# Distribution models

3

# Outline

1. Discrete distributions
   - Binomial distribution
   - Geometric distribution
   - Poisson distribution

2. Continuous distributions
   - Uniform distribution
   - Exponential distribution
   - Normal distribution
     - Central Limit Theorem

# Binomial distribution

A random experiment consists of $n$ trials such that: The trials are independent. Each trial results either in success or failure. The probability of a success, $p$, remains constant.

The random variable that equals the number of trials that result in a success follows a binomial distribution with parameters $0 < p < 1$ and $n = 1, 2, \ldots$

$$X \sim \mathrm{B}(n, p)$$

$X \equiv$ number of (indep.) trials that result in a success

# Binomial distribution

- Probability mass function:

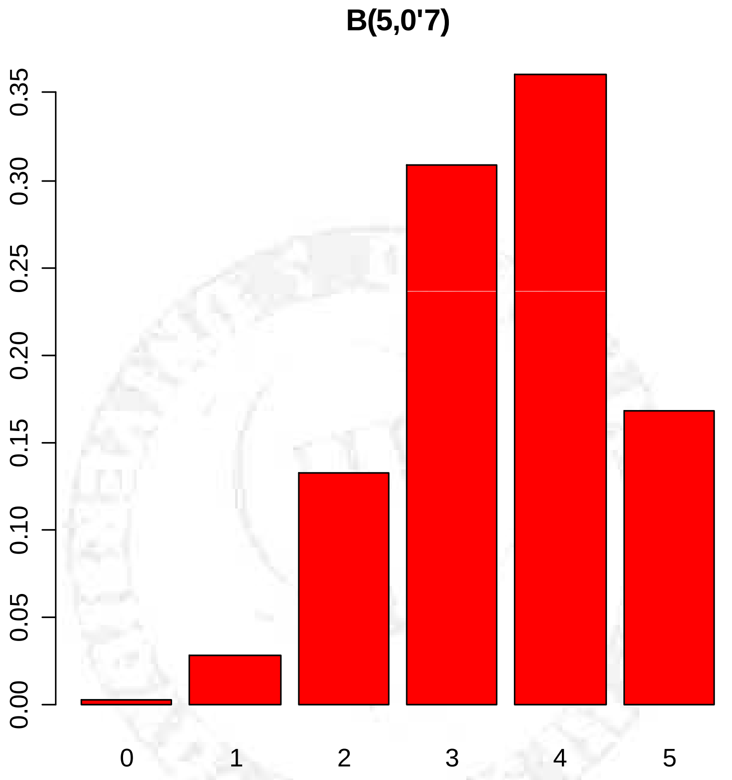
$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{0,1,\dots,n\}.$$

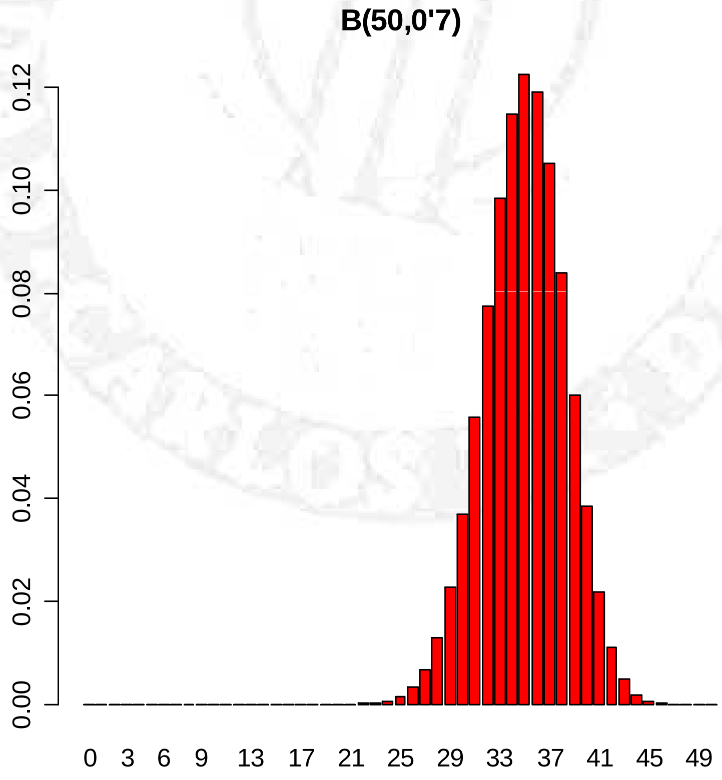
It is possible to write $X = X_1 + \dots + X_n$ where $X_i \sim B(1, p)$ independent random variables.

- Parameters: E[$X$] = $np$ ; Var[$X$] = $np(1-p)$

If $X \sim B(n_1, p)$ and $Y \sim B(n_2, p)$ are indep, $X+Y \sim B(n_1+n_2, p)$

# Binomial distribution



# Example (Four flamates and a die)

Four flatmates Alice, Bob, Charly, and Dave roll a 6-sided die every night in order to decide who washes the dishes after dinner. If the outcome is 1, Alice washes, it is its 2, Bob does, while for 3 or 4 Charly must wash the dishes and for 5 or 6 it is Dave's turn.

a) What is the probability that Charly washes the dishes at most twice in a week (7 days)?

b) How many days (dinners) must we wait on average until it is Alice's turn to wash the dishes?

