

## Βιοστατιστική

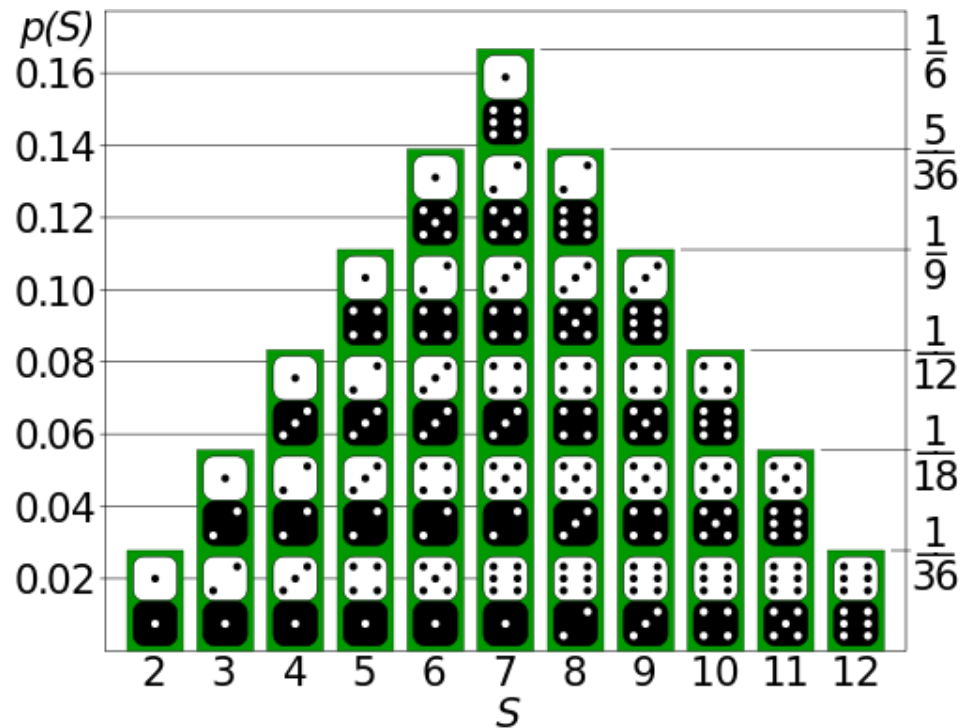
# Κατανομές Πιθανοτήτων

# Πηγές υλικού

- Διαφάνειες και ασκήσεις του Theophanis Tsandilas (National Institute for Research in Digital Science and Technology, INRIA, Γαλλία)
- Διαφάνειες και ασκήσεις του Σπύρου Γαλατσίδα (Τμ. Δασολογίας & Διαχείρισης Περιβάλλοντος & Φυσικών Πόρων, ΔΠΘ)

# Τι είναι μια κατανομή πιθανοτήτων;

- Θεωρούμε τον πληθυσμό όλων των πιθανών αποτελεσμάτων όταν ρίχνουμε δύο ζάρια
- Πόσο πιθανό είναι κάθε άθροισμα  $S$  των μετρήσεων από τα δύο ζάρια;
- Μία **κατανομή πιθανοτήτων** μας δείχνει την πιθανότητα εμφάνισης όλων των πιθανών αποτελεσμάτων σε ένα πείραμα

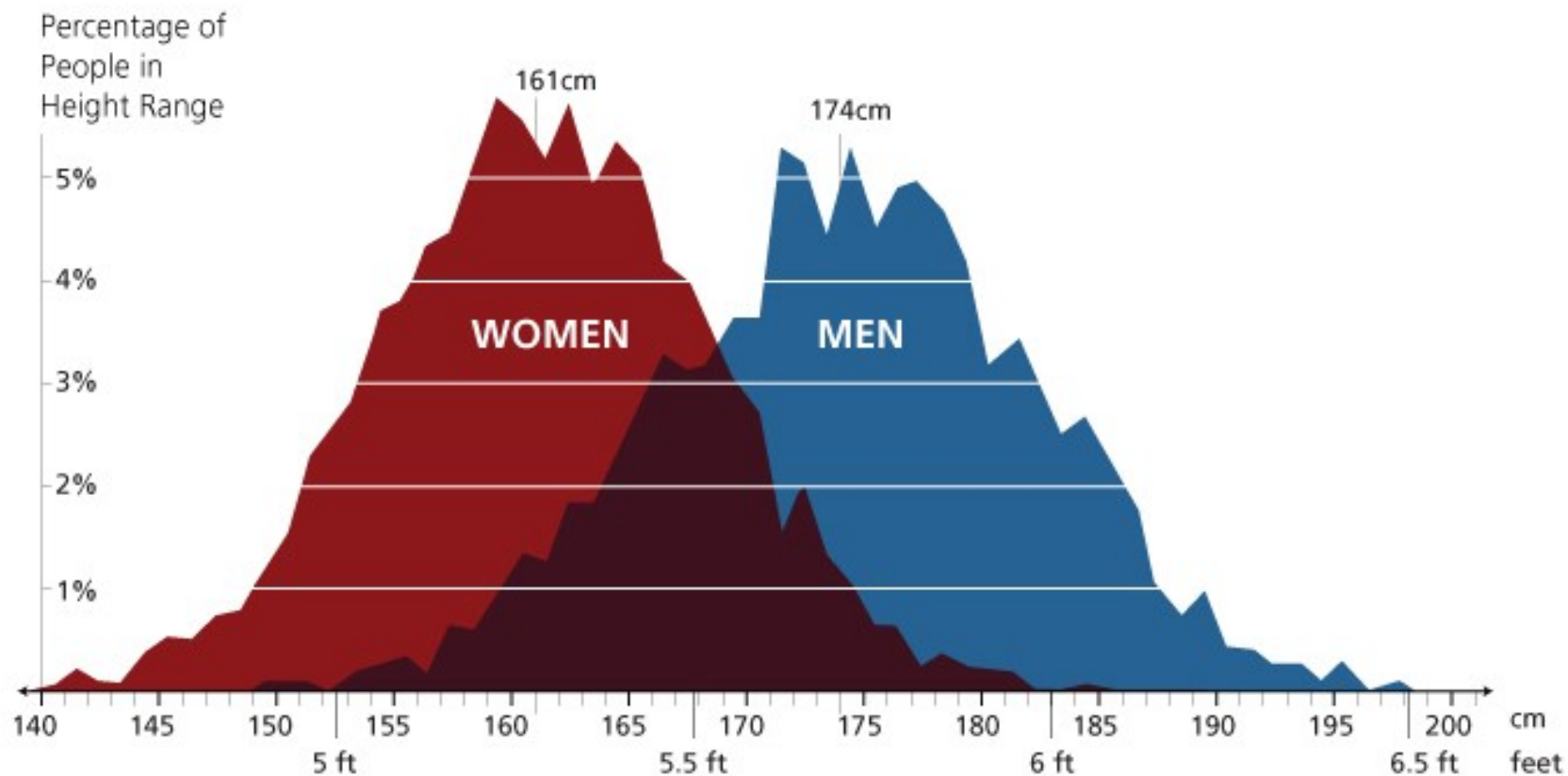


# Κατανομή πιθανοτήτων

- Για έναν πληθυσμό, η κατανομή των πιθανοτήτων του είναι συνήθως άγνωστη
- Προσπαθούμε να εξετάσουμε την κατανομή αυτή σε ένα **δείγμα** από τον πληθυσμό και μετά να επεκτείνουμε το συμπέρασμα που θα βγάλουμε από το δείγμα στον πληθυσμό!

