Crash Course in Statistics
ZNZ 2026
II

Christoph Luchsinger and Zofia Baranczuk

Based on Script by Daniel J. Stekhoven

Random variables and distributions

- Conditional probability
- What is a random variable?
- How is it connected to a distribution?
- What are the most important summaries of a distribution?
- Some important distributions…

Breast cancer

- The probability that a woman above 40 has breast cancer is 1%
- Diagnostics:
  - Mammography *positive*, if sick: 80% (*sensitivity*) – ideal world: 100%
  - Mammography *negative*, if healthy: 90% (*specificity*) – ideal world: 100%
    - i.e. Mammography *positive*, if healthy: 10% **(false positive)**

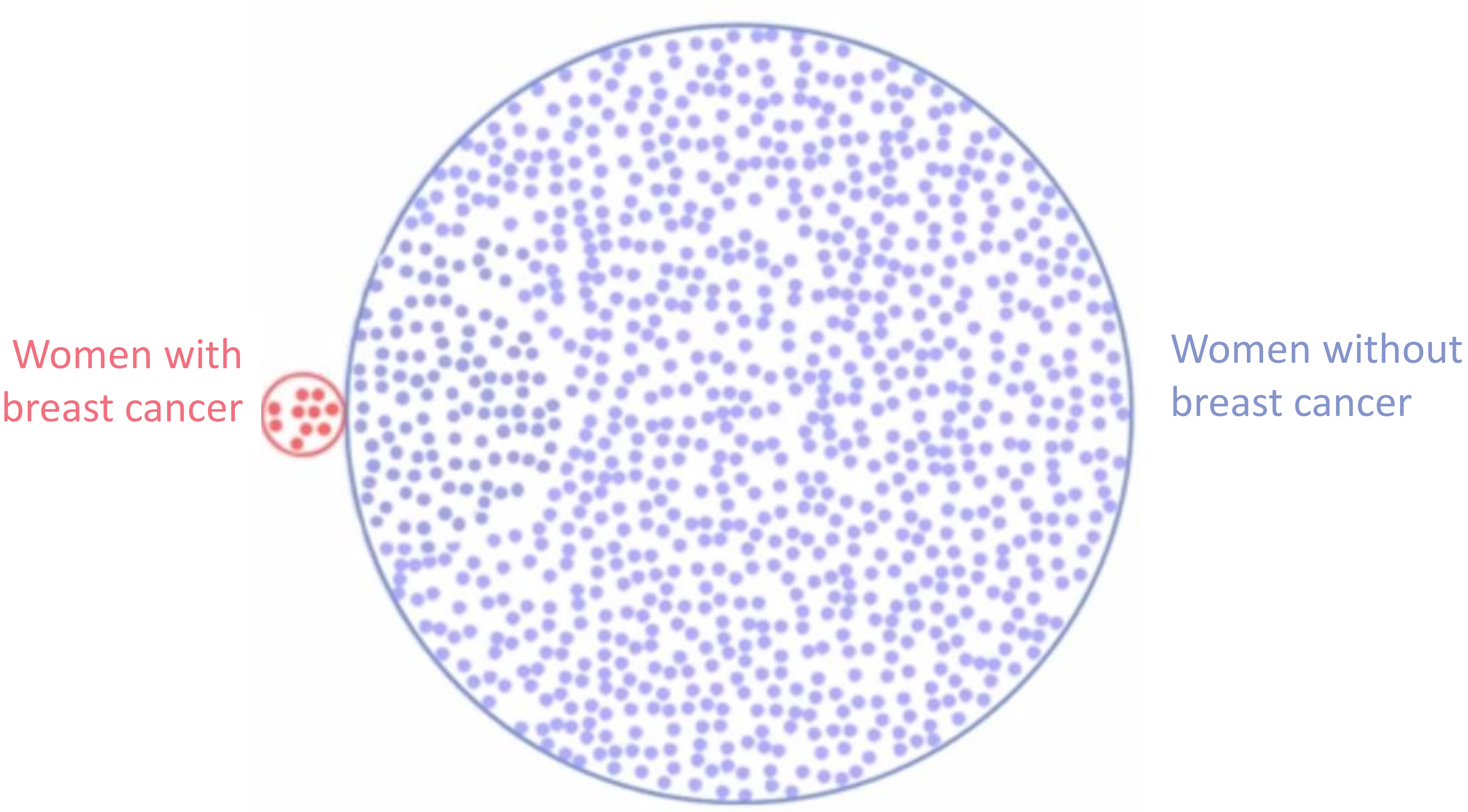
- If you (40, ♀) have a *positive* test result in a mammography. How likely is it that you actually have breast cancer?