# Outline

1. Discrete distributions
   - Binomial distribution
   - Geometric distribution
   - Poisson distribution

2. Continuous distributions
   - Uniform distribution
   - Exponential distribution
   - Normal distribution
     - Central Limit Theorem

# Geometric distribution

In a series of independent trials with constant probability of a success, $0<p<1$, the random variable denoting the number of trials until the first success follows a geometric distribution with parameter $p$

$$X \sim \mathrm{G}(p)$$

$X \equiv$ number of (indep) trials until the first success

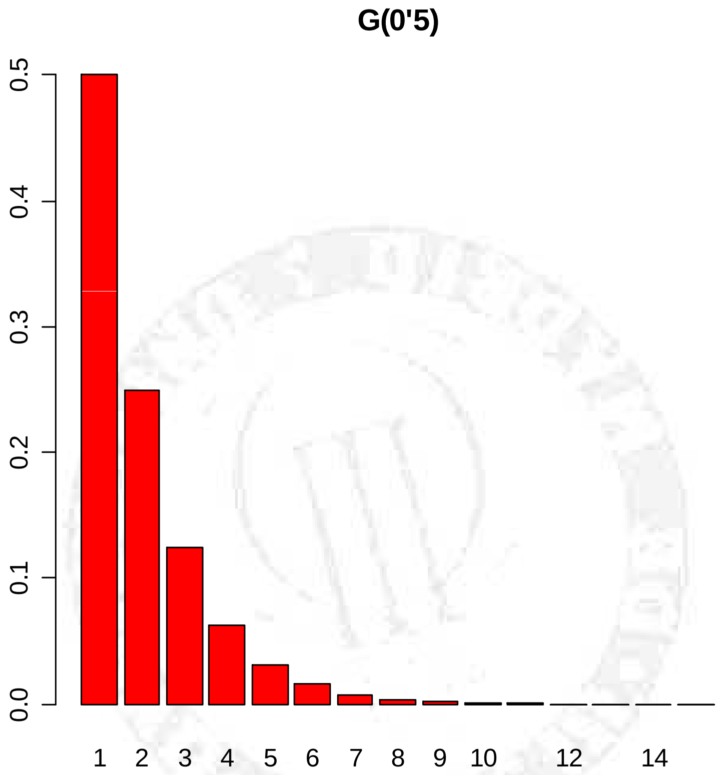
# Geometric distribution

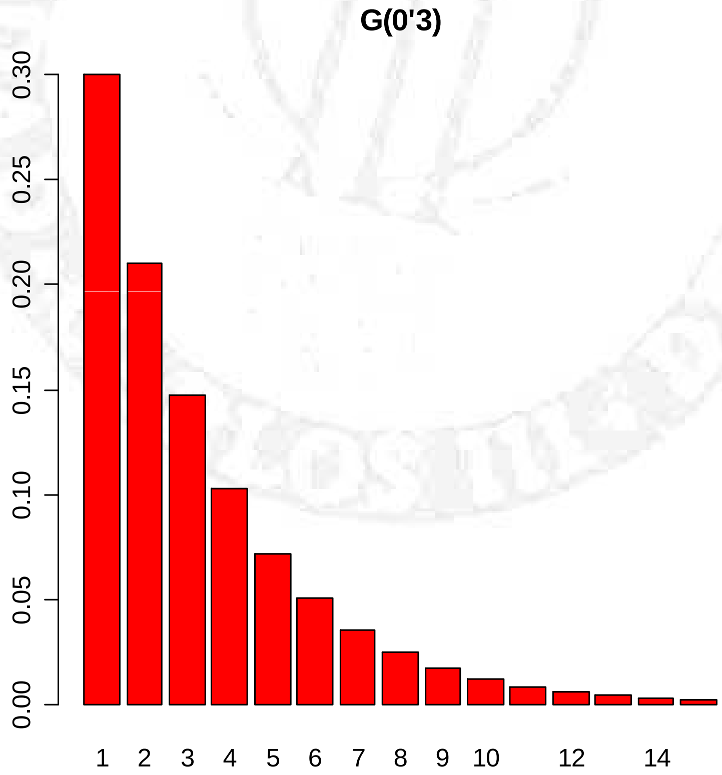
- Probability mass function:

$$P(X=k)=(1-p)^{k-1}p,\quad k\in\{1,2,3,\ldots\}.$$

- Parameters: $\mathrm{E}[X] = 1/p$ ; $\mathrm{Var}[X] = (1-p)/p^2$

# Geometric distribution



# Example (Four flamates and a die)

Four flatmates Alice, Bob, Charly, and Dave roll a 6-sided die every night in order to decide who washes the dishes after dinner. If the outcome is 1, Alice washes, it is its 2, Bob does, while for 3 or 4 Charly must wash the dishes and for 5 or 6 it is Dave's turn.

a) What is the probability that Charly washes the dishes at most twice in a week (7 days)?

b) How many days (dinners) must we wait on average until it is Alice's turn to wash the dishes?

# Outline

1. Discrete distributions
   - Binomial distribution
   - Geometric distribution
   - Poisson distribution

2. Continuous distributions
   - Uniform distribution
   - Exponential distribution
   - Normal distribution
     - Central Limit Theorem

# Poisson distribution

Assume that certain events occur in a fixed interval of real numbers (period of time, area, volume,…) with a known average rate $\lambda>0$ and independently one from the others.

