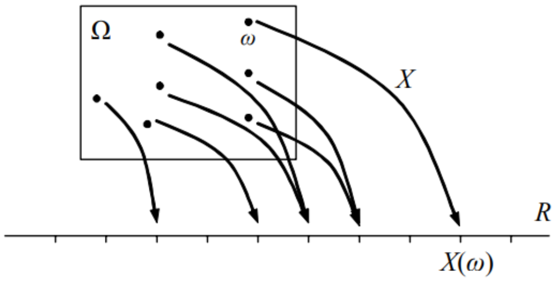
Τυχαία Μεταβλητή (Random variable-variable aléatoire)

Σε πολλούς τύπους πειραμάτων τα αποτελέσματα είναι από τη φύση τους πραγματικοί αριθμοί. Παραδείγματα τέτοιων πειραμάτων αποτελούν οι μετρήσεις των υψών και των βαρών των ατόμων, η παρακολούθηση των τιμών της αγοράς και της ζήτησης ενός προϊόντος, η μέτρηση του αριθμού των χιλιομέτρων που διανύει ένα αυτοκίνητο με 100 λίτρα βενζίνης, η διαφορά πόντων σε ένα παιχνίδι μπάσκετ κ.λ.π. Οι δειγματικοί χώροι αυτών των πειραμάτων είναι υποσύνολα του συνόλου των πραγματικών αριθμών.

Υπάρχουν όμως πειράματα των οποίων τα αποτελέσματα δεν είναι πραγματικοί αριθμοί, με συνέπεια ο δειγματικός τους χώρος να μην είναι υποσύνολο του συνόλου των πραγματικών αριθμών. Παραδείγματα τέτοιων πειραμάτων αποτελούν το στρίψιμο νομίσματος με δυνατά αποτελέσματα κεφάλι ή γράμματα, η διαπίστωση της ποιότητας ενός προϊόντος που επιλέγουμε τυχαία από κάποια γραμμή παραγωγής, η διαδικασία εντόπισης των ελαττωματικών εξαρτημάτων μιας μηχανής, η βλάβη των οποίων προκαλεί βλάβη της μηχανής κ.λ.π. Τα αποτελέσματα αυτών των πειραμάτων δεν είναι πραγματικοί αριθμοί και συνεπώς οι δειγματικοί τους χώροι δεν είναι υποσύνολα του **R**.

Αποδεικνύεται όπως θα δούμε στην πράξη ότι η μετατροπή αυτών των δειγματικών χώρων σε δειγματικούς χώρους με στοιχεία πραγματικούς αριθμούς είναι εξαιρετικά χρήσιμη και επιτυγχάνεται με την αντιστοίχιση ενός πραγματικού αριθμού σε κάθε αποτέλεσμα του αρχικού δειγματικού χώρου. Σε αυτή την περίπτωση η εφαρμογή των κανόνων υπολογισμού πιθανοτήτων είναι ευκολότερη. Η συνάρτηση αυτή λέγεται **τυχαία μεταβλητή (τ.μ)**. Μια λοιπόν τυχαία μεταβλητή είναι ένας κανόνας (μια συνάρτηση) που αναθέτει (αντιστοιχίζει) έναν αριθμό σε κάθε ενδεχόμενο του πειράματος.

Ορισμός: Έστω δ.χ Ω. Ορίζουμε ως **τυχαία μεταβλητή** (τ.μ) X μια συνάρτηση από το Ω στο σύνολο των πραγματικών αριθμών **R**: $X: \Omega \to \mathbf{R}$. Οι τ.μ. συμβολίζονται συνήθως με κεφαλαία γράμματα και οι τιμές τους με τα αντίστοιχα μικρά.



Παράδειγμα: Ρίχνουμε ένα νόμισμα 2 φορές. Ορίζουμε την τυχαία μεταβλητή X την συνάρτηση η οποία σε κάθε ενδεχόμενο ω αντιστοιχεί το πλήθος των κεφαλών. Έτσι προκύπτει ο παρακάτω πίνακας:

Ω	X(ω)	P(ω)
ΚΚ	2	1/4
ΓΓ	0	1/4
ΚΓ	1	1/4
ΓΚ	1	1/4

Πχ: P(ΚΓ)=1/4, ενώ P(X=1)=P(ΚΓ ή ΓΚ)= P(ΚΓ)+P(ΓΚ)=1/2.

