και έχουμε

$x \geq \theta$ αφού $t \geq 1$

Τότε

$$\overline{F}(x)=\left(\frac{x}{\theta}\right)^{c} \Leftrightarrow \overline{F}(x)=\left(\frac{x}{\theta}\right)^{c}$$

$$\Leftrightarrow 1-F(x)=\theta^{-c}x^{c}$$

$$\Leftrightarrow F(x)=1-\theta^{-c}x^{c}$$

Επίσης

$$\lim_{x\to+\infty} F(x)=1 \Leftrightarrow \lim_{x\to+\infty}\left(1-\theta^{c}x^{-c}\right)=1$$

$$\Leftrightarrow \lim_{x\to+\infty}\left(x^{c}\right)=0$$

από όπου συμπεραίνουμε ότι $c<0$ και έστω ότι είναι $c=-\alpha$ με $\alpha>0$.

Τελικά έχουμε

$$F(x)=1-\theta^{\alpha}x^{-\alpha}, \quad x\ge\theta>0,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$.

O Rossberg πρότεινε δυο ενδιαφέροντες χαρακτηρισμούς της κατανομής Pareto που βασίζονται στις πιθανότητες των υπό συνθήκη κατανομών του λόγου των διατεταγμένων στατιστικών. Σχετικά είναι τα θεωρήματα (4.2.3) και (4.2.4) που παρουσιάζουμε στη συνέχεια.

ΘΕΩΡΗΜΑ 4.2.3

Έστω X μια τυχαία μεταβλητή η οποία έχει απολύτως συνεχή συνάρτηση κατανομής F(x) με $x\in[\theta,+\infty)$, $\theta>0$ και $X_{1:n}<X_{2:n}<\cdots<X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n του πληθυσμού. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X\sim P(\theta,\alpha)$ είναι

$$P\left(\frac{X_{k+1:n}}{X_{k:n}}\ge x\right)=x^{-\alpha(n-k)}, \quad x\ge 1 \qquad (4.2.3)$$

για κάποιο αριθμό k με $1\le k\le n-1$.

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$. Τότε **σύμφωνα με το θεώρημα** (3.9.2) **η τυχαία μεταβλητή**

$$R_{k+1} = \frac{X_{k+1:n}}{X_{k:n}}, \quad r_{k+1} \geq 1$$

ακολουθεί κατανομή Pareto με $R_{k+1} \sim P\left(1,(n-k)\alpha\right)$. Άρα έχει συνάρτηση κατανομής

$$F_{r_{k+1}}(x) = 1 - x^{-\alpha(n-k)}, \quad x \geq 1$$

Το πρώτο μέλος του τύπου (4.2.3) γράφεται

$$P\left(\frac{X_{k+1:n}}{X_{k:n}} \geq x\right) = P\left(R_{k+1} \geq x\right)$$

$$= 1 - P\left(R_{k+1} \leq x\right)$$

$$= 1 - F_{r_{k+1}}(x)$$

$$= 1 - \left(1 - x^{-\alpha(n-k)}\right)$$

$$= x^{-\alpha(n-k)}, \quad x \geq 1$$

Αντιστρόφως

Η από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{k:n} < X_{k+1:n}$ **από τον τύπο** (3.5.1) **είναι**

$$f_{k,k+1}(x_k, x_{k+1}) = \frac{n!}{(k-1)!(n-k-1)!} F^{k-1}(x_k)\left[1 - F(x_{k+1})\right]^{n-k-1} f(x_k) f(x_{k+1})$$

Θεωρούμε το μετασχηματισμό

$$U = X_{k:n} \text{ και } V = \frac{X_{k+1:n}}{X_{k:n}}$$

οπότε

$$X_{k:n} = U \text{ και } X_{k+1:n} = UV$$

και

$$x_k = u \text{ και } x_{k+1} = uv, \quad u \geq \theta, \quad v \geq 1$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \dfrac{d(x_k, x_{k+1})}{d(u,v)}$

$$J = \begin{vmatrix} \dfrac{dx_k}{du} & \dfrac{dx_k}{dv} \\ \dfrac{dx_{k+1}}{du} & \dfrac{dx_{k+1}}{dv} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

Η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V είναι

$$f_{uv}(u,v) = f_{k,k+1}(x_k, x_{k+1})|J|$$

$$= \frac{n!}{(m-1)!(n-m-1)!} F^{k-1}(u)\left[1-F(uv)\right]^{n-k-1} f(u)f(uv)u$$

Για τη συνάρτηση πιθανότητας $f_v(v)$ της τυχαίας μεταβλητής $V = \dfrac{X_{k+1:n}}{X_{k:n}}$ έχουμε

$$f_v(v) = \int_\theta^{+\infty} f_{uv}(u,v)du$$

$$= \frac{n!}{(m-1)!(n-m-1)!} \int_\theta^{+\infty} F^{k-1}(u)\left[1-F(uv)\right]^{n-k-1} f(u)f(uv)udu$$

$$\frac{n!}{(m-1)!(n-m-1)!} \int_\theta^{+\infty} F^{k-1}(u) \frac{d}{dv}\left[-\frac{\left[1-F(uv)\right]^{n-k}}{n-k}\right] f(u)du$$

$$= -\frac{n!}{(m-1)!(n-m-1)!(n-k)} \int_\theta^{+\infty} F^{k-1}(u) \frac{d}{dv}\left[1-F(uv)\right]^{n-k} f(u)du$$

Έστω ότι ισχύει ο τύπος (4.2.3). Τότε έχουμε

$$P\left(\frac{X_{k+1:n}}{X_{k:n}} \geq x\right) = x^{-\alpha(n-k)} \Leftrightarrow P(V \geq x) = x^{-\alpha(n-k)}$$

$$\Leftrightarrow 1 - P(V \leq x) = x^{-\alpha(n-k)}$$

$$\Leftrightarrow 1 - F_v(x) = x^{-\alpha(n-k)}$$

$$\Leftrightarrow -F_v(x) = -1 + x^{-\alpha(n-k)}$$

$$\Leftrightarrow F_v(x) = 1 - x^{-\alpha(n-k)}$$

Παραγωγίζουμε και τα δύο μέλη της παραπάνω ισότητας και έχουμε

$$\frac{d}{dx}\left[F_v(x)\right]=\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]$$

$$\Leftrightarrow \quad f_v(x)=\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]$$

$$\Leftrightarrow \quad \frac{f_v(x)}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}=1$$

$$\Leftrightarrow \quad \frac{-\frac{n!}{(m-1)!(n-m-1)!(n-k)}\int_{\theta}^{+\infty}F^{k-1}(u)\frac{d}{dx}\left[1-F(ux)\right]^{n-k}f(u)du}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}=1$$

$$\Leftrightarrow \quad \frac{\int_{\theta}^{+\infty}F^{k-1}(u)\frac{d}{dx}\left[1-F(ux)\right]^{n-k}f(u)du}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}=-\frac{(m-1)!(n-m-1)!(n-k)}{n!}$$

$$\Leftrightarrow \quad \int_{\theta}^{+\infty}\frac{\frac{d}{dx}\left[1-F(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}F^{k-1}(u)f(u)du=-\frac{(m-1)!(n-m-1)!(n-k)}{n!}$$

$$\Leftrightarrow \quad \frac{1}{k}\int_{\theta}^{+\infty}\frac{\frac{d}{dx}\left[\overline{F}(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}dF^{k}(u)=-\frac{(m-1)!(n-m-1)!(n-k)}{n!}$$

$$\Leftrightarrow \quad \int_{\theta}^{+\infty}\frac{\frac{d}{dx}\left[\overline{F}(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}dF^{k}(u)=-\frac{(m-1)!(n-m-1)!(n-k)k}{n!}$$

Στην παραπάνω ισότητα το δεύτερο μέλος είναι ανεξάρτητο από τη μεταβλητή x, συνεπώς θα είναι και το πρώτο μέλος ανεξάρτητο από τη μεταβλητή x. Αυτό σημαίνει ότι η συνάρτηση

$$\frac{\frac{d}{dx}\left[\overline{F}(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}$$

είναι συνάρτηση μόνο του u και έστω ότι είναι H(u). Άρα

$$\frac{\frac{d}{dx}\left[\overline{F}(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]}=H(u) \quad\Leftrightarrow\quad \frac{d}{dx}\left[\overline{F}(ux)\right]^{n-k}=H(u)\frac{d}{dx}\left[1-x^{-\alpha(n-k)}\right]$$

$$\Leftrightarrow\quad \left[\overline{F}(ux)\right]^{n-k}=H(u)\left[1-x^{-\alpha(n-k)}\right]+c$$

Ισχύει

$$\lim_{x\to+\infty}\overline{F}(ux)=0 \quad\Leftrightarrow\quad \lim_{x\to+\infty}\left[H(u)\left[1-x^{-\alpha(n-k)}\right]+c\right]=0$$

$$\Leftrightarrow\quad H(u)\left[1-\lim_{x\to+\infty}x^{-\alpha(n-k)}\right]+c=0$$

Η **παραπάνω ισότητα ισχύει μόνο όταν**

$$\lim_{x\to+\infty}x^{-\alpha(n-k)}=0$$

από όπου συμπεραίνουμε ότι

$$-\alpha(n-k)<0 \quad\Leftrightarrow\quad \alpha>0$$

και τότε είναι

$$H(u)+c=0 \quad\Leftrightarrow\quad c=-H(u)$$

Άρα

$$\left[\overline{F}(ux)\right]^{n-k}=H(u)\left[1-x^{-\alpha(n-k)}\right]+c \quad\Leftrightarrow\quad \left[\overline{F}(ux)\right]^{n-k}=H(u)\left[1-x^{-\alpha(n-k)}\right]-H(u)$$

$$\Leftrightarrow\quad \left[\overline{F}(ux)\right]^{n-k}=H(u)\left[1-x^{-\alpha(n-k)}-1\right]$$

$$\Leftrightarrow\quad \left[\overline{F}(ux)\right]^{n-k}=-H(u)x^{-\alpha(n-k)}$$

$$\Leftrightarrow\quad \left[\overline{F}(ux)\right]^{n-k}=H(u)\left[-x^{-\alpha(n-k)}\right]$$

$$\Leftrightarrow\quad \overline{F}(ux)=\left[-H(u)\right]^{\frac{1}{n-k}}x^{-\alpha}$$

Θέτουμε

$$H_1(u)=\left[-H(u)\right]^{\frac{1}{n-k}}$$

οπότε η **παραπάνω** ισότητα γράφεται

$$\overline{F}(ux)=H_1(u)x^{-\alpha}$$

Με χρήση του τύπου (Α.17) όπως δίνεται στο παράρτημα λύνουμε την παραπάνω συναρτησιακή εξίσωση και έχουμε

$$\overline{F}(t)=pt^{-\alpha} \text{ και } H_1(t)=pt^{-\alpha}, \quad t\geq\theta \text{ αφού } ux\geq\theta$$

Άρα

$$F(x)=1-px^{-\alpha}, \quad x\geq\theta>0,\ \alpha>0$$

Ισχύει

$$F(\theta)=0 \Leftrightarrow 1-p\theta^{-\alpha}=0$$

$$\Leftrightarrow p\theta^{-\alpha}=1$$

$$\Leftrightarrow p=\theta^{\alpha}$$

Τελικά έχουμε

$$F(x)=1-\theta^{\alpha}x^{-\alpha}, \quad x\geq\theta>0,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$.

ΘΕΩΡΗΜΑ 4.2.4

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής $F(x)$ με $x\in[1,+\infty)$. Αν $X_{1:n}<X_{2:n}<X_{3:n}<\cdots<X_{n:n}$ είναι τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n του πληθυσμού, τότε μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X\sim P(1,\alpha)$ είναι

$$P\left(\frac{X_{k+1:n}}{X_{k:n}}\geq x\right)=[1-F(x)]^{n-k}, \quad x\geq 1 \qquad (4.2.4)$$

για κάποιον αριθμό k με $1\leq k\leq n-1$.