The random variable that equals the number of events occurring in the interval follows a Poisson distribution with parameter $\lambda$,

$$X \sim \wp(\lambda)$$

$X \equiv$ number of events in the interval

# Poisson distribution

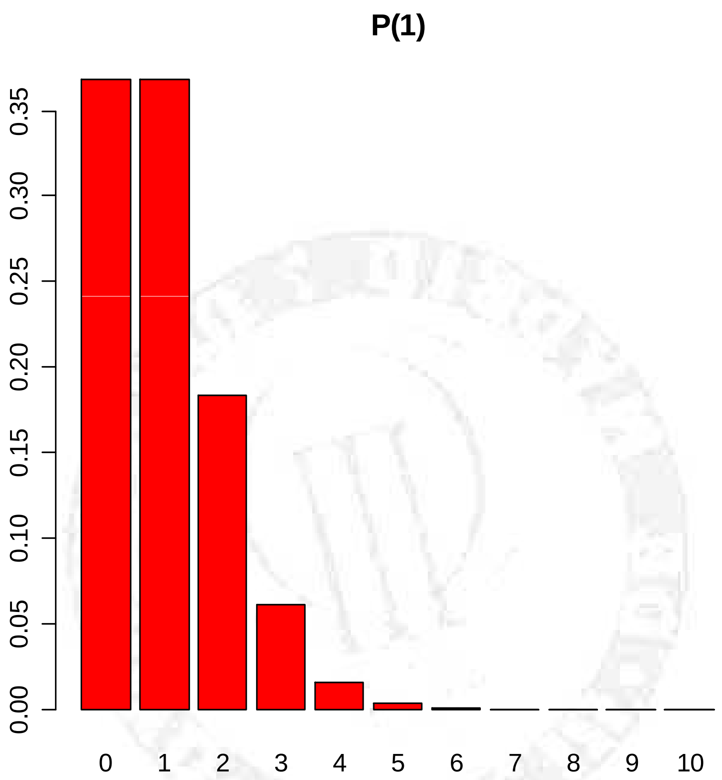
- Probability mass function:

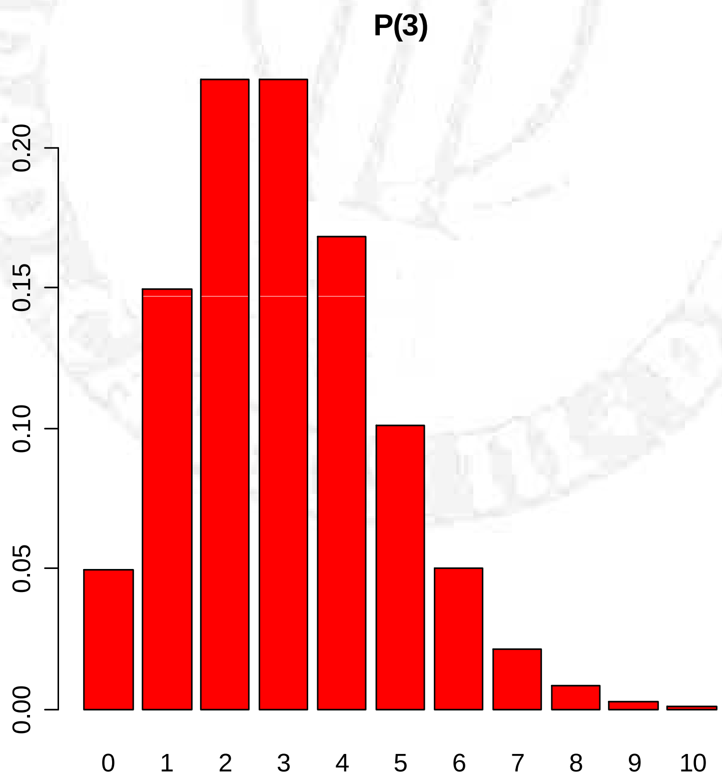
$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!},\quad k\in\{0,1,2,\ldots\}.$$

- Parameters: $E[X]=\lambda$ ; $Var[X]=\lambda$

If $X\sim\wp(\lambda_1)$ and $Y\sim\wp(\lambda_2)$ are indep, $X+Y\sim\wp(\lambda_1+\lambda_2)$

# Poisson distribution



# Example (radioactive material)

A sample of radioactive material emits, on average, 15 alpha particles per minute. If the number of alpha particles emitted follows a Poisson distribution, what is the probability of 10 alpha particles being emitted in:

a) 1 minute ?

b) 2 minutes ?

c) Many years later, the material averages 6 alpha particles emitted per min. What is the probability of at least 6 alpha particles being emitted in 1 minute?

# Outline

1. Discrete distributions
   - Binomial distribution
   - Geometric distribution
   - Poisson distribution

2. Continuous distributions
   - Uniform distribution
   - Exponential distribution
   - Normal distribution
     - Central Limit Theorem

# (Continuous) Uniform distribution

A random variable uniformly distributed in the interval $(a,b)$ represents a number chosen at random between $a$ and $b$. The selection is made in such a way that the probability that the random variable lays in any interval inside $(a,b)$ depends only on the length of such interval,

$$X \sim \mathrm{U}(a,b)$$

# (Continuous) Uniform distribution

- Density mass function:

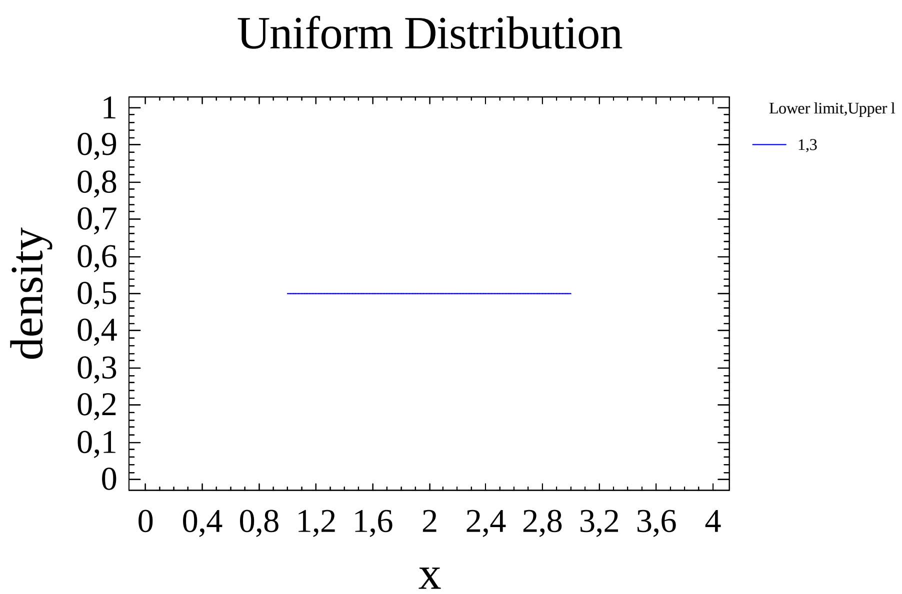
$$f(x)=\begin{cases}\frac{1}{b-a} & if \quad x\in(a,b)\\ 0 & if \quad x\notin(a,b)\end{cases}$$

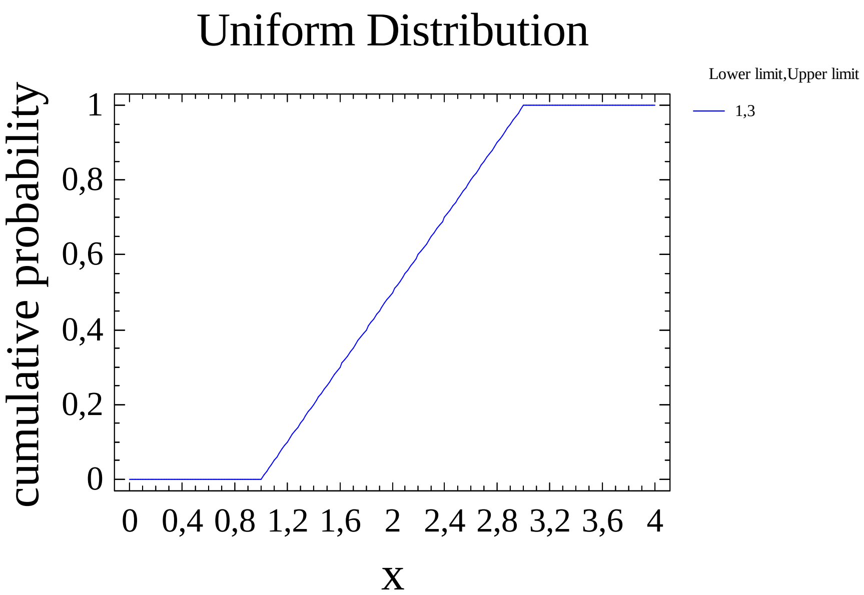
- Cumulative distribution function:

$$F(x)=\begin{cases}0 & if \quad x\le a\\ \frac{x-a}{b-a} & if \quad a<x<b\\ 1 & if \quad x\ge b\end{cases}$$

- Parameters: $E[X] = (a+b)/2$ ; $Var[X] = (b-a)^2/12$

# (Continuous) Uniform distribution



# Outline

1. Discrete distributions
   - Binomial distribution
   - Geometric distribution
   - Poisson distribution

2. Continuous distributions
   - Uniform distribution
   - Exponential distribution
   - Normal distribution
     - Central Limit Theorem

# Exponential distribution

The random variable that equals the distance between successive events in a Poisson process with mean $\lambda>0$ follows an exponential distribution with parameter $\lambda$,

$$X \sim \mathrm{Exp}(\lambda)$$

$X \equiv$ distance between successive events

# Exponential distribution

- Density mass function:

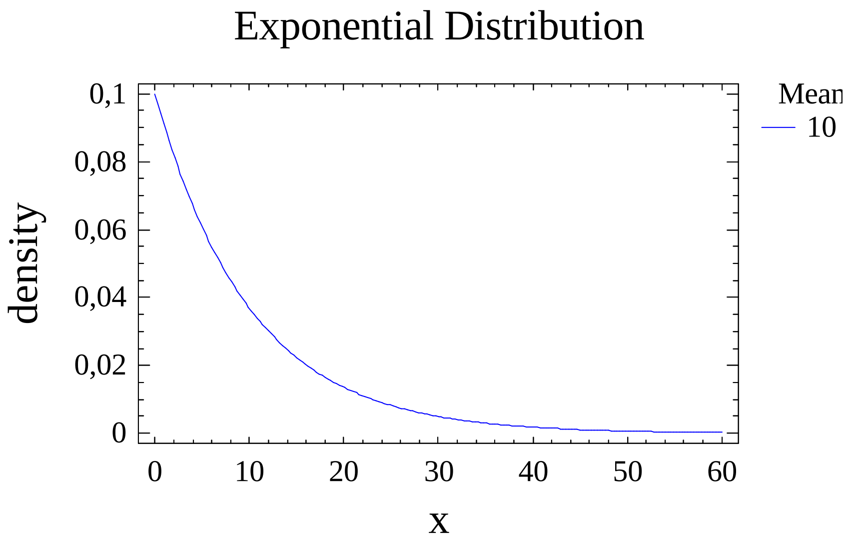
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & if \quad x > 0 \\ 0 & if \quad x \le 0 \end{cases}$$