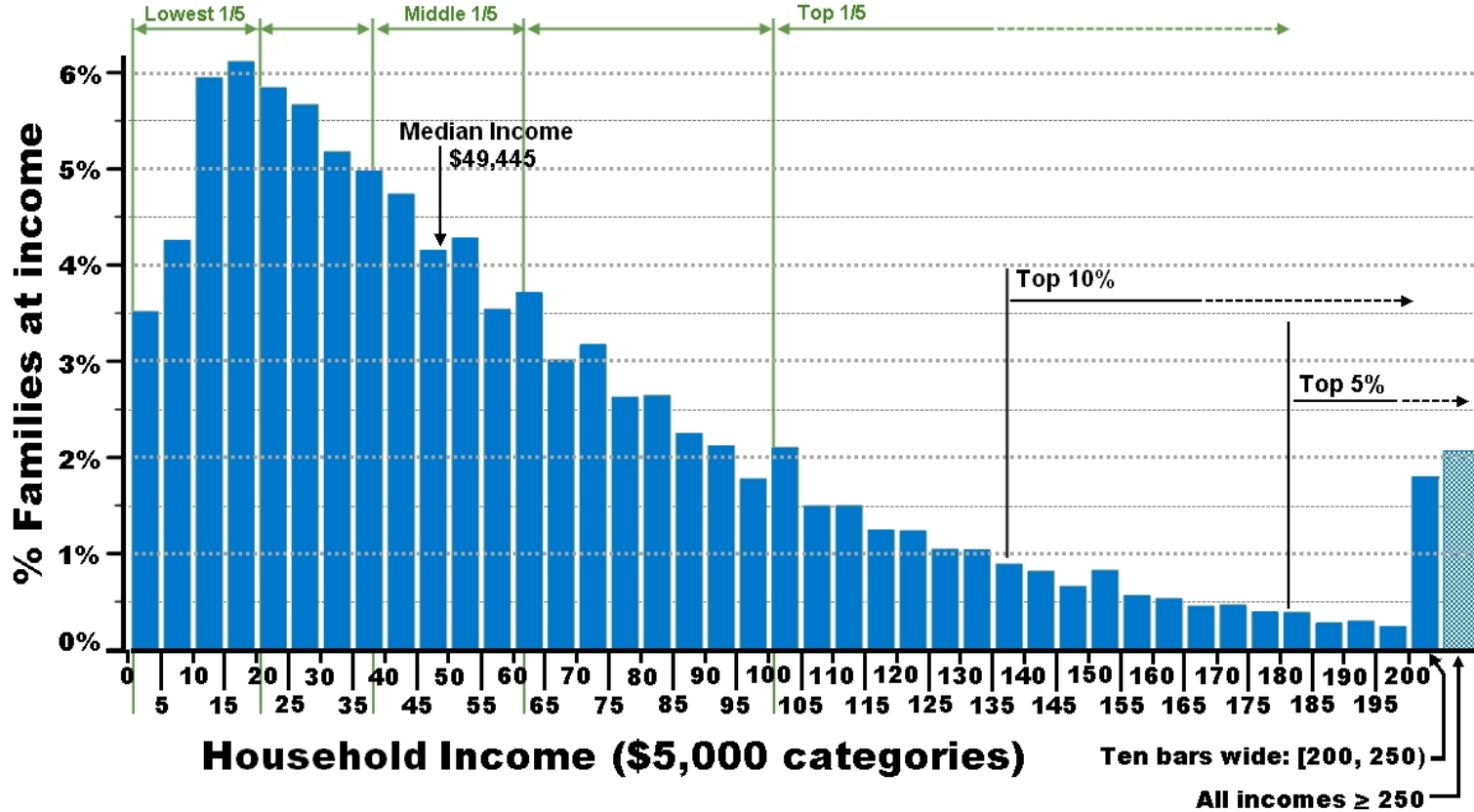
## Height of Adult Women and Men

Within-group variation and between-group overlap are significant



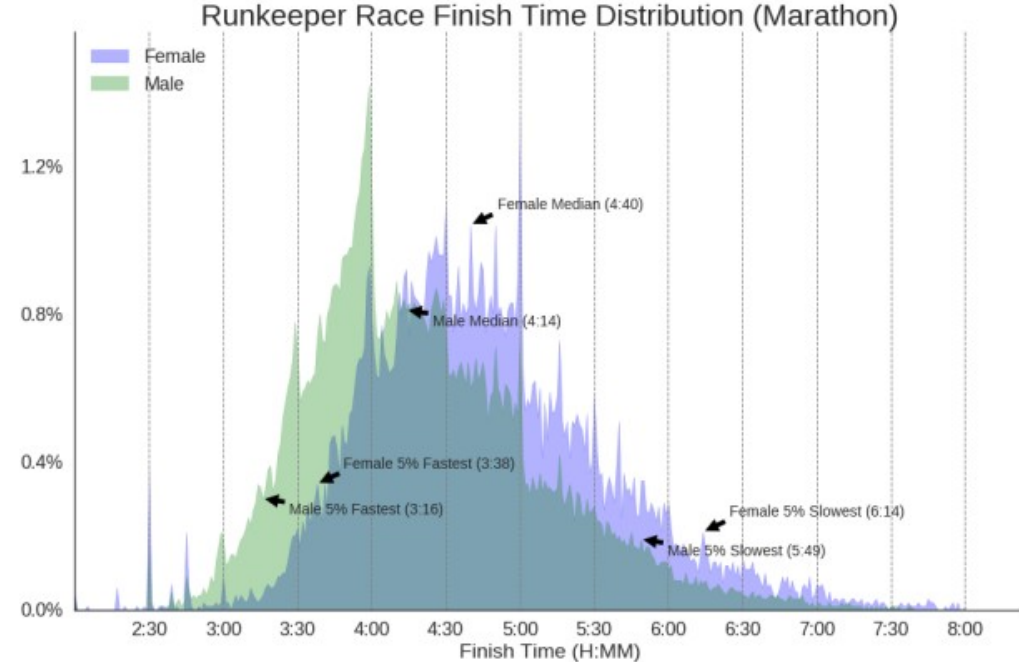
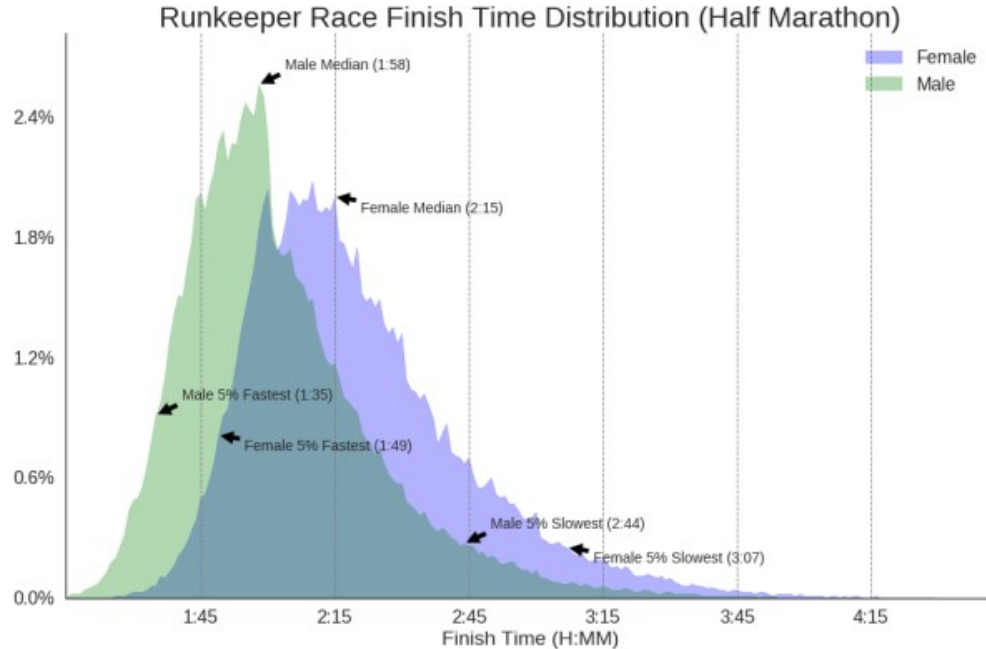
Data from U.S. CDC, adults ages 18-86 in 2007

# US Income distribution (older than 2013)



Data source: [http://www.census.gov/hhes/www/cpstables/032011/hhinc/new06\\_000.htm](http://www.census.gov/hhes/www/cpstables/032011/hhinc/new06_000.htm)

# Half Marathon and Marathon race finish time



Source: <https://medium.com/runkeeper-everyone-every-run/how-long-till-the-finish-line-494361cc901b>

WORD	COUNT	PERCENT	bar graph
the	53.10	7.14%	the
of	30.97	4.16%	of
and	22.63	3.04%	and
to	19.35	2.60%	to
in	16.89	2.27%	in
a	15.31	2.06%	a
is	8.38	1.13%	is
that	8.00	1.08%	that
for	6.55	0.88%	for
it	5.74	0.77%	it
as	5.70	0.77%	as
was	5.50	0.74%	was
with	5.18	0.70%	with
be	4.82	0.65%	be
by	4.70	0.63%	by
on	4.59	0.62%	on
not	4.52	0.61%	not
he	4.11	0.55%	he
i	3.88	0.52%	i
this	3.83	0.51%	this
are	3.70	0.50%	are
or	3.67	0.49%	or
his	3.61	0.49%	his
from	3.47	0.47%	from
at	3.41	0.46%	at
which	3.14	0.42%	which
but	2.79	0.38%	but
have	2.78	0.37%	have
an	2.73	0.37%	an
had	2.62	0.35%	had
they	2.46	0.33%	they
you	2.34	0.31%	you
were	2.27	0.31%	were
their	2.15	0.29%	their
one	2.15	0.29%	one
all	2.06	0.28%	all
we	2.06	0.28%	we
can	1.67	0.22%	can
her	1.63	0.22%	her
has	1.63	0.22%	has
there	1.62	0.22%	there
been	1.62	0.22%	been
if	1.56	0.21%	if
more	1.55	0.21%	more
when	1.52	0.20%	when
will	1.49	0.20%	will
would	1.47	0.20%	would
who	1.46	0.20%	who
so	1.45	0.19%	so
no	1.40	0.19%	no

## Distribution of most frequent English words

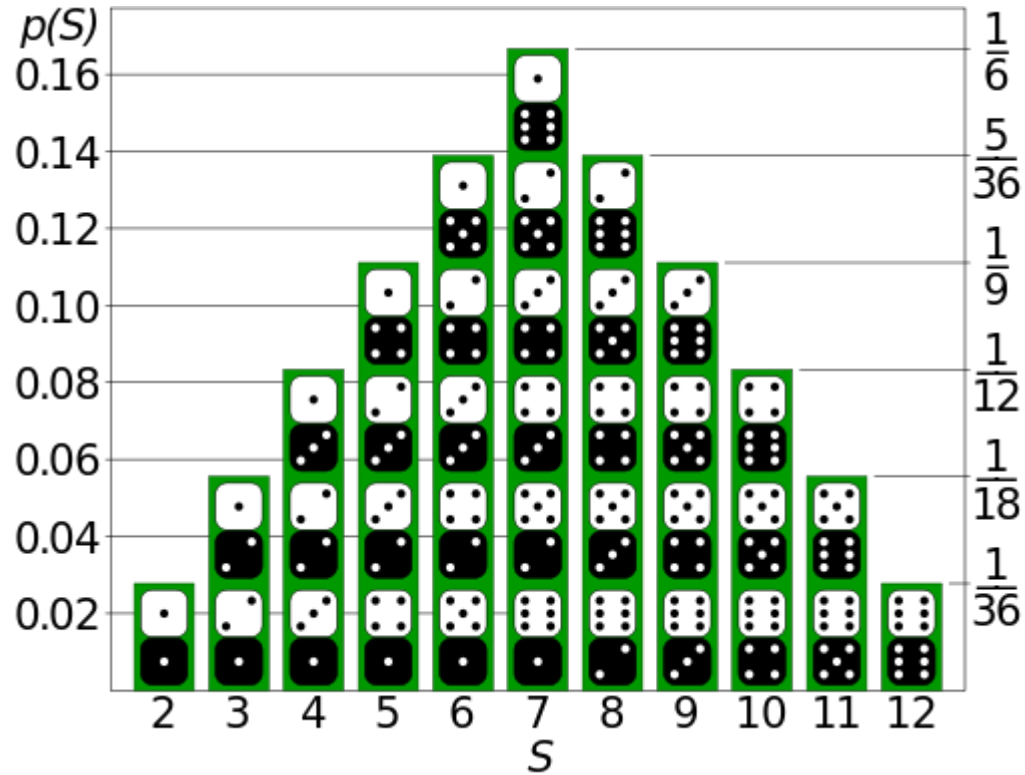
<http://norvig.com/mayzner.html>



# Διακριτές (discrete) κατανομές

- Ο πληθυσμός (και επομένως τα δείγματά του) περιέχουν διακριτές τιμές, είτε πεπερασμένες, είτε άπειρες σε αριθμό
  - π.χ.,  $\{-3, 1, 0, 1, 2\}$ ,  $\{\text{'μπλε'}, \text{'καφέ'}, \text{'πράσινο'}\}$  ή  $\{1, 2, 3, \dots\}$
- Η πιθανότητα μιας τιμής  $x$  σε έναν πληθυσμό μπορεί να εκφράζεται ως συνάρτηση:  $f(x) = p(X = x)$ 
  - η πιθανότητα η τυχαία μεταβλητή  $X$  να πάρει την τιμή  $x$
- Η συνάρτηση  $f(x)$  ονομάζεται συνάρτηση μάζας πιθανότητας (probability mass function - **pmf**)

# Μια διακριτή κατανομή



# Ιδιότητες διακριτών κατανομών

- Το άθροισμα όλων των πιθανοτήτων είναι 1

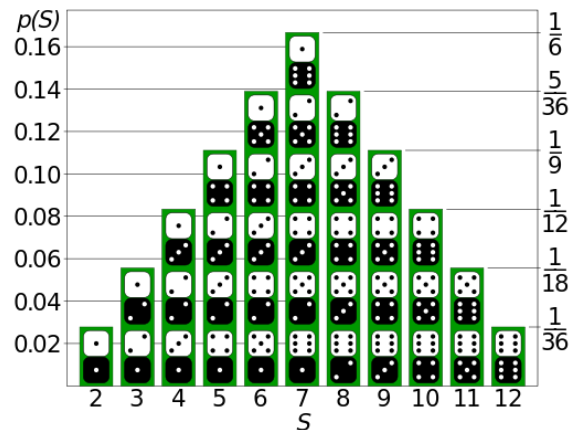
$$\sum_{x \in S} f(x) = 1$$

- Η αναμενόμενη τιμή (μέσος) είναι:

$$E(x) = \sum_{x \in S} f(x)x$$

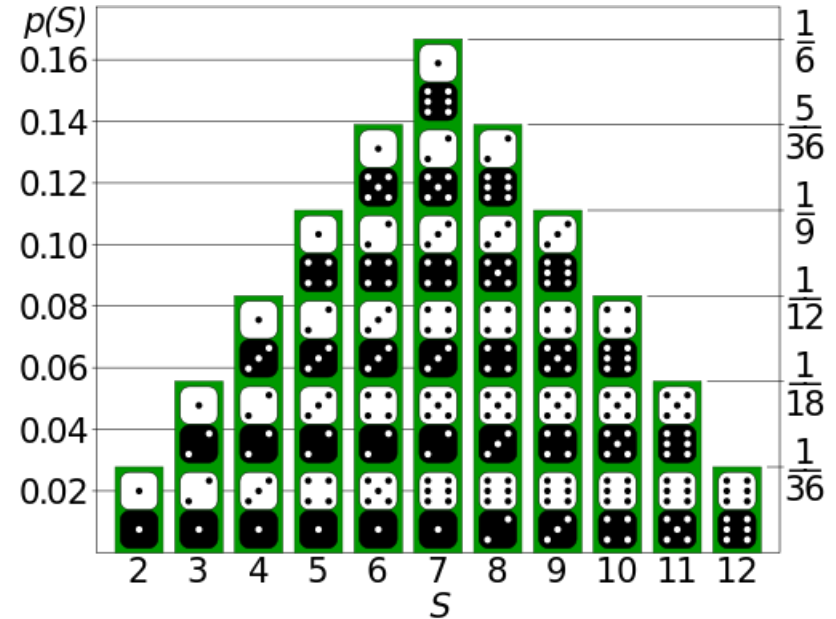
$$= 2 \frac{1}{36} + 3 \frac{1}{18} + 4 \frac{1}{12} + 5 \frac{1}{9} + 6 \frac{5}{36} + 7 \frac{1}{6} + 8 \frac{5}{36} + 9 * \frac{1}{9} + 10 \frac{1}{12} + 11 \frac{1}{18} + 12 \frac{1}{36} = 7$$

- Ποιος είναι ο τύπος εδώ;
- Ποια είναι η διάμεσος;



# Ιδιότητες διακριτών κατανομών

- Όταν ο μέσος ταυτίζεται με τη διάμεσο και με τον τύπο έχουμε συμμετρική και μονοκόρυφη κατανομή πιθανοτήτων
  - symmetrical, unimode distribution