Διακριτές και συνεχείς τ.μ.

Στη θεωρία, αλλά και στην πράξη, εμφανίζονται συχνά καταστάσεις στις οποίες οι τυχαίες μεταβλητές που πρέπει να χρησιμοποιήσουμε είναι άλλοτε **συνεχείς** και άλλοτε **διακριτές.** Στην πρώτη περίπτωση οι τυχαίες μεταβλητές μπορούν να πάρουν οποιαδήποτε τιμή σε ένα δεδομένο διάστημα (η τιμή αυτή εξαρτάται προφανώς από το αποτέλεσμα του τυχαίου πειράματος).

Παραδείγματα συνεχών τ.μ.:

Α) Έστω Τ η τυχαία μεταβλητή, που παριστάνει το χρόνο διάσπασης ενός ραδιενεργού σωματιδίου.

Β) Υ η τ.μ. που παριστάνει το χρόνο ζωής ενός ηλεκτρικού λαμπτήρα.

Γ) W η τυχαία μεταβλητή που παριστάνει το ύψος ενός ατόμου από ένα δεδομένο δείγμα ατόμων.

Δ) Τ ο χρόνος (σε δευτερόλεπτα) που απαιτείται για να ανοίξει ο υπολογιστής ένα πρόγραμμα.

Ε) Χ ο χρόνος που χρειάζεται για ένα κοτόπουλο ώστε να φθάνει τα 1.5kg.

ΣΤ) Ρ το ποσό της βροχής που πέφτει σε μια πόλη κάποιο συγκεκριμένο μήνα κ.ο.κ.

Στις διακριτές μεταβλητές οι τιμές είναι διακεκριμένες (πεπερασμένες ή αριθμήσιμες).

Παραδείγματα διακριτών τ.μ.:

Α) Το πλήθος των επιτυχιών σε μια ακολουθία δοκιμών (π.χ. τη ρίψη ενός νομίσματος) με δύο δυνατά αποτελέσματα.

Β) Το πλήθος των πελατών ενός καταστήματος μια προσεχή συγκεκριμένη μέρα.

<u>*Ακολουθεί λεπτομερής μελέτη κάποιων παραδειγμάτων:*</u>

1) Ω είναι το σύνολο των περιπτώσεων μιας οικογένειας με τρία παιδιά:

Ω={ααα, αακ, ακα, ακκ, καα, κακ, κκα, κκκ} όπου για παράδειγμα ακα σημαίνει ότι η οικογένεια έχει δύο αγόρια και ένα κορίτσι και το κορίτσι γεννήθηκε δεύτερο στη σειρά.

Ορίζουμε ως τ.μ. Χ το πλήθος των αγοριών. Άρα Χ(ααα)=3, Χ(αακ)=2, Χ(ακα)=2, Χ(ακκ)=1, Χ(καα)=2, Χ(κακ)=1, Χ(κκα)=1, Χ(κκκ)=0.

Προφανώς $P(X=0)=P(\{κκκ\})=\frac{1}{8}$, $P(X=1)=P(\{ακκ, κακ, κκα\})=\frac{3}{8}$, $P(X=2)=P(\{αακ, ακα, καα\})=\frac{3}{8}$, $P(X=3)=\frac{1}{8}$.

Παρατηρούμε ότι $P(X=0)+P(X=1)+P(X=2)+P(X=3)=1$.

Αν έχουμε το ενδεχόμενο Α={τουλάχιστον ένα αγόρι} τότε $P(A)=P(1\leq X\leq 3)=\frac{3}{8}+\frac{3}{8}+\frac{1}{8}=\frac{7}{8}$.

2) Έστω Ω το σύνολο των ενδείξεων κατά τη ρήψη ενός «τίμιου» νομίσματος, δηλ. Ω={Κορώνα (Κ), Γράμματα (Γ)}. Έστω $X: \Omega \to R$:Χ(Κ)=1 και Χ(Γ)=0. Ισχύει $P(X=0)+P(X=1)=\frac{1}{2}+\frac{1}{2}=1$.

3) Έστω Ω το σύνολο των ενδείξεων κατά τη ρήψη ενός «τίμιου» ζαριού. Δηλαδή Ω = {1, 2, 3, 4, 5, 6}. Ορίζουμε την τ.μ. X τέτοια ώστε X(ω)=ω, όπου ω∈Ω. Τότε $P(X=1) = P(X=2) = P(X=3)= P(X=4) = P(X=5) = P(X=6) = \frac{1}{6}$.