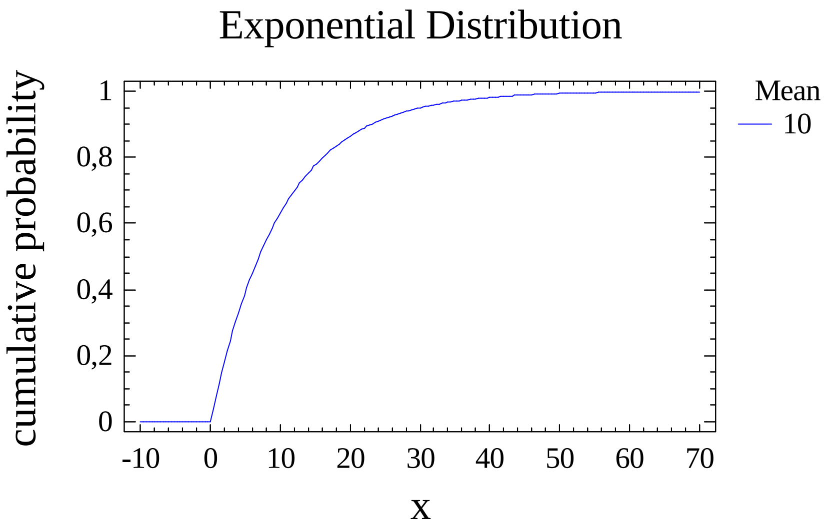
- Cumulative distribution function:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & if \quad x > 0 \\ 0 & if \quad x \le 0 \end{cases}$$

- Parameters: $\mathrm{E}[X] = \lambda^{-1}$ ; $\mathrm{Var}[X] = \lambda^{-2}$

# Exponential distribution



# Exponential distribution

- Lack of memory property.

  For an exponential random variable $T$,

  given $t_1$,$t_2$>0

$$P(T > t_1+t_2 \mid T > t_1) = P(T > t_2)$$

# Example (radioactive material)

On average, a sample of radioactive material emits 15 alpha particles per minute.

a) What is the average time between the emission of two alpha particles?

b) What is the probability that the time between the emission of two alpha particles is longer than 10 sec?

c) Last alpha particle was emitted 10 seconds ago. What is the probability that it still takes longer than 10 seconds until the next particle is emitted?

# Outline

1. Discrete distributions
   - Binomial distribution
   - Geometric distribution
   - Poisson distribution

2. Continuous distributions
   - Uniform distribution
   - Exponential distribution
   - Normal distribution
     - Central Limit Theorem

# Normal distribution

The most widely used model for the distribution of a random variable is a normal (or Gaussian) distribution. Apart from other relevant properties, it appears as the limit distribution in the Central Limit Theorem. A normal distribution is determined by two parameters, the mean $\mu$ and the standard deviation $\sigma>0$,

$$X \sim \mathrm{N}(\mu,\sigma)$$

# Normal distribution

- Standard normal density mass function N(0,1):

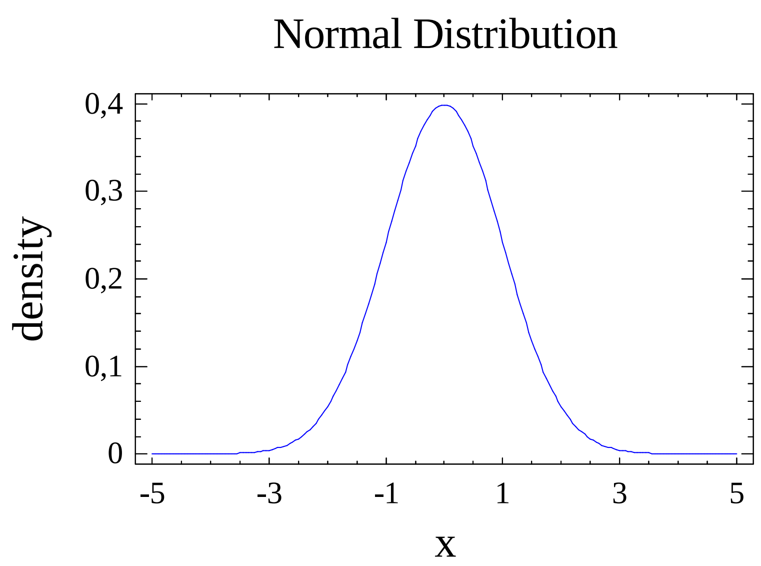
$$f(x)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}$$