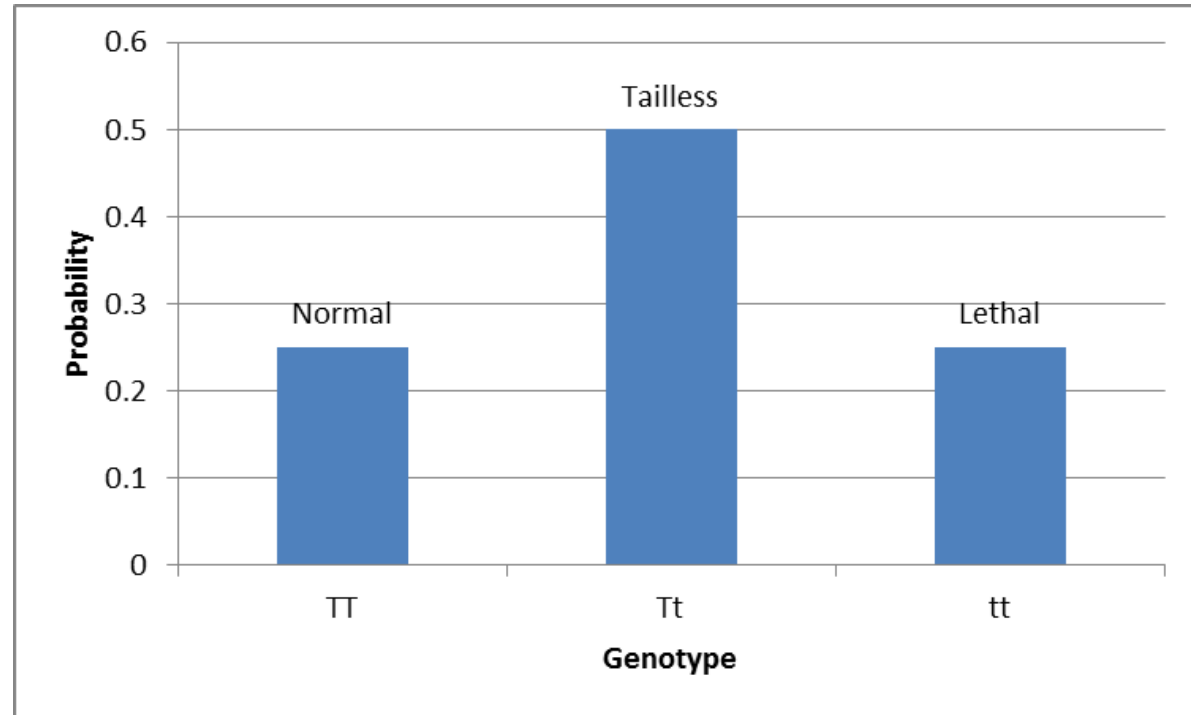
# Διωνυμική (binomial) κατανομή

- Σκεφτείτε έναν πληθυσμό με δύο τιμές: 1 και 0
- Ένα μόνο δείγμα από τον πληθυσμό αυτό ( $n = 1$ ) λέγεται **πείραμα Bernoulli**
  - Ένα νόμισμα που ρίχνουμε είναι ένα πείραμα Bernoulli
    - $Pr(\text{Head}) = 0.5$  και  $Pr(\text{Tail}) = 0.5$
- Σε μια επανάληψη  $n$  ανεξάρτητων πειραμάτων Bernoulli, ο αριθμός των επιτυχών παρατηρήσεων (ή αποτυχημένων) ακολουθεί τη **διωνυμική κατανομή**
  - Πολύ σημαντική για πειράματα μοριακής βιολογίας και γενετικής

# Παράδειγμα διακριτής κατανομής



- Η διασταύρωση δύο γατών Manx
  - Κλασικό παράδειγμα πλειοτροπισμού από τη Γενετική
    - Ένα γονίδιο ελέγχει δύο χαρακτήρες



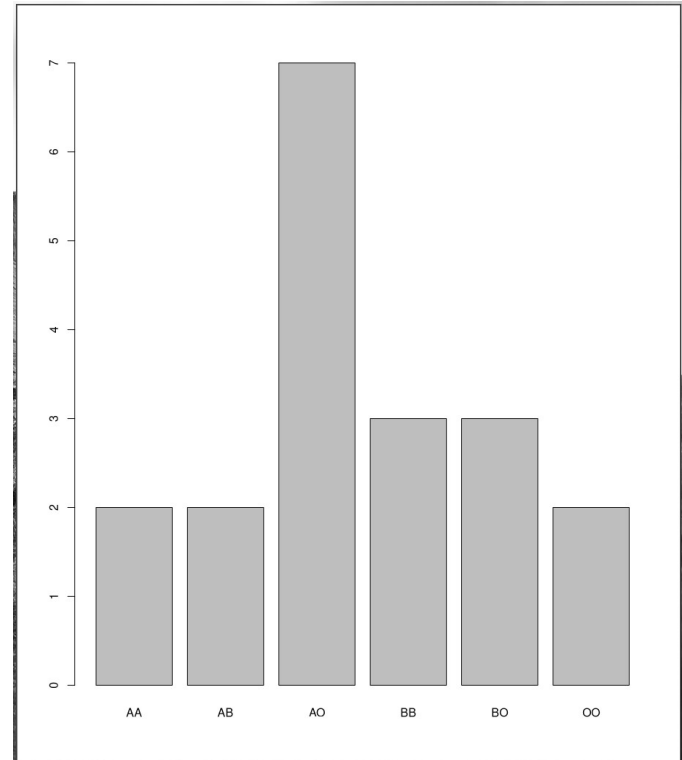
# Παράδειγμα διακριτής κατανομής στο R – ομάδες ABO

```
> genotype = c("AA", "AO", "BB", "AO", "OO", "AO", "AA", "BO", "BO", "AO", "BB", "AO", "BO", "AB", "OO", "AB", "BB", "AO", "AO")
> table(genotype)
genotype
AA AB AO BB BO OO
 2  2  7  3  3  2
```

```
> table(genotype)
genotype
AA AB AO BB BO OO
 2  2  7  3  3  2
```

```
> genotypeF = factor(genotype)
> levels(genotypeF)
[1] "AA" "AB" "AO" "BB" "BO" "OO"
> table(genotypeF)
genotypeF
AA AB AO BB BO OO
 2  2  7  3  3  2
```

```
> plot(genotypeF)
> 
```



# Η διωνυμική κατανομή

- Θεωρούμε έναν πληθυσμό που περιέχει δύο τιμές: 1 και 0
- Ένα μόνο δείγμα ( $n = 1$ ) από έναν τέτοιο πληθυσμό είναι γνωστό ως **πείραμα Bernoulli**
  - Ένα πείραμα ρίψης νομίσματος είναι ένα πείραμα Bernoulli με  $Pr(\text{Head}) = 0,5$  και  $Pr(\text{Tail}) = 0,5$
- Εάν εκτελέσουμε  $n$  ανεξάρτητες δοκιμές Bernoulli, ο αριθμός των επιτυχιών (ή αποτυχιών) ακολουθούν μια **διωνυμική κατανομή**
  - Σημαντική εφαρμογή στη Γενετική



# Η διωνυμική κατανομή

- Όταν μια τυχαία μεταβλητή  $X$  ακολουθεί τη διωνυμική κατανομή, θεωρούμε:  $X \sim B(n, P)$ 
  - $X$  είναι ο αριθμός των επιτυχιών, όταν
    - $n$  είναι ο αριθμός των επαναλήψεων του πειράματος Bernoulli
    - $P$  είναι η πιθανότητα επιτυχίας σε ένα πείραμα
- Αν γνωρίζουμε τα  $n$  και  $P$  μπορούμε να περιγράψουμε την κατανομή με ακρίβεια

# Η διωνυμική κατανομή

PMF of Bernoulli Distribution



$$P(X = x) = \begin{cases} p & \text{for } x = 1 \\ 1-p & \text{for } x = 0 \end{cases}$$

# Η διωνυμική κατανομή

## Probability Mass Function (pmf)

$$f(x; n, P) = \binom{n}{x} P^x (1 - P)^{n-x}$$



Η πιθανότητα ακριβώς  $x$  επιτυχιών



Η πιθανότητα ακριβώς  $n - x$  αποτυχιών



Ο αριθμός των πιθανών τρόπων έτσι ώστε  $n$  δοκιμές Bernoulli να οδηγούν σε  $x$  επιτυχίες

# Παράδειγμα

- Ποια είναι η πιθανότητα για ένα ζευγάρι να έχει τρία παιδιά, και τα τρία κορίτσια;
  - Θεωρούμε  $n=3$ ,  $x=3$  και  $p=0,5$ ,  $q=1-p=0,5$
  - Οπότε έχουμε:

$$P = p^x \cdot q^{n-x}$$

$$P = \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^{(3-3)} = \frac{1}{8}$$

# Παράδειγμα

- Ποια είναι η πιθανότητα για ένα ζευγάρι να έχει τρία παιδιά και τα δύο από αυτά κορίτσια;
  - Θεωρούμε  $n=3$ ,  $x=2$  και  $p=0,5$ ,  $q=1-p=0,5$
  - Όλοι οι δυνατοί συνδυασμοί για  $x=2$  είναι:
  - Οπότε έχουμε:

$$C = \frac{n!}{x! \cdot (n-x)!} = 3$$

$$P = C \cdot p^x \cdot q^{n-x}$$

$$P = 3 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{(3-2)} = \frac{3}{8}$$

# Παράδειγμα

- Και οι δύο γονείς σε ένα ζευγάρι είναι φορείς της δρεπανοκυτταρικής αναιμίας. Ποια είναι η πιθανότητα από τα 4 παιδιά που θα γεννηθούν από το ζευγάρι:
  - τα δύο να έχουν την ασθένεια;
  - τα δύο να είναι φορείς;

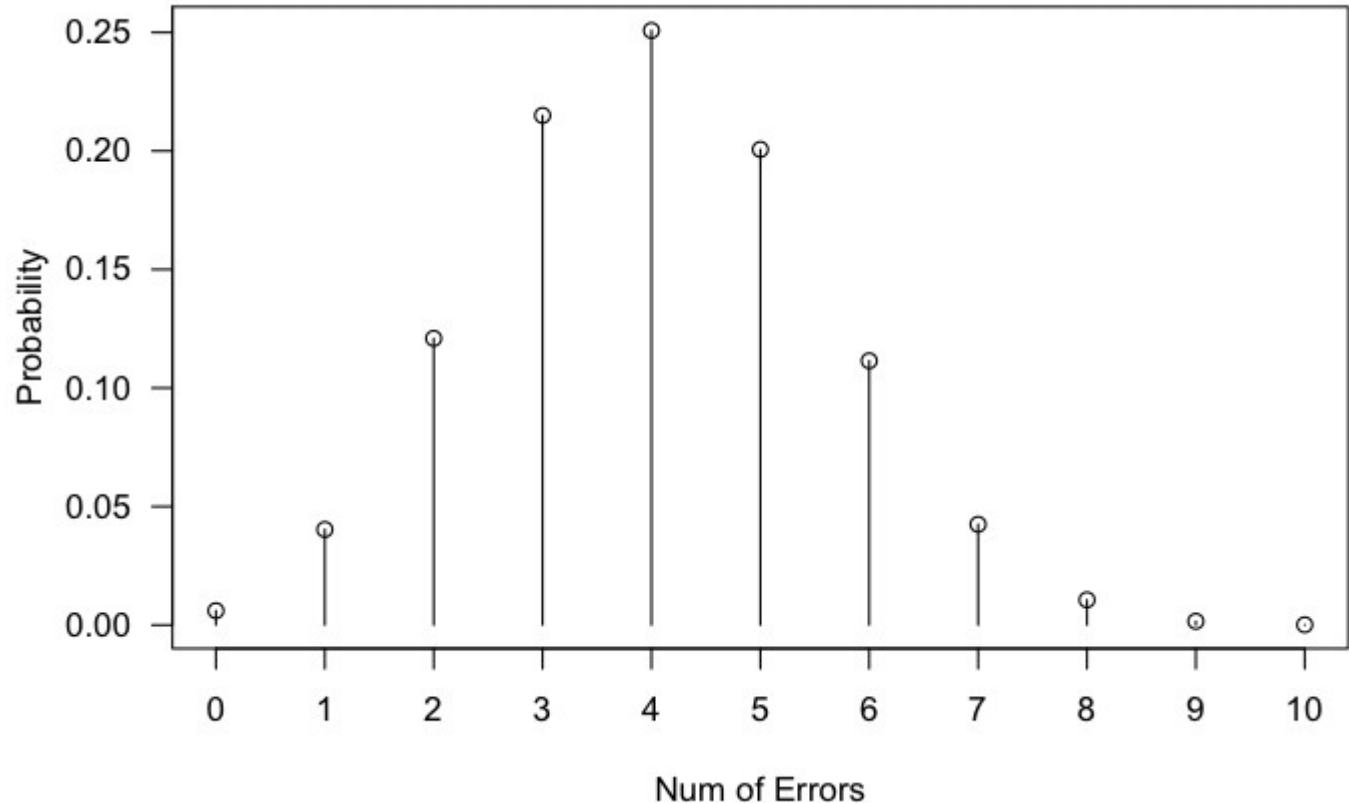
# Παράδειγμα

- Και οι δύο γονείς σε ένα ζευγάρι είναι φορείς της δρεπανοκυτταρικής αναιμίας. Ποια είναι η πιθανότητα από τα 4 παιδιά που θα γεννηθούν από το ζευγάρι:
  - τα δύο να έχουν την ασθένεια;
  - τα δύο να είναι φορείς;

```
> p_girl <- dbinom(2, 3, 0.5)
> p_girl
[1] 0.375
> p_sickle <- dbinom(2, 4, 0.25)
> p_sickle
[1] 0.2109375
> p_het <- dbinom(2, 4, 0.50)
> p_het
[1] 0.375
> _
```

