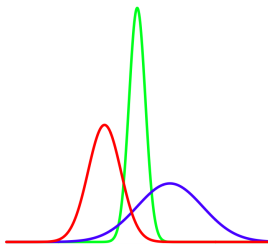


# Common Families of Distributions



# Some known distributions

- Discrete distributions

- Bernoulli
- Binomial
- Geometric
- Poisson

- Continuous distributions

- Uniform
- Exponential
- Normal (Gaussian)
  
- $\chi^2$ -distribution
- $t$ -distribution

...will be introduced later in the course ..  
... also important in Statistics

# Bernoulli distribution

- $X \sim \text{ber}(p)$ 
  - Models one trial in an experiment that can result in either **success** or **failure**.
    - .. The associated  $X$  is only an indicator for the experiment  
Example: tossing a fair coin.
  - A random variable  $X$  has a Bernoulli distribution with parameter  $p$  if:
    - $X$  takes the values **1** (for success) and **0** (for failure).
    - $P(X = 1) = p$  and  $P(X = 0) = 1 - p = q$

	$x$	0	1
pmf	$p(x)$	$1 - p$	$p$
cdf	$F(x)$	$1 - p$	1

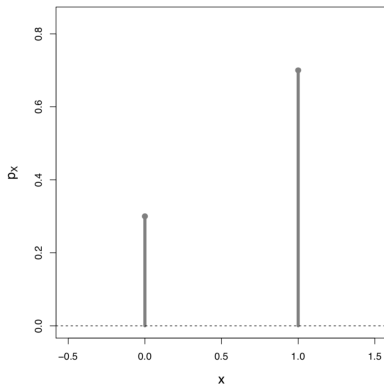
- Expectation/Variance

$$\mathbb{E}(X) = p, \quad \text{Var}(X) = p(1 - p)$$

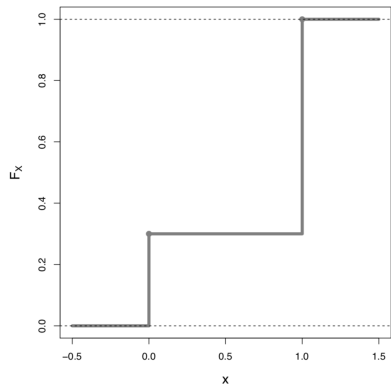
... HW Check!

# Example

- $X \sim \text{ber}(0.7)$



Probability mass function



Cumulative distribution function

# Binomial distribution

- $X \sim B(n, p)$  or  $X \sim b(n, p)$ 
  - The **binomial distribution** models the **number** of successes in  $n$  independent **ber( $p$ )** trials.

...  $B(1, p)$  is the same as  $ber(p)$

Example: The **number of heads in  $n$  flips** of a coin with probability  $p$  of heads follows a  $B(n, p)$  distribution

- $X$  follows a binomial distribution if
  - $X$  takes the values  $0, 1, 2, 3, \dots, n$
  - its probability mass function is given by

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

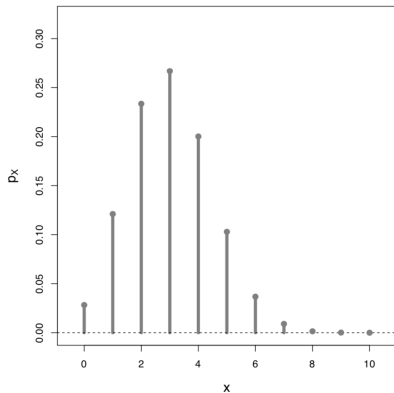
... by the Binomial formula we get

$$\sum_{x=0}^n p(x) = \sum_{x=0}^n \binom{n}{x} p^x \cdot (1 - p)^{n-x} = (p + (1 - p))^n = 1$$

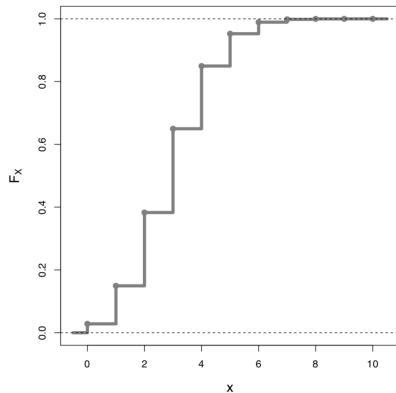
- Expectation/Variance:  $\mathbb{E}(X) = np, \quad \text{Var}(X) = np(1 - p)$

# Example

- $X \sim B(10, 0.3)$



Probability mass function



Cumulative distribution function

# Example

## Dice probabilities

(a) Find the probability of obtaining at least one 6 in four rolls of a die.

- This experiment can be modeled as a sequence of four Bernoulli trials with success probability  $p = \frac{1}{6} = P(\text{die shows 6})$ .

Define the random variable

$X = \text{Total number of sixes in four rolls}$

- Then,  $X \sim B(4, \frac{1}{6})$  and

$$\begin{aligned} P(\text{at least one 6}) &= P(X \geq 1) = 1 - P(X = 0) \\ &= 1 - \binom{4}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^4 \approx 0.518 \end{aligned}$$

- In R: use `dbinom()`

```
1-dbinom(0,4,1/6)
[1] 0.5177469
```

# Example

(b) Consider another game; throw a pair of dice 24 times and ask for the probability of at least two double 6s.

- This experiment can be modeled by the binomial distribution with success probability  $p$ , with  $p = P(\text{roll two sixes}) = \frac{1}{36}$ .
- If  $Y = \text{number of double 6s in 24 rolls}$ , then  $Y \sim B(24, \frac{1}{36})$  and

$$\begin{aligned} P(\text{at least two double sixes}) &= P(Y \geq 2) = 1 - P(Y = 0) - P(Y = 1) \\ &= 1 - \binom{24}{0} \left(\frac{1}{36}\right)^0 \left(\frac{35}{36}\right)^{24} - \binom{24}{1} \left(\frac{1}{36}\right)^1 \left(\frac{35}{36}\right)^{23} \approx 0.1427 \end{aligned}$$

- In R: use `dbinom()`

```
1 - sum(dbinom(0:1, 24, 1/36))  
[1] 0.1426522
```

or `pbinom()`

```
pbinom(1, 24, 1/36, lower.tail = FALSE)  
[1] 0.1426522
```



# More examples

- HW** Standardized tests provide an interesting application of probability theory. Suppose that a test consists of 20 multiple-choice questions, each with 4 possible answers, of which exactly one is correct. If 17 questions are answered correctly, the exam is passed. A student comes unprepared for this test and randomly crosses one of the four possible answers. (If the student guesses on each question, then the taking of the exam can be modeled as a sequence of 20 independent events.) Use R to compute the probability that the student will pass the test.
- HW** From a list of 15 households, 9 are homeowners and 6 households live in rental housing. Four households are randomly selected from these 15. Find the probability that the number of households that do not own a home is at least three.

