ΣΥΝΕΧΕΙΣ ΜΕΤΑΒΛΗΤΕΣ

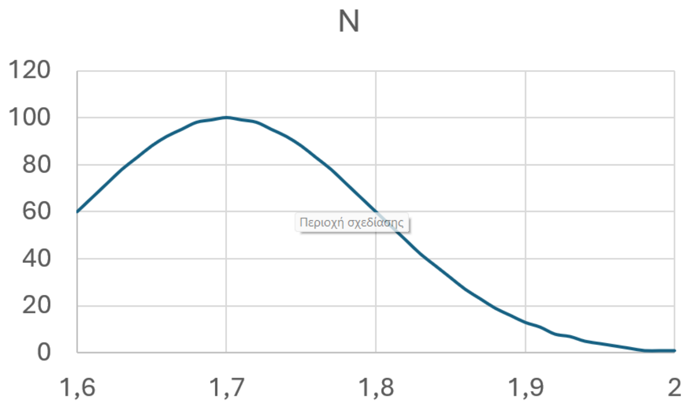
X Παίρνει τιμές από το σύνολο των πραγματικών μεταβλητών κατά συνεχή τρόπο

Παράδειγμα: Το ύψος των ανθρώπων, π.χ. σε ένα δείγμα περίπου $N_{ο\lambda} \approx 2100$ ανθρώπων πάρθηκαν οι εξής μετρήσεις

h ύψος m	N	P = N/Nολ
1,6	60	0,028
1,61	66	0,031
1,62	72	0,034
1,63	78	0,037
1,64	83	0,039
1,65	88	0,042
1,66	92	0,043
1,67	95	0,045
1,68	98	0,046
1,69	99	0,047
1,7	100	0,047
1,71	99	0,047
1,72	98	0,046
1,73	95	0,045
1,74	92	0,043
1,75	88	0,042
1,76	83	0,039
1,77	78	0,037
1,78	72	0,034
1,79	66	0,031
1,8	60	0,028
1,81	54	0,026
1,82	48	0,023
1,83	42	0,020
1,84	37	0,017
1,85	32	0,015
1,86	27	0,013
1,87	23	0,011
1,88	19	0,009
1,89	16	0,008
1,9	13	0,006
1,91	11	0,005
1,92	8	0,004
1,93	7	0,003
1,94	5	0,002
1,95	4	0,002
1,96	3	0,001

1,97	2	0,001
1,98	1	0,000
1,99	1	0,000
2	1	0,000
ΣΥΝΟΛΟ	2116	1

Σε γραφική:

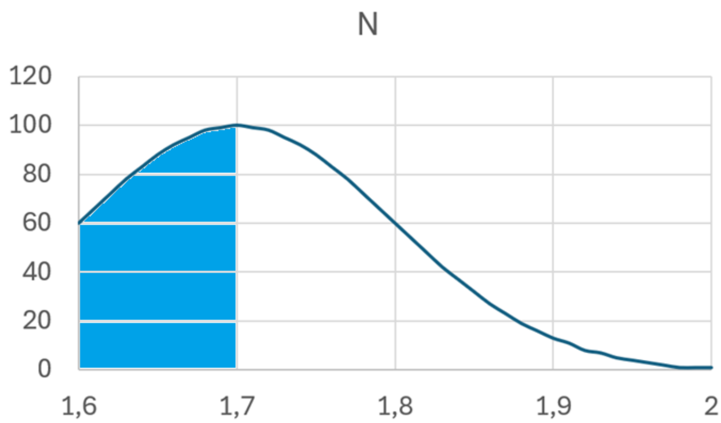


Παρατηρώ ότι η πιθανότητα η οποία είναι ίση με τον αριθμό των μετρήσεων σε κάθε ύψος δια τον ολικό αριθμό $N = 2116$ είναι πολύ μικρή γιατί υπάρχουν πάρα πολλές τιμές από 1,6 έως 2 m