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(1,\alpha)$. Τότε είναι

$$F(x)=1-x^{-\alpha} \text{ και } 1-F(x)=x^{-\alpha}, \quad x\geq 1,\ \alpha>0$$

Σύμφωνα με το θεώρημα (3.9.2) η τυχαία μεταβλητή

$$R_{k+1} = \frac{X_{k+1:n}}{X_{k:n}}, \quad r_{k+1} \geq 1$$

ακολουθεί κατανομή Pareto με $R_{k+1} \sim P(1,(n-k)\alpha)$. Άρα έχει συνάρτηση **κατανομής**

$$F_{r_{k+1}}(x) = 1 - x^{-\alpha(n-k)}, \quad x \geq 1$$

Για την πιθανότητα $P\left(\frac{X_{k+1:n}}{X_{k:n}} \geq x\right)$ έχουμε

$$P\left(\frac{X_{k+1:n}}{X_{k:n}} \geq x\right) = P\left(R_{k+1} \geq x\right)$$

$$= 1 - F_{r_{k+1}}(x)$$

$$= 1 - \left(1 - x^{-\alpha(n-k)}\right) = x^{-\alpha(n-k)} = \left(x^{-\alpha}\right)^{n-k}$$

$$= \left[1 - F(x)\right]^{n-k}, \quad x \geq 1$$

<u>Αντιστρόφως</u>

Η από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{k:n} < X_{k+1:n}$ από τον τύπο (3.5.1) είναι

$$f_{k,k+1}(x_k, x_{k+1}) = \frac{n!}{(k-1)!(n-k-1)!} F^{k-1}(x_k)\left[1 - F(x_{k+1})\right]^{n-k-1} f(x_k) f(x_{k+1})$$

Θεωρούμε το μετασχηματισμό

$$U = X_{k:n}$$

και

$$V = \frac{X_{k+1:n}}{X_{k:n}}$$

οπότε

$$X_{k:n} = U$$

και

$$X_{k+1:n} = UV$$

Τότε είναι

$$x_k = u$$

και

$$x_{k+1} = uv\,, \quad u \geq 1, \quad v \geq 1$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \dfrac{d(x_k, x_{k+1})}{d(u,v)}$

$$J = \begin{vmatrix} \dfrac{dx_k}{du} & \dfrac{dx_k}{dv} \\ \dfrac{dx_{k+1}}{du} & \dfrac{dx_{k+1}}{dv} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

Η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V είναι

$$f_{uv}(u,v) = f_{k,k+1}(x_k, x_{k+1})|J|$$

$$= \frac{n!}{(m-1)!(n-m-1)!} F^{k-1}(u)[1-F(uv)]^{n-k-1} f(u)f(uv)u$$

Για τη συνάρτηση πιθανότητας $f_v(v)$ της τυχαίας μεταβλητής $V = \dfrac{X_{k+1:n}}{X_{k:n}}$ έχουμε

$$f_v(v) = \int_\theta^{+\infty} f_{uv}(u,v)du$$

$$= \frac{n!}{(m-1)!(n-m-1)!} \int_\theta^{+\infty} F^{k-1}(u)[1-F(uv)]^{n-k-1} f(u)f(uv)udu$$

$$\frac{n!}{(m-1)!(n-m-1)!} \int_\theta^{+\infty} F^{k-1}(u) \frac{d}{dv}\left[-\frac{[1-F(uv)]^{n-k}}{n-k}\right] f(u)du$$

$$= -\frac{n!}{(m-1)!(n-m-1)!(n-k)} \int_\theta^{+\infty} F^{k-1}(u) \frac{d}{dv}[1-F(uv)]^{n-k} f(u)du$$

Έστω ότι ισχύει ο τύπος (4.2.4). Τότε έχουμε

$$P\left(\frac{X_{k+1:n}}{X_{k:n}} \geq x\right) = [1-F(x)]^{n-k} \quad \Leftrightarrow \quad P(V \geq x) = [1-F(x)]^{n-k}$$

$$\Leftrightarrow \quad 1 - P(V \leq x) = [1-F(x)]^{n-k}$$

$$\Leftrightarrow \quad 1 - F_v(x) = [1-F(x)]^{n-k}$$

Παραγωγίζουμε και τα δύο μέλη της παραπάνω ισότητας και έχουμε

$$\frac{d}{dx}\left[1-F_v(x)\right]=\frac{d}{dx}\left[1-F(x)\right]^{n-k}$$

$$\Leftrightarrow\quad -f_v(x)=\frac{d}{dx}\left[1-F(x)\right]^{n-k}$$

$$\Leftrightarrow\quad \frac{f_v(x)}{\frac{d}{dx}\left[1-F(x)\right]^{n-k}}=-1$$

$$\Leftrightarrow\quad \frac{-\frac{n!}{(m-1)!(n-m-1)!(n-k)}\int_{\theta}^{+\infty}F^{k-1}(u)\frac{d}{dx}\left[1-F(ux)\right]^{n-k}f(u)du}{\frac{d}{dx}\left[1-F(x)\right]^{n-k}}=-1$$

$$\Leftrightarrow\quad \frac{\int_{\theta}^{+\infty}F^{k-1}(u)\frac{d}{dx}\left[1-F(ux)\right]^{n-k}f(u)du}{\frac{d}{dx}\left[1-F(x)\right]^{n-k}}=\frac{(m-1)!(n-m-1)!(n-k)}{n!}$$

$$\Leftrightarrow\quad \int_{\theta}^{+\infty}\frac{\frac{d}{dx}\left[1-F(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-F(x)\right]^{n-k}}F^{k-1}(u)f(u)du=\frac{(m-1)!(n-m-1)!(n-k)}{n!}$$

$$\Leftrightarrow\quad \frac{1}{k}\int_{\theta}^{+\infty}\frac{\frac{d}{dx}\left[\bar{F}(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-F(x)\right]^{n-k}}dF^{k}(u)=\frac{(m-1)!(n-m-1)!(n-k)}{n!}$$

$$\Leftrightarrow\quad \int_{\theta}^{+\infty}\frac{\frac{d}{dx}\left[\bar{F}(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-F(x)\right]^{n-k}}dF^{k}(u)=\frac{(m-1)!(n-m-1)!(n-k)k}{n!}$$

Στην παραπάνω ισότητα το δεύτερο μέλος είναι ανεξάρτητο από τη μεταβλητή x, συνεπώς θα είναι και το πρώτο μέλος ανεξάρτητο από τη μεταβλητή x. Αυτό σημαίνει ότι η συνάρτηση

$$\frac{\frac{d}{dx}\left[\bar{F}(ux)\right]^{n-k}}{\frac{d}{dx}\left[1-F(x)\right]^{n-k}}$$

είναι συνάρτηση μόνο του u και έστω ότι είναι **H(u)**. Άρα

$$\frac{\frac{d}{dx}\left[\overline{F}(ux)\right]^{n-k}}{\frac{d}{dx}[1-F(x)]^{n-k}} = H(u) \Leftrightarrow \frac{d}{dx}\left[\overline{F}(ux)\right]^{n-k} = H(u)\frac{d}{dx}[1-F(x)]^{n-k}$$

$$\Leftrightarrow \left[\overline{F}(ux)\right]^{n-k} = H(u)[1-F(x)]^{n-k} + c$$

Όμως

$$\lim_{x\to+\infty}\overline{F}(ux)=0 \Leftrightarrow \lim_{x\to+\infty}\left[H(u)[1-F(x)]^{n-k}+c\right]=0$$

$$\Leftrightarrow H(u)\left[1-\lim_{x\to+\infty}F(x)\right]^{n-k}+c=0$$

$$\Leftrightarrow H(u)(1-1)^{n-k}+c=0$$

$$\Leftrightarrow c=0$$

Άρα

$$\left[\overline{F}(ux)\right]^{n-k} = H(u)[1-F(x)]^{n-k} \Leftrightarrow \overline{F}(ux)=H(u)^{\frac{1}{n-k}}[1-F(x)]$$

$$\Leftrightarrow \overline{F}(ux)=H(u)^{\frac{1}{n-k}}\overline{F}(x)$$

$$\Leftrightarrow \overline{F}(ux)=H_1(u)\overline{F}(x)$$

Με χρήση του τύπου (A.16) όπως δίνεται στο παράρτημα λύνουμε την παραπάνω συναρτησιακή εξίσωση και έχουμε

$$\overline{F}(t)=pt^{-\alpha} \text{ και } H_1(t)=pt^{-\alpha}, \quad t\geq 1 \text{ αφού } ux\geq 1$$

Άρα

$$F(x)=1-px^{-\alpha}, \quad x\geq 1,\ \alpha>0$$

Ισχύει

$$F(1)=0 \Leftrightarrow 1-p=0$$

$$\Leftrightarrow p=1$$

Τελικά έχουμε

$$F(x)=1-x^{-\alpha}, \quad x\geq 1,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$.

4.3 Χαρακτηρισμοί με βάση τις συναρτήσεις των διατεταγμένων στατιστικών.

Στην ενότητα αυτή παρουσιάζουμε χαρακτηρισμούς που βασίζονται στην ανεξαρτησία ή την ισονομία συναρτήσεων διατεταγμένων στατιστικών. Το 1970 ο Malik πρότεινε ένα χαρακτηρισμό για την κατανομή Pareto που βασίζεται στην ανεξαρτησία των στατιστικών $X_{m:n}$ και $\dfrac{X_{m+1:n}}{X_{m:n}}$.

ΘΕΩΡΗΜΑ 4.3.1

Έστω X μια τυχαία μεταβλητή η οποία έχει απολύτως συνεχή συνάρτηση κατανομής F(x) με $x \in [\theta, +\infty)$, $\theta > 0$ και $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n του πληθυσμού. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(\theta, \alpha)$ είναι για κάποιον αριθμό m με $1 \le m \le n-1$ τα στατιστικά

$$U = X_{m:n} \text{ και } V = \frac{X_{m+1:n}}{X_{m:n}} \tag{4.3.1}$$

να είναι ανεξάρτητα.

ΑΠΟΔΕΙΞΗ

Ευθέως

Η από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{m:n} < X_{m+1:n}$ από τον τύπο (3.5.1) είναι

$$f_{m,m+1}(x_m, x_{m+1}) = \frac{n!}{(m-1)!(n-m-1)!} F^{m-1}(x_m)\left[1 - F(x_{m+1})\right]^{n-m-1} f(x_m) f(x_{m+1})$$

Θεωρούμε το μετασχηματισμό

$$U = X_{m:n}$$

και

$$V = \frac{X_{m+1:n}}{X_{m:n}}$$

οπότε

$$X_{m:n} = U \text{ και } X_{m+1:n} = UV$$

Τότε είναι

$$x_m = u$$

και

$$x_{m+1} = uv, \quad u \geq \theta, \;\; v \geq 1$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \dfrac{d(x_m, x_{m+1})}{d(u,v)}$

$$J = \begin{vmatrix} \dfrac{dx_m}{du} & \dfrac{dx_m}{dv} \\ \dfrac{dx_{m+1}}{du} & \dfrac{dx_{m+1}}{dv} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

Η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V είναι

$$f_{uv}(u,v) = f_{m,m+1}(x_m, x_{m+1})|J|$$

$$= \frac{n!}{(m-1)!(n-m-1)!} F^{m-1}(u)\left[1 - F(uv)\right]^{n-m-1} f(u)f(uv)u \qquad (4.3.2)$$

Η συνάρτηση πιθανότητας $f_u(u)$ της τυχαίας μεταβλητής $U = X_{m:n}$ από τον τύπο (3.3.1) είναι

$$f_u(u) = m\binom{n}{m} F^{m-1}(u)\left[1 - F(u)\right]^{n-m} f(u)$$

$$= m\frac{n!}{m!(n-m)!} F^{m-1}(u)\left[1 - F(u)\right]^{n-m} f(u)$$

$$= \frac{n!}{(m-1)!(n-m)!} F^{m-1}(u)\left[1 - F(u)\right]^{n-m} f(u) \qquad (4.3.3)$$

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha\theta^{\alpha}x^{-\alpha-1} \;\; \text{και} \;\; F(x) = 1 - \theta^{\alpha}x^{-\alpha}, \quad x \geq \theta > 0, \; \alpha > 0$$

Στην περίπτωση αυτή η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V γράφεται

$$f_{uv}(u,v)=\frac{n!}{(m-1)!(n-m-1)!}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{m-1}\left[\theta^{\alpha}(uv)^{-\alpha}\right]^{n-m-1}\alpha\theta^{\alpha}u^{-\alpha-1}\alpha\theta^{\alpha}(uv)^{-\alpha-1}u$$

$$=\frac{n!}{(m-1)!(n-m-1)!}\alpha^{2}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{m-1}\theta^{\alpha(n-m-1)}(uv)^{-\alpha(n-m-1)}\theta^{2\alpha}u^{-\alpha}(uv)^{-\alpha-1}$$

$$=\frac{n!}{(m-1)!(n-m-1)!}\alpha^{2}\theta^{\alpha(n-m+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{m-1}u^{-\alpha(n-m-1)}v^{-\alpha(n-m-1)}u^{-2\alpha-1}v^{-\alpha-1}$$

$$=\frac{n!}{(m-1)!(n-m-1)!}\alpha^{2}\theta^{\alpha(n-m+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{m-1}u^{-\alpha(n-m+1)-1}v^{-\alpha(n-m)-1}$$

