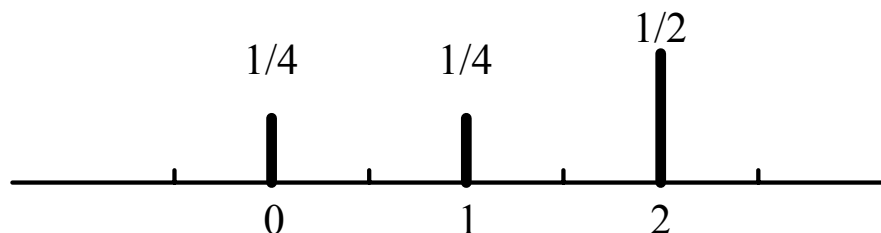

3. Χαρακτηριστικές Παράμετροι Κατανομών

- Αναμενόμενη ή μέση τιμή μιας διακριτής τυχαίας μεταβλητής.

• Θα ήταν αρκετά χρήσιμο να γνωρίζουμε γύρω από ποια τιμή «κυμαίνεται» η τ.μ. X . (γύρω από την οποία «απλώνεται» η κατανομή της τ.μ. X). Ας ονομάσουμε την τιμή αυτή «**μέση τιμή**» της τ.μ. X .

• Ένας ιδανικός ορισμός είναι αυτός που θέτει ως «μέση τιμή» μιας κατανομής το «**κέντρο βάρους**» της κατανομής αυτής (θεωρώντας ότι η συνολική πιθανότητα είναι μάζα που έχει κατανεμηθεί είτε στον R (σε σημεία η σε διάστημα)).

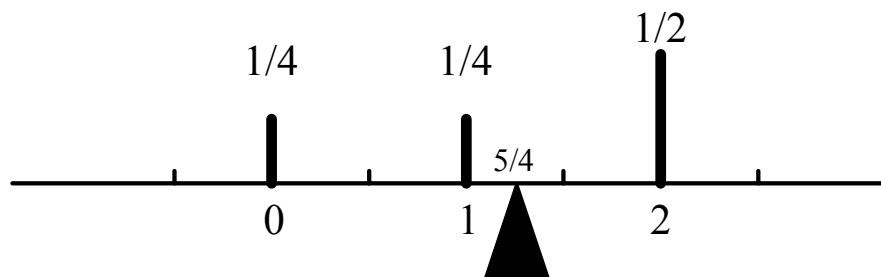
- Για παράδειγμα, αν μία τ.μ. X παίρνει τιμές 0, 1 και 2 με πιθανότητες $1/4$ και $1/4$ και $1/2$ αντίστοιχα, τότε η σ.π. της σχηματικά θα είναι



Θεωρώντας τώρα ότι στα σημείο 0, 1, 2 υπάρχει μάζα (πιθανότητας) μεγέθους $1/4$, $1/4$ και $1/2$ αντίστοιχα, τότε το κέντρο βάρους θα είναι το

$$0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} = \frac{5}{4}.$$

Συνεπώς, θα θεωρούμε ότι η μέση τιμή αυτής της τ.μ. θα είναι το $5/4$:



Έστω X μία διακριτή τ.μ. με τιμές στο $\{a_1, a_2, \dots\}$ και σ.π. f_X . Η ποσότητα

$$E(X) = \sum_{i=1}^{\infty} a_i P(X = a_i) = \sum_{i=1}^{\infty} a_i f_X(a_i) \quad (= \sum_{x \in X(\Omega)} x f_X(x))$$

(υπό τον όρο ότι η παραπάνω σειρά συγκλίνει απόλυτα δηλ. $\sum |a_i| f_X(a_i) < \infty$) θα καλείται **μέση ή αναμενόμενη τιμή** ή πληθυσμιακός μέσος της τ.μ. X .