# Geometric distribution

- Geometric distribution with parameter  $p$ 
  - A geometric distribution models the total number of attempts before a success.

Example: The number of tails **before** the first head in a sequence of coin flips (Bernoulli trials).

- The random variable  $X$  follows a geometric distribution with parameter  $p$  if
  - $X$  takes the values  $0, 1, 2, 3, \dots$
  - its probability mass function is given by

$$p(x) = P(X = x) = (1 - p)^x \cdot p \quad \text{for } x \in \{0, 1, 2, \dots\}$$

... using the geometric series  $\sum_{x=0}^{\infty} k^x = \frac{1}{1-k}$ , for  $|k| < 1$ , we get

$$\sum_{x=0}^{\infty} p(x) = p \cdot \sum_{x=0}^{\infty} (1 - p)^x = \frac{p}{1 - (1 - p)} = 1$$

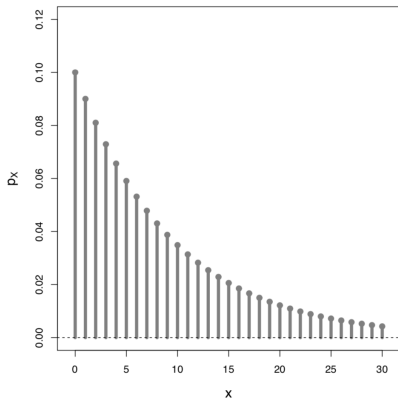
- Expectation/Variance

$$\mathbb{E}(X) = \frac{1-p}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2}$$

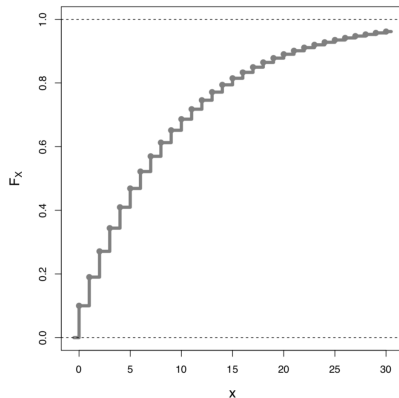
... HW Check!

# Example

- $X \sim \text{Geometric distribution with } p = 0.1$



Probability mass function



Cumulative distribution function

# Example

- Failure times

Geometric distribution is used to model lifetimes of devices or components. For example, if the probability is 0.001 that a light bulb will fail on any given day, then the probability that it will last more than 30 days is

$$\begin{aligned}P(X > 30) &= \sum_{x=31}^{\infty} 0.001(1 - 0.001)^{x-1} = 1 - \sum_{x=1}^{30} 0.001(1 - 0.001)^{x-1} \\&= 1 - \sum_{x=0}^{29} 0.001(1 - 0.001)^x \approx 0.970\end{aligned}$$

- In R: use `dgeom()`

```
1-sum(dgeom(0:29, 0.001))  
[1] 0.970431
```

or `pgeom()`

```
1-pgeom(29, 0.001)  
[1] 0.970431
```

# Example

HW Suppose that the inhabitants of an island plan their families by having babies until the first girl is born. Assume the probability of having a girl with each pregnancy is 0.4 independent of other pregnancies, that all babies survive and there are no multiple births. What is the probability that a family has 6 boys?

# Poisson distribution

- $X \sim \mathcal{P}(\lambda)$  or  $X \sim \text{Poi}(\lambda)$ 
  - $X$  has **Poisson distribution** with parameter  $\lambda > 0$ , if
    - $X$  takes the values  $0, 1, 2, 3, \dots$
    - its probability mass function is given by

$$p(x) = P(X = x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda} \quad \text{for } x \in \{0, 1, 2, \dots\}$$

...  $\lambda$  is the intensity parameter

... using the exponential series  $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda}$ , for all  $\lambda \in \mathbb{R}$ , we obtain

$$\sum_{x=0}^{\infty} p(x) = e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = 1$$

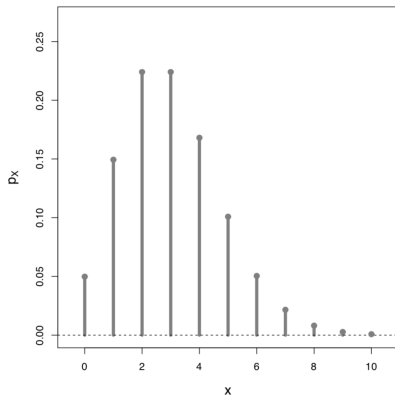
Usually,  $x$  is interpreted as the number of times an event occurs in an interval

- Expectation/Variance:

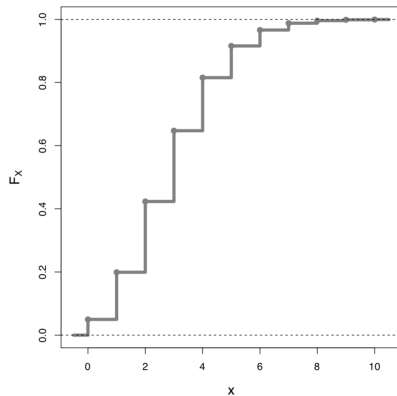
$$\mathbb{E}(X) = \lambda, \quad \text{Var}(X) = \lambda$$

# Example

- $X \sim \mathcal{P}(3)$



Probability mass function



Cumulative distribution function

# Example

- Waiting time

Consider a telephone operator who, on the average, handles five calls every 3 minutes. What is the probability that there will be no calls in the next minute? At least three calls?