και η συνάρτηση **πιθανότητας** $f_u(u)$ **της** τυχαίας μεταβλητής U γράφεται

$$f_u(u)=\frac{n!}{(m-1)!(n-m)!}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{m-1}\left(\theta^{\alpha}u^{-\alpha}\right)^{n-m}\alpha\theta^{\alpha}u^{-\alpha-1}$$

$$=\frac{n!}{(m-1)!(n-m)!}\alpha\theta^{\alpha(n-m+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{m-1}u^{-\alpha(n-m)}u^{-\alpha-1}$$

$$=\frac{n!}{(m-1)!(n-m)!}\alpha\theta^{\alpha(n-m+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{m-1}u^{-\alpha(n-m+1)-1}$$

Για την τυχαία μεταβλητή V ισχύει

$$V=\frac{X_{m+1:n}}{X_{m:n}}=R_{m+1}$$

και σύμφωνα με το θεώρημα (3.9.2) **η μεταβλητή** R_{m+1} **ακολουθεί** κατανομή Pareto με $R_{m+1}\sim P\left(1,(n-m)\alpha\right)$. Άρα

$$f_v(v)=(n-m)\alpha v^{-\alpha(n-m)-1}$$

Για να αποδείξουμε ότι τα στατιστικά

$$U=X_{m:n}\ \text{και}\ V=\frac{X_{m+1:n}}{X_{m:n}}$$

είναι ανεξάρτητα αρκεί να αποδείξουμε ότι

$$f_{uv}(uv)=f_u(u)f_v(v)$$

Πράγματι έχουμε

$$f_u(u)f_v(v) = \frac{n!}{(m-1)!(n-m)!}\alpha\theta^{\alpha(n-m+1)}\left(1-\theta^\alpha u^{-\alpha}\right)^{m-1} u^{-\alpha(n-m+1)-1}(n-m)\alpha v^{-\alpha(n-m)-1}$$

$$= \frac{n!(n-m)}{(m-1)!(n-m)!}\alpha^2\theta^{\alpha(n-m+1)}\left(1-\theta^\alpha u^{-\alpha}\right)^{m-1} u^{-\alpha(n-m+1)-1} v^{-\alpha(n-m)-1}$$

$$= \frac{n!}{(m-1)!(n-m-1)!}\alpha^2\theta^{\alpha(n-m+1)}\left(1-\theta^\alpha u^{-\alpha}\right)^{m-1} u^{-\alpha(n-m+1)-1} v^{-\alpha(n-m)-1}$$

$$= f_{uv}(uv)$$

Αντιστρόφως

Έστω ότι οι τυχαίες μεταβλητές U και V είναι ανεξάρτητες. Τότε

$$f_{uv}(u,v) = f_u(u)f_v(v) \quad \Leftrightarrow \quad f_v(v) = \frac{f_{uv}(u,v)}{f_u(u)}$$

Λαμβάνοντας υπόψη τους **τύπους** (4.3.2) **και** (4.3.3) η **συνάρτηση πιθανότητας** $f_v(v)$ γράφεται

$$f_v(v) = \frac{f_{uv}(u,v)}{f_u(u)}$$

$$= \frac{\dfrac{n!}{(m-1)!(n-m-1)!}F^{m-1}(u)\left[1-F(uv)\right]^{n-m-1} f(u)f(uv)u}{\dfrac{n!}{(m-1)!(n-m)!}F^{m-1}(u)\left[1-F(u)\right]^{n-m} f(u)}$$

$$= \frac{(n-m)!\left[1-F(uv)\right]^{n-m-1} f(uv)u}{(n-m-1)!\left[1-F(u)\right]^{n-m-1}\left[1-F(u)\right]}$$

$$= (n-m)\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m-1}\frac{f(uv)u}{1-F(u)} =$$

$$= -\frac{d}{dv}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m}$$

Από την παραπάνω ισότητα για τη συνάρτηση $\frac{1-F(uv)}{1-F(u)}$ έχουμε

$$\frac{d}{dv}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m} = -f_v(v) \quad \Leftrightarrow \quad \left[\frac{1-F(uv)}{1-F(u)}\right]^{n-m} = -\int f_v(v)dv + c_1$$

$$\Leftrightarrow \quad \frac{1-F(uv)}{1-F(u)} = \left[-\int f_v(v)dv + c_1\right]^{\frac{1}{n-m}}$$

Το δεύτερο μέλος της ισότητας αυτής είναι ανεξάρτητο του u που σημαίνει ότι και το πρώτο μέλος είναι ανεξάρτητο του u. Επομένως η συνάρτηση $\frac{1-F(uv)}{1-F(u)}$ είναι συνάρτηση μόνο του v και έστω ότι είναι G(v). Άρα

$$\frac{1-F(uv)}{1-F(u)} = G(v) \quad \Leftrightarrow \quad \frac{\overline{F}(uv)}{\overline{F}(u)} = G(v)$$

$$\Leftrightarrow \quad \overline{F}(uv) = \overline{F}(u)G(v)$$

Με χρήση του τύπου (Α.1.6) όπως δίνεται στο παράρτημα λύνουμε την παραπάνω συναρτησιακή εξίσωση και έχουμε

$$\overline{F}(x) = kx^c \quad \text{και} \quad G(x) = x^c, \quad x \geq \theta \text{ αφού } uv \geq \theta$$

άρα

$$F(x) = 1 - kx^c, \quad x \geq \theta > 0$$

Ισχύει

$$\lim_{x\to+\infty} F(x) = 1 \quad \Leftrightarrow \quad \lim_{x\to+\infty}\left(1-kx^c\right) = 1$$
$$\Leftrightarrow \quad \lim_{x\to+\infty}\left(x^c\right) = 0$$

από όπου συμπεραίνουμε ότι $c<0$ και έστω ότι είναι $c=-\alpha$ με $\alpha>0$. Τότε

$$F(x) = 1 - kx^{-\alpha}$$

και

$$F(\theta) = 0 \quad \Leftrightarrow \quad 1 - k\theta^{-\alpha} = 0$$
$$\Leftrightarrow \quad k\theta^{-\alpha} = 1$$
$$\Leftrightarrow \quad k = \theta^{\alpha}$$

Τελικά έχουμε

$$F(x)=1-\theta^{\alpha}x^{-\alpha}, \quad x\geq\theta>0,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$.

Το 1972 ο Samanta βελτίωσε τον παραπάνω χαρακτηρισμό. Το 1973 οι Ahsanullah-Kabir γενίκευσαν τον παραπάνω χαρακτηρισμό. Το θεώρημα που πρότειναν έχει ως εξής

ΘΕΩΡΗΜΑ 4.3.2

Έστω X μια τυχαία μεταβλητή η οποία έχει απολύτως συνεχή συνάρτηση κατανομής F(x) με $x\in[\theta,+\infty)$, $\theta>0$ και $X_{1:n}<X_{2:n}<\cdots<X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n του πληθυσμού. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X\sim P(\theta,\alpha)$ είναι για κάποιους αριθμούς r και s με $1\leq r<s\leq n$ τα στατιστικά

$$U=X_{r:n} \text{ και } V=\frac{X_{s:n}}{X_{r:n}} \tag{4.3.4}$$

να είναι ανεξάρτητα.

ΑΠΟΔΕΙΞΗ

Ευθέως

Η από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{r:n}<X_{s:n}$ από τον τύπο (3.5.1) είναι

$$f_{rs}(x_r,x_s)=k\cdot F^{r-1}(x_r)\left[F(x_s)-F(x_r)\right]^{s-r-1}\left[1-F(x_s)\right]^{n-s}f(x_r)f(x_s)$$

$$\text{με } k=\frac{n!}{(r-1)!(s-r-1)!(n-s)!}$$

Θεωρούμε το μετασχηματισμό

$$U=X_{r:n}$$

και

$$V=\frac{X_{s:n}}{X_{r:n}}$$

οπότε

$$X_{r:n} = U$$

και

$$X_{s:n} = UV$$

Τότε είναι

$$x_r = u$$

και

$$x_s = uv, \quad u \geq \theta, \quad v \geq 1$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \dfrac{d(x_r, x_s)}{d(u,v)}$

$$J = \begin{vmatrix} \dfrac{dx_r}{du} & \dfrac{dx_r}{dv} \\ \dfrac{dx_s}{du} & \dfrac{dx_s}{dv} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

Η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V είναι

$$f_{uv}(u,v) = f_{rs}(x_r, x_s)|J|$$

$$= k \cdot F^{r-1}(u)\left[F(uv) - F(u)\right]^{s-r-1}\left[1 - F(uv)\right]^{n-s} f(u)f(v)u \qquad (4.3.5)$$

με $k = \dfrac{n!}{(r-1)!(s-r-1)!(n-s)!}$

Η συνάρτηση πιθανότητας $f_u(u)$ της τυχαίας μεταβλητής $U = X_{r:n}$ από τον τύπο (3.3.1) είναι

$$f_u(u) = r\binom{n}{r}F^{r-1}(u)\left[1 - F(u)\right]^{n-r} f(u)$$

$$= r\frac{n!}{r!(n-r)!}F^{r-1}(u)\left[1 - F(u)\right]^{n-r} f(u)$$

$$= \frac{n!}{(r-1)!(n-r)!}F^{r-1}(u)\left[1 - F(u)\right]^{n-r} f(u) \qquad (4.3.6)$$

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(\theta,\alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha\theta^{\alpha}x^{-\alpha-1} \quad \text{και} \quad F(x) = 1-\theta^{\alpha}x^{-\alpha}, \quad x \geq \theta > 0,\ \alpha > 0$$

Στην περίπτωση αυτή η από κοινού συνάρτηση πιθανότητας $f_{uv}(u,v)$ των τυχαίων μεταβλητών U και V γράφεται

$$f_{uv}(u,v) = k\cdot\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}\left[\theta^{\alpha}u^{-\alpha}-\theta^{\alpha}(uv)^{-\alpha}\right]^{s-r-1}\left[\theta^{\alpha}(uv)^{-\alpha}\right]^{n-s}\alpha\theta^{\alpha}u^{-\alpha-1}\alpha\theta^{\alpha}(uv)^{-\alpha-1}u$$

$$= k\cdot\alpha^{2}\theta^{\alpha(s-r-1+n-s+2)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}\left[u^{-\alpha}-(uv)^{-\alpha}\right]^{s-r-1}(uv)^{-\alpha(n-s)}u^{-\alpha-1}(uv)^{-\alpha-1}u$$

$$= k\cdot\alpha^{2}\theta^{\alpha(n-r+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}\left(1-v^{-\alpha}\right)^{s-r-1}u^{-\alpha(s-r-1+n-s+1+1)-1}v^{-\alpha(n-s+1)-1}$$

$$= k\cdot\alpha^{2}\theta^{\alpha(n-r+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}\left(1-v^{-\alpha}\right)^{s-r-1}u^{-\alpha(n-r+1)-1}v^{-\alpha(n-s+1)-1}$$

και η συνάρτηση πιθανότητας $f_u(u)$ της τυχαίας μεταβλητής U γράφεται

$$f_u(u) = \frac{n!}{(r-1)!(n-r)!}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}\left(\theta^{\alpha}u^{-\alpha}\right)^{n-r}\alpha\theta^{\alpha}u^{-\alpha-1}$$

$$= \frac{n!}{(r-1)!(n-r)!}\alpha\theta^{\alpha(n-r+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}u^{-\alpha(n-r)}u^{-\alpha-1}$$

$$= \frac{n!}{(r-1)!(n-r)!}\alpha\theta^{\alpha(n-r+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}u^{-\alpha(n-r+1)-1}$$

Θα υπολογίσουμε τη συνάρτηση πιθανότητας της τυχαίας μεταβλητής V. Για την τυχαία μεταβλητή V ισχύει

$$V = \frac{X_{s:n}}{X_{r:n}} = \frac{X_{r+1:n}}{X_{r:n}}\frac{X_{r+2:n}}{X_{r+1:n}}\cdots\frac{X_{s:n}}{X_{s-1:n}} = R_{r+1}R_{r+2}\cdots R_{s}$$

