

Μάθημα 7: Συσχέτιση τυχαίων μεταβλητών

Θωμάς Καμαλάκης

31 Αυγούστου 2023

1 Εισαγωγή

Στα επόμενα μαθήματα θα αναφερόμαστε συχνά στην συσχέτιση μεταξύ δύο ή περισσότερων τυχαίων μεταβλητών. Ας υποθέσουμε ότι έχουμε δύο ζάρια τα οποία τα ρίχνουμε και καταγράφουμε το αποτέλεσμα που έρχεται. Αν το πρώτο ζάρι αντιστοιχεί στην τυχαία μεταβλητή X και το δεύτερο στην Y τότε μπορούμε να υπολογίσουμε την πιθανότητα το πρώτο ζάρι να έρθει i και το δεύτερο να έρθει k με $1 \leq i, k \leq 6$,

$$\Pr\{X = i, Y = k\} \quad (1)$$

Ας θεωρήσουμε τώρα την κατά συνθήκη πιθανότητα $\Pr\{X = i|Y = k\}$ του να συμβεί το $X = i$, δεδομένου ότι έχει συμβεί το $Y = k$,

$$\Pr\{X = i|Y = k\} = \frac{\Pr\{X = i, Y = k\}}{\Pr\{Y = k\}} \quad (2)$$

Έχει μία αξία να ερμηνεύσουμε την κατά συνθήκη πιθανότητα στο (2). Αν ρίξουμε N φορές τα δύο ζάρια και μετρήσουμε πόσες φορές N_k έχει έρθει το δεύτερο ζάρι $Y = k$, αυτές θα είναι $N_k = N \Pr\{Y = k\}$. Το πλήθος των φορών N_{ik} που έχει έρθει το πρώτο $X = i$ και το δεύτερο $Y = k$ είναι $N \Pr\{X = i, Y = k\}$. Το πηλίκο N_{ik}/N_k είναι η πιθανότητα να έρθει το $X = i$ δεδομένου ότι έχει έρθει $Y = k$. Στην περίπτωση των ζαριών είναι φυσικό να θεωρήσουμε ότι το ένα ζάρι δεν επηρεάζει το άλλο. Επομένως η πιθανότητα να έρθει $X = i$ δεν εξαρτάται από το τι έχει έρθει το Y και επομένως για κάθε k θα πρέπει να έχουμε απλά:

$$\Pr\{X = i|Y = k\} = \Pr\{X = i\} \quad (3)$$

Καταλήγουμε επομένως στο ίδιο αποτέλεσμα που θα καταλήγαμε αν δεν λαμβάναμε υπόψη το δεύτερο ζάρι. Το δεύτερο ζάρι *δεν επηρεάζει καθόλου* το πρώτο και επομένως λέμε πως οι τυχαίες μεταβλητές X και Y είναι *ανεξάρτητες*. Προσέξτε ότι για να ισχύει αυτό θα πρέπει να ισχύει η (3), ή ισοδύναμα:

$$\Pr\{X = i, Y = k\} = \Pr\{X = i\} \Pr\{Y = k\} \quad (4)$$

που είναι η απαραίτητη συνθήκη για να πούμε ότι οι X και οι Y είναι ανεξάρτητες. Αν δεν ισχύει η (4), τότε οι X και Y λέμε ότι είναι *εξαρτημένες*. Για τις συνεχείς τυχαίες διαδικασίες τα πράγματα είναι παρόμοια, μόνο που αντί για τις πιθανότητες θα πρέπει να θεωρήσουμε τις πυκνότητες πιθανότητας.

Ας δούμε μερικά παραδείγματα:

1. Σε ένα σύστημα επικοινωνιών μεταδίδουμε *bits* δηλαδή σύμβολα “0” ή “1” με την ίδια πιθανότητα και ανεξάρτητα το ένα από το άλλο. Επομένως εάν b_k είναι το k σε σειρά *bit* που έχουμε να μεταδώσουμε θα ισχύουν οι παραπάνω σχέσεις:

$$\Pr\{b_k = 0\} = \Pr\{b_k = 1\} = \frac{1}{2} \quad (5)$$

$$\Pr\{b_k = p, b_m = q\} = \Pr\{b_k = p\} \Pr\{b_m = q\} = \frac{1}{4} \quad (6)$$

Η πρώτη σχέση προκύπτει από το ότι τα *bit* είναι ισοπίθانا ενώ η δεύτερη από το ότι είναι ανεξάρτητα μεταξύ τους. Τώρα ας θεωρήσουμε μία συγκεκριμένη ακολουθία από 5 διαδοχικά *bit*, π.χ. 10010. Ποια είναι η πιθανότητα να εμφανιστεί αυτή η σειρά; Μαθηματικά ψάχνουμε την πιθανότητα:

$$\Pr\{b_1 = 1, b_2 = 0, b_3 = 0, b_4 = 1, b_5 = 0\} \quad (7)$$

Εφόσον τα bit είναι ανεξάρτητα έπεται ότι:

$$\Pr \{b_1 = 1, b_2 = 0, b_3 = 0, b_4 = 1, b_5 = 0\} = \Pr \{b_1 = 1\} \Pr \{b_2 = 0\} \Pr \{b_3 = 0\} \Pr \{b_4 = 1\} \Pr \{b_5 = 0\} = \left(\frac{1}{2}\right)^5 = \frac{1}{32} \quad (8)$$

Εφόσον τα bit είναι ισοπίθανα δεν έχει σημασία η ακριβής τιμή των bit που θεωρούμε. Για παράδειγμα είναι εύκολο να δείξουμε ότι και η ακολουθία 01011 έχει την ίδια πιθανότητα εμφάνισης, $\frac{1}{32}$. Γενικότερα μία ακολουθία από Q bits $v_1, v_2, \dots, v_Q$ θα έχει πιθανότητα εμφάνισης:

$$\Pr \{b_1 = v_1, b_2 = v_2, \dots, b_Q = v_Q\} = \frac{1}{2^Q} \quad (9)$$

2. Σε ένα σύστημα επικοινωνιών συμβαίνουν σφάλματα κατά την μετάδοση των διαδοχικών bit b_k από τον πομπό στον δέκτη. Συμβολίζουμε με E_k είναι το ενδεχόμενο να συμβεί ένα σφάλμα στο k bit b_k που μεταδίδουμε και θεωρούμε ότι η πιθανότητα να συμβεί στο σφάλμα είναι ίση με p ενώ τα σφάλματα μεταξύ διαδοχικών bit συμβαίνουν ανεξάρτητα το ένα από το άλλο. Οπότε έχουμε:

$$\Pr \{E_k\} = p \quad (10)$$

$$\Pr \{E_m E_n\} = \Pr \{E_m\} \Pr \{E_n\} = p^2 \quad (11)$$

Η παραπάνω εξίσωση υπολογίζει την πιθανότητα να συμβεί σφάλμα ταυτόχρονα στο m και στο n bit θεωρώντας βέβαια ότι $m \neq n$. Ας προσπαθήσουμε τώρα να υπολογίσουμε το ενδεχόμενο σε μία ακολουθία Q bit να συμβεί ακριβώς ένα λάθος. Αν συμβολίσουμε με C_k το ενδεχόμενο το k bit να μεταδοθεί σωστά τότε προφανώς έχουμε $\Pr \{C_k\} = 1 - \Pr \{E_k\} = 1 - p$. Το ενδεχόμενο A_1 να γίνει σφάλμα σε 1 bit είναι η ένωση των ενδεχομένων να γίνει σφάλμα στο 1ο bit, στο 2ο bit και ούτω καθεξής. Οπότε:

$$A_1 = E_1 C_2 C_3 \dots C_Q \cup C_1 E_2 C_3 \dots C_Q \cup C_1 C_2 E_3 \dots C_Q \cup \dots \cup C_1 C_2 C_3 \dots E_Q \quad (12)$$

Είναι εύκολο να δούμε ότι τα ενδεχόμενα (12) είναι αμοιβαίως αποκλειόμενα οπότε

$$\Pr \{A_1\} = \Pr \{E_1 C_2 C_3 \dots C_Q\} + \Pr \{C_1 E_2 C_3 \dots C_Q\} + \Pr \{C_1 C_2 E_3 \dots C_Q\} + \dots + \Pr \{C_1 C_2 C_3 \dots E_Q\} = Qp(1-p)^{Q-1} \quad (13)$$