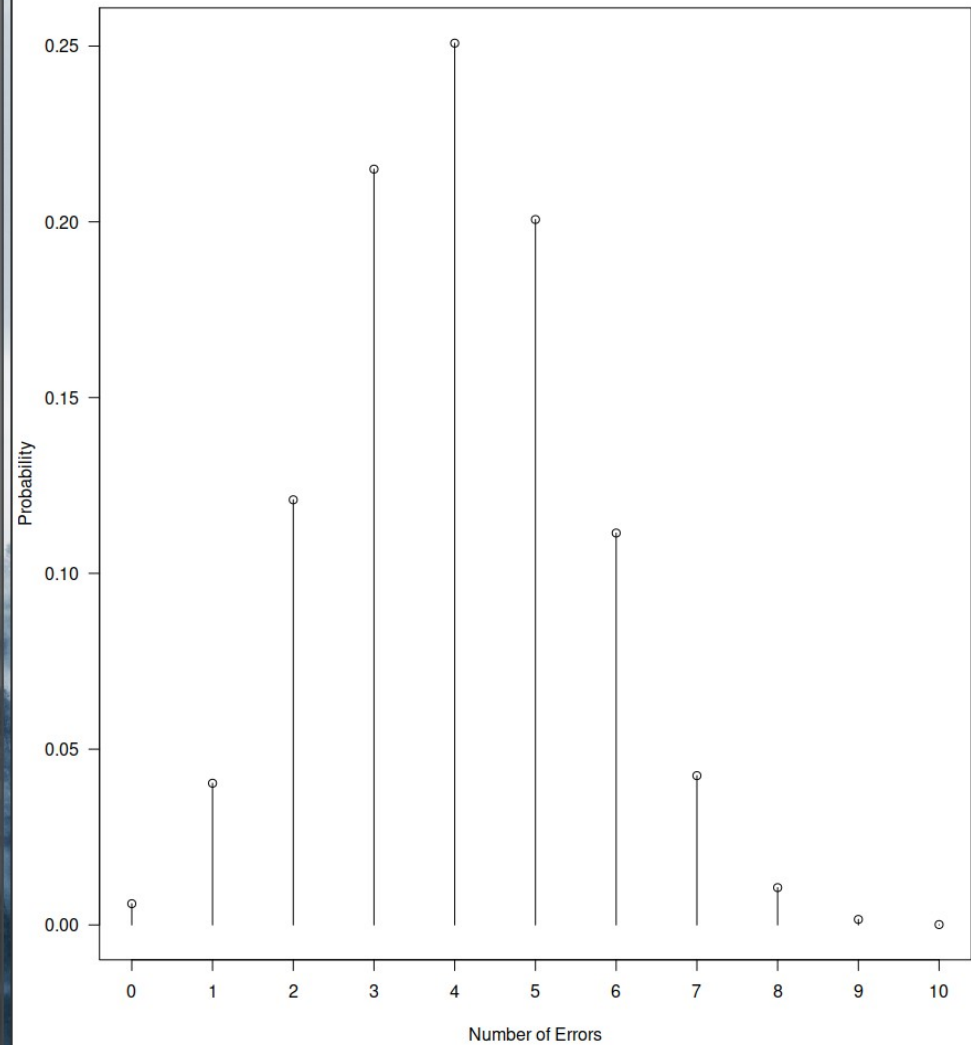
# Διωνυμική κατανομή

- Η κατανομή της πιθανότητας εμφάνισης συγκεκριμένου αριθμού σφαλμάτων, όταν σε ένα πείραμα συμβαίνουν λάθη με πιθανότητα 40%



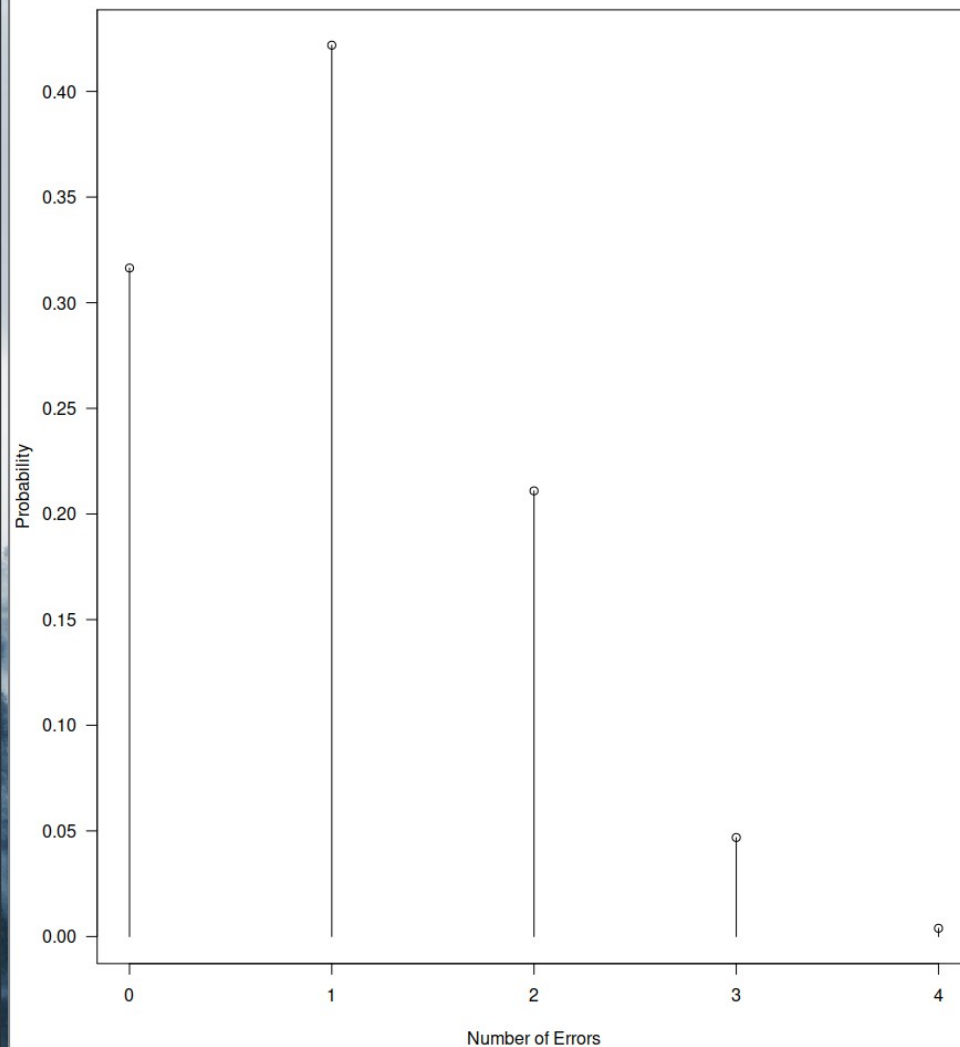


```
> errors <- 0:10
> prob.mass <- dbinom(errors, 10, 0.4)
> plot(errors, prob.mass, pch=1, xlab="Number of Errors", ylab="Probability",
, xaxt="n", yaxt="n")
> segments(x0=errors, y0=0, x1=errors, y1=prob.mass)
> axis(1, at=seq(0,10, by=1), las=1)
> axis(2, at=seq(0,1, by=0.05), las=1)
>
```



```
prob.mass <- dbinom(sickle, 4, 0.25)
plot(sickle, prob.mass, pch=1, xlab="Number of Errors", ylab="Probability",
     xaxt="n", yaxt="n")
segments(x0=sickle, y0=0, x1=sickle, y1=prob.mass)
axis(1, at=seq(0,4, by=1), las=1)
axis(2, at=seq(0,1, by=0.05), las=1)
> _
```

Δημιουργούμε το ίδιο γράφημα για  
Το παράδειγμα της δρεπανοκυτταρικής  
αναιμίας



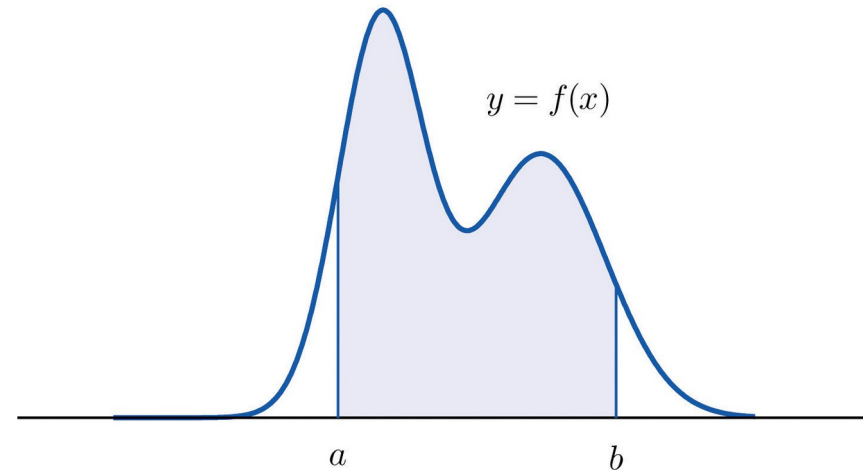
# Άσκηση #1

- Σε έναν πληθυσμό ανθρώπων, πολύ μεγάλου μεγέθους (πρακτικά άπειρου), γνωρίζουμε ότι η κυστική ίνωση εμφανίζεται στα νεογέννητα με συχνότητα περίπου  $1/10.000$  γέννες ( $0,0001$ ).
  - Υπολογίστε την πιθανότητα μιας καταγραφής εμφάνισης της ασθένειας (1) σε ένα δείγμα 100 γεννήσεων
  - Υπολογίστε τις πιθανότητες μιας καταγραφής εμφάνισης της ασθένειας (1) σε όλα τα δείγματα μεγέθους από 100 ως 100.000 γεννήσεις. Απεικονίστε το αποτέλεσμα σας σε ένα απλό γράφημα
    - Τι συμπέρασμα βγάζετε;
  - Προσπαθήστε να υπολογίσετε τις πιθανότητες **τουλάχιστον** μιας καταγραφής εμφάνισης της ασθένειας ( $>0$ ) σε όλα τα δείγματα μεγέθους από 100 ως 100.000 γεννήσεις. Απεικονίστε το αποτέλεσμα σας σε ένα απλό γράφημα
    - Τι συμπέρασμα βγάζετε;

# Συνεχείς κατανομές

- Δεν περιορίζονται σε συγκεκριμένες τιμές
  - Λαμβάνουν οποιαδήποτε τιμή ανάμεσα στο κατώτερο και το ανώτερο όριο ενός πληθυσμού
    - Οι πληθυσμοί μπορεί να είναι και άπειροι (απεριόριστοι)