δηλαδή είναι γινόμενο των τυχαίων μεταβλητών R_{r+1}, R_{r+2}, $\cdots$, R_s οι οποίες σύμφωνα με το θεώρημα (3.9.2) είναι ανεξάρτητες και ακολουθούν κατανομή Pareto με $R_i \sim P\left(1,(n-i+1)\alpha\right)$. Αυτό το γινόμενο όμως των ανεξάρτητων μεταβλητών από κατανομή Pareto σύμφωνα με το θεώρημα (1.9.5) και με χρήση του τύπου (A.2) του παραρτήματος έχει συνάρτηση πιθανότητας

$$f_v(v) = f_v(r_{r+1} r_{r+2} \cdots r_s)$$

$$= \prod_{i=r+1}^{s} (n-i+1)\alpha \sum_{j=r+1}^{s} \frac{1}{\prod\limits_{\substack{i=r+1 \\ i\neq j}}^{s} [(n-i+1)\alpha - (n-j+1)\alpha]} v^{-(n-j+1)\alpha-1}$$

$$= \alpha^{s-r} \prod_{i=r+1}^{s} (n-i+1) \sum_{j=r+1}^{s} \frac{1}{\alpha^{s-r-1} \prod\limits_{\substack{i=r+1 \\ i\neq j}}^{s} (n-i+1-n+j-1)} v^{-(n-j+1)\alpha-1}$$

$$= \alpha^{s-r} \frac{1}{\alpha^{s-r-1}} \prod_{i=r+1}^{s} (n-i+1) \sum_{j=r+1}^{s} \frac{1}{\prod\limits_{\substack{i=r+1 \\ i\neq j}}^{s} (j-i)} v^{-(n-j+1)\alpha-1}$$

$$= \alpha^{s-r} \frac{1}{\alpha^{s-r-1}} \prod_{i=r+1}^{s} (n-i+1) v^{-(n-s+1)\alpha-1} \sum_{j=r+1}^{s} \frac{1}{\prod\limits_{\substack{i=r+1 \\ i\neq j}}^{s} (j-i)} v^{-(s-j)\alpha}$$

$$= \alpha \frac{(n-r)!}{(n-s)!} v^{-(n-s+1)\alpha-1} \sum_{j=r+1}^{s} \frac{1}{\prod\limits_{\substack{i=r+1 \\ i\neq j}}^{s} (j-i)} v^{-(s-j)\alpha}$$

$$= \alpha \frac{(n-r)!}{(n-s)!} v^{-(n-s+1)\alpha-1} \sum_{j=r+1}^{s} \frac{(-1)^{s-j}}{(j-r-1)!(s-j)!} v^{-(s-j)\alpha}$$

$$= \alpha \frac{(n-r)!}{(n-s)!} v^{-(n-s+1)\alpha-1} \sum_{j=r+1}^{s} \frac{(-1)^{s-j}}{(j-r-1)!(s-j)!} v^{-(s-j)\alpha}$$

$$= \alpha \frac{(n-r)!}{(n-s)!(s-r-1)!} v^{-(n-s+1)\alpha-1} \sum_{j=r+1}^{s} \frac{(-1)^{s-j}(s-r-1)!}{(j-r-1)!(s-j)!} v^{-(s-j)\alpha}$$

$$= \alpha(s-r) \frac{(n-r)!}{(n-s)!(s-r)!} v^{-(n-s+1)\alpha-1} \sum_{j=r+1}^{s} (-1)^{s-j} \binom{s-r-1}{j-r-1} \left(v^{-\alpha}\right)^{s-j}$$

$$= \alpha(s-r) \binom{n-r}{n-s} v^{-(n-s+1)\alpha-1} \left(1-v^{-\alpha}\right)^{s-r-1}$$

Για να αποδείξουμε ότι τα στατιστικά

$$U = X_{r:n} \quad \text{και} \quad V = \frac{X_{s:n}}{X_{r:n}}$$

είναι ανεξάρτητα αρκεί να αποδείξουμε ότι

$$f_{uv}(uv) = f_u(u) f_v(v)$$

Πράγματι έχουμε

$$f_u(u)f_v(v) = \left[p \cdot \alpha\theta^{\alpha(n-r+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1} u^{-\alpha(n-r+1)-1}\right]\left[q \cdot v^{-(n-s+1)\alpha-1}\left(1-v^{-\alpha}\right)^{s-r-1}\right]$$

$$= pq \cdot \alpha^2\theta^{\alpha(n-r+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}\left(1-v^{-\alpha}\right)^{s-r-1} u^{-\alpha(n-r+1)-1} v^{-(n-s+1)\alpha-1}$$

όπου

$$p = \frac{n!}{(r-1)!(n-r)!} \quad \text{και} \quad q = (s-r)\binom{n-r}{n-s}$$

Για το γινόμενο pq έχουμε

$$pq = \frac{n!}{(r-1)!(n-r)!}(s-r)\binom{n-r}{n-s} v$$

$$= \frac{n!(s-r)(n-r)!}{(r-1)!(n-r)!(n-s)!(s-r)!}$$

$$= \frac{n!}{(r-1)!(s-r-1)!(n-s)!} = k$$

Άρα

$$f_u(u)f_v(v) = k \cdot \alpha^2\theta^{\alpha(n-r+1)}\left(1-\theta^{\alpha}u^{-\alpha}\right)^{r-1}\left(1-v^{-\alpha}\right)^{s-r-1} u^{-\alpha(n-r+1)-1} v^{-(n-s+1)\alpha-1}$$

$$= f_{uv}(uv)$$

<u>Αντιστρόφως</u>

Έστω ότι οι τυχαίες μεταβλητές U και V είναι ανεξάρτητες. Τότε

$$f_{uv}(u,v) = f_u(u)f_v(v) \quad \Leftrightarrow \quad f_v(v) = \frac{f_{uv}(u,v)}{f_u(u)}$$

Λαμβάνοντας υπόψη τους τύπους (4.3.5) **και** (4.3.6) **η συνάρτηση πιθανότητας** $f_v(v)$ **γράφεται**

$$f_v(v)=\frac{f_{uv}(u,v)}{f_u(u)}$$

$$=\frac{k\cdot F^{r-1}(u)\left[F(uv)-F(u)\right]^{s-r-1}\left[1-F(uv)\right]^{n-s}f(u)f(uv)u}{\frac{n!}{(r-1)!(n-r)!}F^{r-1}(u)\left[1-F(u)\right]^{n-r}f(u)}$$

$$=\frac{k\cdot\left[F(uv)-F(u)\right]^{s-r-1}\left[1-F(uv)\right]^{n-s}f(uv)u}{\frac{n!}{(r-1)!(n-r)!}\left[1-F(u)\right]^{n-r}}$$

και αντικαθιστώντας την τιμή του k με $k=\frac{n!}{(r-1)!(s-r-1)!(n-s)!}$ έχουμε

$$f_v(v)=\frac{\frac{n!}{(r-1)!(s-r-1)!(n-s)!}\left[F(uv)-F(u)\right]^{s-r-1}\left[1-F(uv)\right]^{n-s}f(uv)u}{\frac{n!}{(r-1)!(n-r)!}\left[1-F(u)\right]^{n-r}}$$

$$=\frac{(n-r)!}{(n-s)!(s-r-1)!}\frac{\left[F(uv)-F(u)\right]^{s-r-1}\left[1-F(uv)\right]^{n-s}f(uv)u}{\left[1-F(u)\right]^{s-r-1}\left[1-F(u)\right]^{n-s}\left[1-F(u)\right]}$$

$$=\frac{(n-r)!}{(n-s)!(s-r-1)!}\left[\frac{F(uv)-F(u)}{1-F(u)}\right]^{s-r-1}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-s}\frac{f(uv)u}{1-F(u)}$$

$$=-\frac{(n-r)!}{(n-s)!(s-r-1)!}\left[1-\frac{1-F(uv)}{1-F(u)}\right]^{s-r-1}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-s}\frac{d}{dv}\left[\frac{1-F(uv)}{1-F(u)}\right]$$

Από την παραπάνω ισότητα συμπεραίνουμε ότι η συνάρτηση

$$Q(u,v)=\left[1-\frac{1-F(uv)}{1-F(u)}\right]^{s-r-1}\left[\frac{1-F(uv)}{1-F(u)}\right]^{n-s}\frac{d}{dv}\left[\frac{1-F(uv)}{1-F(u)}\right]$$

είναι ανεξάρτητη του u οπότε σύμφωνα με το λήμμα (A.14) του παραρτήματος η συνάρτηση

$$q(u,v)=\frac{1-F(uv)}{1-F(u)}$$

είναι ανεξάρτητη του u δηλαδή είναι συνάρτηση μόνο του v και έστω ότι είναι G(v). Άρα

$$\frac{1-F(uv)}{1-F(u)}=G(v) \iff \frac{\overline{F}(uv)}{\overline{F}(u)}=G(v)$$

$$\iff \overline{F}(uv)=\overline{F}(u)G(v)$$

Με χρήση του τύπου (Α.16) του παραρτήματος λύνουμε την παραπάνω συναρτησιακή εξίσωση και έχουμε

$$\overline{F}(x)=kx^{c} \text{ και } G(x)=x^{c}, \quad x\geq\theta \text{ αφού } uv\geq\theta$$

άρα

$$F(x)=1-kx^{c}, \quad x\geq\theta>0$$

Ισχύει

$$\lim_{x\to+\infty}F(x)=1 \iff \lim_{x\to+\infty}\left(1-kx^{c}\right)=1$$

$$\iff \lim_{x\to+\infty}\left(x^{c}\right)=0$$

από όπου συμπεραίνουμε ότι $c<0$ και έστω ότι είναι $c=-\alpha$ με $\alpha>0$. Τότε

$$F(x)=1-kx^{-\alpha}$$

και

$$F(\theta)=0 \iff 1-k\theta^{-\alpha}=0$$

$$\iff k\theta^{-\alpha}=1$$

$$\iff k=\theta^{\alpha}$$

Τελικά έχουμε

$$F(x)=1-\theta^{\alpha}x^{-\alpha}, \quad x\geq\theta>0,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(\theta,\alpha)$.

Στη συνέχεια θα παρουσιάσουμε έναν χαρακτηρισμό που βασίζεται στις κατανομές των στατιστικών $\frac{X_{s_1:n}}{X_{r:n}}$ και $X_{s_1-r:n-r}$ όπου $X_{s_1-r:n-r}$ είναι το διατεταγμένο στατιστικό τάξης s_1-r ενός δείγματος μεγέθους $n-r$ για δύο διακριτές τιμές s_1 και s_2. Το σχετικό θεώρημα είναι το παρακάτω (4.3.3) και δόθηκε από τον Ahsanullah.

<u>ΘΕΩΡΗΜΑ 4.3.3</u>

Έστω X μια τυχαία μεταβλητή η οποία έχει απολύτως συνεχή και γνησίως αύξουσα συνάρτηση κατανομής $F(x)$ με $x \in [1,+\infty)$. Έστω επίσης $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n του πληθυσμού και $X_{1:n-r} < X_{2:n-r} < \cdots < X_{n-r:n-r}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους $n-r$ του πληθυσμού με $r<n$. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(1,\alpha)$ είναι τα στατιστικά

$$V_i = \frac{X_{s_i:n}}{X_{r:n}} \text{ και } W_i = X_{s_i-r:n-r}, \quad i=1,2 \text{ και } 1 \le r < s_1 < s_2 \le n \tag{4.3.7}$$

να ακολουθούν ακριβώς την ίδια κατανομή για δυο διακριτούς αριθμούς s_1 και s_2.