Η τελευταία σχέση προκύπτει από το γεγονός ότι κάθε επιμέρους ενδεχόμενο έχει πιθανότητα $p(1-p)^{Q-1}$. Για παράδειγμα το $E_1 C_2 C_3 \dots C_Q$ είναι το ενδεχόμενο να συμβεί σφάλμα στο πρώτο bit και όλα τα άλλα bit να είναι σωστά. Εφόσον τα σφάλματα συμβαίνουν ανεξάρτητα στα διαδοχικά bits θα έχουμε:

$$\Pr \{E_1 C_2 C_3 \dots C_Q\} = \Pr \{E_1\} \times \Pr \{C_2\} \times \Pr \{C_3\} \times \dots \times \Pr \{C_Q\} = p(1-p)^Q \quad (14)$$

3. Μπορούμε να επεκτείνουμε την συλλογιστική για να υπολογίσουμε την πιθανότητα να συμβούν q σφάλματα σε μία ακολουθία από Q bit. Όπως και πριν το ενδεχόμενο A_q του να συμβούν q σφάλματα είναι μία ένωση ενδεχομένων της μορφής $L_1 L_2 \dots L_Q$ όπου θα πρέπει q ενδεχόμενα να αντιστοιχούν σε εσφαλμένα bit και $Q - q$ να αντιστοιχούν σε σωστά μεταδιδόμενα bit. Για παράδειγμα αν $q = 2$ τότε

$$A_2 = E_1 E_2 C_3 \dots C_Q \cup E_1 C_2 E_3 \dots C_Q \cup C_1 E_2 E_3 \cup \dots \quad (15)$$

Κάθε επιμέρους ενδεχόμενο θα έχει πιθανότητα $p^q(1-p)^{Q-q}$ και υπάρχουν $\binom{Q}{q}$ τέτοια ενδεχόμενα όπου

$$\binom{Q}{q} = \frac{Q!}{q!(Q-q)!} \quad (16)$$

είναι το πλήθος των συνδυασμών Q στοιχείων ανά q . Επομένως θα έχουμε:

$$\Pr \{N_e = q\} = \Pr \{A_q\} = \binom{Q}{q} p^q (1-p)^{Q-q} \quad (17)$$

όπου το N_e είναι ο αριθμός των σφαλμάτων στην ακολουθία. Το παραπάνω αποτέλεσμα ονομάζεται και *διωνυμική κατανομή* για το N_e και μπορεί να χρησιμοποιηθεί για να περιγράψει την πιθανότητα να συμβούν ακριβώς q σφάλματα σε μία ακολουθία από Q bits.