- Let  $X$  = number of calls in a minute. Then  $X \sim \mathcal{P}(\lambda)$ , where  $\mathbb{E}X = \lambda = \frac{5}{3}$ . Then,

$$P(\text{no calls in the next minute}) = P(X = 0) = \frac{e^{-\frac{5}{3}} \left(\frac{5}{3}\right)^0}{0!} \approx 0.189$$

$$\begin{aligned} P(\text{at least three calls in the next minute}) &= P(X \geq 3) \\ &= 1 - \sum_{x=0}^2 P(X = x) = 1 - \sum_{x=0}^2 \frac{e^{-\frac{5}{3}} \left(\frac{5}{3}\right)^x}{x!} \approx 0.234. \end{aligned}$$

- In R: we use `dpois()`

```
1 - sum(dpois(0:2, 5/3))  
[1] 0.2340045
```



# Example

## HW Web visitors

A website manager has noticed that during the evening hours, about three people per minute check out from their shopping cart and make an online purchase. She believes that each purchase is independent of the others and wants to model the number of purchases per minute.

- (1) What model might you suggest to model the number of purchases per minute?
- (2) What is the probability that in any one minute at least one purchase is made?
- (3) What is the probability that no one makes a purchase in the next two minutes?

# Binomial-Poisson relationship

- $\mathcal{P}(\lambda)$  distribution as a **limit** of the  $B(n, p)$  distribution
  - For  $X_n \sim B(n, \frac{\lambda}{n})$  it holds

$$\lim_{n \rightarrow \infty} P(X_n = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

- **Interpretation:** If there are many independent and identical Bernoulli experiments with a low probability of success, the number of successes can be approximated with a Poisson distribution
- Rule of thumb

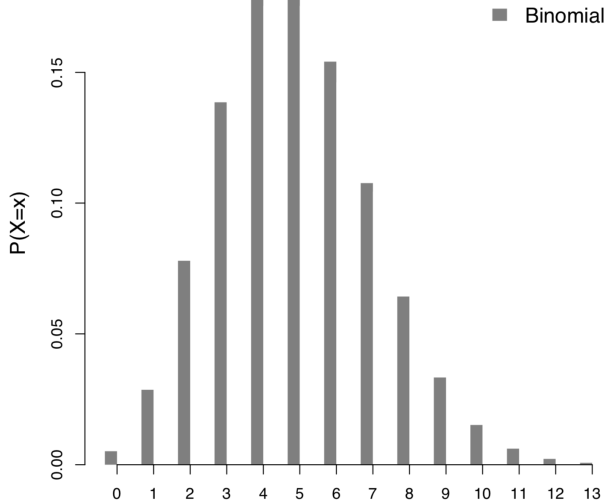
For  $n \geq 50$ ,  $p \leq \frac{1}{10}$  and  $np \leq 10$  a random variable  $X \sim B(n, p)$  can be approximated with  $Y \sim \mathcal{P}(np)$

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \approx P(Y = x) = \frac{(np)^x e^{-np}}{x!}, \quad x = 0, 1, \dots, n$$

... we will **return** to this later in the course...

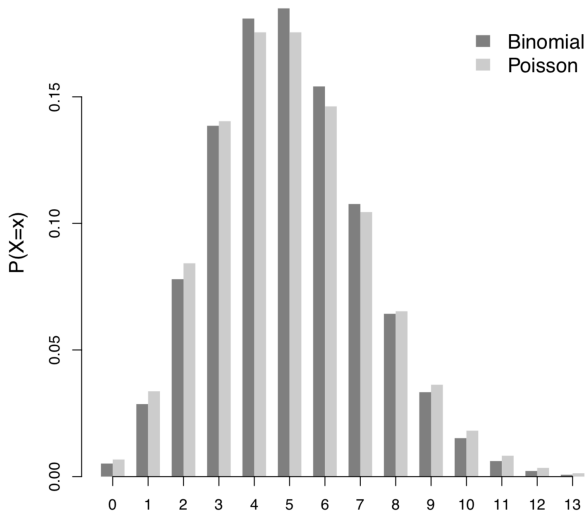
# Example

- $X \sim B(50, \frac{1}{10})$



# Example

- $X \sim B(50, \frac{1}{10}) \approx \mathcal{P}(5)$



# Uniform distribution

- $X \sim \mathcal{U}(a, b)$ 
  - $X$  is a random variable with **uniform** distribution over the interval  $(a, b)$  ( $a < b, a, b \in \mathbb{R}$ ) if its pdf is of the form

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

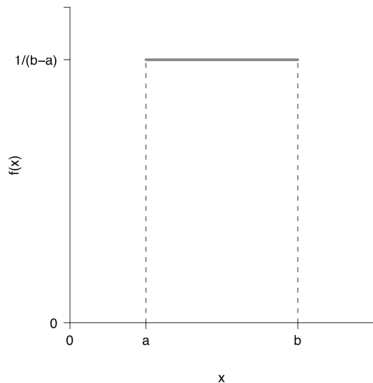
- The cdf is given by

$$F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a}, & a < x < b \\ 1, & x \geq b \end{cases}$$

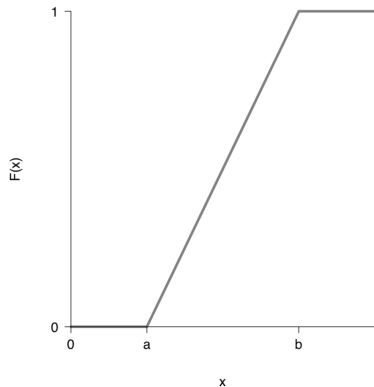
- Expectation/Variance:  $\mathbb{E}(X) = \frac{a+b}{2}, \quad \mathbb{V}\text{ar}(X) = \frac{(b-a)^2}{12}$

# Example

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



Probability density function



Cumulative distribution function

# Exponential distribution

- $X \sim \exp(\lambda)$
- $X$  has an exponential distribution with the parameter  $\lambda > 0$  if its pdf is of the form

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

... models waiting time  
(continuous analogue of geometric distribution)