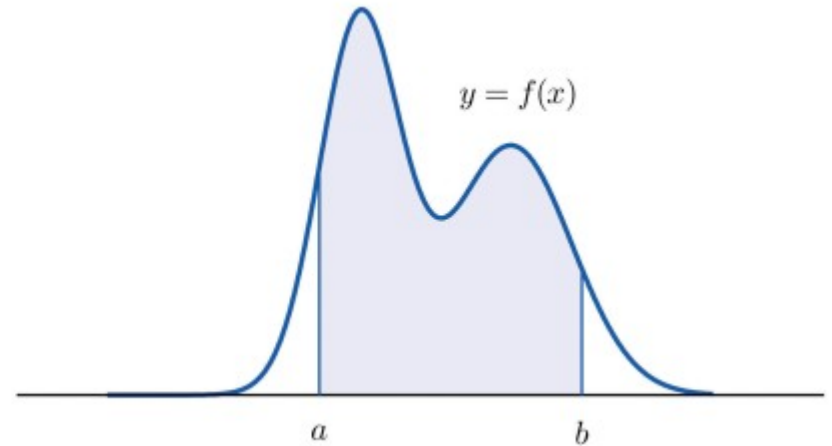
$$P(a < X < b) = \text{area of shaded region}$$



# Συνεχείς κατανομές

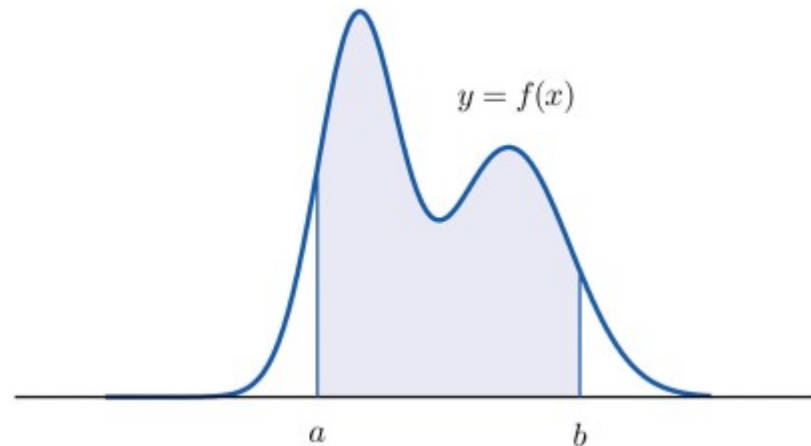
- Η πιθανότητα για οποιοδήποτε σημείο είναι 0. Οι πιθανότητες μπορούν μόνο να υπολογιστούν μόνο για διαστήματα με συγκεκριμένο εύρος (intervals):

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



# Συνεχείς κατανομές

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



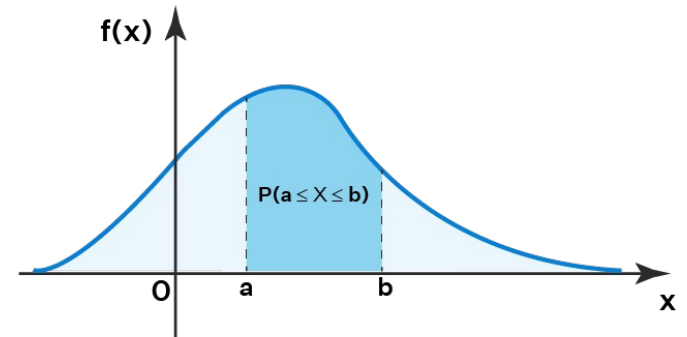
- Όπου  $f(x)$  είναι η **συνάρτηση πιθανότητας πυκνότητας** (probability density function -pdf)
- Μας δίνει τη σχετική (και όχι την απόλυτη) πιθανότητα για μια τυχαία μεταβλητή  $X$  να λάβει την τιμή  $x$

# Συνεχείς κατανομές

- Όταν ο πληθυσμός είναι άπειρος, τότε οποιαδήποτε τιμή  $X$  θα πέσει μέσα στο διάστημα και η πιθανότητα γίνεται 1

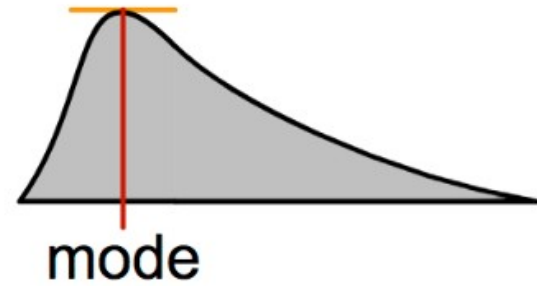
$$Pr(-\infty \leq X \leq \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$$

Graph of Probability Density Function

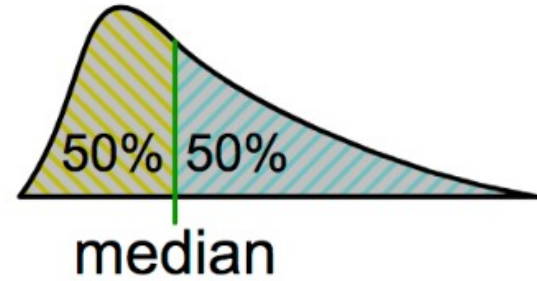


# Συνεχείς κατανομές

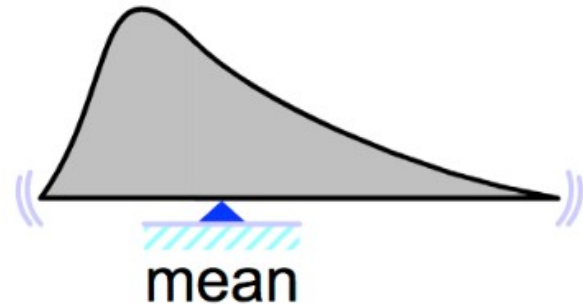
**Mode:** value of highest peak



**Median:** value that divides the area exactly in half

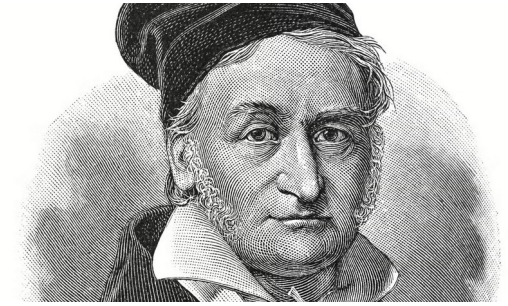


**Mean:**  $\mu = \int_{-\infty}^{\infty} x f(x) dx$

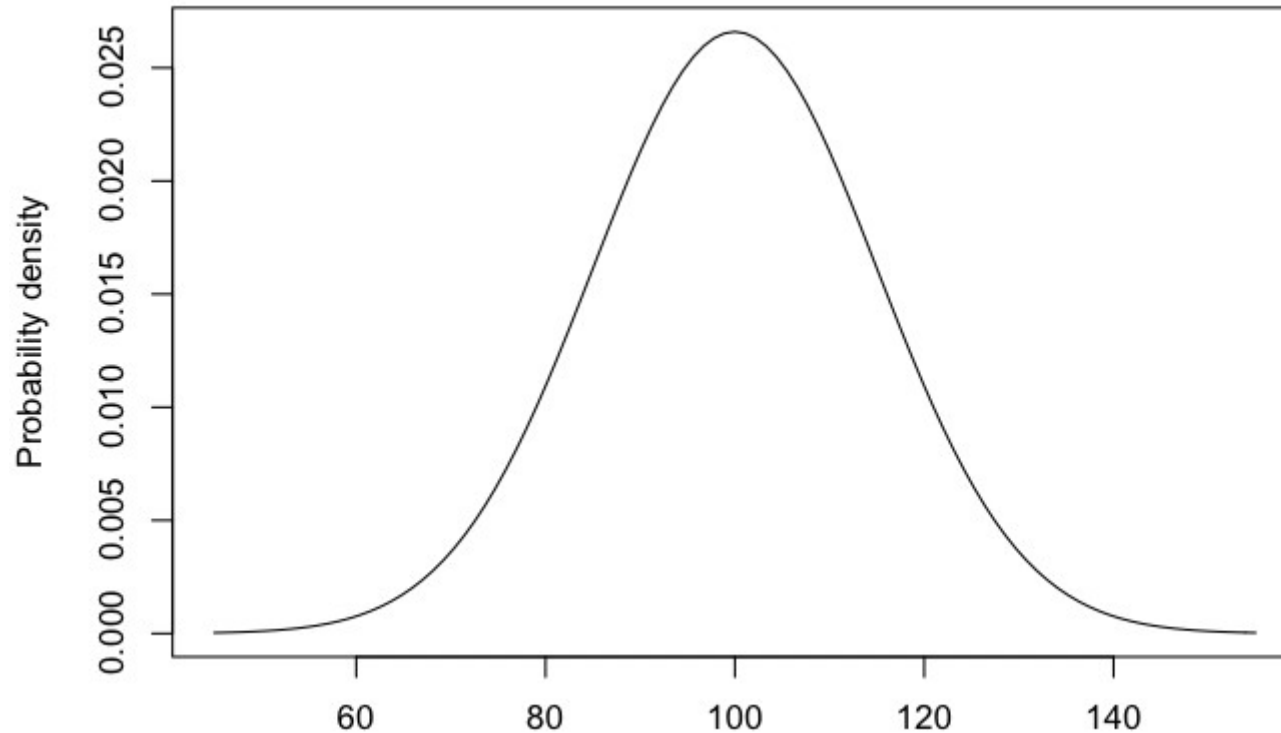




# Κανονική κατανομή

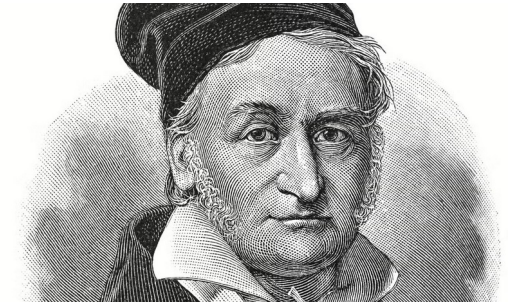


- Γνωστή και ως κατανομή του Gauss



Johann Carl Friedrich Gauß  
(1777-1855)

# Κανονική κατανομή

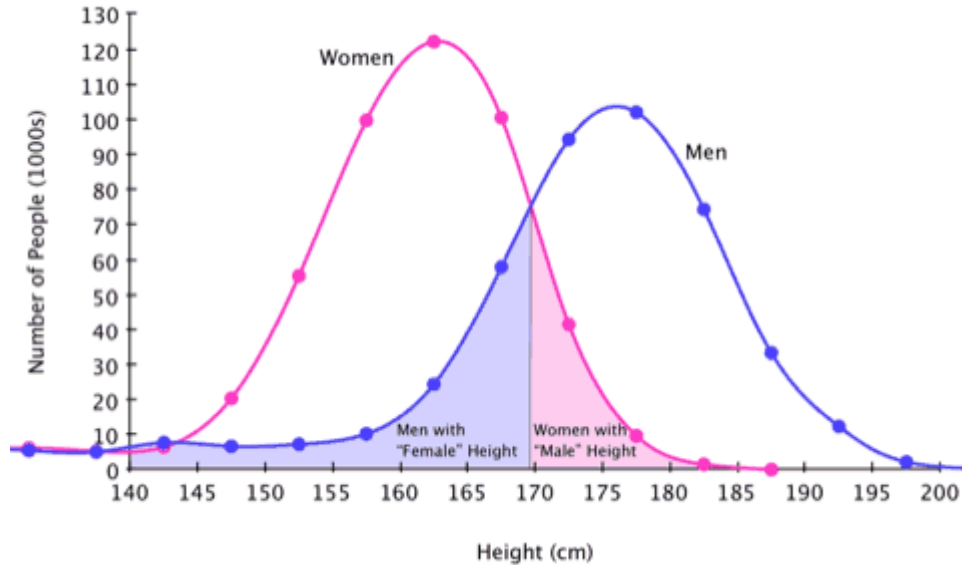


Johann Carl Friedrich Gauß  
(1777-1855)

- Συμμετρική, συνεχής και τυπική
- Μπορεί να προκύψει ως το άθροισμα ενός άπειρου αριθμού από ανεξάρτητες τυχαίες μεταβλητές
- Είναι ιδιαίτερα χρήσιμη για την περιγραφή δεδομένων που οφείλονται στην αθροιστική και συνδυασμένη δράση πολλών ανεξάρτητων και τυχαίων γεγονότων
  - Εφαρμογή σε βιολογικά δεδομένα

# Κανονική κατανομή

- Π.χ., το ύψος των ανθρώπων είναι το αποτέλεσμα της δράσης πολλών ανεξάρτητων γενετικών και περιβαλλοντικών παραγόντων



# Κανονική κατανομή

- Μια κανονική κατανομή μπορεί να περιγραφεί πλήρως με δύο μόνο παραμέτρους: τον μέσο όρο  $\mu$  και τη διακύμανσή  $\sigma^2$
- Μια κανονικά κατανεμημένη μεταβλητή  $X$  μπορεί να γραφτεί ως:

$$X \sim N(\mu, \sigma^2)$$

- Η συνάρτηση πιθανότητας πυκνότητας είναι:

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

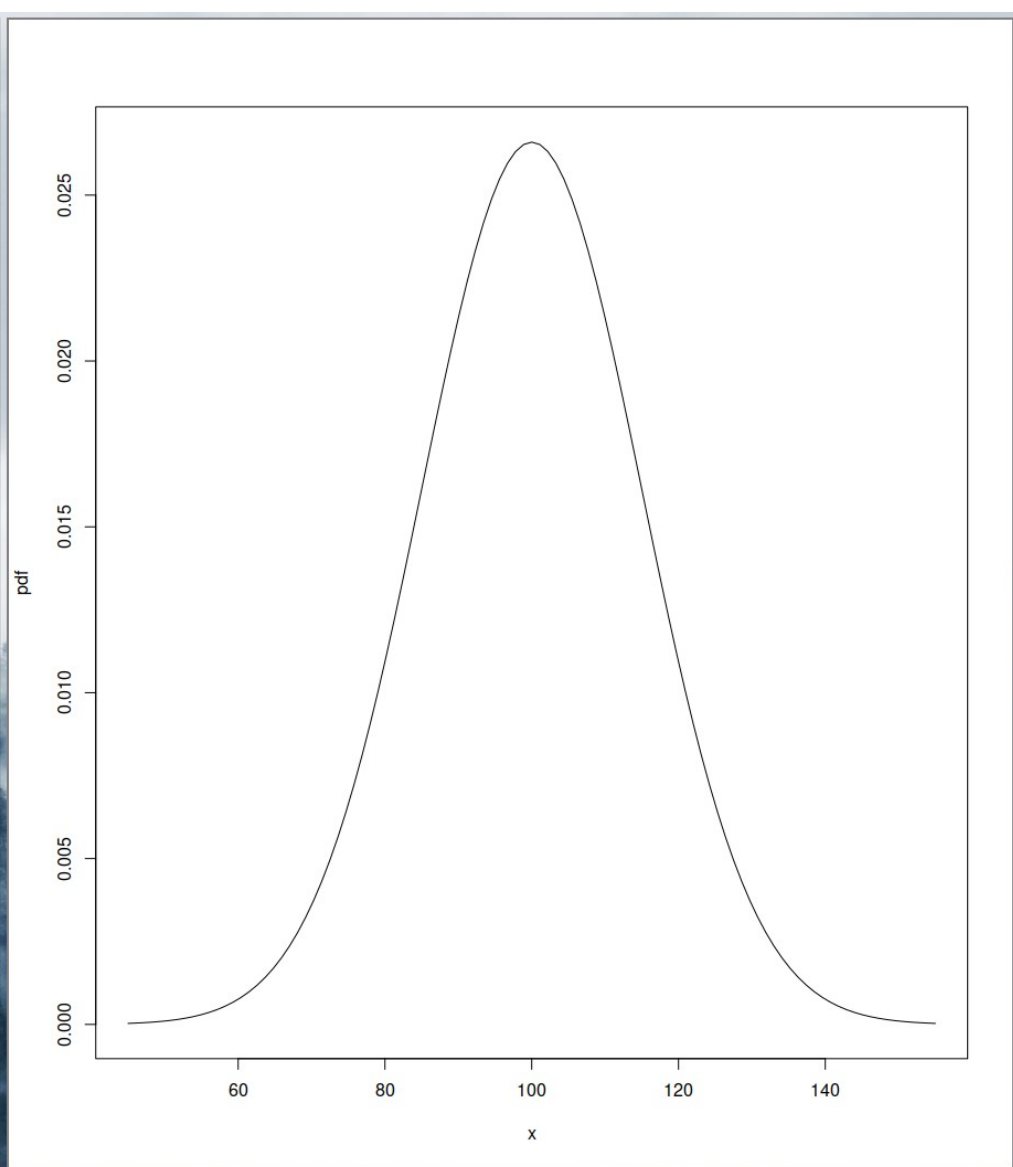
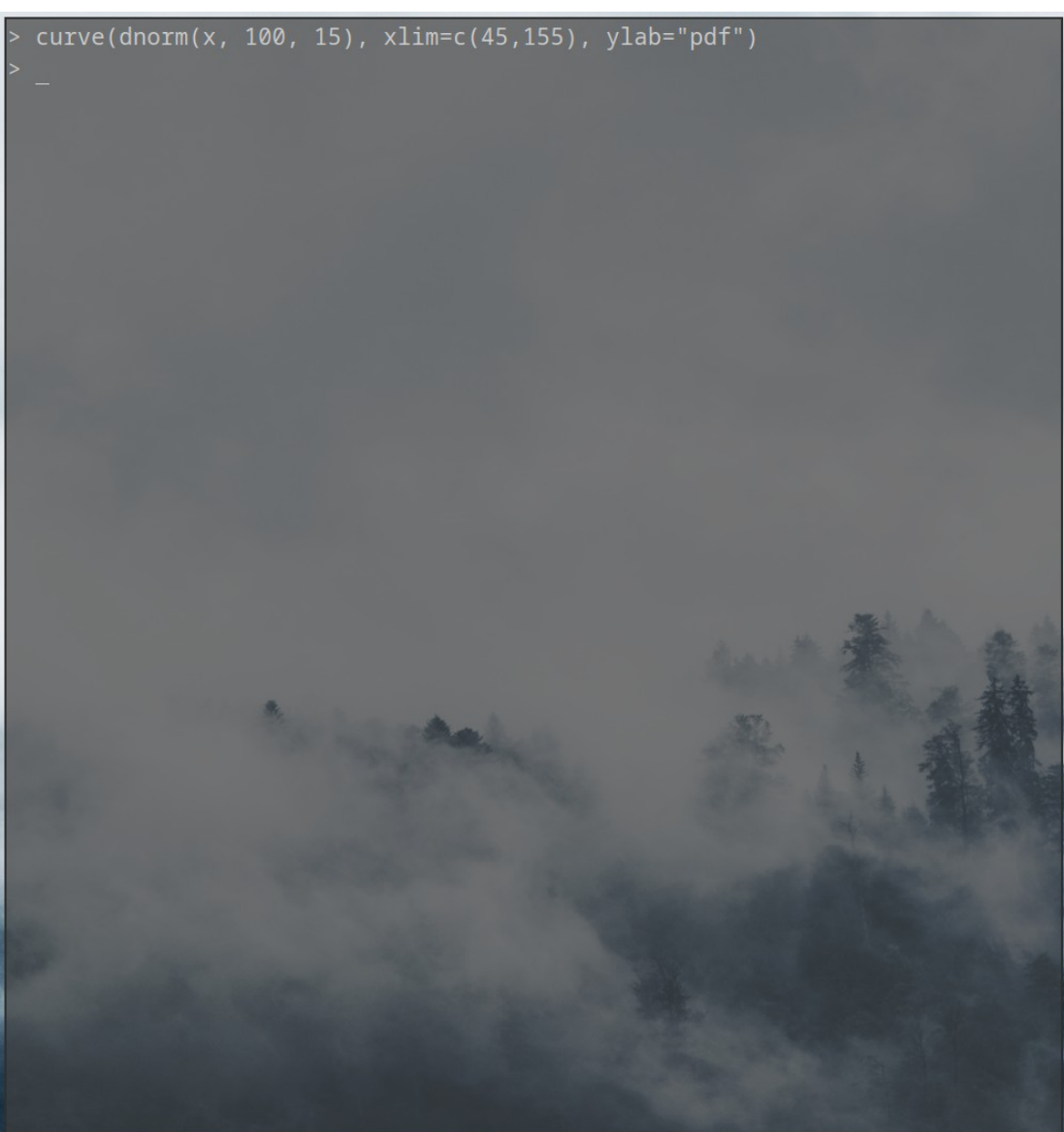
# Παράδειγμα

- Ας σχεδιάσουμε μια κανονική κατανομή με μέσο  $\mu=100$  και τυπική απόκλιση  $\sigma^2=15$

```
> curve(dnorm(x, 100, 15), xlim=c(45,155), ylab="pdf")
```

```
>
```

```
_
```

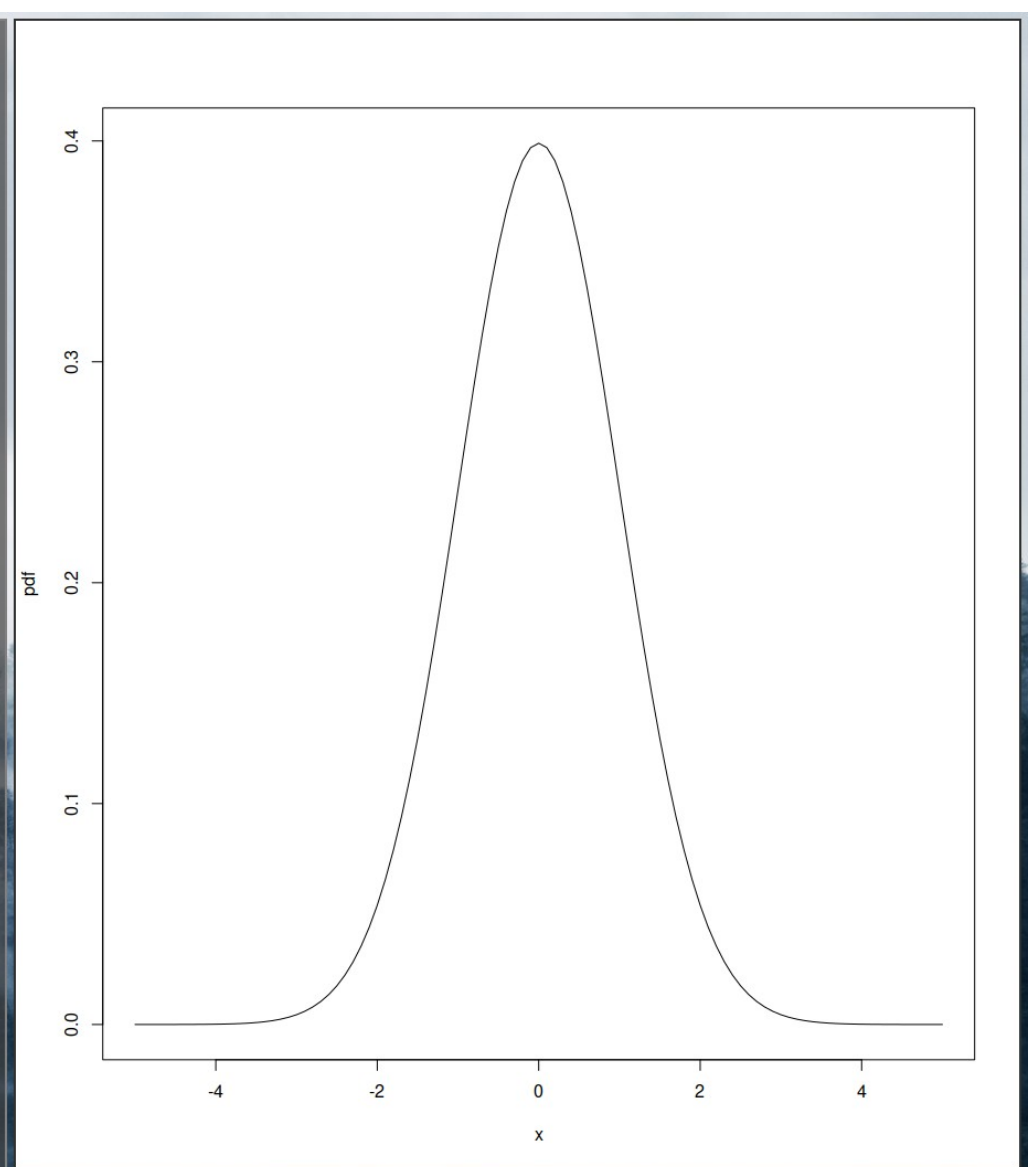
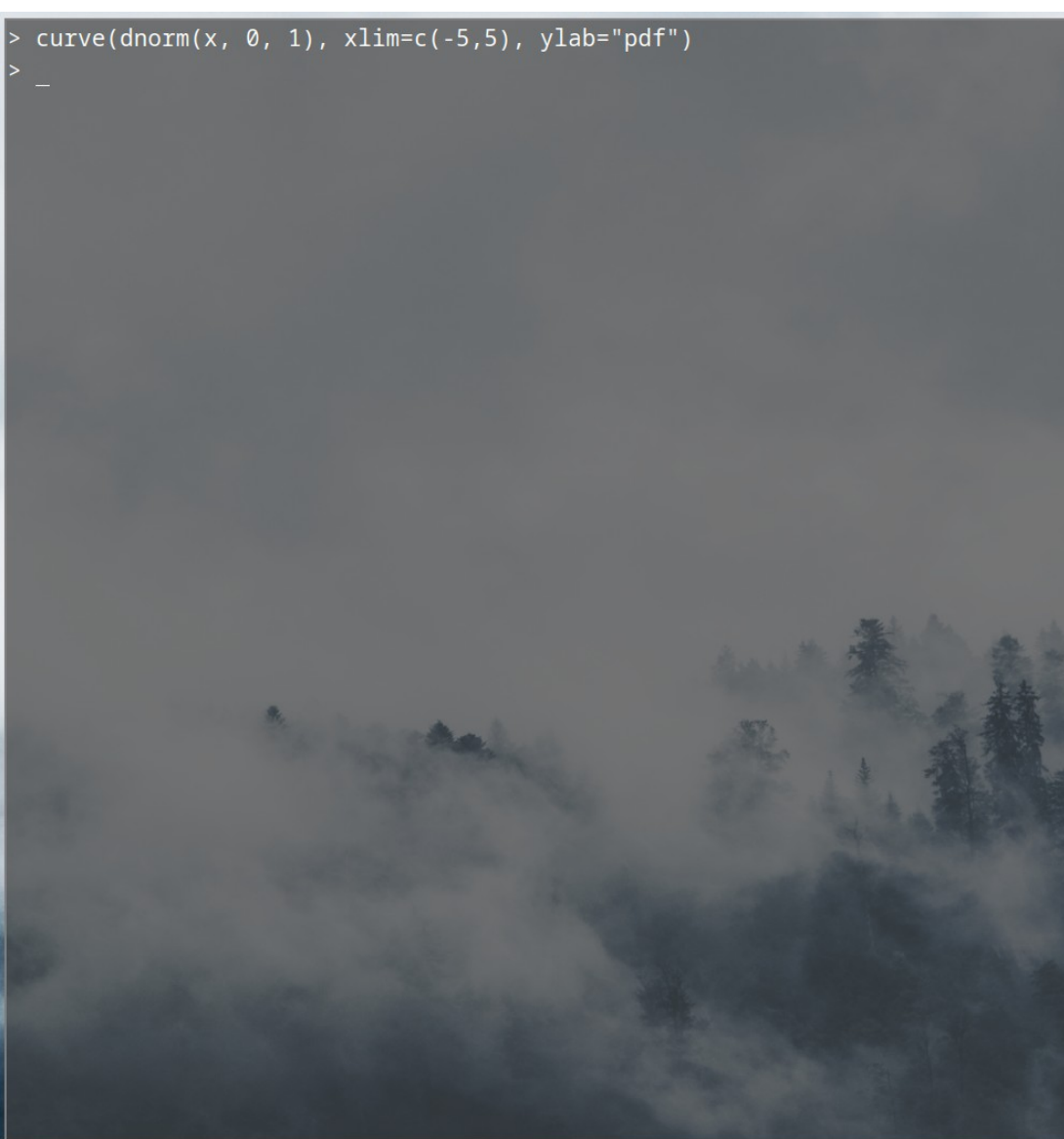