Αν $X \in \{0, 1, 2, \dots\}$ τότε

$$E(X) = \sum_{x=0}^{\infty} x P(X = x) = \sum_{x=0}^{\infty} x f_X(x).$$

Παράδειγμα. Ένας παίκτης ρουλέτας χρησιμοποιεί το ακόλουθο σύστημα. Ποντάρει στο κόκκινο 1χιλ και αν κερδίσει αποχωρεί. Αν χάσει, ποντάρει ξανά στο κόκκινο αυτή τη φορά 2χιλ και ανεξάρτητα με το αποτέλεσμα αποχωρεί. Θεωρώντας ότι η πιθανότητα να έρθει κόκκινο είναι $\frac{1}{2}$, να βρείτε το μέσο κέρδος του παίκτη.

Ο δειγματικός χώρος του συγκεκριμένου πειράματος μπορεί να θεωρηθεί ότι είναι ο

$$\Omega = \{(K,K), (K,M), (M,K), (M,M)\}$$

Έστω X η τ.μ. που εκφράζει το κέρδος του παίκτη. Θα έχουμε ότι

στοιχ. ενδεχ.	X	
(K,K)	1000	πόνταρε την πρώτη φορά 1000 στο K, κέρδισε και αποχώρησε
(K,M)	1000	πόνταρε την πρώτη φορά 1000 στο K, κέρδισε και αποχώρησε
(M,K)	$-1000+2000 = 1000$	πόνταρε την πρώτη φορά 1000 στο K, έχασε και ξαναπόνταρε 2000 στο K και κέρδισε
(M,M)	$-1000-2000 = -3000$	πόνταρε την πρώτη φορά 1000 στο K, έχασε και ξαναπόνταρε 2000 στο K και ξαναέχασε.

Επομένως, $P(X = 1000) = 3/4$, $P(X = -3000) = 1/4$ και

$$E(X) = \sum_{a=1}^2 a_i P(X = a_i) = 1000P(X = 1000) - 3000P(X = 3000) = 1000 \frac{3}{4} - 3000 \frac{1}{4} = 0.$$

Άσκηση. Να βρεθεί η μέση τιμή της τ.μ. X με σ.π. $f(x) = \frac{2x}{v(v+1)}, x = 1, 2, \dots, v$

Λύση

$$E(Y) = \sum_{i=1}^{\infty} a_i f(a_i) = \sum_{i=1}^v i f(i) = \frac{2}{v(v+1)} \sum_{i=1}^v i^2 = \frac{2}{v(v+1)} \frac{v(v+1)(2v+1)}{6} = \frac{2v+1}{3}$$

Άσκηση Η κατανομή των κερδών σε μία σειρά λαχείων «Ξυστό» είναι:

Κατηγορία κέρδους (δρχ)	Κερδίζοντα λαχεία
300	1.530.000
1.000	183.600
5.000	30.600
10.000	10.200

Κατηγορία κέρδους (δρχ)	Κερδίζοντα λαχεία
30.000	1.867
1.000.000	10
3.000.000	6
30.000.000	4
Σύνολο	1.756.287

Λαχεία έκδοσης: 8.160.000 (300δρχ. η τιμή του κάθε δελτίου)

Αν X είναι το κέρδος από την αγορά ενός τυχαία επιλεγμένου λαχείου,

$$E(X) = -300 \left(1 - \frac{1.756.287}{8.160.000} \right) + 0 \cdot \frac{1.530.000}{8.160.000} + 700 \frac{183.600}{8.160.000} + 4700 \frac{30.600}{8.160.000} \\ + \frac{9.700 \cdot 10.200}{8.160.000} + \frac{29.700 \cdot 1.867}{8.160.000} + \frac{999.700 \cdot 10}{8.160.000} + \frac{2.999.700 \cdot 6}{8.160.000} + \frac{29.999.700 \cdot 4}{8.160.000} = -165.$$

- Σε αρκετές περιπτώσεις, ζητείται ο υπολογισμός της μέσης τιμής μιας τ.μ. $Y = g(X)$ που είναι συνάρτηση του X .