- The cdf is of the form

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

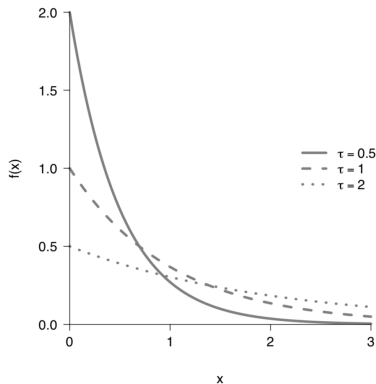
- Expectation/Variance:

$$\mathbb{E}(X) = \tau = \frac{1}{\lambda}, \quad \text{Var}(X) = \tau^2 = \frac{1}{\lambda^2}$$

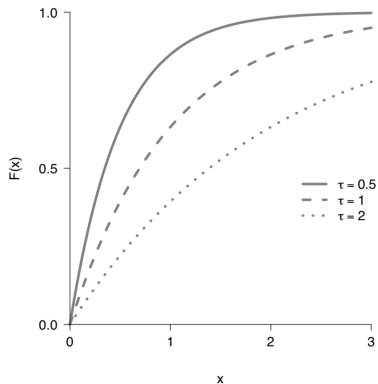
... HW Check!

# Example

- $X \sim \exp(\tau)$ ,  $\tau = \frac{1}{\lambda}$  for  $\tau = 0.5, \tau = 1$  and  $\tau = 2$



Probability density function



Cumulative distribution function



# Example

- Waiting time

Anna noticed that taxis drive past her street on the average of once every 10 minutes. Suppose time spent waiting for a taxi is modeled by an exponential random variable

$$X \sim \exp\left(\frac{1}{10}\right); \quad f(x) = \frac{1}{10}e^{-\frac{x}{10}}, \quad \text{for } x \geq 0$$

Calculate the probability of waiting for a taxi between 3 and 7 minutes.

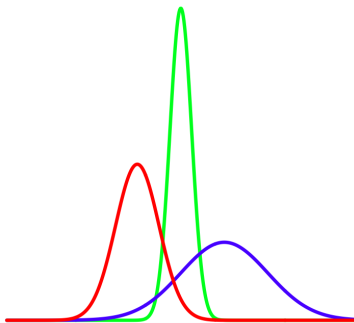
- $P(3 < X < 7) = \int_3^7 f(x) dx = F(7) - F(3) \approx 0.244$

- In R: we use `pexp()`

```
pexp(7, 0.1) - pexp(3, 0.1)
[1] 0.2442329
```

# Normal (Gaussian) distribution

- $X \sim \mathcal{N}(\mu, \sigma^2)$ 
  - One of the most important distributions in statistics and probability theory.



Normal distributions are symmetric with relatively more values at the center of the distribution and relatively few in the tails.

# Normal distribution

- $X \sim \mathcal{N}(\mu, \sigma^2)$ 
  - $X$  follows a **Normal distribution** (or **Gaussian distribution**) with the parameters  $\mu \in \mathbb{R}$  and  $\sigma > 0$ , if its pdf is of the form

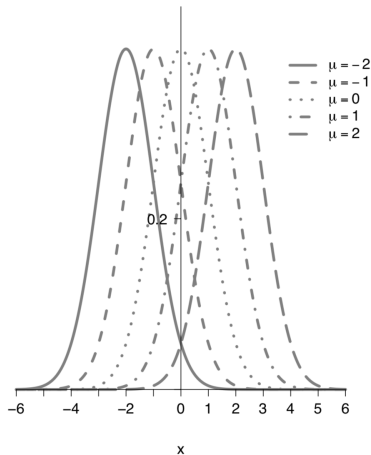
$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for } x \in \mathbb{R}$$

- Expectation/Variance

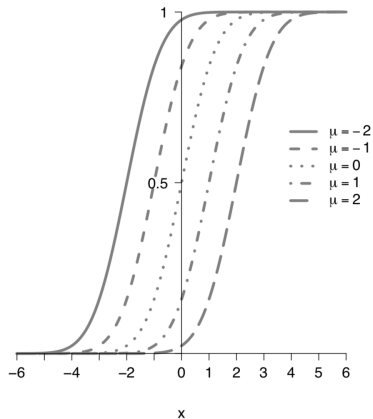
$$\mathbb{E}(X) = \mu, \quad \text{Var}(X) = \sigma^2$$

# Examples

- $X \sim \mathcal{N}(\mu, 1)$ , for  $\mu = -2, \mu = -1, \mu = 0, \mu = 1$  and  $\mu = 2$



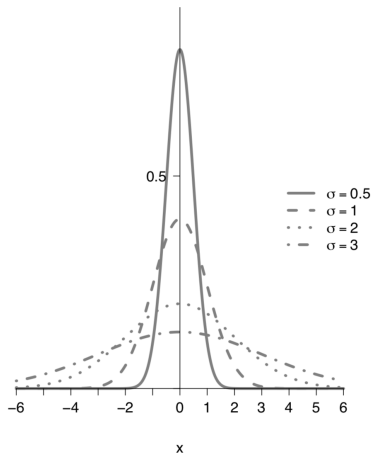
Probability density function



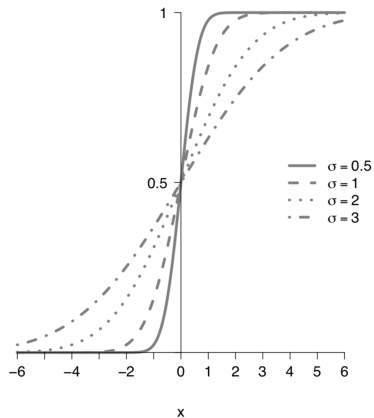
Cumulative distribution function

# Examples

- $X \sim \mathcal{N}(0, \sigma^2)$ , for  $\sigma = 0.5$ ,  $\sigma = 1$ ,  $\sigma = 2$  and  $\sigma = 3$



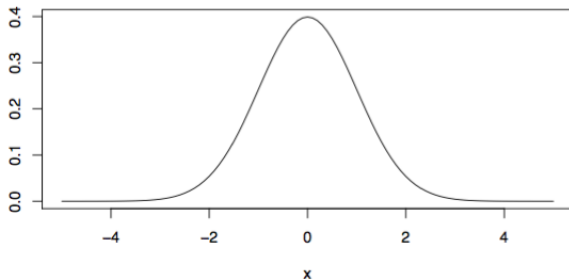
Probability density function



Cumulative distribution function

# Standard Normal distribution

- In particular,
  - $Z \sim \mathcal{N}(0, 1)$  is the **Standard Normal distribution**