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 from scipy.special import comb
4
5 plt.rcParams.update({
6     "text.usetex": True,
7     "font.family": "serif",
8     "font.serif": ["Palatino"],
9     "font.size" : 12,
10    "lines.linewidth" : 2,
11 })
12
13 def binomial(q, Q, p):
14     return comb(Q, q) * p ** q * (1-p) ** (Q-q)
15
16 Q = 10
17 errors = np.zeros(Q - 1)
18 p_error = 0.1
19 Niterations = 10000
20
21
22 for m in range(Niterations):
23     b = np.random.randint(0, 2, Q)
24     e = np.random.rand(Q)
25     i = np.where(e <= p_error)
26     c = np.copy(b)
27     c[i] = np.logical_not( b[i] )
28     no_errors = np.sum(np.abs(c-b))
29     if no_errors > 0:
30         errors[no_errors-1] += 1
31
32 Ne = np.arange(1, Q)
33 plt.close('all')
34 plt.figure(1)
35
36 plt.plot(Ne, errors / Niterations, 's', label = 'Monte-Carlo')
37 plt.plot(Ne, binomial(Ne, Q, p_error), label = 'Binomial')
38 plt.xlabel('$k$')
39 plt.ylabel('$\Pr\{N_e=k\}$')
40 plt.legend()

```

Listing 1: binomial.py

Στο [listing 1](#) δείχνουμε ένα παράδειγμα όπου προσπαθούμε να υπολογίσουμε την πιθανότητα (17) με την βοήθεια της Python τόσο θεωρητικά όσο πραγματοποιώντας το πείραμα της μετάδοσης μίας ακολουθίας **bit**. Έχει ένα ενδιαφέρον να μελετήσουμε το [listing](#) σε λίγο μεγαλύτερη λεπτομέρεια καθώς χρησιμοποιεί μία τεχνική που θα εφαρμόσουμε και σε άλλες περιπτώσεις.

Το κεντρικό σημείο είναι ο βρόγχος `for` όπου σε κάθε επανάληψη δημιουργούμε ένα πίνακα `b` με μέγεθος ίσο με τον αριθμό των **bits** `Q` που μεταδίδουμε. Στη συνέχεια υπολογίζουμε ένα πίνακα `e` με τυχαίους αριθμούς που λαμβάνονται από μία ομοιόμορφη κατανομή στο $[0, 1]$ με χρήση της έτοιμης συνάρτησης `rand`. Ο πίνακας αυτός χρησιμοποιείται για να αποφασίσουμε αν κάποιο **bit** είναι εσφαλμένο ή όχι. Τα λάθη πρέπει να συμβαίνουν με πιθανότητα `p_error` η οποία περιγράφει την πιθανότητα p στην (17).

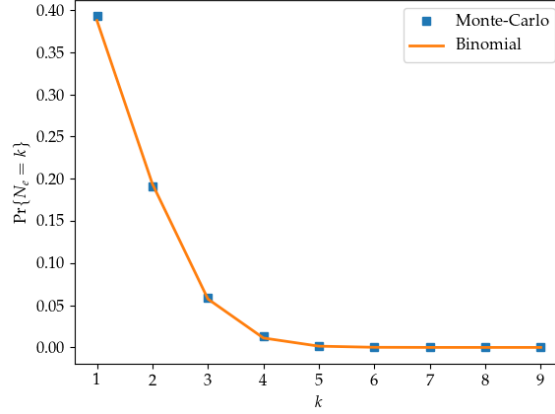
Εάν έχουμε μία ομοιόμορφη τυχαία μεταβλητή e στο $[0, 1]$ τότε η **PDF** της είναι $f_e(x) = 1$ και η πιθανότητα να έχουμε $e \leq p$ είναι:

$$\Pr\{e \leq p\} = \int_0^p f_e(x) dx = \int_0^p dx = p \quad (18)$$

Επομένως ένας τρόπος για να παράγουμε σφάλματα με πιθανότητα p είναι να χρησιμοποιούμε την ομοιόμορφη τυχαία μεταβλητή e και να θεωρούμε ότι έχουμε σφάλμα όταν $e \leq p$. Επομένως ψάχνουμε με την `where` ποια στοιχεία του `e` είναι μικρότερα του `p_error` και εκεί αντιστρέφουμε το **bit** με την `logical_not`. Η τεχνική αυτή όπου διεξάγουμε το πείραμα πολλές φορές γεννώντας κάθε φορά διαφορετικά δείγματα και υπολογίζουμε την πιθανότητα να συμβεί ένα ενδεχόμενο ονομάζεται μέθοδος *Monte Carlo* και εφαρμόζεται συχνά, ειδικά όταν ο υπολογισμός της πιθανότητας δεν είναι εύκολο να γίνει μαθηματικά. Στη συγκεκριμένη περίπτωση έχουμε ένα πίνακα, τον `errors` ο οποίος μετράει τον αριθμό των φορών $N(k)$ που έγιναν $k = 1, k = 2$, κ.ο.κ. σφάλματα. Στο τέλος της προσομοίωσης υπολογίζουμε την πιθανότητα να συμβούν k σφάλματα ως εξής

$$\Pr\{N_e = k\} \cong \frac{N(k)}{N} \quad (19)$$

όπου N ο αριθμός των επαναλήψεων. Παρεμπιπτόντως, αυτή η μέθοδος υπολογισμού των πιθανοτήτων με προσομοίωση μπορεί να εφαρμοστεί και σε άλλους τομείς της επιστήμης όπως τη φυσική και τα οικονομικά και αναπτύχθηκε



Εικόνα 1: Υπολογισμός του $\Pr\{N_e = k\}$ με προσομοίωση Monte Carlo

από τους Stanislaw Ulam και John von Neumann στο πρόγραμμα Manhattan που αφορούσε την κατασκευή της πρώτης ατομικής βόμβας στις ΗΠΑ. Ένας συνάδερφος τους ο οποίος μάλιστα είχε και Ελληνική καταγωγή ο Νικόλαος Μητρόπουλος (Nicholas Metropolis στα αγγλικά) και είχε σημαντική συνεισφορά στην περαιτέρω ανάπτυξη της μεθόδου πρότείνει να ονομαστεί “Monte Carlo” ίσως επειδή ο θεός του Ulam συνέχεια δανείζονταν χρήματα για να παίξει στο ομώνυμο καζίνο στο Μονακό.

Στην εικόνα 1 συγκρίνουμε τα αποτελέσματα της μεθόδου Monte-Carlo με το θεωρητικό αποτέλεσμα της (17). Παρατηρούμε ότι η συμφωνία είναι πολύ καλή κάτι που σημαίνει ότι το θεωρητικό μας μοντέλο που βασίζεται στην διωνυμική κατανομή είναι σωστό.

Η μέθοδος Monte-Carlo δεν είναι πανάκεια διότι έχει ένα σημαντικό πρόβλημα που μπορεί να αποτρέψει την εφαρμογή της: αν η πιθανότητα εμφάνισης των σφαλμάτων p είναι πολύ μικρή τότε χρειάζεται ένας πολύ μεγάλος αριθμός επαναλήψεων για να υπολογιστούν σωστά οι πιθανότητες βάσει της (19) διαφορετικά το $N(k)$ θα είναι πολύ μικρό ώστε η $\Pr\{N_e = k\}$ να εκτιμηθεί σωστά. Ένας πρακτικός κανόνας είναι το $N(k)$ να είναι της τάξης του 10 ώστε να έχουμε σχετικά καλή προσέγγιση της πιθανότητας.

2 Συνδυασμένη πυκνότητα πιθανότητας

Στο προηγούμενο μάθημα είχαμε ορίσει την πυκνότητα πιθανότητας μίας τυχαίας μεταβλητής X ως εξής:

$$f_X(x) = \lim_{\Delta x \rightarrow 0} \frac{\Pr\left\{x - \frac{\Delta x}{2} \leq X \leq x + \frac{\Delta x}{2}\right\}}{\Delta x} \quad (20)$$

Σε αναλογία με την (20), ορίζουμε την συνδυασμένη πυκνότητα πιθανότητας $f_{XY}(x, y)$ ως εξής:

$$f_{XY}(x, y) = \lim_{\Delta x, \Delta y \rightarrow 0} \frac{\Pr\left\{x - \frac{1}{2}\Delta x \leq X \leq x + \frac{1}{2}\Delta x, y - \frac{1}{2}\Delta y \leq Y \leq y + \frac{1}{2}\Delta y\right\}}{\Delta x \Delta y} \quad (21)$$

Σύμφωνα με τον παραπάνω ορισμό, για πολύ μικρά Δx και Δy μπορούμε να γράψουμε με καλή ακρίβεια,

$$f_{XY}(x, y) \approx \frac{\Pr\left\{x - \frac{1}{2}\Delta x \leq X \leq x + \frac{1}{2}\Delta x, y - \frac{1}{2}\Delta y \leq Y \leq y + \frac{1}{2}\Delta y\right\}}{\Delta x \Delta y} \quad (22)$$

Αν χωρίσουμε όλο το διάστημα όπου η τυχαία μεταβλητή Y μπορεί να πάρει τιμές σε ένα πολύ μεγάλο αριθμό διαστημάτων $J_i = [y_i - \Delta y/2, y_i + \Delta y/2]$ με κέντρο το y_i και μήκος το Δy . Μπορούμε να δούμε ότι τα ενδεχόμενα $Y \in J_p$ και $Y \in J_q$ δηλαδή η μεταβλητή Y να βρίσκεται εντός του J_p και εντός του J_q είναι αμοιβαίως αποκλειόμενα εάν $p \neq q$. Επίσης τα ενδεχόμενα $Y \in J_i$ είναι πλήρη εφόσον το Y με πιθανότητα 1 θα πάρει κάποια τιμή από το πεδίο τιμών της. Οπότε:

$$\sum_k \Pr\{Y \in J_k\} = 1 \quad (23)$$

Αν συμβολίσουμε με A το ενδεχόμενο να έχουμε $x - \frac{1}{2}\Delta x \leq X \leq x + \frac{1}{2}\Delta x$ τότε $f_X(x) = \Pr\{A\}/\Delta x$ και σύμφωνα με την (??) μπορούμε να γράψουμε:

$$\Pr\{A\} = \sum_k \Pr\{A|Y \in J_k\} \Pr\{Y \in J_k\} = \sum_k \Pr\{A, Y \in J_k\} \quad (24)$$

οπότε και

$$f_X(x)\Delta x = \sum_k \Pr\{A, Y \in J_k\} \quad (25)$$

Επίσης μπορούμε να δούμε ότι

$$\Pr\{A, Y \in J_k\} = f_{XY}(x, y_k)\Delta x\Delta y \quad (26)$$

και συνδυάζοντας τις τρεις παραπάνω σχέσεις,

$$f_X(x) = \sum_k f_{XY}(x, y_k)\Delta y \quad (27)$$

Παίρνοντας το όριο για $\Delta y \rightarrow 0$ θα έχουμε:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y)dy \quad (28)$$

Η συνδυασμένη πυκνότητα πιθανότητας χρησιμοποιείται για τον υπολογισμό αναμενόμενων τιμών της μορφής $\mathbb{E}\{g(X, Y)\}$, ως εξής:

$$\mathbb{E}\{g(X, Y)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y)f_{XY}(x, y)dxdy \quad (29)$$

Για να είναι ανεξάρτητες δύο συνεχείς τυχαίες μεταβλητές X και Y θα πρέπει να ισχύει:

$$f_{XY}(x, y) = f_X(x)f_Y(y) \quad (30)$$

όπου $f_X(x)$ και $f_Y(y)$ είναι οι πυκνότητες πιθανότητας των X και Y , αντίστοιχα. Για τις αναμενόμενες τιμές της μορφής $\mathbb{E}\{g(X)h(Y)\}$ όπου τα X και Y είναι ανεξάρτητες τυχαίες μεταβλητές μπορούμε να γράψουμε:

$$\begin{aligned} \mathbb{E}\{g(X)h(Y)\} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x)h(y)f_{XY}(x, y)dxdy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x)h(y)f_X(x)f_Y(y)dxdy = \\ &= \int_{-\infty}^{+\infty} g(x)f_X(x) \int_{-\infty}^{+\infty} h(y)f_Y(y)dxdy = \left(\int_{-\infty}^{+\infty} g(x)f_X(x)dx \right) \left(\int_{-\infty}^{+\infty} h(y)f_Y(y)dy \right) = \\ &= \mathbb{E}\{g(X)\} \mathbb{E}\{h(Y)\} \end{aligned} \quad (31)$$

Επομένως η αναμενόμενη τιμή $\mathbb{E}\{g(X)h(Y)\}$ ενός γινομένου συναρτήσεων $g(X)h(Y)$ ανεξαρτήτων μεταβλητών X και Y ισούται με το γινόμενο των αναμενόμενων τιμών $\mathbb{E}\{g(X)\} \mathbb{E}\{h(Y)\}$.

Στο [listing 2](#), δείχνουμε ένα παράδειγμα τέτοιων υπολογισμών. Θεωρούμε δύο τυχαίες μεταβλητές X και Y που έχουν ομοιόμορφη κατανομή στο $[0, 1]$. Εφόσον τα περιμένουμε οι αντίστοιχες μεταβλητές X και Y να είναι ανεξάρτητες. Στη συνέχεια υπολογίζουμε δύο νέες τυχαίες μεταβλητές Z και W ως εξής:

$$Z = X + 2Y \quad (32a)$$

$$W = 2X + Y \quad (32b)$$

Ορίζουμε τις συναρτήσεις $g(x)$ και $h(y)$ ως εξής στο παράδειγμα μας:

$$g(x) = x^2 \quad (33a)$$

$$h(y) = \sqrt{y} \quad (33b)$$

Στο [listing 2](#) υπολογίζουμε αριθμητικά την αναμενόμενη τιμή $\mathbb{E}\{g(X)h(Y)\}$ και την συγκρίνουμε με το γινόμενο $\mathbb{E}\{g(X)\} \mathbb{E}\{h(Y)\}$. Βρίσκουμε $\mathbb{E}\{g(X)h(Y)\} \approx \mathbb{E}\{g(X)\} \mathbb{E}\{h(Y)\} \approx 0.220$ οπότε στην περίπτωση των τυχαίων μεταβλητών X και Y που είναι ανεξάρτητες ισχύει η (31). Στην περίπτωση των μεταβλητών W και Z όμως, οι αριθμητικές τιμές που υπολογίζονται από το [listing 2](#) λαμβάνουμε $\mathbb{E}\{g(Z)h(W)\} \approx 3.565$ και $\mathbb{E}\{g(Z)\} \mathbb{E}\{h(W)\} \approx 3.144$ οπότε είναι φανερό ότι $\mathbb{E}\{g(Z)h(W)\} \neq \mathbb{E}\{g(Z)\} \mathbb{E}\{h(W)\}$ επειδή οι Z και W είναι εξαρτημένες.