4) Ρίχνουμε ένα «τίμιο» νόμισμα 3 φορές. Τότε Ω = {ΚΚΚ, ΚΚΓ, ΚΓΚ, ΚΓΓ, ΓΓΓ, ΓΚΓ, ΓΓΚ, ΓΚΚ}.

Για κάθε εμφάνιση γράμματος σε μία ρίψη κερδίζουμε 1€ ενώ για κάθε κεφάλι χάνουμε 1€. Έστω η τ.μ. X: το συνολικό κέρδος μετά από κάθε ρίψη. Ισχύει ο παρακάτω πίνακας:

ω	ΓΓΓ	ΓΓΚ	ΓΚΓ	ΓΚΚ	ΚΓΓ	ΚΓΚ	ΚΚΓ	ΚΚΚ
x	3	1	1	-1	1	-1	-1	-3
P(X=x)	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

Επομένως $P(X=1)=3/8$, $P(-1\leq X\leq 1)=6/8$.

5) Ρίχνουμε ένα «τίμιο» νόμισμα δύο φορές. Ορίζουμε 2 τ.μ.:

X: ο αριθμός των κορωνών που εμφανίζονται,

Y: ο αριθμός των κορωνών μείον τον αριθμό των γραμμάτων.

Ισχύουν οι παρακάτω πίνακες:

ω	ΚΚ	ΚΓ	ΓΚ	ΓΓ
X	2	1	1	0
P(x)	1/4	1/4	1/4	1/4

ω	ΚΚ	ΚΓ	ΓΚ	ΓΓ
Y	2	0	0	-2
P(y)	1/4	1/4	1/4	1/4

6) Παρατηρούμε δύο ειδών βακτήρια και καταγράφουμε τη διάρκεια της ζωής τους. Ο δ.χ. αποτελείται από διατεταγμένα ζεύγη (t_1,t_2) όπου $t_1,t_2\in[0,+\infty)$ (t_1,t_2 σε min). Έστω X συνεχής τ.μ. που δηλώνει τη μικρότερη από τις δύο ζωές, π.χ. X(3.1, 7.6)=3.1. Δηλαδή $X(\omega_1,\omega_2)=\min\{\omega_1,\omega_2\}$.

Συνοψίζοντας, παρατηρούμε ότι κατά τη μελέτη ενός πειράματος τύχης ενδιαφερόμαστε (συνήθως) για κάποια συνάρτηση του αποτελέσματος και όχι για το αποτέλεσμα αυτό καθ’ εαυτό.

Χαρακτηριστικά Τυχαίων Μεταβλητών

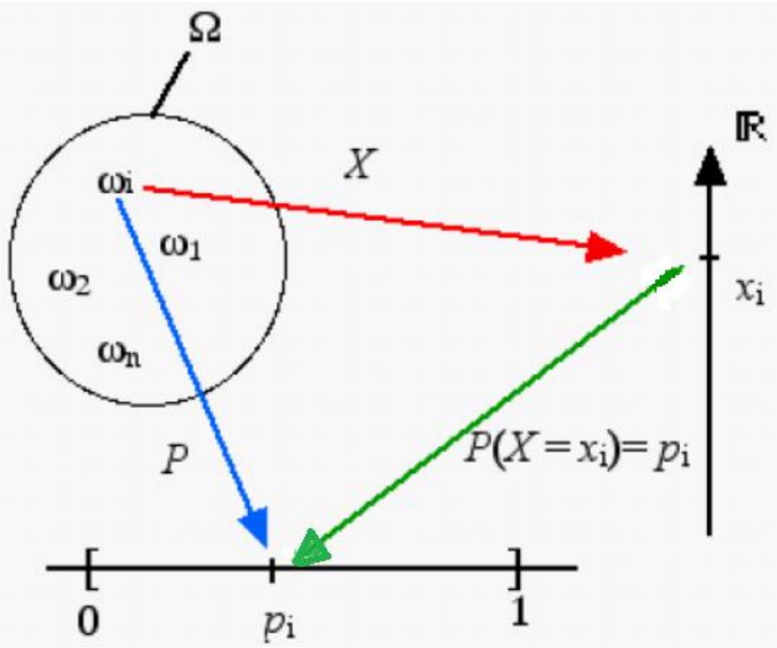
Συνάρτηση πιθανότητας και Συνάρτηση πυκνότητας πιθανότητας

Α) Έστω *Χ διακριτή τ.μ*.