<u>ΑΠΟΔΕΙΞΗ</u>

<u>Ευθέως</u>

Η από κοινού συνάρτηση πιθανότητας των διατεταγμένων στατιστικών $X_{r:n} < X_{s_i:n}$ με $i=1,2$ από τον τύπο (3.5.1) είναι

$$f_{r,s_i}(x_r, x_{s_i}) = k \cdot F^{r-1}(x_r)\left[F(x_{s_i}) - F(x_r)\right]^{s_i-r-1}\left[1-F(x_{s_i})\right]^{n-s_i} f(x_r)f(x_{s_i})$$

$$\text{με } k = \frac{n!}{(r-1)!(s_i-r-1)!(n-s_i)!}$$

Θεωρούμε το μετασχηματισμό

$$U = X_{r:n}$$

και

$$V_i = \frac{X_{s_i:n}}{X_{r:n}}, \quad i=1,2$$

οπότε

$$X_{r:n} = U$$

και

$$X_{s_i:n} = UV_i$$

Τότε είναι

$$x_r = u \text{ και } x_{s_i} = uv_i, \quad u \ge 1,\ v \ge 1,\ i=1,2$$

Υπολογίζουμε τη Ιακωβιανή ορίζουσα $J = \dfrac{d(x_r, x_{s_i})}{d(u,v)}$

$$J = \begin{vmatrix} \dfrac{dx_r}{du} & \dfrac{dx_r}{dv_i} \\ \dfrac{dx_{s_i}}{du} & \dfrac{dx_{s_i}}{dv_i} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

Η από κοινού συνάρτηση πιθανότητας $f_{u,v_i}(u,v_i)$ των τυχαίων μεταβλητών U και V_i είναι

$$f_{u,v_i}(u,v_i) = f_{r,s_i}(x_r, x_{s_i})|J|$$

$$= k \cdot F^{r-1}(u)\left[F(uv_i) - F(u)\right]^{s_i-r-1}\left[1 - F(uv_i)\right]^{n-s_i} f(u)f(uv_i)u \qquad (4.3.8)$$

Για τη συνάρτηση πιθανότητας $f_u(u)$ της τυχαίας μεταβλητής $U = X_{r:n}$ από τον τύπο (3.3.1) είναι

$$f_u(u) = r\binom{n}{r} F^{r-1}(u)\left[1 - F(u)\right]^{n-r} f(u)$$

$$= r\frac{n!}{r!(n-r)!} F^{r-1}(u)\left[1 - F(u)\right]^{n-r} f(u)$$

$$= \frac{n!}{(r-1)!(n-r)!} F^{r-1}(u)\left[1 - F(u)\right]^{n-r} f(u) \qquad (4.3.9)$$

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(1,\alpha)$. Τότε από τους τύπους (1.2.1) και (1.2.2) έχουμε

$$f(x) = \alpha x^{-\alpha-1} \text{ και } F(x) = 1 - x^{-\alpha}, \quad x \geq 1,\ \alpha > 0$$

Στην περίπτωση αυτή η από κοινού συνάρτηση πιθανότητας $f_{u,v_i}(u,v_i)$ των τυχαίων μεταβλητών U και V_i γράφεται

$$f_{u,v_i}(u,v_i) = k \cdot \left(1 - u^{-\alpha}\right)^{r-1}\left[u^{-\alpha} - (uv_i)^{-\alpha}\right]^{s_i-r-1}\left[(uv_i)^{-\alpha}\right]^{n-s_i} \alpha u^{-\alpha-1}\alpha(uv_i)^{-\alpha-1}u$$

$$= k \cdot \alpha^2\left(1 - u^{-\alpha}\right)^{r-1}\left[u^{-\alpha}(1 - v_i)^{-\alpha}\right]^{s_i-r-1}(uv_i)^{-\alpha(n-s_i)}u^{-\alpha-1}(uv_i)^{-\alpha-1}u$$

$$= k \cdot \alpha^2\left(1 - u^{-\alpha}\right)^{r-1}\left(1 - v_i^{-\alpha}\right)^{s_i-r-1} u^{-\alpha(s_i-r-1+n-s_i+1+1)-1} v_i^{-\alpha(n-s_i+1)-1}$$

$$= k \cdot \alpha^2\left(1 - u^{-\alpha}\right)^{r-1}\left(1 - v^{-\alpha}\right)^{s_i-r-1} u^{-\alpha(n-r+1)-1} v_i^{-\alpha(n-s_i+1)-1}$$

και η συνάρτηση **πιθανότητας** $f_u(u)$ της τυχαίας μεταβλητής U γράφεται

$$f_u(u)=\frac{n!}{(r-1)!(n-r)!}\left(1-u^{-\alpha}\right)^{r-1}\left(u^{-\alpha}\right)^{n-r}\alpha u^{-\alpha-1}$$

$$=\frac{n!}{(r-1)!(n-r)!}\alpha\left(1-u^{-\alpha}\right)^{r-1}u^{-\alpha(n-r+1)-1}$$

Λόγω της παραπάνω ισότητας η **από κοινού συνάρτηση πιθανότητας** $f_{u,v_i}(u,v_i)$ των τυχαίων μεταβλητών U και V_i **γράφεται γινόμενο ως εξής**

$$f_{u,v_i}(u,v_i)=\frac{n!}{(r-1)!(s_i-r-1)!(n-s_i)!}\alpha^2\left(1-u^{-\alpha}\right)^{r-1}\left(1-v_i^{-\alpha}\right)^{s_i-r-1}u^{-\alpha(n-r+1)-1}v_i^{-\alpha(n-s_i+1)-1}$$

$$=\left[\frac{n!}{(r-1)!(n-r)!}\alpha\left(1-u^{-\alpha}\right)^{r-1}u^{-\alpha(n-r+1)-1}\right]\left[\frac{(n-r)!}{(s_i-r-1)!(n-s_i)!}\alpha\left(1-v_i^{-\alpha}\right)^{s_i-r-1}v_i^{-\alpha(n-s_i+1)-1}\right]$$

$$=f_u(u)\left[\frac{(n-r)!}{(s_i-r-1)!(n-s_i)!}\alpha\left(1-v_i^{-\alpha}\right)^{s_i-r-1}v_i^{-\alpha(n-s_i+1)-1}\right]$$

Στην παραπάνω ισότητα έχουμε την από κοινού συνάρτηση πιθανότητας $f_{u,v_i}(u,v_i)$ ως γινόμενο της συνάρτησης πιθανότητας $f_u(u)$ επί μια παράσταση που περιέχει μόνο τη μεταβλητή v_i. **Αυτό σημαίνει ότι** αυτή η **παράσταση είναι** η **συνάρτηση πιθανότητας της τυχαίας μεταβλητής** V_i. Δηλαδή

$$f_{v_i}(v_i)=\frac{(n-r)!}{(s_i-r-1)!(n-s_i)!}\alpha\left(1-v_i^{-\alpha}\right)^{s_i-r-1}v_i^{-\alpha(n-s_i+1)-1} \qquad (4.3.10)$$

Θέτουμε

$$W_i=X_{s_i-r:n-r} \qquad i=1,2$$

Η συνάρτηση πιθανότητας **της μεταβλητής** W_i **λόγω του** τύπου (3.3.1) **είναι**

$$f_{w_i}(w_i)=(s_i-r)\binom{n-r}{s_i-r}F^{s_i-r-1}(w_i)[1-F(w_i)]^{n-s_i}f(w_i)$$

$$=(s_i-r-1)\frac{(n-r)!}{(s_i-r)!(n-s_i)!}F^{s_i-r-1}(w_i)[1-F(w_i)]^{n-s_i}f(w_i)$$

$$=\frac{(n-r)!}{(s_i-r-1)!(n-s_i)!}F^{s_i-r-1}(w_i)[1-F(w_i)]^{n-s_i}f(w_i) \qquad (4.3.11)$$

και αφού πρόκειται για κατανομή Pareto με $X \sim P(1,\alpha)$ είναι

$$f_{w_i}(w_i) = \frac{(n-r)!}{(s_i-r-1)!(n-s_i)!}\left(1-w_i^{-\alpha}\right)^{s_i-r-1}\left(w_i^{-\alpha}\right)^{n-s_i}\alpha w_i^{-\alpha-1}$$

$$= \frac{(n-r)!}{(s_i-r-1)!(n-s_i)!}\alpha\left(1-w_i^{-\alpha}\right)^{s_i-r-1} w_i^{-\alpha(n-s_i+1)-1} \qquad (4.3.12)$$

Από τους τύπους (4.3.10) **και** (4.3.12) **συμπεραίνουμε** ότι οι τυχαίες μεταβλητές $V_i = \frac{X_{s_i:n}}{X_{r:n}}$ και $W_i = X_{s_i-r:n-r}$ **με** $i=1,2$ **ακολουθούν** ακριβώς την ίδια κατανομή.

Αντιστρόφως

Έστω ότι οι τυχαίες μεταβλητές $V_i = \frac{X_{s_i:n}}{X_{r:n}}$ και $W_i = X_{s_i-r:n-r}$ **με** $i=1,2$ ακολουθούν ακριβώς την ίδια κατανομή. Αυτό σημαίνει ότι ισχύει

$$f_{v_i}(x) = f_{w_i}(x), \quad i=1,2 \qquad (4.3.13)$$

Η από κοινού **συνάρτηση** πιθανό**τητας** $f_{u,v_i}(u,v_i)$ **των** τυχαίων μεταβλητών $U = X_{r:n}$ **και** $V_i = \frac{X_{s_i:n}}{X_{r:n}}$ **με** $i=1,2$ **από τον τύπο** (4.3.8) **είναι**

$$f_{u,v_i}(u,v_i) = k\cdot F^{r-1}(u)\left[F(uv_i)-F(u)\right]^{s_i-r-1}\left[1-F(uv_i)\right]^{n-s_i} f(u)f(uv_i)u$$

$$\text{με } k = \frac{n!}{(r-1)!(s_i-r-1)!(n-s_i)!}$$

Η περιθώρια συνάρτηση πιθανότητας της τυχαίας μεταβλητής $V_i = \frac{X_{s_i:n}}{X_{r:n}}$ με $i=1,2$ είναι

$$f_{v_i}(v_i) = \int_1^{+\infty} f_{u,v_i}(u,v_i)du$$

$$= k\int_1^{+\infty} F^{r-1}(u)\left[F(uv_i)-F(u)\right]^{s_i-r-1}\left[1-F(uv_i)\right]^{n-s_i} f(u)f(uv_i)udu$$

Η συνάρτηση π**ιθανότητας της τυχαίας μεταβλητής** $W_i = X_{s_i-r:n-r}$ με $i=1,2$ από τον τύπο (4.3.11) είναι

$$f_{w_i}(w_i)=\frac{(n-r)!}{(s_i-r-1)!(n-s_i)!}F^{s_i-r-1}(w_i)[1-F(w_i)]^{n-s_i}f(w_i)$$

$$=\rho\cdot F^{s_i-r-1}(w_i)[1-F(w_i)]^{n-s_i}f(w_i)$$

όπου $\rho=\dfrac{(n-r)!}{(s_i-r-1)!(n-s_i)!}$

Η ισότητα (4.3.13) γράφεται

$$f_{v_i}(x)=f_{w_i}(x) \quad\Leftrightarrow\quad \frac{f_{w_i}(x)}{f_{v_i}(x)}=1$$

$$\Leftrightarrow\quad \frac{k\int_1^{+\infty}F^{r-1}(u)[F(ux)-F(u)]^{s_i-r-1}[1-F(ux)]^{n-s_i}f(u)f(ux)u\,du}{\rho\cdot F^{s_i-r-1}(x)[1-F(x)]^{n-s_i}f(x)}=1$$

$$\Leftrightarrow\quad \int_1^{+\infty}\frac{F^{r-1}(u)[F(ux)-F(u)]^{s_i-r-1}[1-F(ux)]^{n-s_i}f(u)f(ux)u}{F^{s_i-r-1}(x)[1-F(x)]^{n-s_i}f(x)}du=\frac{\rho}{k}$$

$$\Leftrightarrow\quad \int_1^{+\infty}F^{r-1}(u)\frac{[F(ux)-F(u)]^{s_i-r-1}}{F^{s_i-r-1}(x)}\frac{[1-F(ux)]^{n-s_i}}{[1-F(x)]^{n-s_i}}\frac{f(u)f(ux)u}{f(x)}du=\frac{\rho}{k}$$

Όμως κάνοντας πράξεις στο περιεχόμενο του παραπάνω ολοκληρώματος έχουμε

$$\frac{[F(ux)-F(u)]^{s_i-r-1}}{F^{s_i-r-1}(x)}\frac{[1-F(ux)]^{n-s_i}}{[1-F(x)]^{n-s_i}}=\left[\frac{F(ux)-F(u)}{F(x)}\right]^{s_i-r-1}\left[\frac{1-F(ux)}{1-F(x)}\right]^{n-s_i}$$

$$=\left[\frac{\dfrac{F(ux)-F(u)}{F(x)}}{\dfrac{1-F(ux)}{1-F(x)}}\right]^{s_i}\left[\frac{F(x)}{F(ux)-F(u)}\right]^{r+1}\left[\frac{1-F(ux)}{1-F(x)}\right]^{n}$$