Στις συνεχείς μεταβλητές έχει περισσότερο νόημα να πάρω την πιθανότητα μέσα σε ένα διάστημα, π.χ. πόση είναι η πιθανότητα να έχω ύψος από 1,6 έως 1.7 m

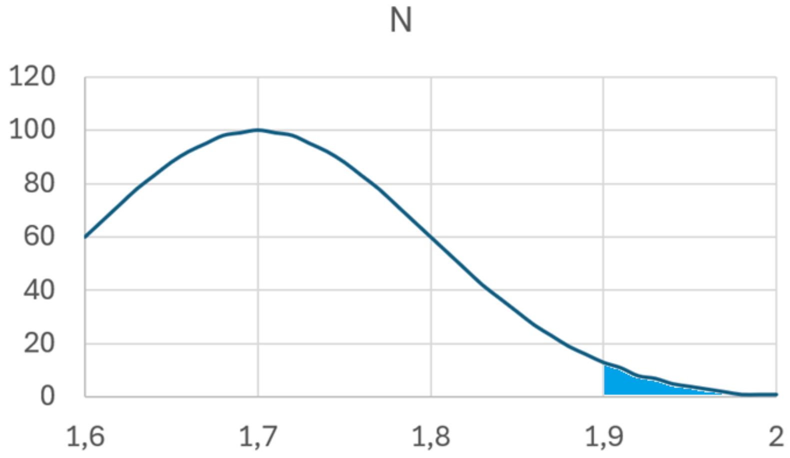
$$P(1.6 < x < 1.7)$$

Η πιθανότητα είναι ανάλογη του εμβαδού κάτω από την καμπύλη που ως γνωστόν είναι ίση με το ολοκλήρωμα της καμπύλης μέσα σε αυτά τα όρια



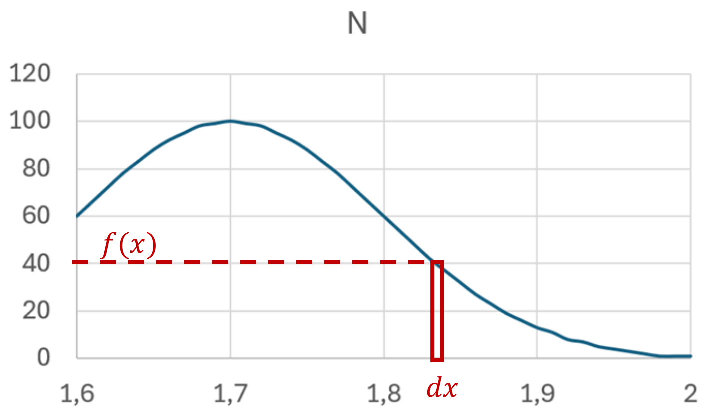
Πχ. για την πιθανότητα να έχω ύψος από 1,9 έως 2.0 m

$$P(1.9 < x < 2.0) < P(1.6 < x < 1.7)$$



Έστω ότι επιλέγω ένα στοιχειώδες ποσό dx της μεταβλητής μου x, όπως στο παρακάτω σχήμα. Η στοιχειώδης πιθανότητα dP να είναι οι τιμές μου μέσα στο εύρος dx είναι ίση με το στοιχειώδες εμβαδό δηλαδή