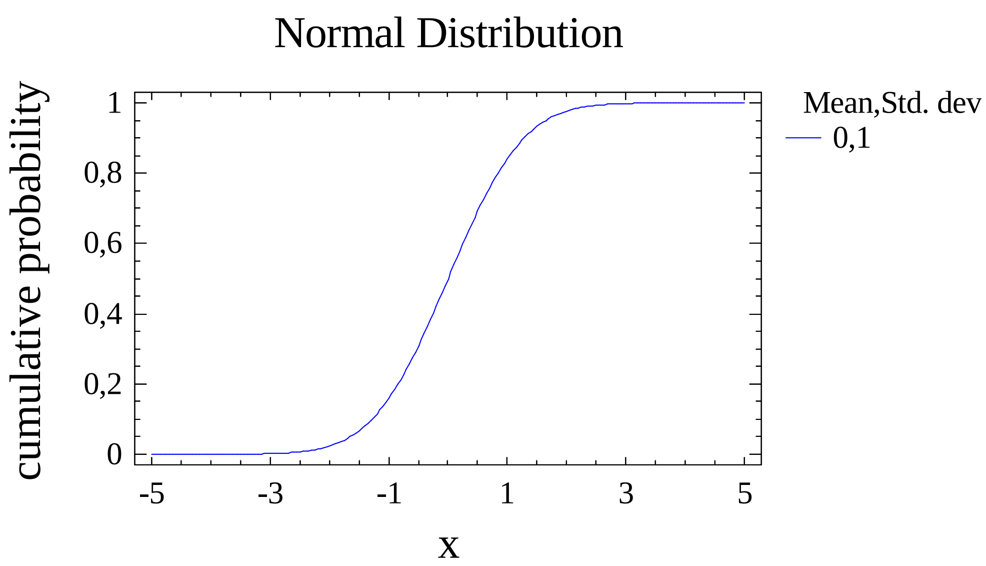
- Density mass function N(μ,σ):

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

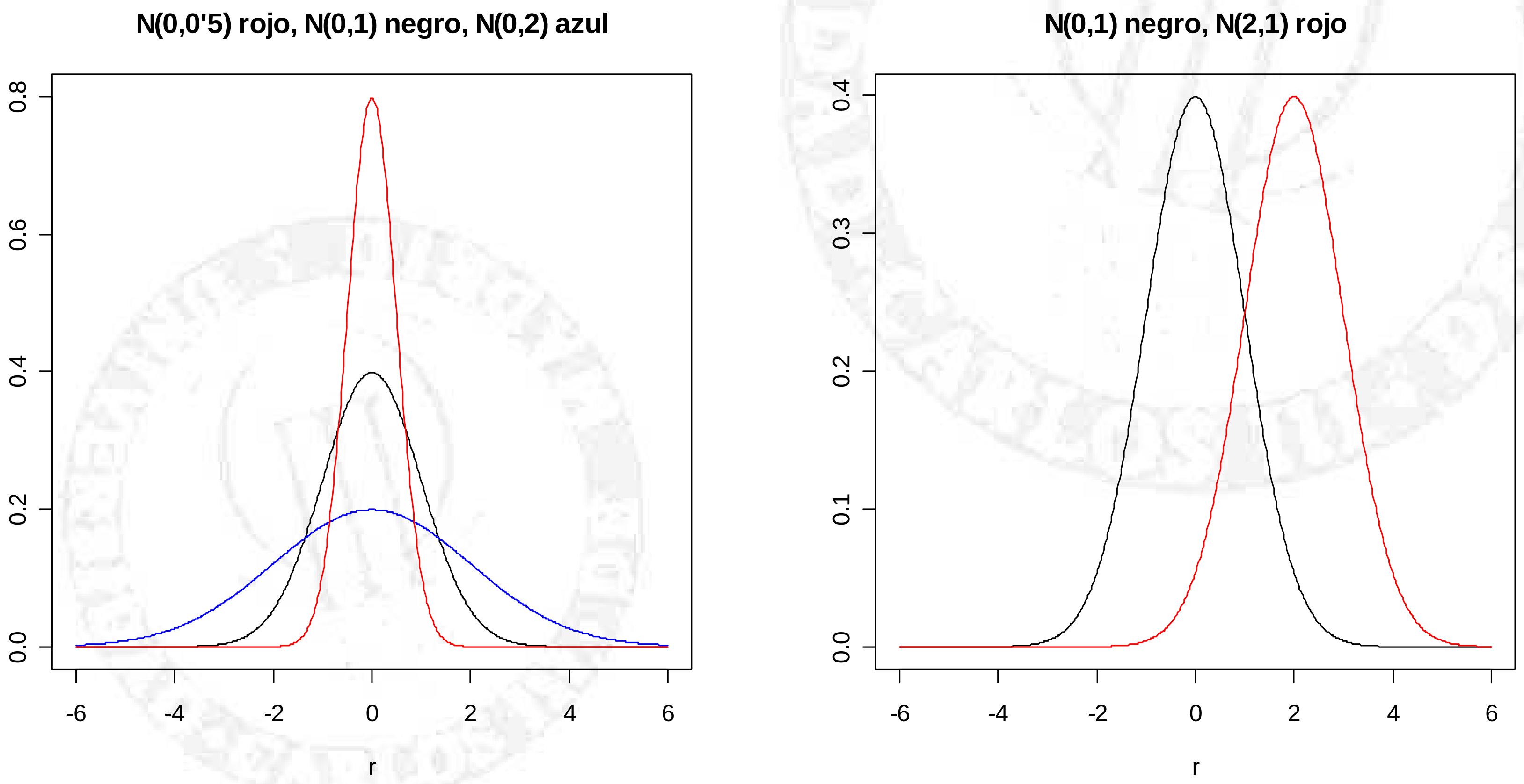
- Parameters: $E[X]=\mu$ ; $\mathrm{Var}[X]=\sigma^2$

# Normal distribution



# Normal distribution



# Normal distribution

- Properties.