Πρόταση Έστω X μία διακριτή τ.μ. με τιμές στο $\{a_1, a_2, \dots\}$ και σ.π. f_X . Αν $Y = g(X)$, τότε

$$E(Y) = E(g(X)) = \sum_{i=1}^{\infty} g(a_i) f_X(a_i).$$

- Π.χ. αν $X \in \{0, 1, 2, \dots\}$,

$$E(X^2) = \sum_{i=1}^{\infty} i^2 P(X = i) = \sum_{i=1}^{\infty} i^2 f_X(i)$$

Άσκηση 3.3. α) Να βρεθεί η μέση τιμή $E(X)$, όπου η τ.μ. X έχει την κατανομή

x	0	1	2	3
$f(x)$	0.5	0.1	0.3	0.1

β) Ποιά είναι η μέση τιμή των τ.μ. X^2 , $5X + 3$, $\sqrt{X}$;

Λύση. (α) Η μέση τιμή της τ.μ. X θα είναι

$$E(X) = \sum_{i=0}^3 iP(X=i) = 0P(X=0) + 1P(X=1) + 2P(X=2) + 3P(X=3) = 0.1 + 0.6 + 0.3 = 1$$

$$\begin{aligned} E(Y) = E(X^2) &= \sum_{i=0}^3 i^2 P(X=i) = 0^2 P(X=0) + 1^2 P(X=1) + 2^2 P(X=2) + 3^2 P(X=3) \\ &= 0.1 + 1.2 + 0.9 = 2.2 \end{aligned}$$

$$E(Z) = E(5X + 3) = \sum_{i=0}^3 (5i + 3)P(X=i) = \dots = 8$$

$$E(W) = E(\sqrt{X}) = \sum_{i=0}^3 \sqrt{i}P(X=i) = 0.1 + \sqrt{2} \cdot 0.3 + \sqrt{3} \cdot 0.1 = 0.697$$

Μέση τιμή συνεχών τυχαίων μεταβλητών (κατανομών)

- Η θεωρία που αντιστοιχεί στις συνεχείς τ.μ. είναι ανάλογη με τη θεωρία που αφορά διακριτές τ.μ. (όπου στις διακριτές κατανομές εμφανίζεται άθροισμα, στις συνεχείς κατανομές εμφανίζεται ολοκλήρωμα)

Ορισμός Έστω X μία συνεχής τ.μ. με σ.π.π. f_X . Η ποσότητα

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

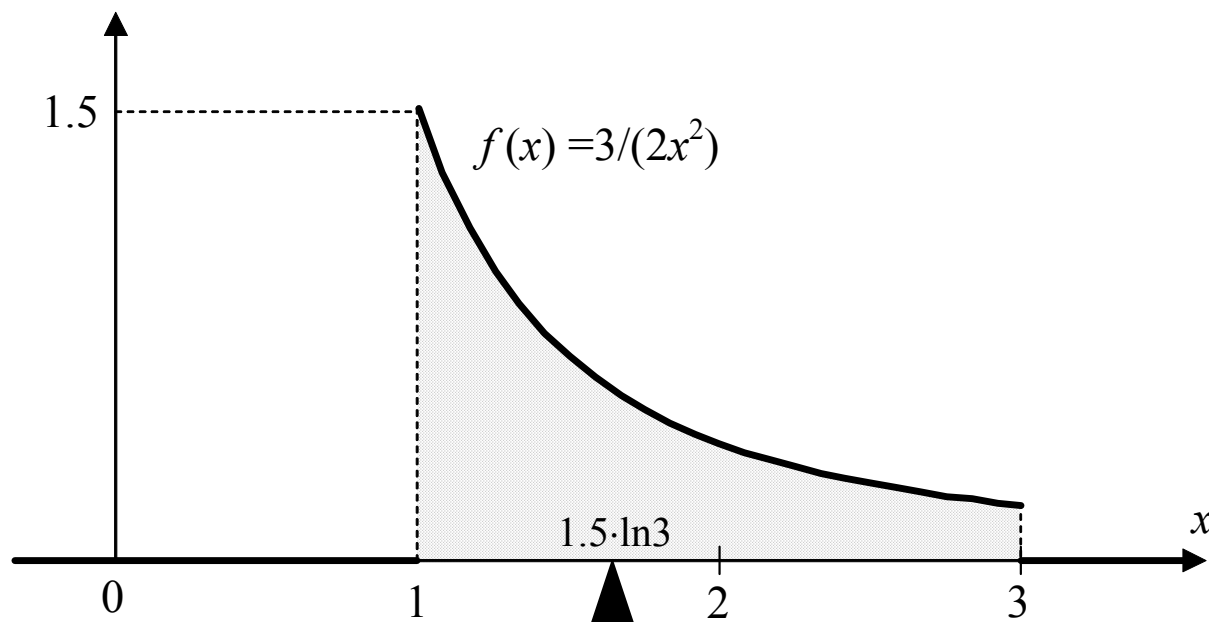
(υπό τον όρο ότι $\int |x| f_X(x) < \infty$) θα καλείται **μέση ή αναμενόμενη τιμή** ή **πληθυσμιακός μέσος της τ.μ. X** .