# Τυπική κανονική κατανομή

- Είναι μια κανονική τιμή με μέσο ίσο με 0 και τυπική απόκλιση ίση με 1
  - Συχνά συμβολίζεται με  $z$

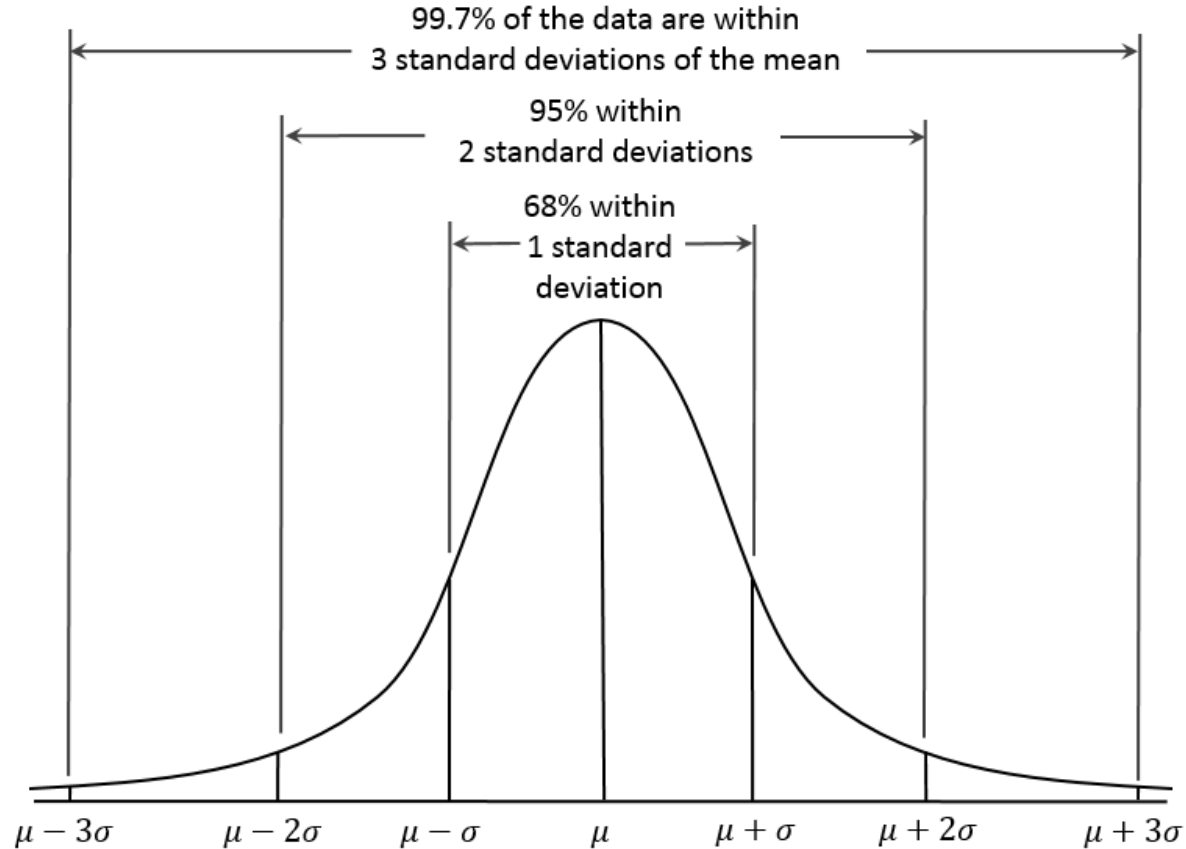
$$z \sim N(0, 1)$$

```
> curve(dnorm(x, 0, 1), xlim=c(-5,5), ylab="pdf")  
> _
```





# Ιδιότητες μιας κανονικής κατανομής



# Παράδειγμα

- Για μια μεταβλητή ενός πληθυσμού με μέσο όρο 10 και τυπική απόκλιση 2, που ακολουθεί την κανονική τιμή, ας πάρουμε ένα δείγμα 10 ατόμων και ας υπολογίσουμε την τυπική απόκλιση του δείγματος

```
> n<-10  
> sd(rnorm(n, mean=10, sd=2))  
[1] 1.918979  
> _
```

- Μπορούμε να επαναλάβουμε τη διαδικασία; Πόσο σταθερή μένει η τιμή;
- Ακολουθούμε τα ίδια βήματα για δείγμα 10.000 ατόμων
- Μπορούμε να κάνουμε το ίδιο και για τους μέσους;

# Παράδειγμα

- Τι συμπέρασμα μπορούμε να βγάλουμε από την άσκηση αυτή, σχετικά με το μέγεθος του δείγματος και την αξιοπιστία των στατιστικών;

# Παράδειγμα

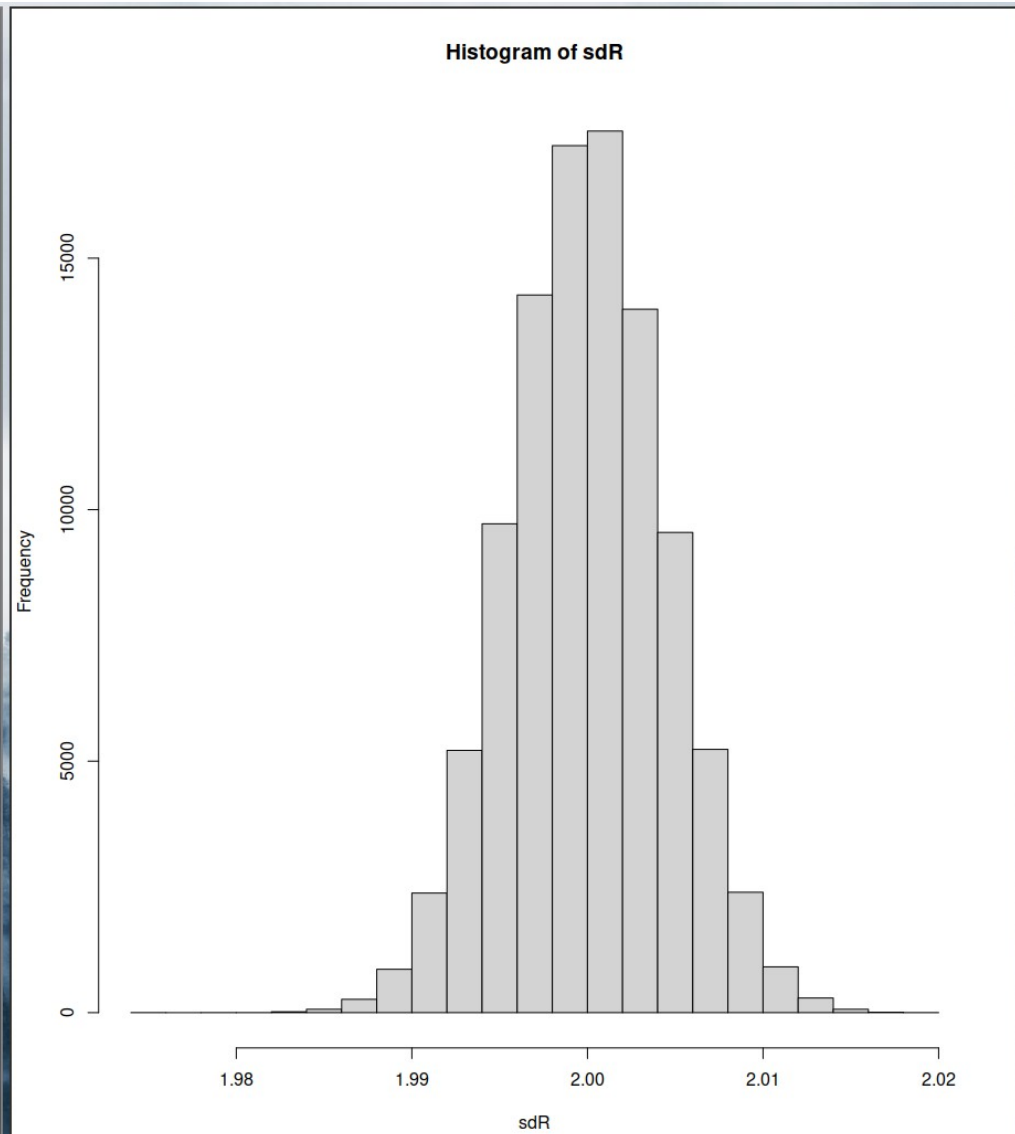
- Μπορούμε να αυτοματοποιήσουμε τη διαδικασία με την R και την εντολή *replicate*
  - `replicate(100, sd(rnorm(10, mean = 10, sd = 2)))`
  - Με την εντολή αυτή επαναλαμβάνουμε μια δειγματοληψία **100** φορές σε έναν πληθυσμό που η μεταβλητή μας ακολουθεί την κανονική κατανομή και έχει μέσο 10 και τυπική απόκλιση 2, και υπολογίζουμε την τυπική απόκλιση για κάθε δείγμα

```
> R<-100000
> sdR <- replicate(R, sd(rnorm(n, mean = 10, sd = 2)))
>
_
```

- Για 100000 επαναλήψεις θέλετε δυνατό Η/Υ και χρόνο...