Έχουμε την εξής συνάρτηση:

$$\boldsymbol{\omega_i \to x_i \to P(x_i)=p_i,}$$

ή σύντομα την συνάρτηση: $\mathbf{P(x_i)=P(X=x_i)}$.



Η συνάρτηση αυτή λέγεται συνήθως **συνάρτηση πιθανότητας** της τ.μ Χ ή **κατανομή πιθανότητας** της τ.μ. Χ (probability distribution – fonction de distribution)

Παραδείγματα κατανομής πιθανότητας είναι τα προηγούμενα παραδείγματα 1)-5)

Επομένως αν

$$X: \Omega \to \mathbf{R} \text{ με } X(\Omega)=\{x_1, x_2,\ldots,x_\nu\}$$

Η κατανομή πιθανότητας δίνεται από τον παρακάτω πίνακα:

x_i	x_1	x_2	…	x_ν	
p_i	p_1	p_2	…	p_ν	1

Όπου: $0 \le p_i \le 1$ με $p_i=P(x=x_i)$ και $\sum_{i=1}^{\nu} p_i = 1$ **(Ι)**.

Γενικεύοντας έχουμε τον συμβολισμό $P(\alpha \leq X \leq \beta)$ που παριστάνει την πιθανότητα του ενδεχομένου που αποτελείται από όλα τα στοιχειώδη ενδεχόμενα των οποίων η εικόνα μέσω της Χ είναι μεγαλύτερη ή ίση του α και μικρότερη ή ίση του β.

Έτσι για παράδειγμα $P(X \leq \alpha) = 1-P(X>\alpha)$.

Παράδειγμα:

Έστω Χ μια τ.μ. ορισμένη από τον παρακάτω πίνακα:

x_i	1	2	3	4	5
p_i	0,1	0,2	0,4	0,1	0,2

Τότε:

$P(2 \leq X \leq 4) = P(X=2 \text{ ή } X=3 \text{ ή } X=4) = P(X=2)+P(X=3)+P(X=4) = 0{,}2+0{,}4+0{,}1 = 0{,}7.$

$P(X \leq 3) = 1-P(X>3) = 1-P(X=4)-P(X=5) = 1-0{,}1-0{,}2 = 0{,}7.$

Β) Έστω *Χ συνεχής τ.μ.*

Αντίστοιχα για μια συνεχή τ.μ. Χ ορίζεται η **συνάρτηση πυκνότητας πιθανότητας f(x)** (probability density function – densité de probabilité) με ιδιότητες αντίστοιχες της **(Ι)**:

$$f(x) \geq 0 \quad \text{και} \quad \int_{-\infty}^{+\infty} f(x)dx = 1.$$

Ορίζεται επίσης η συνάρτηση $F(x)=P(X\leq x)$ η οποία λέγεται **αθροιστική συνάρτηση κατανομής** της μεταβλητής Χ (comulative distribution function). Ισχύει:

$$P(X\leq x)=F(x) = \int_{-\infty}^{x} f(u)du \quad \text{και} \quad P(a\leq x\leq b) = \int_{a}^{b} f(x)dx\,.$$

Παράδειγμα:

Δίνεται τ.μ. Χ με συνάρτηση πυκνότητας πιθανότητας $f(x)=\frac{3}{2}x-\frac{3}{4}x^2$, $x\in[0,2]$.

Ισχύει $\int_0^2 \left(\frac{3}{2}x-\frac{3}{4}x^2\right)dx = \left[\frac{3}{4}x^2-\frac{1}{4}x^3\right]_0^2 = 1$. Επίσης

$$P(1<x<\frac{3}{2}) = \int_1^{3/2} f(x)dx = \int_1^{3/2} \left(\frac{3}{2}x-\frac{3}{4}x^2\right)dx = \left[\frac{3}{4}x^2-\frac{1}{4}x^3\right]_1^{3/2} = \frac{11}{32}.$$

Μέση τιμή, διασπορά τ.μ.