```

1 import numpy as np
2
3 Ni = 10000
4 X = np.random.rand(Ni)
5 Y = np.random.rand(Ni)
6
7 Z = X + 2 * Y
8 W = 2 * X + Y
9
10 g = lambda x : x ** 2.0
11 h = lambda y : np.sqrt(y)
12
13 EgX = np.mean( g(X) )
14 EhY = np.mean( h(Y) )
15 EgXhY = np.mean( g(X) * h(Y) )
16
17 print('Expected Value of g(X)h(Y) : ', EgXhY)
18 print('Expected Value of g(X) times expected value of h(Y) : ', EgX * EhY)
19
20 EgZ = np.mean( g(Z) )
21 EhW = np.mean( h(W) )
22 EgZhW = np.mean( g(Z) * h(W) )
23
24 print('Expected Value of g(Z)h(W) : ', EgZhW)
25 print('Expected Value of g(Z) times expected value of h(W) : ', EgZ * EhW)

```

Listing 2: expectedvalues.py

3 Συντελεστής Συσχέτισης Τυχαίων Μεταβλητών

Ο συντελεστής συσχέτισης δύο τυχαίων μεταβλητών X και Y είναι ένα μέτρο πόσο συσχετισμένες είναι. Αν υπολογίσουμε τις μέσες τιμές και τις διακυμάνσεις των μεταβλητών ως εξής:

$$\mu_X = \mathbb{E}\{X\} \quad (34a)$$

$$\mu_Y = \mathbb{E}\{Y\} \quad (34b)$$

$$\sigma_X^2 = \mathbb{E}\{(X - \mu_X)^2\} \quad (34c)$$

$$\sigma_Y^2 = \mathbb{E}\{(Y - \mu_Y)^2\} \quad (34d)$$

τότε ο συντελεστής συσχέτισης ρ δίνεται από την σχέση:

$$\rho = \frac{\mathbb{E}\{(X - \mu_X)(Y - \mu_Y)\}}{\sigma_X \sigma_Y} \quad (35)$$

Στην περίπτωση όπου οι X και Y είναι ανεξάρτητες τότε μπορούμε να δείξουμε ότι:

$$\mathbb{E}\{(X - \mu_X)(Y - \mu_Y)\} = \mathbb{E}\{(X - \mu_X)\} \mathbb{E}\{(Y - \mu_Y)\} = 0 \quad (36)$$

Επίσης μπορεί να αποδειχθεί ότι ο συντελεστής ρ θα είναι πάντα $|\rho| \leq 1$.

4 Πολλαπλές Κανονικές Τυχαίες Μεταβλητές

Στην περίπτωση όπου οι X και η Y είναι κανονικές τυχαίες μεταβλητές τότε η συνδυασμένη πυκνότητα πιθανότητας τους θα δίνεται από την σχέση:

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} \right] \right) \quad (37)$$

όπου τα μ_X , μ_Y είναι οι μέσες τιμές των X και Y ενώ σ_X^2 και σ_Y^2 οι διακυμάνσεις τους όπως καθορίζονται από την (34). Η παράμετρος ρ ονομάζεται υπολογίζεται ως εξής από την (35).

Στην περίπτωση όπου οι μεταβλητές είναι ανεξάρτητες, θα έχουμε:

$$\begin{aligned}\mathbb{E}\{(x - \mu_X)(Y - \mu_Y)\} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(Y - \mu_Y) f_{XY}(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(Y - \mu_Y) f_X(x) f_Y(y) dx dy = \\ &= \int_{-\infty}^{+\infty} (x - \mu_X) f_X(x) dx \int_{-\infty}^{+\infty} (y - \mu_Y) f_Y(y) dy = 0\end{aligned}\quad (38)$$

Στην παραπάνω σχέση χρησιμοποιήσαμε ότι:

$$\mu_X = \int_{-\infty}^{+\infty} x f_X(x) dx \quad (39a)$$

$$\mu_Y = \int_{-\infty}^{+\infty} y f_Y(y) dy \quad (39b)$$

οπότε:

$$\int_{-\infty}^{+\infty} (x - \mu_X) f_X(x) dx = 0 \quad (40a)$$

$$\int_{-\infty}^{+\infty} (y - \mu_Y) f_Y(y) dy = 0 \quad (40b)$$

Επομένως ο συντελεστής συσχέτισης για δύο ανεξάρτητες μεταβλητές θα είναι $\rho = 0$. Αν οι X και Y είναι κανονικές ανεξάρτητες τυχαίες μεταβλητές τότε η PDF (37) θα γραφτεί ως εξής:

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \mathbb{E}p\left(-\frac{(x - \mu_X)^2}{2\sigma_X^2} - \frac{(y - \mu_Y)^2}{2\sigma_Y^2}\right) \quad (41)$$

Οι πυκνότητες πιθανότητας των X και Y δίνονται από τις σχέσεις:

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} \mathbb{E}p\left(-\frac{(x - \mu_X)^2}{2\sigma_X^2}\right) \quad (42a)$$

$$f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} \mathbb{E}p\left(-\frac{(y - \mu_Y)^2}{2\sigma_Y^2}\right) \quad (42b)$$

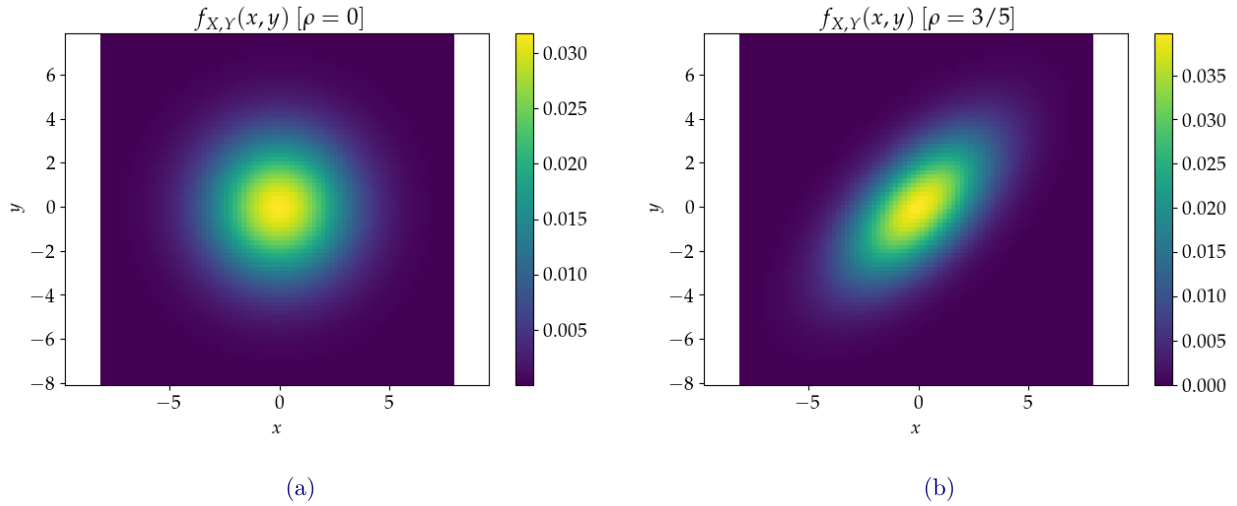
Συνδυάζοντας τις (42) και (41) βλέπουμε ότι το γεγονός ότι $\rho = 0$ συνεπάγεται ότι:

$$f_{XY}(x, y) = f_X(x) f_Y(y) \quad (43)$$

Επομένως προκύπτει ότι:

- στην περίπτωση όπου δύο μεταβλητές X και Y είναι ανεξάρτητες τότε έπεται ότι $\rho = 0$.
- ειδικά για την περίπτωση των κανονικών τυχαίων μεταβλητών, αν $\rho = 0$ τότε οι μεταβλητές X και Y είναι ανεξάρτητες.