- the cdf

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^z f(t) dt = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt$$

- Moreover,

$$\Phi(-z) = 1 - \Phi(z), \quad \text{for } z \in \mathbb{R}$$

... the values of  $\Phi$  can be read from the Table  
or obtained in R by applying `pnorm()`

$$\Phi(z) = P(Z \leq z)$$

- Table of the cdf of  $N(0,1)$ -distribution

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9958	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

# Table of $\Phi(z)$ : Examples

<b>Z</b>	<b>0.00</b>	<b>0.01</b>	<b>0.02</b>	<b>0.03</b>	<b>0.04</b>	<b>0.05</b>	<b>0.06</b>	<b>0.07</b>
<b>0.0</b>	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279
<b>0.1</b>	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675
<b>0.2</b>	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064
<b>0.3</b>	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443
<b>0.4</b>	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808
<b>0.5</b>	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157
<b>0.6</b>	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486
<b>0.7</b>	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794
<b>0.8</b>	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079

- $\Phi(0.43) = 0.6664$                        $\Phi(0.76) = 0.7764$
- $\Phi(-0.43) = 1 - 0.6664 = 0.3336$                        $\Phi(-0.76) = 1 - 0.7764 = 0.2236$
- $P(-0.76 \leq Z \leq 0.43) = \Phi(0.43) - \Phi(-0.76) = 0.4428$



# Table of $\Phi(z)$

<b>Z</b>	<b>0.00</b>	<b>0.01</b>	<b>0.02</b>	<b>0.03</b>	<b>0.04</b>	<b>0.05</b>	<b>0.06</b>	<b>0.07</b>
<b>0.0</b>	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279
<b>0.1</b>	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675
<b>0.2</b>	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064
<b>0.3</b>	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443
<b>0.4</b>	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808
<b>0.5</b>	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157
<b>0.6</b>	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486
<b>0.7</b>	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794
<b>0.8</b>	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079

- Compute  $P(|Z| \leq \frac{1}{2})$

$$\begin{aligned}P(|Z| \leq \frac{1}{2}) &= P(-0.5 \leq Z \leq 0.5) \\&= \Phi(0.5) - \Phi(-0.5) \\&= \Phi(0.5) - (1 - \Phi(0.5)) \\&= 2 \cdot \Phi(0.5) - 1 \\&= 2 \cdot 0.6915 - 1 = 0.3830\end{aligned}$$

# $\Phi(z)$ in R

`pnorm(x, mean, sd)`

... `pnorm(x,  $\mu$ ,  $\sigma$ )` uses  $\sigma$  while in the notation of the normal distribution  $\mathcal{N}(\mu, \sigma^2)$  is used  $\sigma^2$ ...

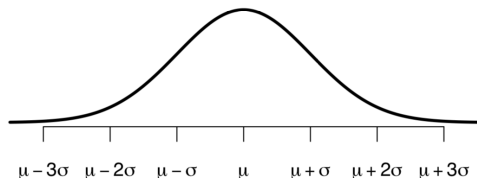
```
#Values of the cdf of N(0,1) distribution
pnorm(0.5)
[1] 0.6914625
# it is the same as
pnorm(0.5, mean=0, sd=1)
[1] 0.6914625
pnorm(1.3)
[1] 0.9031995
# Values of the cdf of N(3,4) distribution
pnorm(2.5, mean=3, sd=2)
[1] 0.4012937
# it is the same as pnorm((2.5-3)/2)
pnorm(-0.25)
[1] 0.4012937
```

- We compute  $P(|Z| \leq \frac{1}{2}) \approx 0.3830$  using R:

```
diff(pnorm(c(-0.5, 0.5)))
[1] 0.3829249
```

# Questions

- (1) What is the probability that  $Z \sim \mathcal{N}(0, 1)$  is at most one /two/three standard deviations away from the expectation  $\mu = 0$ ?
- (2) What is the probability that  $X \sim \mathcal{N}(0, 3^2)$  is at most one /two/three standard deviations away from the expectation  $\mu = 0$ ?
- (3) How do the answers to question (2) change if we change  $\mu$ ?



... The following probabilities should be calculated

$$P(X \in [\mu - \sigma, \mu + \sigma])$$

$$P(X \in [\mu - 2\sigma, \mu + 2\sigma])$$

$$P(X \in [\mu - 3\sigma, \mu + 3\sigma])$$

# Probabilities of the normal distribution

(1) Let  $Z \sim \mathcal{N}(0,1)$ . Then,

$$P(-1 \leq Z \leq 1) \approx 0.68$$

$$P(-2 \leq Z \leq 2) \approx 0.95$$

$$P(-3 \leq Z \leq 3) \approx 0.997$$

```
# one sigma-away from mu=0
diff(pnorm(c(-1,1)))
[1] 0.6826895
# two sigmas-away from mu=0
diff(pnorm(c(-2,2)))
[1] 0.9544997
# three sigmas-away from mu=0
diff(pnorm(c(-3,3)))
[1] 0.9973002
```

# Probabilities of the normal distribution

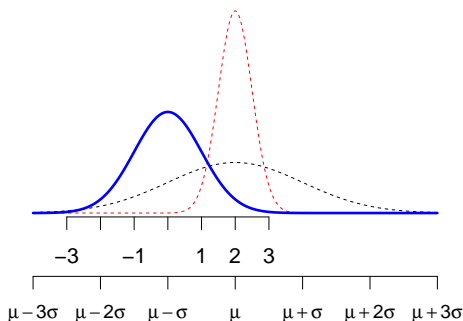
(2) Let  $X \sim \mathcal{N}(\mu, \sigma^2)$ .

- Standardization

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

... Standardizing any normal random variable  $X$  produces the standard normal  $Z$

... We use  $Z$  for a  $\mathcal{N}(0, 1)$



# Probabilities of the normal distribution

(2) Let  $X \sim \mathcal{N}(\mu, \sigma^2)$ .