και αν θέσουμε

$$H(u,x)=\frac{F(ux)-F(u)}{1-F(ux)}\frac{1-F(x)}{F(x)}$$

τότε

$$\int_1^{+\infty}F^{r-1}(u)H^{s_i}(u,x)\left[\frac{F(x)}{F(ux)-F(u)}\right]^{r+1}\left[\frac{1-F(ux)}{1-F(x)}\right]^{n}\frac{f(u)f(ux)u}{f(x)}du=\frac{\rho}{k}$$

Η παραπάνω ισότητα ισχύει για $i=1$ και $i=2$. Αφαιρώντας κατά μέλη προκύπτει

$$\int_{1}^{+\infty} F^{r-1}(u)\left[H^{s_1}(u,x)-H^{s_2}(u,x)\right]\left[\frac{F(x)}{F(ux)-F(u)}\right]^{r+1}\left[\frac{1-F(ux)}{1-F(x)}\right]^{n}\frac{f(u)f(ux)u}{f(x)}du=0$$

Η συνάρτηση F είναι γνησίως αύξουσα. Άρα συμπεραίνουμε ότι

$$\left[H^{s_1}(u,x)-H^{s_2}(u,x)\right]=0 \quad \Leftrightarrow \quad H^{s_1}(u,x)\left[1-H^{s_2-s_1}(u,x)\right]=0$$

$$\Leftrightarrow \quad H^{s_1}(u,x)=0 \quad ή \quad H^{s_2-s_1}(u,x)=1$$

$$\Leftrightarrow \quad H(u,x)=0 \quad ή \quad H(u,x)=1$$

Επειδή όμως καμία από τις συναρτήσεις $F(ux)-F(x)$ και $1-F(x)$ δεν είναι μηδενική αποκλείεται να είναι $H(u,x)=0$. Άρα

$$H(u,x)=1 \quad \Leftrightarrow \quad \frac{F(ux)-F(u)}{1-F(ux)}\frac{1-F(x)}{F(x)}=1$$

$$\Leftrightarrow \quad [F(ux)-F(u)][1-F(x)]=F(x)[1-F(ux)]$$

$$\Leftrightarrow \quad F(ux)-F(x)F(ux)-F(u)+F(x)F(u)=F(x)-F(x)F(ux)$$

$$\Leftrightarrow \quad 1-F(ux)=1-F(x)-F(u)+F(x)F(u)$$

$$\Leftrightarrow \quad 1-F(ux)=[1-F(x)][1-F(u)]$$

$$\Leftrightarrow \quad \overline{F}(ux)=\overline{F}(x)\overline{F}(u)$$

Η λύση της παραπάνω συναρτησιακής εξίσωσης με βάση τον τύπο (Α.15) του παραρτήματος είναι

$$\overline{F}(x)=x^{c}$$

και

$$F(x)=1-x^{c}$$

Ισχύει

$$\lim_{x\to+\infty}F(x)=1 \quad \Leftrightarrow \quad 1-\lim_{x\to+\infty}x^{c}=1$$

$$\Leftrightarrow \quad \lim_{x\to+\infty}x^{c}=0$$

οπότε συμπεραίνουμε ότι $c<0$ και έστω ότι είναι $c=-\alpha$ με $\alpha>0$.

Τελικά έχουμε

$$F(x)=1-x^{-\alpha},\quad x\geq 1,\ \alpha>0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(1,\alpha)$.

Ο Desu πρότεινε ένα χαρακτηρισμό για την εκθετική κατανομή, ο οποίος έδωσε την ιδέα της διατύπωσης του παρακάτω θεωρήματος που αφορά την κατανομή Pareto.

<u>ΘΕΩΡΗΜΑ 4.3.4</u>

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x\in[1,+\infty)$. Αν $X_{1:n}<X_{2:n}<\cdots<X_{n:n}$ είναι τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n του πληθυσμού, τότε μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X\sim P(1,\alpha)$ είναι οι τυχαίες μεταβλητές

$$X \text{ και } Z=X_{1:n}^{n} \tag{4.3.14}$$

να ακολουθούν ακριβώς την ίδια κατανομή.

<u>ΑΠΟΔΕΙΞΗ</u>

<u>Ευθέως</u>

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X\sim P(1,\alpha)$. Τότε από τον τύπο (1.2.2) έχουμε

$$F(x)=1-x^{-\alpha},\quad x\geq 1,\ \alpha>0$$

Σύμφωνα με τους τύπους (3.9.6) και (3.9.7) η τυχαία μεταβλητή

$$Z=X_{1:n}^{n}=\left(\frac{X_{1:n}}{X_{0:n}}\right)^{n}=R_{1:n}^{n}\qquad (X_{0:n}=1)$$

ακολουθεί κατανομή Pareto με $Z\sim P(1,\alpha)$. Άρα

$$F_z(z)=1-z^{-\alpha},\quad z\geq 1,\ \alpha>0$$

πράγμα που σημαίνει ότι οι τυχαίες μεταβλητές X και Z ακολουθούν ακριβώς την ίδια κατανομή.

Αντιστρόφως

Έστω ότι οι τυχαίες μεταβλητές X και Z ακολουθούν ακριβώς την ίδια κατανομή. Αυτό σημαίνει ότι ισχύει

$$F(z) = F_z(z)$$

Λόγω του τύπου (3.4.3) η συνάρτηση κατανομής του μικρότερου διατεταγμένου στατιστικού $X_{1:n}$ είναι

$$F_1(x) = 1 - [1 - F(x)]^n$$

και αν θέσουμε $x = z^{\frac{1}{n}}$ αφού $Z = X_{1:n}^n$ τότε έχουμε

$$F_z(z) = F_1\left(z^{\frac{1}{n}}\right) = 1 - \left[1 - F\left(z^{\frac{1}{n}}\right)\right]^n$$

Με δεδομένο ότι $z = x^n$ η ισότητα $F(z) = F_z(z)$ γράφεται

$$\begin{aligned} F(z) = 1 - \left[1 - F\left(z^{\frac{1}{n}}\right)\right]^n \quad &\Leftrightarrow \quad F(x^n) = 1 - [1 - F(x)]^n \\ &\Leftrightarrow \quad 1 - F(x^n) = [1 - F(x)]^n \\ &\Leftrightarrow \quad \overline{F}(x^n) = \overline{F}^n(x) \\ &\Leftrightarrow \quad \log \overline{F}(x^n) = \log \overline{F}^n(x) \\ &\Leftrightarrow \quad \log \overline{F}(x^n) = n \log \overline{F}(x) \end{aligned}$$

Θέτουμε

$$H(x) = \log \overline{F}(x)$$

οπότε έχουμε

$$H(x^n) = nH(x)$$

Η λύση της συναρτησιακής εξίσωσης λόγω του τύπου (A.18) του παραρτήματος είναι

$$H(x) = c \log x$$

οπότε έχουμε

$$\begin{aligned} \log \overline{F}(x) = c \log x \quad &\Leftrightarrow \quad \log \overline{F}(x) = \log x^c \\ &\Leftrightarrow \quad \overline{F}(x) = x^c \end{aligned}$$

και

$$F(x) = 1 - x^c$$

Ισχύει

$$\lim_{x\to+\infty} F(x) = 1 \Leftrightarrow 1 - \lim_{x\to+\infty} x^c = 1$$

$$\Leftrightarrow \lim_{x\to+\infty} x^c = 0$$

οπότε συμπεραίνουμε ότι $c < 0$ και έστω ότι είναι $c = -\alpha$ με $\alpha > 0$.

Τελικά έχουμε

$$F(x) = 1 - x^{-\alpha}, \quad x \geq 1, \ \alpha > 0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(1,\alpha)$.

4.4 Χαρακτηρισμοί με βάση τις ροπές των διατεταγμένων στατιστικών.

Στην ενότητα αυτή παρουσιάζουμε χαρακτηρισμούς που βασίζονται στις σχέσεις των ροπών των διατεταγμένων στατιστικών.

ΘΕΩΡΗΜΑ 4.4.1

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [1,+\infty)$ και $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n της κατανομής αυτής. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(1,\alpha)$ είναι

$$\mu'_{1:n} = \frac{n\alpha}{n\alpha - 1}, \quad \alpha > 1 \tag{4.4.1}$$

όπου $\mu'_{1:n}$ είναι η ροπή πρώτης τάξης γύρω από το 0 του διατεταγμένου στατιστικού $X_{1:n}$.

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(1,\alpha)$. Τότε το πρώτο μέλος του τύπου (4.4.1) λόγω του τύπου (3.10.1) γράφεται

$$\mu'_{1:n} = \frac{n!}{(n-1)!} \frac{\Gamma\left(n-\frac{1}{\alpha}\right)}{\Gamma\left(n+1-\frac{1}{\alpha}\right)}$$

$$= n \frac{\Gamma\left(n-\frac{1}{\alpha}\right)}{\left(n-\frac{1}{\alpha}\right)\Gamma\left(n-\frac{1}{\alpha}\right)}$$

$$= \frac{n\alpha}{n\alpha-1}$$

<u>Αντιστρόφως</u>

Ο τύπος (4.4.1) λόγω του τύπου (3.6.2) γράφεται

$$\mu'_{1:n} = \frac{n\alpha}{n\alpha-1} \Leftrightarrow n\int_0^1 F^{-1}(u)(1-u)^{n-1}du = \frac{n\alpha}{n\alpha-1}$$

$$\Leftrightarrow \int_0^1 F^{-1}(u)(1-u)^{n-1}du = \frac{\alpha}{n\alpha-1}$$

$$\Leftrightarrow \int_0^1 F^{-1}(u)(1-u)^{n-1}du = \frac{\alpha}{n\alpha-1}\left[-\int_0^1 d(1-u)^{\frac{n\alpha-1}{\alpha}}\right]$$

$$\Leftrightarrow \int_0^1 F^{-1}(u)(1-u)^{n-1}du = \int_0^1 (1-u)^{\frac{n\alpha-1}{\alpha}-1}du$$

$$\Leftrightarrow \int_0^1 F^{-1}(u)(1-u)^{n-1}du - \int_0^1 (1-u)^{n-1-\frac{1}{\alpha}}du = 0$$

$$\Leftrightarrow \int_0^1 \left[F^{-1}(u)(1-u)^{n-1}du - (1-u)^{n-1-\frac{1}{\alpha}}\right]du = 0$$

$$\Leftrightarrow \int_0^1 \left[F^{-1}(u) - (1-u)^{-\frac{1}{\alpha}}\right](1-u)^{n-1}du = 0$$

Από την ισότητα αυτή προκύπτει ότι

$$F^{-1}(u) - (1-u)^{-\frac{1}{\alpha}} = 0$$

Θέτουμε

$$u = F(x)$$

και έχουμε

$$F^{-1}(F(x))-[1-F(x)]^{-\frac{1}{\alpha}}=0 \quad \Leftrightarrow \quad x-[1-F(x)]^{-\frac{1}{\alpha}}=0$$

$$\Leftrightarrow \quad x=[1-F(x)]^{-\frac{1}{\alpha}}$$

$$\Leftrightarrow \quad x^{-\alpha}=1-F(x)$$

$$\Leftrightarrow \quad F(x)=1-x^{-\alpha}, \quad x \geq 1,\ \alpha>0$$

που σημαίνει ότι **η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto** με $X \sim P(1,\alpha)$.

ΘΕΩΡΗΜΑ 4.4.2

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής F(x) με $x \in [1,+\infty)$ και $X_{1:n}<X_{2:n}<\cdots<X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n της κατανομής αυτής. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(1,\alpha)$ είναι

$$\mu'_{i+1:n}(2)-\mu'_{i:n}(2)=\frac{2}{\alpha(n-i)}\mu'_{i+1:n}(2), \quad 1 \leq i \leq n \tag{4.4.2}$$

όπου $\mu'_{i:n}(2)$ και $\mu'_{i+1:n}(2)$ είναι οι ροπές δεύτερης τάξης γύρω από το 0 των διατεταγμένων στατιστικών $X_{i:n}<X_{i+1:n}$.