```
1 import matplotlib.pyplot as plt
2
3 import numpy as np
4
5 plt.rcParams.update({
6     "text.usetex": True,
7     "font.family": "serif",
8     "font.serif": ["Palatino"],
9     "font.size" : 14,
10    "lines.linewidth" : 2,
11 })
12
13 def gausspdf2D(x, y, r, s2):
14     expo = -1.0/2.0/(1 - r**2) * (x**2/s2 + y**2/s2 - 2.0*r*x*y/s2)
15     pdf = 1.0/2.0/np.sqrt(1 - r**2)/ np.pi /s2 * np.exp(expo)
16     return pdf
17
18 def plot_2D(x, y, pdf, title):
```



Εικόνα 2: Συνδυασμένη PDF $f_{X,Y}(x,y)$ για δύο κανονικές μεταβλητές X και Y με $\mu_X = \mu_Y = 0$, $\sigma_X^2 = \sigma_Y^2 = 5$ και (a) $\rho = 0$ και (b) $\rho = 3/5$.

```

19 plt.pcolor(x,y,pdf, shading='auto')
20 plt.colorbar()
21 plt.axis('equal')
22 plt.title(title)
23 plt.xlabel('$x$')
24 plt.ylabel('$y$')
25 plt.tight_layout()
26
27 dx = 0.2
28 dy = 0.2
29 r = 3.0/5.0
30 s2 = 5
31
32 # generate 2d grids for the x & y bounds
33 y = np.arange(-8, +8, dy)
34 x = np.arange(-8, +8, dx)
35
36
37 yy, xx = np.meshgrid(y, x)
38 plt.close('all')
39
40 pdf1 = gausspdf2D(xx,yy, r, s2)
41 plt.figure(1)
42 plot_2D(x, y, pdf1, r'$f_{X,Y}(x,y)$ $\rho=3/5$')
43
44 pdf2 = gausspdf2D(xx,yy, 0, s2)
45 plt.figure(2)
46 plot_2D(x, y, pdf2, r'$f_{X,Y}(x,y)$ $\rho=0$')

```

Listing 3: bivariate.py

Μπορούμε να χρησιμοποιήσουμε την Python για να καταλάβουμε λίγο καλύτερα την συσχέτιση δύο τυχαίων κανονικών μεταβλητών. Ας αρχίσουμε με την PDF που δίνεται από την (37). Χρησιμοποιούμε την βιβλιοθήκη matplotlib για να κάνουμε γραφική παράσταση της $f_{X,Y}(x,y)$ στις τρεις διαστάσεις. Χρησιμοποιούμε τον κώδικα που δείχνουμε στο listing 3. Πριν κάνουμε κάποια γραφική παράσταση χρησιμοποιούμε την matplotlib.pyplot.rcParams.update για να επιφέρουμε κάποιες αισθητικές αλλαγές στο πως κάνουμε γραφικές παραστάσεις - κυρίως για να χρησιμοποιούμε Latex. Στη συνέχεια η gausspdf2D, χρησιμοποιείται για να υπολογίζει την $f_{X,Y}(x,y)$ σύμφωνα και με την (37). Η plot_2D, κάνει γραφική παράσταση συναρτήσεων $z = g(x,y)$ δύο μεταβλητών x και y χρησιμοποιώντας μία κλίμακα χρωμάτων για τις τιμές του z . Στη συνέχεια φτιάχνουμε δύο άξονες, έναν για την μεταβλητή x και μία για την μεταβλητή y :

```

1 y = np.arange(-8, +8, dy)
2 x = np.arange(-8, +8, dx)

```

Χρησιμοποιώντας την numpy.meshgrid φτιάχνουμε δύο διδιάστατα array, τα xx και yy που χρησιμοποιούνται για τον υπολογισμό της PDF σε δύο διαστάσεις. Οι πίνακες αυτοί έχουν στοιχεία:

```

1 yy[i,k] = y[k]

```


και χρησιμοποιούνται για τον υπολογισμό της PDF. Στην εικόνα 2a δείχνουμε $f_{XY}(x,y)$ για $\mu_X = \mu_Y = 0$, $\sigma_X^2 = \sigma_Y^2 = 5$ και $\rho = 0$ ενώ στην εικόνα 2b την PDF για τις ίδιες παραμέτρους αλλά $\rho = 3/5$. Έχει ενδιαφέρον να ερμηνεύσουμε τα αποτελέσματα. Στην εικόνα 2a, η $f_{XY}(x,y)$ έχει μία κυκλική συμμετρία ενώ στην 2b έχουμε μία στραμμένη έλλειψη, η οποία οφείλεται στη συσχέτιση των X και Y ($\rho \neq 0$). Η μορφή αυτή της PDF στην 2b έχει ως αποτέλεσμα για παράδειγμα ότι εάν $X > 0$ τότε είναι πιο πιθανό να έχουμε $Y > 0$ (δηλαδή να είμαστε στο πρώτο τεταρτημόριο, $X > 0$ και $Y > 0$) παρά $Y < 0$ (δηλαδή να είμαστε στο τέταρτο τεταρτημόριο, $X > 0$ και $Y < 0$) επειδή η PDF στο πρώτο τεταρτημόριο είναι πιο “φωτεινή” από ότι στο τέταρτο.

5 Μετασχηματισμοί τυχαίων μεταβλητών

Πολλές φορές χρειάζεται να υπολογίσουμε τη συνδυασμένη πυκνότητα πιθανότητας τυχαίων μεταβλητών που εξαρτώνται από άλλες τυχαίες μεταβλητές. Ας ξεκινήσουμε με μία τυχαία μεταβλητή Y που εξαρτάται από μία άλλη μεταβλητή X η οποία έχει PDF την $f_X(x)$ ως εξής:

$$Y = g(X) \quad (44)$$

Στην απλή περίπτωση όπου η g είναι παραγωγίσιμη, η παράγωγος της $g'(x)$ είναι παντού μη μηδενική, $g'(x) \neq 0$ και η $y = g(x)$ έχει πάντα μόνο μία λύση, δηλαδή για κάθε τιμή της $Y = y$ αντιστοιχεί μία και μόνο τιμή της $X = x = g^{-1}(y)$ τότε θα έχουμε:

$$g(x + \Delta x) \cong g(x) + \Delta x g'(x) \quad (45)$$

από όπου προκύπτει ότι η αλλαγή Δy που οφείλεται όταν το x αυξάνεται κατά Δx είναι:

$$\Delta y = g(x + \Delta x) - g(x) \cong \Delta x g'(x) \quad (46)$$

Μπορούμε εύκολα να δούμε ότι εάν το X είναι στο διάστημα $[x - \Delta x, x + \Delta x]$ τότε εάν $g'(x) > 0$ το Y θα είναι στο διάστημα $[y - \Delta x g'(x), y + \Delta x g'(x)]$ ενώ εάν $g'(x) < 0$ θα είναι στο διάστημα $[y + \Delta x g'(x), y - \Delta x g'(x)]$. Σε κάθε περίπτωση επομένως θα έχουμε

$$x - \Delta x \leq X \leq x + \Delta x \iff y - \Delta x |g'(x)| \leq Y \leq y + \Delta x |g'(x)| \quad (47)$$

Εναλλακτικά θα μπορούσαμε να ορίσουμε $\Delta y = \Delta x / |g'(x)|$ και να γράψουμε:

$$x - \frac{\Delta y}{|g'(x)|} \leq X \leq x + \frac{\Delta y}{|g'(x)|} \iff y - \Delta y \leq Y \leq y + \Delta y \quad (48)$$

οπότε θα έχουμε:

$$\frac{1}{\Delta y} \Pr \{y - \Delta y \leq Y \leq y + \Delta y\} = \frac{1}{\Delta y} \Pr \left\{ x - \frac{\Delta y}{|g'(x)|} \leq X \leq x + \frac{\Delta y}{|g'(x)|} \right\} \quad (49)$$

Το πρώτο μέλος είναι η PDF $f_Y(y)$ του Y όταν $\Delta y \rightarrow 0$. Το δεύτερο μέλος γράφεται:

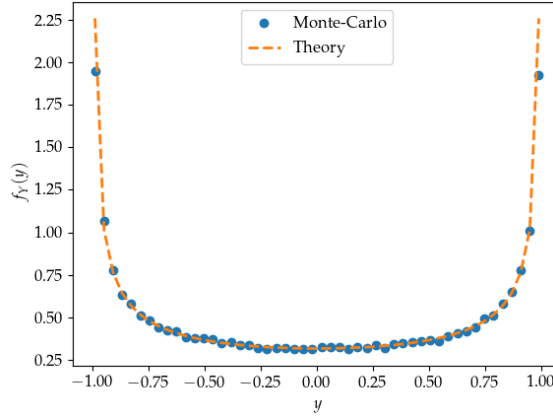
$$\frac{1}{|g'(x)|} \frac{1}{\frac{\Delta y}{|g'(x)|}} \Pr \left\{ x - \frac{\Delta y}{|g'(x)|} \leq X \leq x + \frac{\Delta y}{|g'(x)|} \right\} \quad (50)$$

και τείνει στο $\frac{1}{|g'(x)|} f_X(x)$. Επομένως,

$$f_Y(y) = \frac{1}{|g'(x)|} f_X(x), \text{ όπου } x = g^{-1}(y) \quad (51)$$

Μπορεί κανείς να δείξει ότι εάν η εξίσωση $g(x) = y$ έχει N λύσεις $x_i(y)$ με $g'(x) \neq 0$, τότε η πυκνότητα πιθανότητας γενικεύεται ως εξής:

$$f_Y(y) = \sum_i \frac{1}{|g'(x_i)|} f_X(x_i), \text{ όπου } x_i = g_i^{-1}(y) \quad (52)$$



Εικόνα 3: Επιβεβαίωση της (57) με προσομοίωση Monte Carlo

5.1 Παράδειγμα: μετασχηματισμένη κανονική μεταβλητή

Για ένα πρώτο παράδειγμα ας θεωρήσουμε μία κανονική τυχαία μεταβλητή X με μέση τιμή $\mu_X = 0$, διακύμανση $\sigma_X^2 = 1$ και μία δεύτερη μεταβλητή Y που σχετίζεται με την X ως εξής:

$$Y = \mu + \sigma X \quad (53)$$

Η PDF της $f_X(x)$ θα είναι:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (54)$$

Στην περίπτωση αυτή η συνάρτηση $g(x) = \mu + \sigma x$ είναι γραμμική με $g'(x) = \sigma$ και η $g(x) = y$ έχει μόνο μία λύση, την $x = (y - \mu)/\sigma$. Επομένως θα έχουμε:

$$f_Y(y) = \frac{1}{g'(x)} f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (55)$$

Επομένως εάν έχουμε μία κανονική μεταβλητή X με μηδενική μέση τιμή και διακύμανση ίση με 1, την πολλαπλασιάσουμε με σ και προσθέσουμε μ τότε η μετασχηματισμένη μεταβλητή Y είναι μία κανονική μεταβλητή με μέση τιμή μ και διακύμανση σ^2 .

5.2 Παράδειγμα: συνημίτονο ομοιόμορφης κατανομής

Για δεύτερο παράδειγμα ας υποθέσουμε ότι έχουμε μία μεταβλητή Φ με ομοιόμορφη κατανομή $f_\Phi(\phi) = \frac{1}{2\pi}$ για $|\phi| \leq \pi$ και ψάχνουμε να βρούμε την PDF $f_Y(y)$ της $Y = \cos \Phi$. Αν θυμηθούμε από την τριγωνομετρία, η εξίσωση $y = \cos \phi$ έχει δύο λύσεις $\phi_1(y) = \arccos(y)$ και $\phi_2(y) = -\arccos(y)$ όπου η $\phi = \arccos(y)$ είναι η γωνία $0 \leq \phi \leq \pi$ η οποία έχει συνημίτονο ίσο με y . Από τους τύπους του διαφορικού λογισμού έχουμε:

$$\frac{d}{d\phi} \cos(\phi) = -\sin \phi \quad (56)$$

Επομένως η PDF θα είναι:

$$f_Y(y) = \sum_{i=1}^2 \frac{1}{|\sin(\phi_i)|} f_\Phi(\phi_i) = \frac{1}{\pi \sqrt{1-y^2}} \quad (57)$$

Χρησιμοποιούμε το [listing 4](#) για να επιβεβαιώσουμε την (57). Στην ουσία γεννάμε ομοιόμορφα τυχαία δείγματα της Φ στον πίνακα `phi` και εφαρμόζουμε τον μετασχηματισμό $Y = \cos(\Phi)$. Μετά χρησιμοποιούμε την `calc_pdfcdf` που είχαμε φτιάξει στα προηγούμενα για να υπολογίσουμε την PDF της Y . Πρόκειται στην ουσία πάλι για προσομοίωση Monte-Carlo μόνο που αυτή την φορά δεν χρειάζεται κάποιο `for loop` αφού η `numpy` μας δίνει την δυνατότητα να υπολογίσουμε όλα τα δείγματα ταυτόχρονα. Τα αποτελέσματα φαίνονται στην εικόνα 3 που επιβεβαιώνουν την (57).

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
```

```

4
5
6 plt.rcParams.update({
7     "text.usetex": True,
8     "font.family": "serif",
9     "font.serif": ["Palatino"],
10    "font.size" : 12,
11    "lines.linewidth" : 2,
12 })
13
14 def pdfY(y):
15     pdf = np.zeros(y.size)
16     i = np.where ( np.abs(y) < 1 )
17     pdf[i] = 1 / np.pi / np.sqrt(1 - y[i] ** 2)
18     return pdf
19
20 def calc_pdfcdf(x, y):
21     N = x.size
22     Dy = y[1] - y[0]
23     pdf = np.zeros( y.size )
24     cdf = np.zeros( y.size )
25
26     for i, yy in enumerate(y):
27         Ny = np.sum( (x >= yy-Dy/2) & (x < yy+Dy/2) )
28         pdf[i] = Ny / N / Dy
29         cdf[i] = np.sum( x <= yy ) / N
30
31     return pdf, cdf
32
33 Nsamples = 100000
34 a = -np.pi
35 b = np.pi
36 Nplot = 1000
37
38
39 phi = a + (b-a) * np.random.rand(Nsamples)
40 y = np.cos(phi)
41
42 yy = np.linspace(-0.99, 0.99, 50)
43 pdf, cdf = calc_pdfcdf(y, yy)
44
45 plt.figure()
46 plt.plot(yy, pdf, 'o', label = 'Monte-Carlo')
47 plt.plot(yy, pdfY(yy), '--', label = 'Theory')
48 plt.xlabel('$y$')
49 plt.ylabel('$f_Y(y)$')
50 plt.legend()

```

Listing 4: transform.py

5.3 Μετασχηματισμοί πολλαπλών μεταβλητών

Στην περίπτωση όπου έχουμε μετασχηματισμούς πολλαπλών τυχαίων μεταβλητών υπάρχει ένας αντίστοιχος τρόπος να υπολογίσουμε την συνδυασμένη PDF. Έστω για παράδειγμα ότι οι τυχαίες μεταβλητές Z και W σχετίζονται με τις X και Y ως εξής:

$$Z = g(X, Y) \quad (58)$$

$$W = h(X, Y) \quad (59)$$

Τότε εάν το σύστημα των εξισώσεων

$$g(x, y) = z \quad (60)$$

$$h(x, y) = w \quad (61)$$

έχει διάφορες λύσεις x_i και y_i και ορίσουμε τον πίνακα

$$\mathbf{J}(x, y) = \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} \quad (62)$$

Ο πίνακας $\mathbf{J}$ ονομάζεται ο πίνακας *Jacobi* του συστήματος. Θα έχουμε:

$$f_{Z,W}(z, w) = \sum_i \frac{f_{XY}(x_i, y_i)}{|\det \mathbf{J}(x_i, y_i)|} \quad (63)$$

όπου με $\det \mathbf{J}(x_i, y_i)$ συμβολίζουμε την ορίζουσα του πίνακα $\mathbf{J}(x_i, y_i)$. Ας δούμε ένα παράδειγμα. Έστω

$$Z = X + Y \quad (64)$$

$$W = X - Y \quad (65)$$

οπότε

$$g(X, Y) = X + Y \quad (66)$$

$$h(X, Y) = X - Y \quad (67)$$

Ο πίνακας **Jacobi** θα είναι

$$\mathbf{J}(x, y) = \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (68)$$

που έχει ορίζουσα:

$$\det \mathbf{J} = 1 \times (-1) - 1 \times 1 = -2 \quad (69)$$

Το σύστημα

$$z = x + y \quad (70)$$

$$w = x - y \quad (71)$$

έχει λύση

$$x = \frac{1}{2}(z + w) \quad (72)$$

$$y = \frac{1}{2}(z - w) \quad (73)$$

οπότε σύμφωνα με την (63) θα έχουμε

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY} \left(\frac{z+w}{2}, \frac{z-w}{2} \right) \quad (74)$$

Η παραπάνω σχέση υπολογίζει την συνδυασμένη PDF των Z, W .