$$dP = f(x)dx$$



Για πεπερασμένες τιμές μεταξύ x_1 και x_2 τότε θα έχουμε

$$P(x_1 < x < x_2) = \int_{x_1}^{x_2} dP = \int_{x_1}^{x_2} f(x)dx$$

Για να είναι η παραπάνω ποσότητα πιθανότητα πρέπει

$$f(x)dx \geq 0 => f(x) \geq 0$$

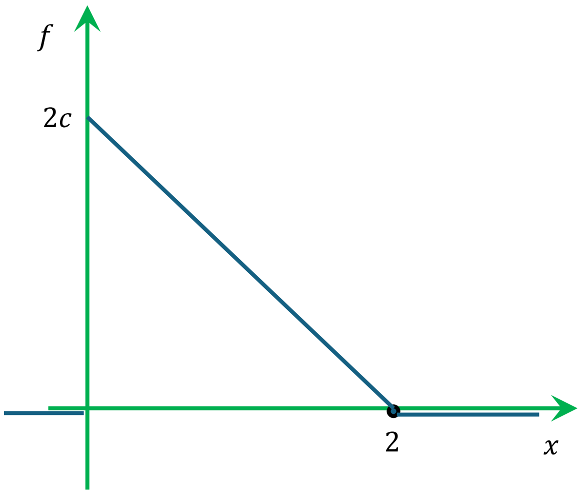
$$P(-\infty < x < \infty) = 1 => \int_{-\infty}^{\infty} f(x)dx = 1$$

Η συνάρτηση $f(x)$ ονομάζεται πυκνότητα πιθανότητας επειδή

$$f = \frac{dP}{dx}$$

<u>Παράδειγμα 4.1</u> του βιβλίου, το X είναι η καθυστέρηση σε ώρες μίας πτήσης και από στατιστικές που κρατάμε η πιθανότητα πέφτει γραμμικά με το X

$$f(x) = c(2 - x)$$



για $0 \leq x \leq 2$ και μηδέν οπουδήποτε αλλού. Να βρεθούν (α) η τιμή της σταθεράς c και (β) η πιθανότητα ή καθυστέρηση της πτήσης να είναι (i) το πολύ 1 ώρα ή (ii) μεταξύ μιας και μιάμιση ώρας

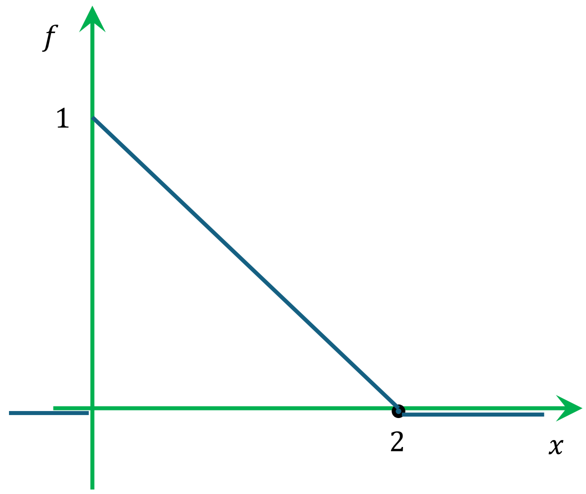
<u>Λύση</u>

(α)

$$\int_{-\infty}^{\infty} f(x)dx = 1 => \int_{0}^{2} f(x)dx = 1$$

$$c\int_{0}^{2} (2 - x)dx = 1$$

Τελικά $c = 1/2$



Η παραπάνω συνάρτηση ικανοποιεί τα κριτήρια της πυκνότητας πιθανότητας, δηλαδή είναι θετική και το ολοκλήρωμα της σε όλο το εύρος είναι 1

(β) (i) το πολύ 1 ώρα

Αναζητάμε την ποσότητα

$$P(0 \leq x \leq 1) = \int_0^1 f(x)dx = \frac{1}{2}\int_0^1 (2-x)dx = 0.75$$

(ii) μεταξύ μιας και μιάμιση ώρας

$$P(1 \leq x \leq 1.5) = \int_1^{1.5} f(x)dx = \frac{1}{2}\int_1^{1.5} (2-x)dx = \frac{3}{16} \approx 18.5\%$$