Μια συνοπτική περιγραφή της συμπεριφοράς της τ.μ. δίνεται από τη μελέτη συγκεκριμένων ποσοτήτων που προκύπτουν από τη συνάρτηση πιθανότητας και λέγονται **παράμετροι** της κατανομής.

Οι παράμετροι που θα μελετηθούν είναι η **μέση τιμή** και η **διασπορά**.

Ορισμός: Έστω X διακριτή τυχαία μεταβλητή που παίρνει τις τιμές x_i με αντίστοιχες πιθανότητες p_i. **Μέση τιμή** της X ή **αναμενόμενη τιμή της** X ή **μαθηματική ελπίδα** της X ονομάζεται το άθροισμα:

$$\mathbf{E(X) = \sum_i x_i p_i} \,.$$

Αν π.χ οι τιμές της μεταβλητής είναι οι $x_1, x_2,...,x_9$ με αντίστοιχες πιθανότητες $p_1, p_2,...,p_9$ τότε η μέση τιμή $E(X) = \sum_{i=1}^{9} x_i p_i = x_1p_1+p_2x_2+...+x_9p_9$.

Για X συνεχή τυχαία μεταβλητή με συνάρτηση πυκνότητος f(x) η μέση τιμή δίνεται στη μορφή ολοκληρώματος:

$$\mathbf{E(X) = \int_{-\infty}^{+\infty} xf(x)dx} \,.$$

Η μέση τιμή δίνει μια ένδειξη για τη θέση γύρω από την οποία είναι τοποθετημένες οι τιμές της τ.μ. X. Εκτός από το συμβολισμό E(X) χρησιμοποιείται και το μ.

Ορισμός: Έστω X διακριτή τυχαία μεταβλητή που παίρνει τις τιμές x_i με αντίστοιχες πιθανότητες p_i. Ονομάζουμε **διασπορά** της τ.μ X την ποσότητα:

$$\mathbf{Var(X) = \sum_i p_i x^2 - \left(\sum_i p_i x_i\right)^2}$$

Εκτός από το Var(X) χρησιμοποιείται και το σύμβολο σ^2.

Ονομάζουμε **τυπική απόκλιση** της διακριτής τ.μ X την ποσότητα $\sigma = \sqrt{Var(X)}$.

Εφαρμογή:

Για την διακριτή τ.μ. X με συνάρτηση πιθανότητας $P(X=x)=\frac{1}{30}x^2$ και x=1,2,3,4 να βρεθούν η μέση τιμή και η τυπική απόκλιση.

Λύση:

Αφού $P(X=1)=\frac{1}{30}$, $P(X=2)=\frac{4}{30}$, $P(X=3)=\frac{9}{30}$ και $P(X=4)=\frac{16}{30}$, ισχύει ο παρακάτω πίνακας:

x_i	p_i	$x_i p_i$	$x_i^2 p_i$
1	1/30	1/30	1/30
2	4/30	8/30	16/30
3	9/30	27/30	81/30
4	16/30	64/30	256/30
Σύνολο	1	10/3	354/30

Επομένως $E(X)=\frac{10}{3}$, $V(X)=\frac{354}{30}-\left(\frac{10}{3}\right)^2=\frac{31}{45}$, άρα $\sigma=\sqrt{\frac{31}{45}}$.

Ορισμός: Έστω <u>X μία τυχαία συνεχής μεταβλητή</u> με μ=E(X). Η ποσότητα

$$Var(X)=\int (x-\mu)^2 f(x)dx$$

λέγεται **διασπορά ή διακύμανση** της τ.μ. X. Η τετραγωνική ρίζα της διασποράς δηλαδή $\sqrt{Var(X)}$ λέγεται **τυπική απόκλιση** της X και συμβολίζεται με σ_X ή απλώς σ.

Η σημασία της τυπικής απόκλισης είναι μεγάλη, επειδή μετράει την διασπορά των τιμών της μεταβλητής γύρω από τη μέση τιμή. Χάρη σε αυτή μπορούμε να διακρίνουμε αν οι τιμές της μεταβλητής απέχουν σημαντικά από τον μέσο όρο. <u>Όσο μικρότερη είναι η τιμή της τυπικής απόκλισης, τόσο ο μέσος όρος αποτελεί αντιπροσωπευτικό στατιστικό μέτρο για την κατανομή της μεταβλητής.</u>