- Ο παραπάνω ορισμός της μέσης τιμής συμπίπτει και πάλι με τον ορισμό του κέντρου βάρους μιας μάζας (πιθανότητας) κατανεμημένης στον άξονα των πραγματικών με πυκνότητα στο σημείο x ίση με $f_X(x)$, $x \in \mathbb{R}$.

Άσκηση. Να βρεθεί η μέση τιμή της τ.μ. X με σ.π.π. $f(x) = \frac{3}{2x^2}$, $x \in [1,3]$,

Σύμφωνα με τον ορισμό θα είναι

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_1^3 x \frac{3}{2x^2} dx = \frac{3}{2} \int_1^3 \frac{1}{x} dx = \frac{3}{2} (\ln x)_1^3 = \frac{3}{2} (\ln 3 - \ln 1) = \frac{3}{2} \ln 3 \approx 1.65.$$



- Σε αρκετές περιπτώσεις ζητείται ο υπολογισμός της μέσης τιμής μιας τ.μ. $Y = g(X)$

Πρόταση Αν $Y = g(X)$ και X είναι μία συνεχής τ.μ. με σ.π.π. f_X τότε

$$E(Y) = E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx .$$

Π.χ.

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx$$

Γενικότερος Ορισμός Μέσης Τιμής - Ολοκλήρωμα Lebesgue

- Για οποιαδήποτε μη αρνητική τ.μ. X (όχι κατ'ανάγκη διακριτή ή συνεχή) αποδεικνύεται εύκολα ότι μπορεί να κατασκευασθεί μια αύξουσα ακολουθία μη αρνητικών διακριτών τ.μ. $X_1, X_2, \dots$ (λαμβάνουν πεπερασμένο πλήθος τιμών) η οποία συγκλίνει στην X , δηλαδή,

$$X_n(\omega) \rightarrow X(\omega) \text{ για κάθε } \omega$$

Το ολοκλήρωμα *Lebesgue* της μη αρνητικής τ.μ. X ως προς την πιθανότητα P (συμβ. με $\int X dP$) και ορίζεται ως το όριο

$$\int X dP = \lim_{n \rightarrow \infty} \sum_x x P(X_n = x)$$

Το όριο αυτό υπάρχει ($< \infty$ ή ∞ διότι πρόκειται για όριο αύξουσας ακολουθίας), και είναι μοναδικό.