1. If $X \sim N(\mu,\sigma)$ , for every $a$ and $b$,

$$aX+b \sim N(a\mu+b\ ,\ |a|\sigma)$$

2. If $X \sim N(\mu_1,\sigma_1)$ , $Y \sim N(\mu_2,\sigma_2)$ indep, for $a$, $b$

$$aX+bY \sim N(a\mu_1+b\mu_2\ ,\ (a^2\sigma_1{}^2+b^2\sigma_2{}^2)^{1/2})$$

- Standardization. Given $X$~N($\mu,\sigma$), the random variable $(X-\mu)/\sigma$ follows a standard normal distribution, N(0,1).

# Table for the N(0,1) cdf

Table entry for $z$ is the area under the standard normal curve to the left of $z$.

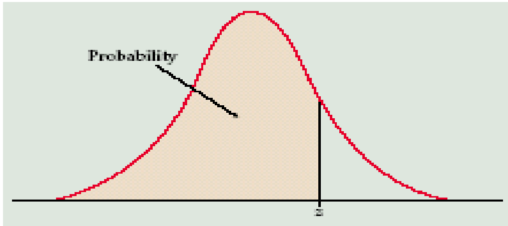



TABLE A Standard normal probabilities (*continued*)

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .5000 | .5040 | .5080 | .5120 | .5160 | .5199 | .5239 | .5279 | .5319 | .5359 |
| 0.1 | .5398 | .5438 | .5478 | .5517 | .5557 | .5596 | .5636 | .5675 | .5714 | .5753 |
| 0.2 | .5793 | .5832 | .5871 | .5910 | .5948 | .5987 | .6026 | .6064 | .6103 | .6141 |
| 0.3 | .6179 | .6217 | .6255 | .6293 | .6331 | .6368 | .6406 | .6443 | .6480 | .6517 |
| 0.4 | .6554 | .6591 | .6628 | .6664 | .6700 | .6736 | .6772 | .6808 | .6844 | .6879 |
| 0.5 | .6915 | .6950 | .6985 | .7019 | .7054 | .7088 | .7123 | .7157 | .7190 | .7224 |
| 0.6 | .7257 | .7291 | .7324 | .7357 | .7389 | .7422 | .7454 | .7486 | .7517 | .7549 |
| 0.7 | .7580 | .7611 | .7642 | .7673 | .7704 | .7734 | .7764 | .7794 | .7823 | .7852 |
| 0.8 | .7881 | .7910 | .7939 | .7967 | .7995 | .8023 | .8051 | .8078 | .8106 | .8133 |
| 0.9 | .8159 | .8186 | .8212 | .8238 | .8264 | .8289 | .8315 | .8340 | .8365 | .8389 |
| 1.0 | .8413 | .8438 | .8461 | .8485 | .8508 | .8531 | .8554 | .8577 | .8599 | .8621 |
| 1.1 | .8643 | .8665 | .8686 | .8708 | .8729 | .8749 | .8770 | .8790 | .8810 | .8830 |
| 1.2 | .8849 | .8869 | .8888 | .8907 | .8925 | .8944 | .8962 | .8980 | .8997 | .9015 |
| 1.3 | .9032 | .9049 | .9066 | .9082 | .9099 | .9115 | .9131 | .9147 | .9162 | .9177 |
| 1.4 | .9192 | .9207 | .9222 | .9236 | .9251 | .9265 | .9279 | .9292 | .9306 | .9319 |
| 1.5 | .9332 | .9345 | .9357 | .9370 | .9382 | .9394 | .9406 | .9418 | .9429 | .9441 |
| 1.6 | .9452 | .9463 | .9474 | .9484 | .9495 | .9505 | .9515 | .9525 | .9535 | .9545 |
| 1.7 | .9554 | .9564 | .9573 | .9582 | .9591 | .9599 | .9608 | .9616 | .9625 | .9633 |
| 1.8 | .9641 | .9649 | .9656 | .9664 | .9671 | .9678 | .9686 | .9693 | .9699 | .9706 |
| 1.9 | .9713 | .9719 | .9726 | .9732 | .9738 | .9744 | .9750 | .9756 | .9761 | .9767 |
| 2.0 | .9772 | .9778 | .9783 | .9788 | .9793 | .9798 | .9803 | .9808 | .9812 | .9817 |
| 2.1 | .9821 | .9826 | .9830 | .9834 | .9838 | .9842 | .9846 | .9850 | .9854 | .9857 |
| 2.2 | .9861 | .9864 | .9868 | .9871 | .9875 | .9878 | .9881 | .9884 | .9887 | .9890 |
| 2.3 | .9893 | .9896 | .9898 | .9901 | .9904 | .9906 | .9909 | .9911 | .9913 | .9916 |
| 2.4 | .9918 | .9920 | .9922 | .9925 | .9927 | .9929 | .9931 | .9932 | .9934 | .9936 |
| 2.5 | .9938 | .9940 | .9941 | .9943 | .9945 | .9946 | .9948 | .9949 | .9951 | .9952 |
| 2.6 | .9953 | .9955 | .9956 | .9957 | .9959 | .9960 | .9961 | .9962 | .9963 | .9964 |
| 2.7 | .9965 | .9966 | .9967 | .9968 | .9969 | .9970 | .9971 | .9972 | .9973 | .9974 |
| 2.8 | .9974 | .9975 | .9976 | .9977 | .9977 | .9978 | .9979 | .9979 | .9980 | .9981 |
| 2.9 | .9981 | .9982 | .9982 | .9983 | .9984 | .9984 | .9985 | .9985 | .9986 | .9986 |
| 3.0 | .9987 | .9987 | .9987 | .9988 | .9988 | .9989 | .9989 | .9989 | .9990 | .9990 |
| 3.1 | .9990 | .9991 | .9991 | .9991 | .9992 | .9992 | .9992 | .9992 | .9993 | .9993 |
| 3.2 | .9993 | .9993 | .9994 | .9994 | .9994 | .9994 | .9994 | .9995 | .9995 | .9995 |
| 3.3 | .9995 | .9995 | .9995 | .9996 | .9996 | .9996 | .9996 | .9996 | .9996 | .9997 |
| 3.4 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9998 |