5.4 Αθροίσματα τυχαίων μεταβλητών

Μπορούμε να υπολογίσουμε την PDF της Z βάσει της (28). Θα έχουμε:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{ZW}(z, w) dw = \frac{1}{2} \int_{-\infty}^{+\infty} f_{XY} \left(\frac{z+w}{2}, \frac{z-w}{2} \right) dw \quad (75)$$

Αν θεωρήσουμε ότι οι X και Y είναι ανεξάρτητες τότε θα έχουμε $f_{XY}(x, y) = f_X(x)f_Y(y)$ οπότε η παραπάνω σχέση θα γραφτεί:

$$f_Z(z) = \frac{1}{2} \int_{-\infty}^{+\infty} f_X \left(\frac{z+w}{2} \right) f_Y \left(\frac{z-w}{2} \right) dw \quad (76)$$

Ας θέσουμε στην παραπάνω σχέση $w' = (z+w)/2$ τότε θα έχουμε $dw = 2dw'$ και $(z-w)/2 = z - w'$. Οπότε το ολοκλήρωμα θα γραφτεί και ως εξής:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(w') f_Y(z - w') dw' = f_X(z) \otimes f_Y(z) \quad (77)$$

Δείξαμε δηλαδή ότι το άθροισμα Z δύο ανεξάρτητων τυχαίων μεταβλητών X, Y έχει PDF $f_Z(z)$ που προκύπτει από τον συγκερασμό των επιμέρους PDF f_X και f_Y .

Υπάρχει μία ακόμα χρήσιμη ιδιότητα του αθροίσματος δύο τυχαίων κανονικών μεταβλητών. Αν X, Y δύο τυχαίες μεταβλητές με μέσες τιμές μ_X και μ_Y αντίστοιχα τότε το άθροισμα τους $Z = X + Y$ είναι και αυτό κατανομή με μέση τιμή μ_Z που δίνονται αντίστοιχα από τις σχέσεις:

$$\mu_Z = \mu_X + \mu_Y \quad (78)$$

Για να αποδείξουμε τις παραπάνω σχέσεις χρησιμοποιούμε το γεγονός ότι σύμφωνα με την (29):

$$\begin{aligned} \mu_Z = \mathbb{E}\{X + Y\} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) f_{XY}(x, y) dx dy = \\ &= \int_{-\infty}^{+\infty} x \left\{ \int_{-\infty}^{+\infty} f_{XY}(x, y) dy \right\} dx + \int_{-\infty}^{+\infty} y \left\{ \int_{-\infty}^{+\infty} f_{XY}(x, y) dx \right\} dy \end{aligned} \quad (79)$$

Σύμφωνα με την (28) θα έχουμε

$$\int_{-\infty}^{+\infty} f_{XY}(x, y) dy = f_Y(y) \quad (80)$$

$$\int_{-\infty}^{+\infty} f_{XY}(x, y) dx = f_X(x) \quad (81)$$

οπότε

$$\mu_Z = \mathbb{E}\{X + Y\} = \int_{-\infty}^{+\infty} x f_X(x) dx + \int_{-\infty}^{+\infty} y f_Y(y) dy = \mu_X + \mu_Y \quad (82)$$

Επομένως η αναμενόμενη τιμή του αθροίσματος δύο τυχαίων μεταβλητών είναι το άθροισμα των επιμέρους αναμενόμενων τιμών. Μπορούμε γενικότερα να δείξουμε ότι για οποιεσδήποτε σταθερές c_1 και c_2 θα ισχύει:

$$\mathbb{E}\{c_1 X + c_2 Y\} = c_1 \mathbb{E}\{X\} + c_2 \mathbb{E}\{Y\} \quad (83)$$

δηλαδή το $\mathbb{E}$ είναι γραμμικό.

Αν οι X και Y είναι ανεξάρτητες μπορούμε να δείξουμε ότι το άθροισμα τους έχει διακύμανση σ_Z^2 που είναι το άθροισμα των επιμέρους διακυμάνσεων σ^2 και σ_Y^2 , δηλαδή:

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \quad (84)$$

Για να το δείξουμε αυτό ξεκινάμε από τον ορισμό του σ_Z^2 ,

$$\sigma_Z^2 = \mathbb{E}\{(Z - \mu_Z)^2\} = \mathbb{E}\{(X - \mu_X + Y - \mu_Y)^2\} = \mathbb{E}\{(X - \mu_X)^2 + (Y - \mu_Y)^2 + 2(X - \mu_X)(Y - \mu_Y)\} \quad (85)$$

Εφόσον όμως το $\mathbb{E}$ είναι γραμμικό θα έχουμε:

$$\sigma_Z^2 = \mathbb{E}\{(X - \mu_X)^2\} + \mathbb{E}\{(Y - \mu_Y)^2\} + 2\mathbb{E}\{(X - \mu_X)(Y - \mu_Y)\} \quad (86)$$

Οι μεταβλητές X και Y είναι ανεξάρτητες άρα:

$$\mathbb{E}\{(X - \mu_X)(Y - \mu_Y)\} = \mathbb{E}\{(X - \mu_X)\} \mathbb{E}\{(Y - \mu_Y)\} = 0 \quad (87)$$

Τελικά το σ_Z^2 θα δίνεται από την σχέση:

$$\sigma_Z^2 = \mathbb{E}\{(X - \mu_X)^2\} + \mathbb{E}\{(Y - \mu_Y)^2\} = \sigma_X^2 + \sigma_Y^2 \quad (88)$$

Επομένως η διακύμανση του αθροίσματος δύο τυχαίων μεταβλητών X και Y ισούται με το άθροισμα των διακυμάνσεων τους.

Στο [listing 5](#) δείχνουμε ένα παράδειγμα υπολογισμού της [PDF](#) του αθροίσματος δύο τυχαίων ανεξάρτητων μεταβλητών X και Y οι οποίες είναι ομοιόμορφες στο διάστημα $[0, 1]$,

$$f_X(u) = f_Y(u) = \begin{cases} 1 & , 0 \leq u \leq 1 \\ 0 & , \text{διαφορετικά} \end{cases} \quad (89)$$

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3
4
5
6 plt.rcParams.update({
7     "text.usetex": True,
8     "font.family": "serif",
9     "font.serif": ["Palatino"],
10    "font.size" : 12,
11    "lines.linewidth" : 2,
12 })
13
14 def uniform_pdf(y, a, b):
15     pdf = np.zeros(y.size)
16     i = np.where ( (y <= b ) & ( y>= a ) )
17     pdf[i] = 1
18     return pdf
19
20 def trig_pdf(z):
21     pdf = np.zeros(z.size)
```