- Ως ολοκλήρωμα *Lebesgue* μιας τ.μ. X με τιμές στο R (ως προς την P) ορίζεται

$$\int X dP = \int X_+ dP - \int X_- dP .$$

(όπου $X_+ = \max\{X, 0\}$ και $X_- = \max\{-X, 0\}$) αρκεί να είναι $< \infty$ τα παραπάνω ολοκληρώματα

- **Μέση τιμή** τώρα μιας τ.μ. X πολύ απλά καλείται το ολοκλήρωμα Lebesgue:

$$E(X) = \int X \, dP.$$

(Ορίζεται όταν $E(X_+), E(X_-) < \infty$ ή ισοδύναμα όταν $E(|X|) < \infty$)

- Αν η X είναι διακριτή τότε αποδεικνύεται ότι

$$E(X) = \int X \, dP = \sum_x xP(X = x)$$

ενώ αν η X είναι συνεχής τότε αποδεικνύεται ότι

$$E(X) = \int X \, dP = \int_{-\infty}^{\infty} xf_X(x)dx$$

Ιδιότητες μέσης τιμής:

Για την μέση τιμή αποδεικνύονται τα ακόλουθα:

(1) $E(a) = a$,

(2) $E(aX + b) = aE(X) + b$

(3) $E(X + Y) = E(X) + E(Y)$.

(4) Αν $X \leq Y$ τότε $E(X) \leq E(Y)$.

(5) (Θεώρημα μονότονης σύγκλισης) Αν $0 \leq X_1 \leq X_2 \leq \dots$ είναι μία ακολουθία μη αρνητικών τ.μ. που συγκλίνει στην τ.μ. X τότε

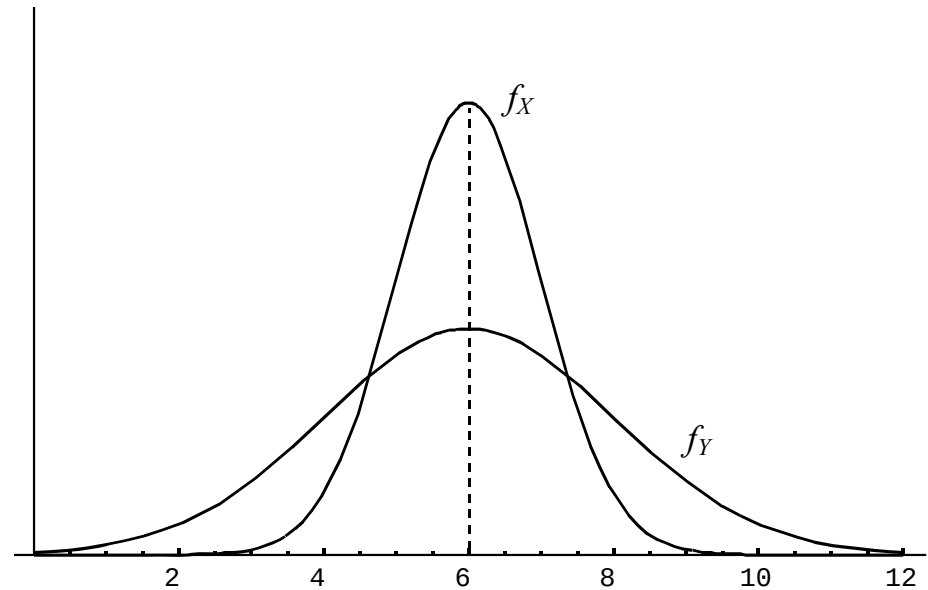
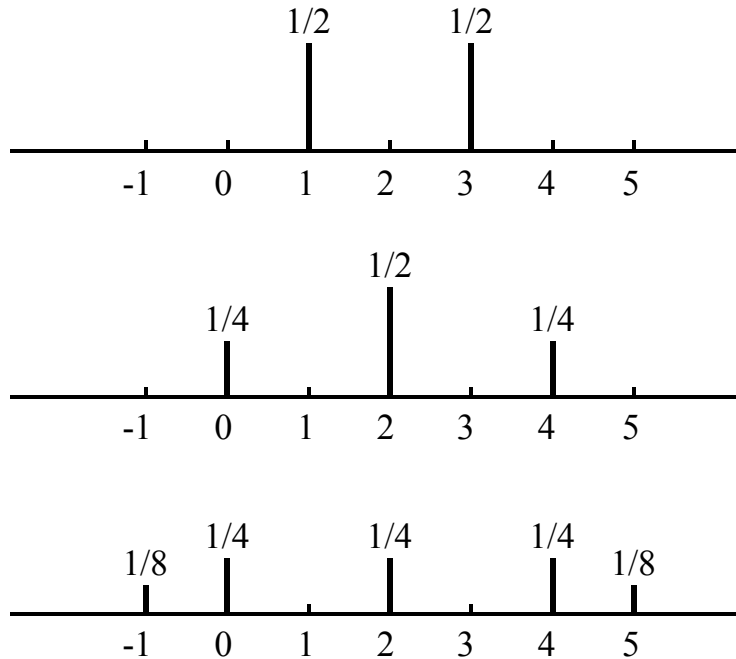
$$E(X) = \lim_{n \rightarrow \infty} E(X_n) \quad (\text{δηλαδή, } E(\lim_{n \rightarrow \infty} X_n) = \lim_{n \rightarrow \infty} E(X_n))$$