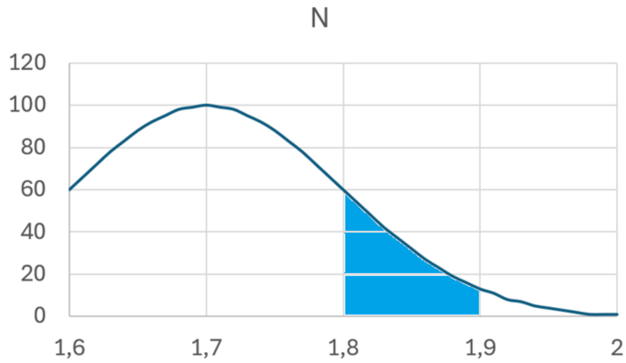
<u>Συνάρτηση αθροιστικής κατανομής ή κατανομής</u>

$$F(x) = \int_{-\infty}^{x} f(\bar{x})d\bar{x}$$

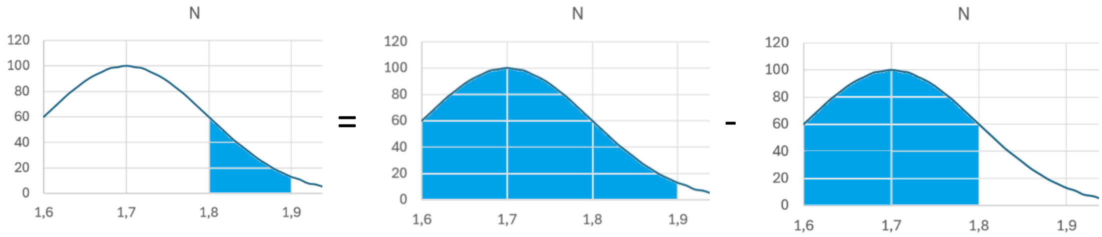
Όπως είδαμε εμείς ενδιαφερόμαστε για πιθανότητες οπότε συνήθως θέλουμε να υπολογίσουμε το εξής

$$P(a \leq x \leq b) = \int_a^b f(x)dx$$

Παράδειγμα ύψος από 1.8 έως 1.9 m



Το παραπάνω εμβαδό είναι ίσο με την διαφορά δυο άλλων εμβαδών που ξεκινούν από την αρχή των τιμών $x = 1.6\ m$ όπως φαίνεται στο παρακάτω σχήμα



Με την ίδια λογική, και μια πιθανότητα που είναι ίση με το ολοκλήρωμα μεταξύ των τιμών a και b, μπορεί εναλλακτικά να γραφτεί και ως η διαφορά δυο άλλων ολοκληρωμάτων με αρχικό όριο το $-\infty$:

$$P(a \leq x \leq b) = \int_a^b f(x)dx = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx$$

Εξ ορισμού όμως αυτά τα ολοκληρώματα είναι ίσα με την αθροιστική κατανομή πιθανότητας F και έτσι εάν γνωρίζουμε αυτήν τη συνάρτηση, μπορούμε απλά να υπολογίσουμε την πιθανότητα παίρνοντας διαφορές:

$$P(a \leq x \leq b) = F(b) - F(a)$$

Παράδειγμα 4.5

Έστω x ο χρόνος σε μήνες μέχρι να παρουσιάσει βλάβη ένα εξάρτημα

Η πιθανότητα να παρουσιάσει βλάβη περιγράφεται από την συνάρτηση Weibull

$$f(x) = \frac{3x^2}{4} e^{-\left(\frac{x}{4}\right)^3}$$

όπου $x > 0$ (πυκνότητα πιθανότητας)

(α) Να βρεθεί η αθροιστική συνάρτηση κατανομής $F(x)$

(β) Η διάμεσος (εκεί όπου $F(x) = 0.5$)

Λύση:

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{0}^{x} f(x)dx$$

Χρειάζομαι το αόριστο ολοκλήρωμα

$$\int f(x)dx = \frac{3}{4}\int x^2 e^{-\left(\frac{x}{4}\right)^3} dx = -16e^{-\left(\frac{x}{4}\right)^3} + c$$

$$F(x) = \int_{-\infty}^{x} f(x)dx = \left(-16e^{-\left(\frac{x}{4}\right)^3} + c\right) - \left(-16e^{-\left(\frac{0}{4}\right)^3} + c\right) =>$$

$$=> F(x) = 16 - 16e^{-\left(\frac{x}{4}\right)^3} = 16\left(1 - e^{-\left(\frac{x}{4}\right)^3}\right)$$