```

22     i = np.where ( (z>=0) & ( z<=1 ) )
23     pdf[i] = z[i]
24     i = np.where ( (z>=1) & ( z<=2 ) )
25     pdf[i] = 2-z[i]
26     return pdf
27
28
29 def calc_pdfcdf(x, y):
30     N = x.size
31     Dy = y[1] - y[0]
32     pdf = np.zeros( y.size )
33     cdf = np.zeros( y.size )
34
35     for i, yy in enumerate(y):
36         Ny = np.sum( (x >= yy-Dy/2) & ( x<yy+Dy/2 ) )
37         pdf[i] = Ny / N / Dy
38         cdf[i] = np.sum( x <= yy ) / N
39
40     return pdf, cdf
41
42 Nsamples = 100000
43 X = np.random.rand(Nsamples)
44 Y = np.random.rand(Nsamples)
45 Z = X + Y
46
47 yy = np.linspace(0, 1, 50)
48 pdfX, cdfX = calc_pdfcdf(X, yy)
49 pdfY, cdfY = calc_pdfcdf(Y, yy)
50
51 zz = np.linspace(-1, 3, 50)
52 pdfZ, cdfZ = calc_pdfcdf(Z, zz)
53
54
55 pdfZ2 = trig_pdf(zz)
56 plt.close('all')
57 plt.figure()
58 plt.plot(yy, pdfX, '-', label = '$f_X$')
59 plt.plot(yy, pdfY, '--', label = '$f_Y$')
60 plt.xlabel('$x$')
61 plt.legend()
62
63 plt.figure()
64 plt.plot(zz, pdfZ, 's', label = 'Monte-Carlo')
65 plt.plot(zz, pdfZ2, '-', label = 'Theoretical')
66 plt.xlabel('$z$')
67 plt.ylabel('$f_Z(z)$')
68
69 plt.legend()

```

Listing 5: sums.py

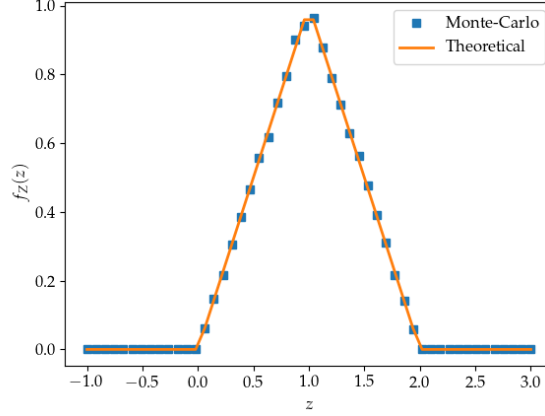
Σύμφωνα με την (77) η PDF του αθροίσματος προκύπτει από τον συγκερασμό των επιμέρους PDF,

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(u)f_Y(z-u)du = \int_0^1 f_Y(z-u)du \quad (90)$$

Σύμφωνα με την παραπάνω σχέση ο συγκερασμός είναι στην ουσία το εμβαδό της $f_Y(z-u)$ η οποία είναι ίση με 1 για $0 \leq z-u \leq 1$ δηλαδή για $z-1 \leq u \leq z$ και μηδέν διαφορετικά. Την ακριβή απόδειξη την αφήνουμε για σπαζοκεφαλιά αλλά προκύπτει ότι το εμβαδόν ισούται με το μήκος του διαστήματος όπου τα $z-1 \leq u \leq z$ και $0 \leq u \leq 1$ συμπίπτουν. Αν $z < 0$ ή $z > 2$ δεν συμπίπτουν πουθενά οπότε το εμβαδόν είναι μηδέν. Αν $0 \leq z < 1$ τότε το μήκος θα είναι z ενώ εάν $1 \leq z < 2$ θα είναι $2-z$. Οπότε θα έχουμε

$$f_Z(z) = \begin{cases} z & , 0 \leq z < 1 \\ 2-z & , 1 \leq z < 2 \\ 0 & , \text{διαφορετικά} \end{cases} \quad (91)$$

Στην εικόνα 4 δείχνουμε το αποτέλεσμα του listing. Παρατηρούμε ότι η (91) ταιριάζει πολύ καλά με την PDF που υπολογίζεται μέσω της προσομοίωσης Monte Carlo.



Εικόνα 4: Επιβεβαίωση της (91) με προσομοίωση Monte Carlo

5.5 Άθροισμα κανονικών τυχαίων μεταβλητών

Η (74) έχει μία ενδιαφέρουσα συνέπεια όταν οι X και Y είναι Gaussian οπότε ισχύει η (37). Ας περιοριστούμε αρχικά στην περίπτωση όπου οι δύο μεταβλητές έχουν μηδενική μέση τιμή $\mu_X = \mu_Y = 0$ και διακύμανση ίση με 1, $\sigma_X^2 = \sigma_Y^2 = 1$. Προκύπτει ότι η συσχέτιση τους θα είναι:

$$\rho = \frac{\mathbb{E}\{(X - \mu_X)(Y - \mu_Y)\}}{\sigma_X \sigma_Y} = \mathbb{E}\{XY\} \quad (92)$$

ενώ θα έχουμε:

$$f_{XY}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)}\right) \quad (93)$$

Χρησιμοποιώντας την (74) και κάνοντας τις πράξεις:

$$f_{ZW}(z, w) = \frac{1}{4\pi\sqrt{1-\rho^2}} \exp\left(-\frac{z^2}{4(1+\rho)} - \frac{w^2}{4(1-\rho)}\right) \quad (94)$$

Η μέση τιμή του Z θα είναι μηδέν επειδή $\mu_Z = \mathbb{E}\{X + Y\} = \mu_X + \mu_Y = 0$. Ομοίως μπορούμε να δείξουμε ότι $\mu_W = 0$. Επίσης η τυπική απόκλιση του Z θα είναι

$$\sigma_Z^2 = \mathbb{E}\{Z^2\} = \mathbb{E}\{X^2\} + \mathbb{E}\{Y^2\} + 2\mathbb{E}\{XY\} = 2 + 2\rho \quad (95)$$

ενώ του W θα είναι:

$$\sigma_W^2 = \mathbb{E}\{W^2\} = \mathbb{E}\{X^2\} + \mathbb{E}\{Y^2\} - 2\mathbb{E}\{XY\} = 2 - 2\rho \quad (96)$$

οπότε η (94) γράφεται:

$$f_{ZW}(z, w) = \frac{1}{2\pi\sigma_Z\sigma_W} \exp\left(-\frac{z^2}{2\sigma_Z^2} - \frac{w^2}{2\sigma_W^2}\right) = \frac{1}{\sigma_Z\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma_Z^2}\right) \times \frac{1}{\sigma_W\sqrt{2\pi}} \exp\left(-\frac{w^2}{2\sigma_W^2}\right) \quad (97)$$

Η παραπάνω σχέση υποδηλώνει ότι οι Z και W είναι ανεξάρτητες κανονικές τυχαίες μεταβλητές. Δείξαμε δηλαδή ότι το άθροισμα και η διαφορά δύο κανονικών τυχαίων μεταβλητών με μέση τιμή 0 και διακύμανση 1 είναι πάλι κανονικές τυχαίες μεταβλητές.

Στη γενικότερη περίπτωση όπου οι X και Y έχουν μέσες τιμές μ_X και μ_Y αντίστοιχα και διακυμάνσεις σ_X^2 και σ_Y^2 αντίστοιχα τότε θεωρούμε τις μεταβλητές $\bar{X}$ και $\bar{Y}$ που συνδέονται με τις αρχικές βάσει των σχέσεων

$$X = \sigma_X \bar{X} + \mu_X \quad (98)$$

$$Y = \sigma_Y \bar{Y} + \mu_Y \quad (99)$$

Είδαμε παραπάνω πως οι $\bar{X}$ και $\bar{Y}$ θα πρέπει να έχουν μέση τιμή μηδέν και διακύμανση 1. Το άθροισμα $Z = X + Y$ και η διαφορά $Z = X - Y$ γράφονται:

$$Z = X + Y = \sigma_X \bar{X} + \sigma_Y \bar{Y} + \mu_X + \mu_Y \quad (100a)$$

$$W = X + Y = \sigma_X \bar{X} - \sigma_Y \bar{Y} + \mu_X - \mu_Y \quad (100b)$$

Μπορούμε να δείξουμε ότι και στην περίπτωση αυτή το Z και το W είναι κανονικές τυχαίες μεταβλητές. Η μέση τιμή και η διακύμανση του Z δίνονται από τις σχέσεις:

$$\mu_Z = \mu_X + \mu_Y \quad (101)$$

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \quad (102)$$

ενώ η PDF του Z προκύπτει ως εξής:

$$f_Z(z) = \frac{1}{\sigma_Z \sqrt{2\pi}} \exp\left(-\frac{(z - \mu_Z)^2}{2\sigma_Z^2}\right) \quad (103)$$

Αφήνουμε την απόδειξη για τις σπαζοκεφαλίες. Πρόκειται για μία χρήσιμη ιδιότητα που θα χρησιμοποιήσουμε και παρακάτω.

6 Ώρα για σπαζοκεφαλίες

1. Αν το X ακολουθεί την διωνυμική κατανομή, δηλαδή

$$\Pr\{X = q\} = \binom{Q}{q} p^q (1-p)^{Q-q} \quad (104)$$

να δείξετε ότι η μέση τιμή του X είναι

$$\mathbb{E}\{X\} = np \quad (105)$$

2. Αποδείξτε την (91) με χαρτί και με μολύβι.
3. Γράψτε ένα Python script που να υπολογίζει αριθμητικά την PDF του αθροίσματος $X + Y$ δύο ανεξάρτητων κανονικών μεταβλητών X και Y και συγκρίνεται την με την (103)
4. Με χαρτί και μολύβι υπολογίστε την συνδυασμένη PDF των Z και W όπως δίνονται από την (100) και δείξτε ότι η PDF του Z καθορίζεται από την (103).