(6) (Θεώρημα κυριαρχημένης σύγκλισης) Αν $X_1, X_2, \dots$ είναι μία ακολουθία τ.μ. που συγκλίνει στην τ.μ. X και $|X_n| < Y$ για κάποια τ.μ. Y με $E|Y| < \infty$ τότε

$$E(X) = \lim_{n \rightarrow \infty} E(X_n) \quad (\text{δηλαδή, } E(\lim_{n \rightarrow \infty} X_n) = \lim_{n \rightarrow \infty} E(X_n))$$

Διακύμανση ή διασπορά διακριτής τ.μ.

- Υπάρχουν κατανομές που ενώ έχουν ίσες μέσες τιμές, είναι μεταξύ τους αρκετά διαφορετικές. Π.χ.



- Θα ήταν λοιπόν αρκετά χρήσιμο αν μπορούσαμε να ποσοτικοποιήσουμε τη «διασπορά» αυτή των τιμών της τ.μ. X ορίζοντας κάποια νέα παράμετρο της κατανομής F_X .

Εστω X μία διακριτή τ.μ. Η ποσότητα

$$V(X) = E[(X - E(X))^2] ,$$

καλείται διασπορά (ή διακύμανση) της τ.μ. X (ή της κατανομής της).

- Η διασπορά ορίζεται όταν $E(X) < \infty$.
- Η μέση τιμή και η διασπορά μιας τ.μ. X συμβολίζονται συνήθως και με μ_X, σ_X^2 .
- Η τετραγωνική ρίζα της διασποράς $\sigma_X = \sqrt{V(X)}$ καλείται τυπική απόκλιση της τ.μ. X

Πρόταση (1) $V(X) = E(X^2) - (E(X))^2$.

(2) $V(a) = 0$,

(3) $V(aX + b) = a^2 V(X)$.

Άσκηση Να βρεθεί η διασπορά της κατανομής που περιγράφεται από τον πίνακα

x	0	1	2	3
$f(x)$	0.5	0.1	0.3	0.1

Ποιά είναι η διασπορά των τυχαίων μεταβλητών X^2 , $\sqrt{X}$;

Λύση. Από παραπάνω άσκηση βρήκαμε ότι

$$E(X) = 1, \quad E(X^2) = 2.2, \quad E(5X + 3) = 8, \quad E(\sqrt{X}) \approx 0.697$$

(α) Προκύπτει άμεσα ότι $V(X) = 2.2 - (1)^2 = 1.2$

(β) Επίσης, $V(X^2) = V(Y) = E(Y^2) - E(Y)^2 = E(X^4) - E(X^2)^2$ όπου

$$E(X^4) = \sum_{i=0}^3 i^4 P(X = i) = 0^4 \cdot 0.5 + 1^4 \cdot 0.1 + 2^4 \cdot 0.3 + 3^4 \cdot 0.1 = 13.$$