ΑΠΟΔΕΙΞΗ

Ευθέως

Έστω ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(1,\alpha)$. Λόγω του τύπου (3.10.1) έχουμε

$$\mu'_{i+1:n}(2)=\frac{n!}{(n-i-1)!}\frac{\Gamma\left(n-i-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)}$$

και

$$\mu'_{i:n}(2)=\frac{n!}{(n-i)!}\frac{\Gamma\left(n-i+1-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)}$$

Τότε το πρώτο μέλος του τύπου (4.4.2) γράφεται

$$A_{\mu\acute{\epsilon}\lambda o\varsigma} = \frac{n!}{(n-i-1)!}\frac{\Gamma\left(n-i-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)} - \frac{n!}{(n-i)!}\frac{\Gamma\left(n-i+1-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)}$$

$$= \frac{n!}{(n-i-1)!}\frac{\Gamma\left(n-i-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)} - \frac{n!}{(n-i)!}\frac{\left(n-i-\frac{2}{\alpha}\right)\Gamma\left(n-i-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)}$$

$$= \frac{n!}{(n-i-1)!}\frac{\Gamma\left(n-i-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)} - \frac{n!}{(n-i-1)!(n-i)}\frac{\left(n-i-\frac{2}{\alpha}\right)\Gamma\left(n-i-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)}$$

$$= \left(1-\frac{n-i-\frac{2}{\alpha}}{n-i}\right)\frac{n!}{(n-i-1)!}\frac{\Gamma\left(n-i-\frac{2}{\alpha}\right)}{\Gamma\left(n+1-\frac{2}{\alpha}\right)}$$

$$= \frac{n-i-n+i+\frac{2}{\alpha}}{n-i}\mu'_{i+1:n}(2)$$

$$= \frac{2}{\alpha(n-i)}\mu'_{i+1:n}(2)$$

<u>Αντιστρόφως</u>

Για τη ροπή $\mu'_{i+1:n}(2)$ **λόγω του τύπου** (3.6.1) έχουμε

$$\mu'_{i+1:n}(2) = (i+1)\binom{n}{i+1}\int_0^1\left[F^{-1}(u)\right]^2 u^i(1-u)^{n-i-1}du$$

$$= (i+1)\frac{n!}{(i+1)!(n-i-1)!}\int_0^1\left[F^{-1}(u)\right]^2 u^i(1-u)^{n-i-1}du$$

$$= \frac{n!}{i!(n-i-1)!}\int_0^1\left[F^{-1}(u)\right]^2 u^i(1-u)^{n-i-1}du$$

Με ολοκλήρωση κατά παράγοντες το παραπάνω ολοκλήρωμα

$$I=\int_0^1\left[F^{-1}(u)\right]^2u^i(1-u)^{n-i-1}du$$

γράφεται

$$I=\int_0^1\left[F^{-1}(u)\right]^2u^i(1-u)^{n-i-1}du$$

$$=-\frac{1}{n-i}\int_0^1\left[F^{-1}(u)\right]^2u^i d(1-u)^{n-i}$$

$$=-\frac{1}{n-i}\left[\left[\left[F^{-1}(u)\right]^2u^i(1-u)^{n-i}\right]_0^1-\int_0^1(1-u)^{n-i}d\left[\left[F^{-1}(u)\right]^2u^i\right]\right]$$

$$=\frac{1}{n-i}\int_0^1(1-u)^{n-i}d\left[\left[F^{-1}(u)\right]^2u^i\right]$$

$$=\frac{1}{n-i}\left[i\int_0^1\left[F^{-1}(u)\right]^2u^{i-1}(1-u)^{n-i}du+2\int_0^1F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)\right]$$

Τελικά για τη ροπή $\mu'_{i+1:n}(2)$ **έχουμε**

$$\mu'_{i+1:n}(2)=\frac{n!}{i!(n-i-1)!}\frac{1}{n-i}\left[i\int_0^1\left[F^{-1}(u)\right]^2u^{i-1}(1-u)^{n-i}du+2\int_0^1F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)\right]$$

$$=\frac{n!}{i!(n-i)!}\left[i\int_0^1\left[F^{-1}(u)\right]^2u^{i-1}(1-u)^{n-i}du+2\int_0^1F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)\right]$$

$$=i\frac{n!}{i!(n-i)!}\int_0^1\left[F^{-1}(u)\right]^2u^{i-1}(1-u)^{n-i}du+\frac{2n!}{i!(n-i)!}\int_0^1F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)$$

$$=i\binom{n}{i}\int_0^1\left[F^{-1}(u)\right]^2u^{i-1}(1-u)^{n-i}du+\frac{2n!}{i!(n-i)!}\int_0^1F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)$$

$$=\mu'_{i:n}(2)+\frac{2n!}{i!(n-i)!}\int_0^1F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)$$

Άρα

$$\mu'_{i+1:n}(2)-\mu'_{i:n}(2)=\frac{2n!}{i!(n-i)!}\int_0^1F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)$$

Ο τύπος (4.4.2) γράφεται

$$\mu'_{i+1:n}(2)-\mu'_{i:n}(2)=\frac{2}{\alpha(n-i)}\mu'_{i+1:n}(2)$$

$$\Leftrightarrow\ \frac{2n!}{i!(n-i)!}\int_0^1 F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)=\frac{2}{\alpha(n-i)}(i+1)\binom{n}{i+1}\int_0^1\left[F^{-1}(u)\right]^2u^i(1-u)^{n-i-1}du$$

$$\Leftrightarrow\ \frac{2n!}{i!(n-i)!}\int_0^1 F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)=\frac{2(i+1)n!}{\alpha(n-i)(i+1)!(n-i-1)!}\int_0^1\left[F^{-1}(u)\right]^2u^i(1-u)^{n-i-1}du$$

$$\Leftrightarrow\ \int_0^1 F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)=\frac{1}{\alpha}\int_0^1\left[F^{-1}(u)\right]^2u^i(1-u)^{n-i-1}du$$

$$\Leftrightarrow\ \int_0^1 F^{-1}(u)u^i(1-u)^{n-i}dF^{-1}(u)-\frac{1}{\alpha}\int_0^1\left[F^{-1}(u)\right]^2u^i(1-u)^{n-i-1}du=0$$

$$\Leftrightarrow\ \int_0^1 F^{-1}(u)u^i(1-u)^{n-i-1}\left[(1-u)\frac{d}{du}F^{-1}(u)-\frac{1}{\alpha}F^{-1}(u)\right]=0$$

Από την ισότητα αυτή προκύπτει ότι

$$(1-u)\frac{d}{du}F^{-1}(u)-\frac{1}{\alpha}F^{-1}(u)=0$$

Λύνουμε την παραπάνω **διαφορική εξίσωση**

$$(1-u)\frac{d}{du}F^{-1}(u)-\frac{1}{\alpha}F^{-1}(u)=0\ \Leftrightarrow\ (1-u)\frac{d}{du}F^{-1}(u)=\frac{1}{\alpha}F^{-1}(u)$$

$$\Leftrightarrow\ \frac{1}{F^{-1}(u)}\frac{d}{du}F^{-1}(u)=\frac{1}{\alpha}\frac{1}{1-u}$$

$$\Leftrightarrow\ \frac{d}{du}\log\left[F^{-1}(u)\right]=\frac{d}{du}\left[-\frac{1}{\alpha}\log(1-u)\right]$$

$$\Leftrightarrow\ \log\left[F^{-1}(u)\right]=-\frac{1}{\alpha}\log(1-u)+\log c$$

$$\Leftrightarrow\ \log\left[F^{-1}(u)\right]=\log\left[c(1-u)^{-\frac{1}{\alpha}}\right]$$

$$\Leftrightarrow\ F^{-1}(u)=c(1-u)^{-\frac{1}{\alpha}}$$

Θέτουμε

$$u = F(x)$$

και έχουμε

$$F^{-1}(u) = c(1-u)^{-\frac{1}{\alpha}} \Leftrightarrow F^{-1}(F(x)) = c(1-F(x))^{-\frac{1}{\alpha}}$$

$$\Leftrightarrow x = c(1-F(x))^{-\frac{1}{\alpha}}$$

$$\Leftrightarrow c^{-1}x = (1-F(x))^{-\frac{1}{\alpha}}$$

$$\Leftrightarrow c^{\alpha}x^{-\alpha} = 1-F(x)$$

$$\Leftrightarrow F(x) = 1-c^{\alpha}x^{-\alpha}$$

Όμως

$$F(1) = 0 \Leftrightarrow 1-c^{\alpha} = 0$$

$$\Leftrightarrow c^{\alpha} = 1$$

$$\Leftrightarrow c = 1$$

και

$$\lim_{x\to+\infty} F(x) = 1 \Leftrightarrow 1-\lim_{t\to+\infty} x^{-\alpha} = 1$$

$$\Leftrightarrow \lim_{x\to+\infty} x^{-\alpha} = 0$$

οπότε συμπεραίνουμε ότι $\alpha > 0$.

Τελικά έχουμε

$$F(x) = 1-x^{-\alpha}, \quad x \geq 1, \ \alpha > 0$$

που σημαίνει ότι η τυχαία μεταβλητή X ακολουθεί κατανομή Pareto με $X \sim P(1,\alpha)$.

ΘΕΩΡΗΜΑ 4.4.3

Έστω X μια τυχαία μεταβλητή η οποία έχει συνεχή συνάρτηση κατανομής $F(x)$ με $x \in [1,+\infty)$ και $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n της κατανομής αυτής. Μια ικανή και αναγκαία συνθήκη για να ακολουθεί η τυχαία μεταβλητή X κατανομή Pareto με $X \sim P(1,\alpha)$ είναι

$$\mu'_{i+1:n}(2) - \mu'_{i:n-1}(2) = \frac{2}{n\alpha}\mu'_{i+1:n}(2), \quad 1 \leq i < n \tag{4.4.3}$$

ΑΠΟΔΕΙΞΗ

Από την αναδρομική σχέση (3.8.1) έχουμε

$$i \cdot \mu'_{i+1:n}(2) + (n-i)\mu'_{i:n}(2) = n \cdot \mu'_{i:n-1}(2) \quad \Leftrightarrow \quad \mu'_{i:n-1}(2) = \frac{1}{n}\left[i \cdot \mu'_{i+1:n}(2) + (n-i)\mu'_{i:n}(2)\right]$$

$$\Leftrightarrow \quad \mu'_{i:n-1}(2) = \frac{i}{n} \cdot \mu'_{i+1:n}(2) + \frac{n-i}{n}\mu'_{i:n}(2)$$

Ο τύπος (4.4.3) γράφεται

$$\mu'_{i+1:n}(2) - \mu'_{i:n-1}(2) = \frac{2}{n\alpha}\mu'_{i+1:n}(2) \Leftrightarrow \mu'_{i+1:n}(2) - \frac{i}{n} \times \mu'_{i+1:n}(2) - \frac{n\text{-}i}{n}\mu'_{i:n}(2) = \frac{2}{n\alpha}\mu'_{i+1:n}(2)$$

$$\Leftrightarrow \frac{n-i}{n}\mu'_{i+1:n}(2) - \frac{n-i}{n}\mu'_{i:n}(2) = \frac{2}{n\alpha}\mu'_{i+1:n}(2)$$

$$\Leftrightarrow \frac{n-i}{n}\left[\mu'_{i+1:n}(2) - \mu'_{i:n}(2)\right] = \frac{2}{n\alpha}\mu'_{i+1:n}(2)$$

$$\Leftrightarrow (n-i)\left[\mu'_{i+1:n}(2) - \mu'_{i:n}(2)\right] = \frac{2}{\alpha}\mu'_{i+1:n}(2)$$

$$\Leftrightarrow \mu'_{i+1:n}(2) - \mu'_{i:n}(2) = \frac{2}{\alpha(n-i)}\mu'_{i+1:n}(2)$$

Από το θεώρημα (4.4.2) συμπεραίνουμε ότι αυτή η σχέση είναι ικανή και αναγκαία συνθήκη για να ακολουθεί η μεταβλητή X κατανομή Pareto με $X \sim P(1,\alpha)$.

ΠΑΡΑΡΤΗΜΑ

Παραθέτουμε εδώ κάποιες βασικές γνώσεις από τα μαθηματικά και τη θεωρία κατανομών και πιθανοτήτων που χρησιμοποιήθηκαν στην παραπάνω εργασία.

Α.1 Συνδυασμοί των n στοιχείων ανά m.

Αν m και n είναι φυσικοί αριθμοί με $0 \leq m \leq n$, τότε το πλήθος των συνδυασμών των n στοιχείων ανά m είναι

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

Α.2 Διώνυμο του Newton.

Αν n είναι φυσικός αριθμός τότε

$$(\alpha+\beta)^n = \sum_{k=0}^{n} \binom{n}{k} \alpha^{n-k} \beta^k$$

A.3 Εκθετική κατανομή.