# Example (Capacitors)

A machine makes capacitors with a mean value of 25 μF and a standard deviation of 6 μF. Assuming that capacitance follows a Gaussian distribution, find the probability that the value of capacitance exceeds 31 μF.

# Central Limit Theorem

Given $n$ independent random variables $X_1, X_2, \ldots, X_n$, with finite means and variances $E[X_i]=\mu_i$ and $Var[X_i]=\sigma_i^2$, the limiting distribution ($n \to \infty$) of their sum is normal

$$X_1+X_2+\ldots+X_n \approx N\left(\Sigma_{i=1,n}\mu_i\,,\,(\Sigma_{i=1,n}\sigma_i^2)^{1/2}\right)$$

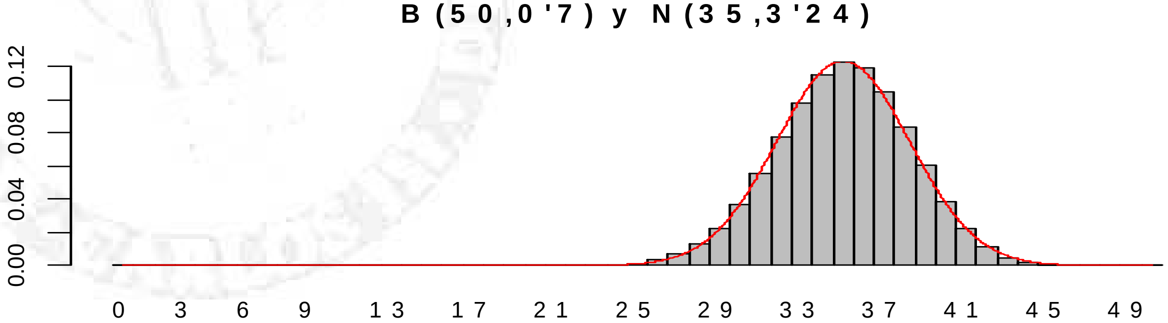
The approximation is usually good for $n > 30$.

If the variables are discrete, we will use a correction factor called continuity correction.

# Normal approximations

- **Normal approximation to the Binomial distribution.** A Binomial distribution B($n$,$p$) with $n > 30$ and $np(1-p) > 5$, is approximately

$$\mathrm{N}\left(np, \sqrt{np(1-p)}\right)$$



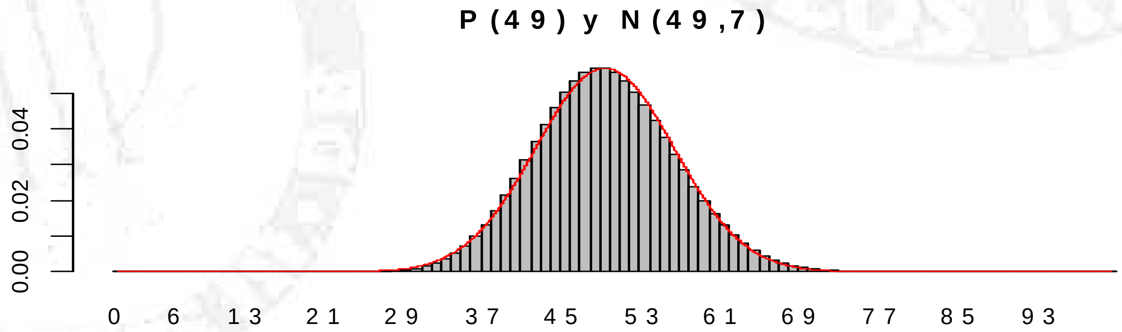
# Example (Four flamates and a die)

Four flatmates Alice, Bob, Charly, and Dave roll a 6-sided die every night in order to decide who washes the dishes after dinner. If the outcome is 1, Alice washes, it is its 2, Bob does, while for 3 or 4 Charly must wash the dishes and for 5 or 6 it is Dave's turn.

c) How many days (dinners) must we wait until Bob washes the dishes at least 11 times with probability 0.95?

# Normal approximations

- Normal approximation to the Poisson distribution. A Poisson distribution $\wp(\lambda)$ with $\lambda>5$ is approximately $N(\lambda, \lambda^{1/2})$



# Example (radioactive material)

On average, a sample of radioactive material emits 15 alpha particles per minute. What is the approximate probability of 10 alpha particles being emitted in:

a) 1 minute ?

b) Many years later, the material averages 6 alpha particles emitted per min. What is the probability of at least 6 alpha particles being emitted in 1 minute?