- **Standardization**

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

... Standardizing any normal random variable  $X$   
produces the standard normal  $Z$

... We use  $Z$  for a  $\mathcal{N}(0, 1)$

- In general, let  $Y$  be a random variable with  $\mathbb{E}(Y) = \mu$  and standard deviation  $\sigma = \sqrt{\text{Var}(Y)}$ . Then, the standardized random variable

$$U = \frac{Y - \mu}{\sigma}$$

has  $\mathbb{E}(U) = 0$  and  $\sqrt{\text{Var}(U)} = 1$ .

... HW

# Probabilities of the normal distribution

(2) Let  $X \sim \mathcal{N}(\mu, \sigma^2)$ .

- Standardization

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

... Standardizing any normal random variable  $X$   
produces the standard normal  $Z$

... We use  $Z$  for a  $\mathcal{N}(0, 1)$

- Then,

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = P(-1 \leq \frac{X - \mu}{\sigma} \leq 1) = P(|Z| \leq 1) \approx 0.68$$

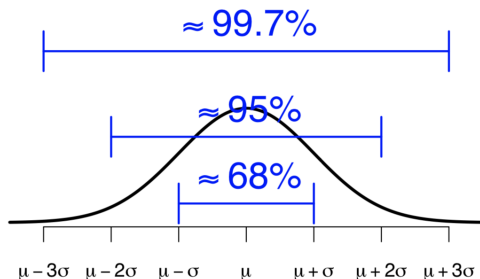
$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(-2 \leq \frac{X - \mu}{\sigma} \leq 2) = P(|Z| \leq 2) \approx 0.95$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq \frac{X - \mu}{\sigma} \leq 3) = P(|Z| \leq 3) \approx 0.997$$

... the 68-95-99.7-Rule

# The 68-95-99.7-Rule

- The 68-95-99.7-Rule applies to probabilities of the normal distribution
  - 68% of the area of a normal distribution is within one standard deviation of the mean.
  - Approximately 95% of the area of a normal distribution is within two standard deviations of the mean.
  - Approximately 99.7% of the area of a normal distribution is within three standard deviations of the mean.





# Properties

- Let  $X \sim \mathcal{N}(\mu, \sigma^2)$  and let  $F_X$  be its cdf.
  - Affine transformation

$$Y = a + bX \sim \mathcal{N}(a + b\mu, b^2\sigma^2)$$

- $p$ -quantile

$$x_p = \mu + \sigma z_p$$

where  $z_p = \Phi^{-1}(p)$  denotes the  $p$ -quantile of  $\mathcal{N}(0, 1)$

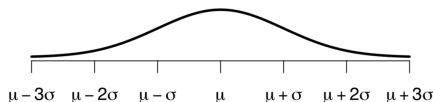
- Symmetry of the  $\mathcal{N}(0, 1)$  with respect to zero

$$z_p = -z_{1-p}$$

- in R: `qnorm()`

# Properties

- Let  $X_1, X_2, \dots, X_n$  be independent identically distributed (iid) random variables with  $X_i \sim \mathcal{N}(\mu, \sigma^2)$   
...  $(X_1, X_2, \dots, X_n)$  is called a **random sample**



- For the **(sample) mean** it holds

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

- $\bar{X}$  is also **normally distributed**
- the expectation of  $\bar{X}$  is  $\mu_{\bar{X}} = \mu$  (the same as the expectation of  $X_i$ )
- the standard deviation of  $\bar{X}$  equals  $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$  (reduction by factor  $\frac{1}{\sqrt{n}}$ )

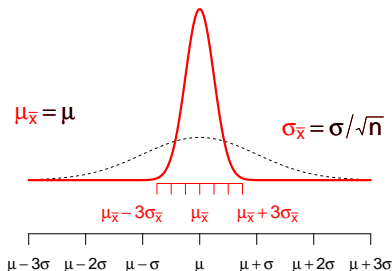
... we are going to prove this!

# The mean of iid normally distributed rvs

- For the **mean** it holds

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

- $\bar{X}$  is also normally distributed
- the expectation of  $\bar{X}$  is  $\mu_{\bar{X}} = \mu$
- the standard deviation of  $\bar{X}$  equals  $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$



# Examples

- (1) As a group, the Dutch are among the tallest people in the world. The average Dutch man is 184 cm tall. If a normal model is appropriate and the standard deviation for men is about 8 cm, what percentage of all Dutch men will be over two meters?

- Let  $X$  denote the height of a Dutch man. Then,  $X \sim \mathcal{N}(184, 8^2)$ .
- We want to calculate  $P(X > 200)$ .

We also know

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 184}{8} \sim \mathcal{N}(0, 1)$$

I way: We use the table and calculate

$$\begin{aligned} P(X > 200) &= P\left(\frac{X - 184}{8} > \frac{200 - 184}{8}\right) \\ &= P(Z > 2) \\ &= 1 - P(Z \leq 2) \\ &= 1 - \Phi(2) \\ &= 1 - 0.9772 = 0.0228 = 2.28\% \end{aligned}$$

# Examples

- (1) As a group, the Dutch are among the tallest people in the world. The average Dutch man is 184 cm tall. If Normal model is appropriate and the standard deviation for men is about 8 cm, what percentage of all Dutch men will be over two meters?
- Let  $X$  denotes the height of a Dutch man. Then,  $X \sim \mathcal{N}(184, 8^2)$ .
  - We want to calculate  $P(X > 200)$ .

II way: Use R

```
pnorm(184, 200, 8)
[1] 0.02275013
# or
1-pnorm(200, 184, 8, lower.tail=FALSE)
[1] 0.02275013
# or
pnorm(184, 200, 8, lower.tail=TRUE)
[1] 0.02275013
# or
pnorm(-2)
[1] 0.02275013
```

# Examples

## Note:

- `pnorm(184,200,8) = 1-pnorm(200,184,8)`
- `pnorm(184,200,8) = pnorm(184,200,8,lower.tail=TRUE)`  
`= 1-pnorm(184,200,8,lower.tail=FALSE)`  
`= pnorm(200,184,8,lower.tail=FALSE)`  
`= 1-pnorm(200,184,8)`
- `pnorm(-2) = 1-pnorm(2)`  
`= 1-pnorm(2,0,1)`

# Examples

(2) Suppose you need 20 minutes, on average, to drive to work, with a standard deviation of 2 minutes. Suppose a Normal model is appropriate for the distribution of driving times.