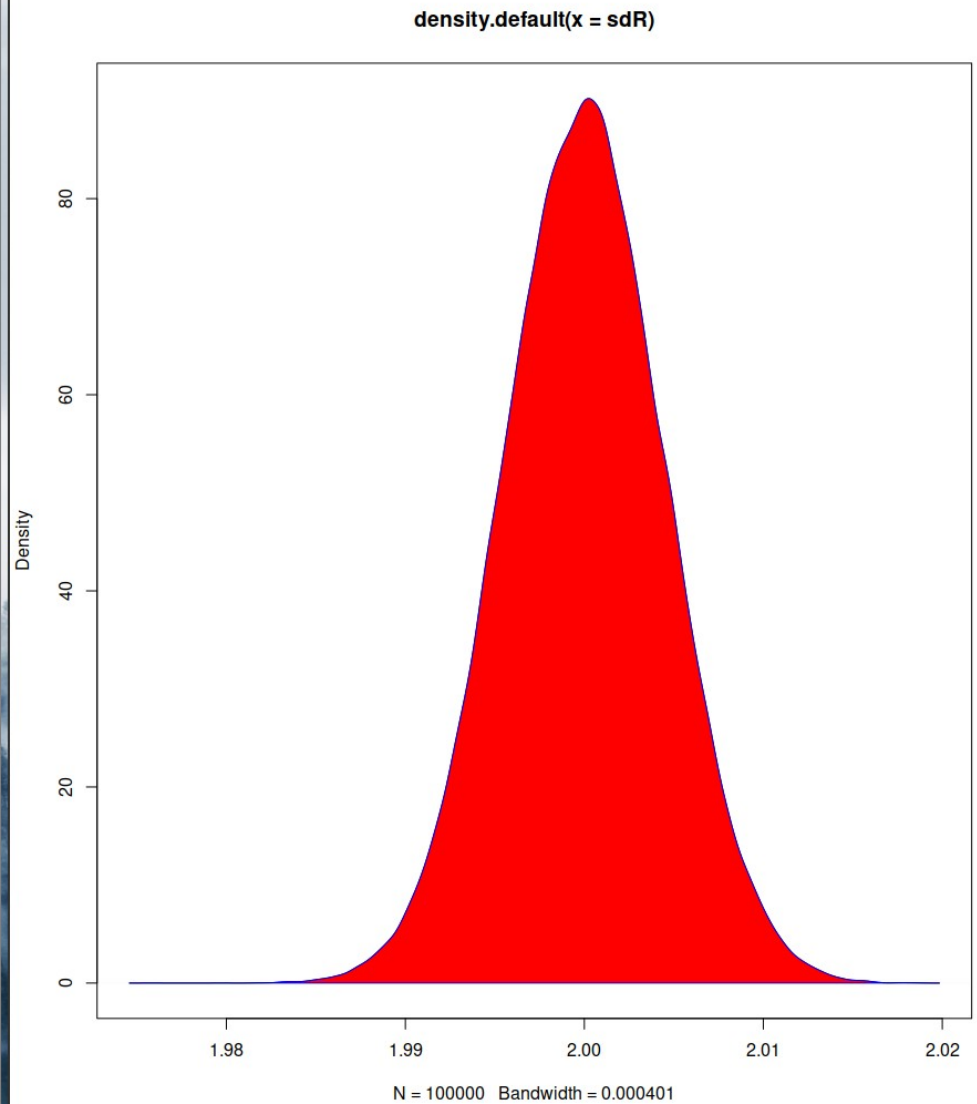
```
> hist(sdR)
>
_
```

ιστόγραμμα



```
> d<-density(sdR)
> plot(d)
> polygon(d, col="red", border="blue")
>
```

## Διάγραμμα πυκνότητας



# Άσκηση #2

- Έχουμε μια μεταβλητή σε έναν πληθυσμό με μέσο όρο 100 και τυπική απόκλιση 22, που ακολουθεί την κανονική κατανομή.
  - Κάντε 10000 δειγματοληψίες με μέγεθος δείγματος **10** και υπολογίστε τον **μέσο όρο** των δειγμάτων:
    - Ποιος είναι ο μέσος όρος των μέσων όρων των δειγμάτων;
    - Ποια είναι η διακύμανση των μέσων όρων των δειγμάτων;
    - Κατασκευάστε ένα ιστόγραμμα και ένα γράφημα πυκνότητας για τους μέσους όρους των δειγμάτων
  - Επαναλάβετε όλη τη διαδικασία με μέγεθος δείγματος **1000**.
  - Συγκρίνετε τις τιμές (μέσο όρο και διακύμανση των μέσων των δειγμάτων) για τα δύο διαφορετικά μεγέθη δείγματος σε έναν πίνακα.
  - Τι συμπέρασμα βγάξετε;

# Κατανομή δειγματοληψίας ενός στατιστικού

- Είναι η κατανομή που παίρνουμε από τον υπολογισμό ενός στατιστικού σε έναν άπειρο αριθμό δειγμάτων μεγέθους  $n$ 
  - Μια τέτοια κατανομή ορίζει την πιθανότητα εμφάνισης μιας συγκεκριμένης τιμής ενός στατιστικού (π.χ. του μέσου)
- Προσοχή: μπορεί να είναι πολύ διαφορετική από την κατανομή των δεδομένων από τα οποία υπολογίζουμε το στατιστικό
- Ο μέσος είναι μια εκτίμηση του μέσου του αρχικού πληθυσμού από όπου πήραμε τα δείγματα
- Η τυπική απόκλιση είναι γνωστή και ως **τυπικό σφάλμα** (standard error of the mean - SEM)



# Το θεώρημα του κεντρικού ορίου

- The central limit theorem (**CLT**)
- Δηλώνει πως η κατανομή της δειγματοληψίας ενός στατιστικού προσεγγίζει την κανονική κατανομή, όσο το μέγεθος του δείγματος ( $n$ ) προσεγγίζει το άπειρο
  - Ισχύει για τον μέσο και τη διακύμανση

# Το θεώρημα του κεντρικού ορίου

- Ο De Moivre το 1733 χρησιμοποίησε την κανονική κατανομή για να προσεγγίσει τη μέση εμφάνιση “κεφαλής” στη ρίζη νομισμάτων
  - Αν και ακολουθεί τη διωνυμική κατανομή
- Ο Laplace βρίσκει την απόδειξη το 1809

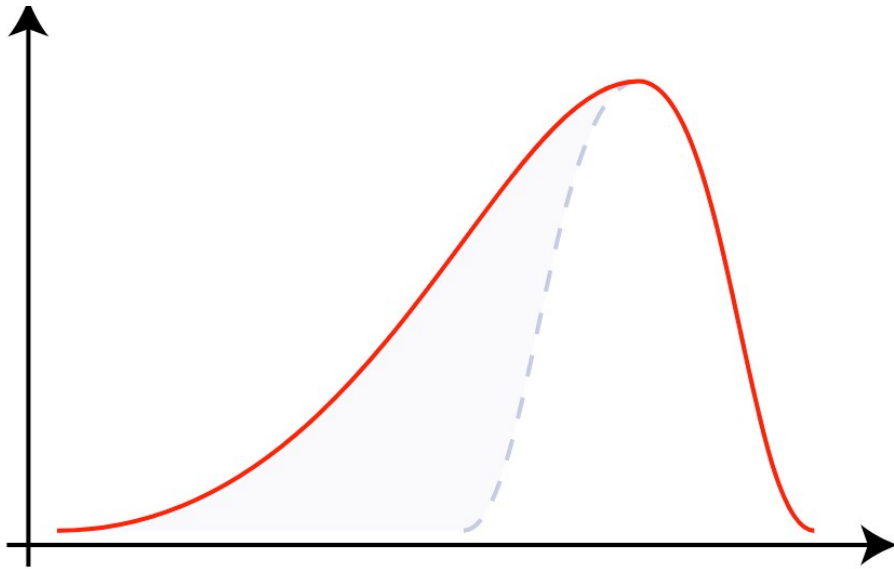


# Πρακτική σημασία του CLT

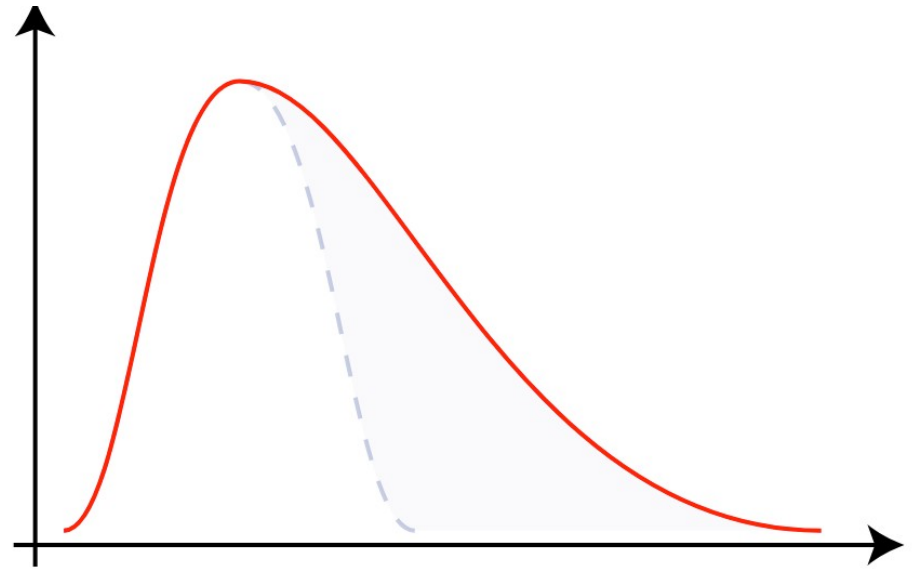
- Εάν το μέγεθος του δείγματος είναι αρκετά **μεγάλο**, τότε η κατανομή δειγματοληψίας του στατιστικού θα είναι περίπου **κανονική**
  - ανεξάρτητα από την κατανομή του πρωτοτύπου πληθυσμού
- Αλλά ποιο μέγεθος δείγματος είναι αρκετά μεγάλο;
  - Αν η αρχική κατανομή του πληθυσμού είναι κανονική, τότε και ένα δείγμα με μέγεθος  $n=1$  αρκεί
  - Όσο απομακρύνεται ο αρχικός πληθυσμός από την κανονική κατανομή, τόσο πιο μεγάλο πρέπει να είναι το δείγμα

# Ασύμμετρες κατανομές

- Οι ασύμμετρες κατανομές (skewed) μπορεί να είναι αρνητικά ή θετικά ασύμμετρες



Negative Skew



Positive Skew

# R distribution functions

	Distribution function (pmf or cdf)	Cumulative distr. function	Random sampling
<b>Binomial</b>	<code>dbinom(x, n, P)</code>	<code>pbinom(x, n, P)</code>	<code>rbinom(size, n, P)</code>
<b>Uniform</b>	<code>dunif(x)</code>	<code>punif(x)</code>	<code>runif(size)</code>
<b>Normal</b>	<code>dnorm(x, mean, sd)</code>	<code>pnorm(x, mean, sd)</code>	<code>rnorm(size, mean, sd)</code>
<b>Log-normal</b>	<code>dlnorm(x, meanlog, sdlog)</code>	<code>plnorm(x, meanlog, sdlog)</code>	<code>rlnorm(size, meanlog, sdlog)</code>
<b>chi-squared</b>	<code>dchisq(x, k)</code>	<code>pchisq(x, k)</code>	<code>rchisq(size, k)</code>

# Οδηγίες για τις ασκήσεις αυτού του κεφαλαίου

- Οι ασκήσεις δεν είναι υποχρεωτικές, όμως δίνουν **bonus** για τον τελικό βαθμό του μαθήματος.
- Μπορείτε να κάνετε τις ασκήσεις σε ομάδες των 2 ή 3 ατόμων.
  - Όχι ατομικά και όχι πάνω από τρία άτομα στην ομάδα!
- Οι δύο αυτές ασκήσεις θα μπουν σε ένα αρχείο **.pdf** που θα περιέχει τον κώδικα, τα γραφήματα και μια περιγραφή της διαδικασίας. Το αρχείο θα έχει τίτλο τα επώνυμα των μελών της ομάδας.
- Κάθε ομάδα θα ανεβάσει μία φορά το αρχείο με τις λυμένες ασκήσεις, από ένα μέλος της. Το αρχείο θα ανέβει στις “Εργασίες” του eclass του μαθήματος, στην **Εργασία Α**.
  - Δεν θα δεχτούμε ασκήσεις με email
- *Αυστηρή προθεσμία: Κυριακή 14 Απριλίου!*

**Thank you**