και άρα $V(X^2) = E(X^4) - E(X^2)^2 = 13 - 2.2^2 = 8.16$

(γ)

$$V(\sqrt{X}) = V(W) = E(W^2) - E(W)^2 = E((\sqrt{X})^2) - E(\sqrt{X})^2 = E(X) - E(\sqrt{X})^2 \approx 0.514$$

Άσκηση Να βρεθούν οι $E(X)$, $V(X)$ της τ.μ. X με σ.π.π.

$$f(x) = \lambda e^{-\lambda x}, x > 0, \lambda > 0$$

Λύση.

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \lambda \int_0^{\infty} xe^{-\lambda x} dx = \frac{1}{\lambda}.$$

Επίσης,

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx = \lambda \int_0^{\infty} x^2 \left(\frac{e^{-\lambda x}}{-\lambda} \right)' dx = \lambda \left(x^2 \frac{e^{-\lambda x}}{-\lambda} \right)_0^{\infty} - \lambda \int_0^{\infty} (x^2)' \frac{e^{-\lambda x}}{-\lambda} dx \\ &= \int_0^{\infty} (x^2)' e^{-\lambda x} dx = \int_0^{\infty} 2xe^{-\lambda x} dx = \frac{2}{\lambda} E(X) = \frac{2}{\lambda^2} \end{aligned}$$

και τελικά,

$$V(X) = E(X^2) - E(X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Τυποποιημένες τ.μ. ή κατανομές

- Μια τ.μ. η οποία έχει μέση τιμή 0 και διασπορά 1 καλείται τυπική κατανομή. Είναι ενδιαφέρουσα η παρατήρηση ότι κάθε τ.μ. X μπορεί να μετασχηματισθεί έτσι ώστε να είναι τυπική.
- Συγκεκριμένα, η μετασχηματισμένη τ.μ.

$$Y = \frac{X - \mu_X}{\sigma_X} = \frac{X - E(X)}{\sqrt{V(X)}}$$

παρατηρούμε ότι έχει

$$E(Y) = E\left(\frac{X - \mu_X}{\sigma_X}\right) = 0, \quad V(Y) = \left(\frac{X - \mu_X}{\sigma_X}\right)^2 = \frac{1}{\sigma_X^2} V(X) = 1$$

- Για το λόγο αυτό η τ.μ.

$$\frac{X - E(X)}{\sqrt{V(X)}}$$

καλείται και τυποποιημένη τ.μ.

Μικτές κατανομές

- Υπάρχουν όμως τ.μ. (αντ. κατανομές) οι οποίες δεν είναι ούτε συνεχείς ούτε διακριτές, αλλά *μικτές*.
- Στις μικτές κατανομές, η συνολική πιθανότητα μπορεί να θεωρηθεί ότι χωρίζεται σε δύο μέρη, ένα διακριτό και ένα συνεχές.
- Το ένα μέρος της συνολικής πιθανότητας κατανέμεται σε *κάποια σημεία* (αριθμήσιμου πλήθους) του $\mathbb{R}$ (διακριτό μέρος) και το υπόλοιπο «απλώνεται» σε *ένα συνεχές διάστημα* του $\mathbb{R}$ (συνεχές μέρος).
- Μία μικτή τ.μ. W θα έχει σ.κ. της μορφής

$$F_W(x) = cF_1(x) + (1-c)F_2(x)$$

όπου F_1 είναι η σ.κ. μιας συνεχούς τ.μ. (έστω W_1),

F_2 είναι η σ.κ. μιας διακριτής τ.μ. (έστω W_2) και $c \in [0,1]$

- Για την μέση τιμή της W ισχύει ότι

$$E(W) = cE(W_1) + (1-c)E(W_2) .$$