- (a) How often will you arrive at work in less than 22 minutes?
- (b) How often will it take you more than 24 minutes?

- Let  $X$  be the time you need to arrive to work. Then,  $X \sim \mathcal{N}(20, 2^2)$ .
- We want to compute  $P(X < 22)$  and  $P(X > 24)$ .
- First we standardise  $X$

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 20}{2} \sim \mathcal{N}(0, 1)$$

and then compute the probabilities

$$(a) \quad P(X < 22) = P\left(\frac{X - 20}{2} < \frac{22 - 20}{2}\right) = P(Z < 1) = \Phi(1) = 0.8413$$

$$(b) \quad P(X > 24) = P(Z > 2) = 1 - \Phi(2) = 1 - 0.9772 = 0.0228$$

# Examples

(3) Let  $Y \sim \mathcal{N}(10, 3^2)$ .

(a) Find 0.7-quantile of  $Y$ .

(b) Find the value  $y_0$  such that it holds  $P(Y \geq y_0) = 0.1$

(a)  $Z = \frac{Y - \mu}{\sigma} = \frac{Y - 10}{3} \sim \mathcal{N}(0, 1)$

If  $F$  is the distribution function of  $Y$ , then 0.7-quantile  $y_{0.7}$  satisfies

$$F(y_{0.7}) = 0.7.$$

This means  $P(Y \leq y_{0.7}) = 0.7$ . Then,

$$0.7 = P(Y \leq y_{0.7}) = P\left(Z \leq \frac{y_{0.7} - 10}{3}\right) = \Phi\left(\frac{y_{0.7} - 10}{3}\right)$$

and

$$\frac{y_{0.7} - 10}{3} = \Phi^{-1}(0.7) \implies y_{0.7} = 3 \cdot \Phi^{-1}(0.7) + 10$$



# Examples

- We find the value  $\Phi^{-1}(0.7)$  in Table of standard normal distribution.

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079

We see  $\Phi(0.53) = 0.7019$  und conclude  $\Phi^{-1}(0.7) \approx 0.53$ .

Then, 0.7-quantil of  $Y$  is

$$y_{0.7} \approx 3 \cdot \Phi^{-1}(0.7) + 10 = 11.59$$

- in R we use `qnorm()`  
 $\Phi^{-1}(0.7) = \text{qnorm}(0.7, \text{mean}=0, \text{sd}=1) = 0.5244$

# Examples

- (b) We look for  $y_0$  such that it satisfies  $P(Y \geq y_0) = 0.1$ , i.e.

$$\begin{aligned} 0.1 &= P(Y \geq y_0) = 1 - P(Y < y_0) = 1 - P\left(Z < \frac{y_0 - 10}{3}\right) \\ &= 1 - \Phi\left(\frac{y_0 - 10}{3}\right) \end{aligned}$$

Then,

$$\Phi\left(\frac{y_0 - 10}{3}\right) = 0.9$$

which leads to

$$y_0 = 3 \cdot \Phi^{-1}(0.9) + 10.$$

- We use R:

$$y_0 = 3 \cdot \text{qnorm}(0.9, 0, 1) + 10 = 13.8447$$

or

$$y_0 = \text{qnorm}(0.1, 10, 3, \text{lower.tail} = \text{FALSE}) = 13.84465$$

# Examples

HW Miraculin is a protein naturally produced in a rare tropical fruit. It can convert a sour taste into a sweet taste. Consequently, miraculin has the potential to be an alternative low-calorie sweetener.

A group of Japanese environmental scientists investigated the ability of a hybrid tomato plant to produce miraculin. For a particular generation of the potato plant, the amount  $X$  of miraculin produced (measured in micrograms per gram of fresh weight) had a mean 105.3 and a standard deviation of 8.0. Assume that  $X$  is normally distributed.

- (a) Find  $P(X > 120)$ .
- (b) Find  $P(100 < X < 110)$ .
- (c) Find the values  $a$  for which  $P(X < a) = 0.25$ .
- (d) Compute the interquartile range of the miraculin production, i.e. compute the difference between the upper and the lower quartile of  $X$ .

# Addition Rules

- Let  $X_1, \dots, X_k, k \in \mathbb{N}$  be independent random variables.

Let  $S = X_1 + \dots + X_k$ .

- If  $X_i \sim \text{bern}(p)$  then  $S \sim B(k, p)$
- If  $X_i \sim B(n_i, p)$  then  $S \sim B(n_1 + \dots + n_k, p)$
- If  $X_i \sim \text{Poi}(\lambda_i)$  then  $S \sim \text{Poi}(\lambda_1 + \dots + \lambda_n)$
- If  $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$  then  $S \sim \mathcal{N}(\mu_1 + \dots + \mu_k, \sigma_1^2 + \dots + \sigma_k^2)$
- If  $X_i \sim \exp(\lambda)$  then  $S \sim \text{Gamma}(k, \lambda)$

Note: The pdf of a *Gamma* ( $\alpha, \lambda$ ),  $\alpha, \lambda > 0$  random variable is given by

$$f(X) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0$$

where  $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$  is the Euler gamma function.

# R-functions

- The following functions in R may be used:

<code>pbinom()</code>	<code>dbinom()</code>	<code>qbinom()</code>	<code>rbinom()</code>
<code>pexp()</code>	<code>dexp()</code>	<code>qexp()</code>	<code>rexp()</code>
<code>pnorm()</code>	<code>dnorm()</code>	<code>qnorm()</code>	<code>rnorm()</code>
<code>punif()</code>	<code>dunif()</code>	<code>qunif()</code>	<code>runif()</code>
<code>ppois()</code>	<code>dpois()</code>	<code>qpois()</code>	<code>rpois()</code>
<code>pgeom()</code>	<code>dgeom()</code>	<code>qgeom()</code>	<code>rgeom()</code>

- The prefix **d** means the probability density function or the probability mass function, **p** means the cumulative distribution function, **q** stays for the quantile value and **r** returns a random simulation.
- The root **binom** stays for binomial distribution, while **exp** stays for exponential distribution, **norm** for normal distribution, **unif** for uniform distribution, **pois** for Poisson distribution and **geom** for geometric distribution.
  - For example, **dnorm** is the height of the density of a normal curve while **dbinom** gives the probability of an outcome of a binomial distribution.