Μια τυχαία μεταβλητή Χ λέμε ότι ακολουθεί εκθετική κατανομή με παράμετρο $\theta > 0$ και συμβολίζουμε $X \sim \exp(\theta)$, αν η συνάρτηση πιθανότητας ορίζεται από τη σχέση

$$f(x) = \begin{cases} \theta e^{-\theta x} & \text{αν } x > 0, \quad \theta > 0 \\ 0 & \text{αν } x \leq 0 \end{cases}$$

A.4 Power function κατανομή.

Μια τυχαία μεταβλητή Χ λέμε ότι ακολουθεί power function κατανομή με παραμέτρους $\alpha > 0$, $\beta > 0$ και συμβολίζουμε $X \sim pf(\alpha,\beta)$, αν η συνάρτηση πιθανότητας ορίζεται από τη σχέση

$$f(x) = \begin{cases} \alpha\beta^{-\alpha}x^{\alpha-1} & \text{αν } 0 \leq x \leq \beta \\ 0 & \text{για κάθε άλλη τιμή του x} \end{cases}$$

A.5 Κατανομή Burr.

Μια τυχαία μεταβλητή Χ λέμε ότι ακολουθεί κατανομή Burr με παραμέτρους $\alpha > 0$, $\beta > 1$ και συμβολίζουμε $X \sim Burr(\alpha,\beta)$, αν η συνάρτηση πιθανότητας ορίζεται από τη σχέση

$$f(x) = \begin{cases} \dfrac{\alpha\beta x^{\beta-1}}{\left(1+x^{\beta}\right)^{\alpha+1}} & \text{αν } x \geq 0, \quad \alpha > 0, \ \beta > 1 \\ 0 & \text{για κάθε άλλη τιμή του x} \end{cases}$$

A.6 Λογιστική κατανομή.

Μια τυχαία μεταβλητή Χ λέμε ότι ακολουθεί λογιστική κατανομή με παραμέτρους $\theta > 0$, d και συμβολίζουμε $X \sim L(\theta,d)$, αν η συνάρτηση πιθανότητας ορίζεται από τη σχέση

$$f(x) = \begin{cases} \theta^{-1}e^{-\frac{x-d}{\theta}}\left(1+e^{-\frac{x-d}{\theta}}\right)^{-2} & \text{αν } -\infty < x < +\infty, \quad \theta > 0 \\ 0 & \text{για κάθε άλλη τιμή του x} \end{cases}$$

Α.7 Συνάρτηση γάμμα.

Η συνάρτηση γάμμα συμβολίζεται **με $\Gamma(\alpha)$ και ορίζεται από τη σχέση**

$$\Gamma(\alpha)=\int_0^{+\infty} t^{\alpha-1}e^{-t}dt\,, \quad \alpha>0$$

Βασικές ιδιότητες της συνάρτησης γάμμα

1) $\Gamma(1)=1$

2) $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$

3) $\Gamma(\alpha+1)=\alpha\Gamma(\alpha), \quad \alpha>0$

4) $\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-\kappa)=\dfrac{\Gamma(\alpha+1)}{\Gamma(\alpha-\kappa)}, \quad \alpha>0$

5) $\Gamma(n+1)=n!, \quad n\in N^*$

Α.8 Κατανομή γάμμα.

Μια τυχαία μεταβλητή Χ λέμε ότι ακολουθεί κατανομή γάμμα με παραμέτρους $\alpha>0$, $\beta>0$ και συμβολίζουμε $X\sim\gamma(\alpha,\beta)$, αν η συνάρτηση πιθανότητας ορίζεται από τη σχέση

$$f(x)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}} & \alpha\nu\ x>0,\quad \alpha>0,\ \beta>0\\ 0 & \alpha\nu\ x\le 0\end{cases}$$

όπου $\Gamma(\alpha)$ είναι η συνάρτηση γάμμα.

Α.9 Συνάρτηση βήτα.

Η συνάρτηση βήτα συμβολίζεται με $B(\alpha,\beta)$ και ορίζεται από τη σχέση

$$B(\alpha,\beta)=\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt\,, \quad \alpha>0\,,\ \beta>0$$

Βασικές ιδιότητες της συνάρτησης βήτα

1) $B(\alpha,\beta)=B(\beta,\alpha)\,, \quad \alpha>0\,,\ \beta>0$

2) $B(\alpha,1)=\dfrac{1}{\alpha}, \quad \alpha>0$

3) $B\left(\frac{1}{2},\frac{1}{2}\right)=\pi$

4) $B(1,1)=1$

5) $B(\alpha,\beta)=\dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, $\alpha>0$, $\beta>0$

<u>Α.10 Κατανομή βήτα.</u>

Μια τυχαία μεταβλητή Χ λέμε ότι ακολουθεί κατανομή βήτα με παραμέτρους $\alpha>0$, $\beta>0$ και συμβολίζουμε $X\sim\beta(\alpha,\beta)$, αν η συνάρτηση πιθανότητας ορίζεται από τη σχέση

$$f(x)=\begin{cases}\dfrac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1} & \text{αν } 0\le x\le 1,\quad \alpha>0,\ \beta>0\\ 0 & \text{για κάθε άλλη τιμή του } x\end{cases}$$

όπου $B(\alpha,\beta)$ είναι η συνάρτηση βήτα.

<u>Α.11 Άθροισμα τυχαίων μεταβλητών από εκθετική κατανομή.</u>

Αν $T_1,T_2,\dots,T_m$ είναι ανεξάρτητες **τυχαίες μεταβλητές** που ακολουθούν εκθετική κατανομή με $T_1\sim\exp(\alpha_1)$, $T_2\sim\exp(\alpha_2),\dots,T_m\sim\exp(\alpha_m)$ τότε η τυχαία μεταβλητή

$$V=T_1+T_2+\dots+T_m=\sum_{j=1}^{m}T_j$$

έχει συνάρτηση κατανομής

$$H_v(v)=1-\prod_{i=1}^{m}\alpha_i\sum_{j=1}^{m}\frac{e^{-\alpha_j v}}{\alpha_j\prod_{\substack{i=1\\ i\neq j}}^{m}(\alpha_i-\alpha_j)}$$

<u>Α.12 Άθροισμα τυχαίων μεταβλητών από εκθετική κατανομή.</u>

Αν $T_1,T_2,\dots,T_n$ είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή με $T_i\sim\exp(\alpha)$, $i=1,2,\dots,n$ τότε η τυχαία μεταβλητή

$$V=T_1+T_2+\dots+T_n=\sum_{i=1}^{n}T_i$$

έχει συνάρτηση πιθανότητας

$$g_v(v)=\frac{\alpha^n}{\Gamma(n)}v^{n-1}e^{-\alpha v}$$

A.13 Θεώρημα Renyi.

Έστω T μια τυχαία μεταβλητή που ακολουθεί εκθετική κατανομή και $T_{1:n} < T_{2:n} < \cdots < T_{n:n}$ τα διατεταγμένα στατιστικά ενός τυχαίου δείγματος μεγέθους n του πληθυσμού. Τότε το διατεταγμένο στατιστικό $Z_{k:n}$ τάξης κ μπορεί να γραφεί στη μορφή

$$Z_{k:n} = \sum_{j=1}^{k} \frac{d_j}{n-j+1}$$

όπου

$$d_j = (n-j+1)\left(Z_{j:n} - Z_{j-1:n}\right), \quad j = 1,2,\ldots,n \text{ και } Z_{0:n} = 0$$

είναι ανεξάρτητες μεταβλητές που ακολουθούν εκθετική κατανομή.

A.14 Λήμμα

Έστω F(x) μια απολύτως συνεχής συνάρτηση κατανομής και η συνάρτηση q(u,v) με

$$q(u,v) = \frac{1-F(uv)}{1-F(u)}$$

Έστω επίσης ότι

$$0 < q(u,v) < 1$$

Αν για κάποιες τιμές r, s με $r < s$ η συνάρτηση Q(u,v) με

$$Q(u,v) = \left[1-q(u,v)\right]^{s-r-1}\left[q(u,v)\right]^{n-s}\frac{d}{dv}q(u,v)$$

είναι ανεξάρτητη του u, τότε και η συνάρτηση q(u,v) είναι ανεξάρτητη του u, δηλαδή είναι συνάρτηση μόνο του v.

A.15 Η συναρτησιακή εξίσωση f(xy)=f(x)f(y).

Η γενική λύση της συναρτησιακής εξίσωσης

$$f(xy) = f(x)f(y), \quad x > 0, \; y > 0$$

είναι

$$f(x) = x^c, \quad x > 0$$

A.16 Η συναρτησιακή εξίσωση f(xy)=f(x)h(y).

Η γενική λύση της συναρτησιακής εξίσωσης

$$f(xy) = f(x)h(y), \quad x > 0, \; y > 0$$

είναι

$$f(t) = \alpha t^c$$

$$h(t) = t^c, \quad t > 0$$

A.17 Η συναρτησιακή εξίσωση f(xy)=g(x)h(y).

Η γενική λύση της συναρτησιακής εξίσωσης

$$f(xy) = g(x)h(y), \quad x > 0, \; y > 0$$

είναι

$$f(t) = \alpha\beta t^c$$

$$g(t) = \alpha t^c$$

$$h(t) = \beta t^c, \quad t > 0$$

A.18 Η συναρτησιακή εξίσωση f(x^n)=nf(x).

Η γενική λύση της συναρτησιακής εξίσωσης

$$f\left(x^n\right) = nf(x), \quad x > 0$$

είναι

$$f(x) = c\log x, \quad x > 0$$

ΒΙΒΛΙΟΓΡΑΦΙΑ

Arnold, Barry C. (1983). Pareto Distributions, International Co-operative Publishing House, Maryland.

Bertsekas, Dimitri P. and Tsitsiklis, John N. (2002). Introduction to Probability, Athens Scientific, Belmont, Massachusetts.

Bronson, R. (1973). Modern Introductory Differential Equations, Department of Mathematics and Computer Science Farleigh Dickinson University, McGraw-Hill, New York.

Cirillo, R. (1979). The Economics of Vilfredo Pareto, Frank Cass and Company Limited, London.

Cramer, H. (1946). Mathematical Methods of Statistics, Princeton University Press, New York.

DeGroot, Morris H. (1989). Probability and Statistics, Addison-Wesley, Reading, Mass.

Dimaki, C. (1977). Characterizations and Estimation Procedures for the Continuous Pareto Distributions, Department of Statistics, University of Warwick, England.

Dimaki, C. and Xekalaki, **E. (1990).** Identifiability of Income Distributions in the Context of Damage and Generating Models, Communications in Statistics-Theory and Methods, Department of Statistics, Athens University of Economics.

Dimaki, C. and **Xekalaki, E. (1992).** Towards a Unification of Certain Characterizations by Conditional Expectations, Department of Statistics, Athens University of Economics.

Durrett, R. (1996). Probability Theory and Examples, Duxbury, Belmont.

Ferguson, **N. S. (1967).** Mathematical Statistics – A Decision Theoretic Approach, Academic, New York.

Johnson, **N. L. and Kotz, S.** and **Balakrishnan,** N. **(1994).** Continuous Univariate Distributions, Wiley, New York.

Kagan, A. and Linnik, Y. and **Rao,** C. **(1973).** Characterization Problems in Mathematical Statistics, Wiley, new York.

Kotz, S. (1974). Characterization of Statistical Distributions – Asupplement to Recent Surveys, International Statistical Review.

Reiss, R. D. (1989). Approximate Distributions of Order Statistics with Applications to Nonparametric Statistics, Springer, New York.

Spiegel, M. (1975). Theory and Problems of Probability and Statistics, Rensselaer Polytechnic Institute of Connecticut, McGraw-Hill, New York.

Spiegel, M. (1962). Theory and Problems of Advanced Calculus, Rensselaer Polytechnic Institute of Connecticut, McGraw-Hill, New York.

Springer, Melvin D. (1979). The Algebra of Random Variables, Wiley, New York.

Κουνιάς, **Σ. και Μωυσιάδης Χ.** (1995). Θεωρία Πιθανοτήτων Ι, Εκδόσεις Ζήτη, Θεσσαλονίκη.

Λαμπράκης, Δ. (1980). Στατιστική, Αθήνα.

Ξεκαλάκη, Ε. (1993). Εισαγωγή στη Θεωρητική Στατιστική, Αθήνα.

Ξεκαλάκη, Ε. και Πανάρετος, Ι. (1993). Πιθανότητες και Στοιχεία Στοχαστικών Ανελίξεων, Αθήνα.

Πανάρετος, Ι. και Ξεκαλάκη, Ε. (2000). Εισαγωγή στη Στατιστική Σκέψη, Τόμος ΙΙ, Αθήνα.