# Various examples

- HW** High temperatures in Vienna for the month of August follow a uniform distribution over the interval 22 to 27 degrees Celsius. Find the temperature which 90% of the August days exceed.
- HW** A construction zone on a highway has a posted speed limit of 40 miles per hour. The speeds of vehicles passing through this construction zone are normally distributed with a mean of 46 miles per hour and a standard deviation of 4 miles per hour. Find the percentage of vehicles passing through this construction zone that is exceeding the posted speed limit.
- HW** The length of a workpiece is normally distributed with  $\mu = 3$  and  $\sigma = 0.2$ . All workpieces that are shorter than 2.8 or longer than 3.2 are considered to be rejects. What is the probability that a workpiece is rejected?

# Various examples

**HW** It is known from a certain mailbox advertisement that in two out of 1000 people a purchase contract is concluded based on this advertisement. What is the probability that out of 800 people who find the advertisement in your mailbox,

- (1) no one
- (2) at most three
- (3) at least four

will conclude a purchase contract?

**HW** Let  $Z$  a standard normal random variable and let  $X = 5Z + 1$ .

- (1) Calculate  $P(|X| \leq 1)$ .
- (2) Recall that the probability that  $Z$  is within one standard deviation of its mean is approximately 68%. What is the probability that  $X$  is within one standard deviation of its mean?

# Various examples

**HW** It is known that the time  $X$  (in hours), that a technician needs to repair a machine, follows an exponential distribution with parameter  $\lambda = 2$ .

- (1) Calculate the associated distribution function  $F$  and sketch it as well as its density.
- (2) What is the probability that the technician needs
  - (i) at most half an hour
  - (ii) between 0.2 and 0.4 hours
  - (iii) more than 12 minutesfor the repair?
- (3) How many hours are required on average for the repair of a machine? Also determine the variance of the repair time.



# A few multiple-choice questions

- (1) Jan wants to compute in R the probability of obtaining at least one 6 when rolling a fair dice 4 times. He should use the command
- a. `1-pbinom(0,4,1/6)`
  - b. `dbinom(0,4,1/6)`
  - c. `qbinom(0,4,1/6)`
  - d. `rbinom(0,4,1/6)`
- (2) The distribution of cholesterol levels for patients in a cardiology practice follows a normal distribution with a mean of 220 and a standard deviation of 40. In this practice, the probability that a patient has a cholesterol level less than 140 is the same as the probability that a patient has a cholesterol level reading more than \_\_\_\_\_.
- a. 320
  - b. 300
  - c. 340
  - d. 400

## A few multiple-choice questions

- (3) Let  $Z$  be a standard normal random variable and let  $X = -4Z + 0.5$ . Calculate

$$P(|X| \leq 0.5).$$

Use the values given in the table below.

$z$	0.00	0.01	0.02	0.03	0.04	0.05	0.06
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051

Table 1: Cumulative distribution function of the standard normal distribution

- a. 0.0987
- b. 0.1915
- c. 0.1987
- d. 0.4013

## A few multiple-choice questions

- (4) In a list of 15 households, 9 own homes and 6 do not own homes. Four households are randomly selected from these 15 households. Find the probability that the number of households in these four who own homes is at most one.
- a. 0.1536
  - b. 0.1792
  - c. 0.3456
  - d. 0.4752
- (5) A literature professor decides to give a 20-question true-false quiz. Each question is worth 1 point. He wants to choose the passing grade so that the probability of passing a student who guesses on every question is less than 0.1. What score should be set as the lowest passing grade?
- a. 12
  - b. 13
  - c. 14
  - d. 15

## A few multiple-choice questions

- (6) A coating machine coats film between 120 and 210 microns with a uniform random distribution. If any section of film with a coating greater than 200 microns cannot be sold, approximately how much of the product must be scrapped?
- a.  $1/3$
  - b.  $1/2$
  - c.  $1/7$
  - d.  $1/9$
- (7) Two (fair six-sided) die are rolled. Which R-command approximates the distribution of the difference of the outcome of the first and the second die?
- a. `table(sample(1:6,1000,replace = TRUE) - sample(1:6,1000,replace = TRUE))/1000`
  - b. `table(sample(1:6,500,replace = TRUE) - sample(1:6,500,replace = TRUE))/1000`
  - c. `sum(sample(1:6,1000,replace = TRUE) - sample(1:6,1000,replace = TRUE))/1000`
  - d. `sum(sample(1:6,500,replace = TRUE) - sample(1:6,500,replace = TRUE))/1000`

## A few multiple-choice questions

(8) Let  $X$  be a random variable with the probability density function

$$f_X(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{else} \end{cases}.$$

If  $Z = -\ln X$  is a random variable, then the median of  $Z$

- a. is smaller than 0.5.
- b. is smaller than 0.85 but bigger than 0.5.
- c. is bigger than 0.85.
- d. is not unique.

The following table of approximate values of the natural logarithm is useful

x	1	1.5	2	4	8
$\approx \ln(x)$	0	0.4	0.7	1.4	2.1

## A few multiple-choice questions

- (9) Which of the following R commands should be used for approximating the interquartile range of a random variable  $X \sim \mathcal{N}(3, 4)$
- a. `diff(qnorm(c(0.25, 0.75), 3, 4))`
  - b. `diff(rnorm(c(0.25, 0.75), 3, 4))`
  - c. `diff(qnorm(c(0.25, 0.75), 3, 2))`
  - d. `diff(pnorm(c(0.25, 0.75), 3, 2))`
- (10) Let  $X \sim \mathcal{N}(-1, 1)$  and  $Y \sim \mathcal{N}(1, 4)$ . Then the distribution of their linear combination  $2X - Y + 3$  is
- a.  $\mathcal{N}(0, 6)$
  - b.  $\mathcal{N}(0, 8)$
  - c.  $\mathcal{N}(0, 9)$
  - d. not normal.

Thank you for your attention!