- Luana Micallef - https://www.youtube.com/watch?v=D8VZqxcu0I0

https://www.youtube.com/watch?v=D8VZqxcu0I0

Breast cancer

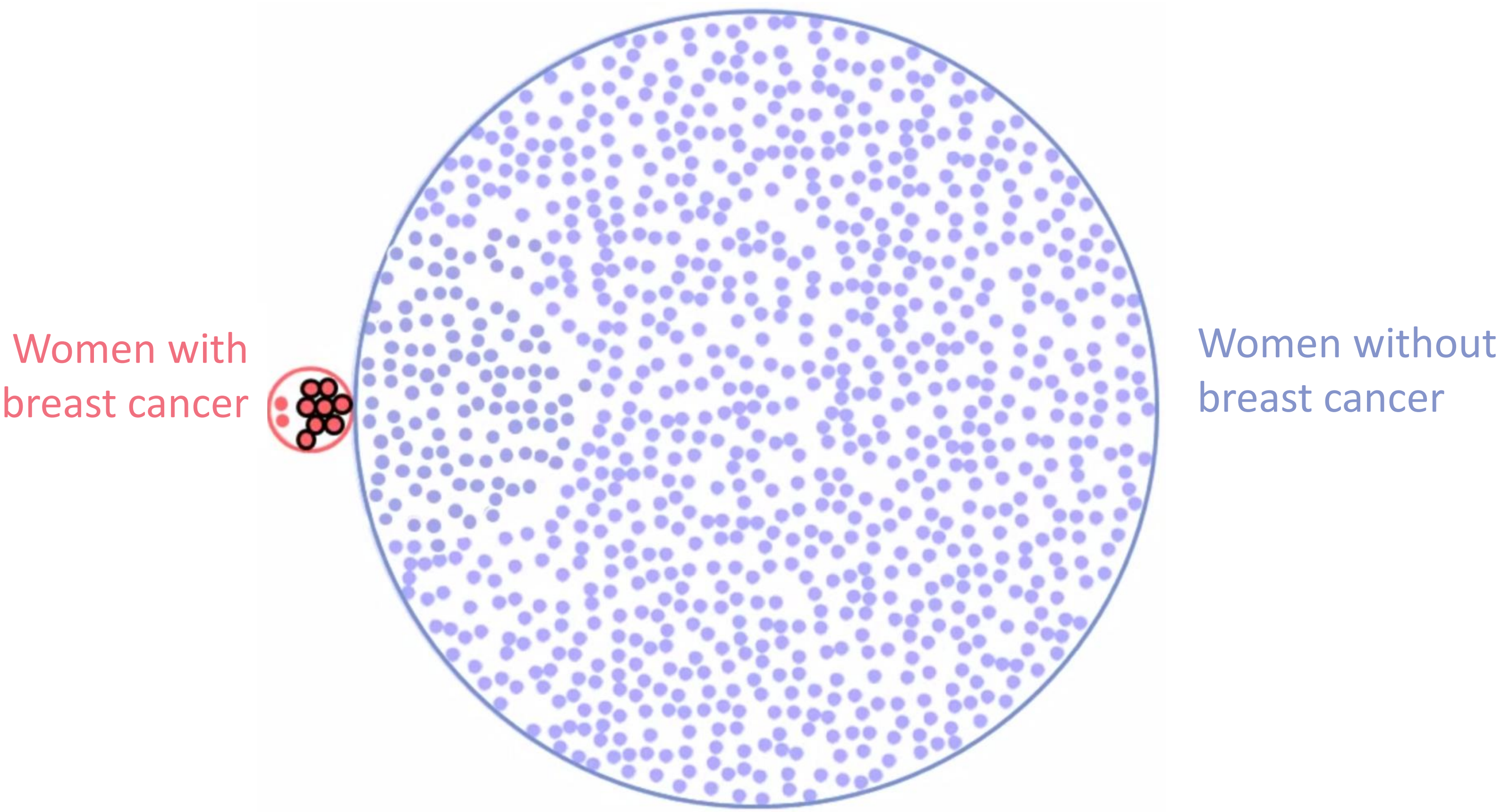
- Probability that a woman of 40 has breast cancer: 1%



https://www.youtube.com/watch?v=D8VZqxcu0I0

Breast cancer

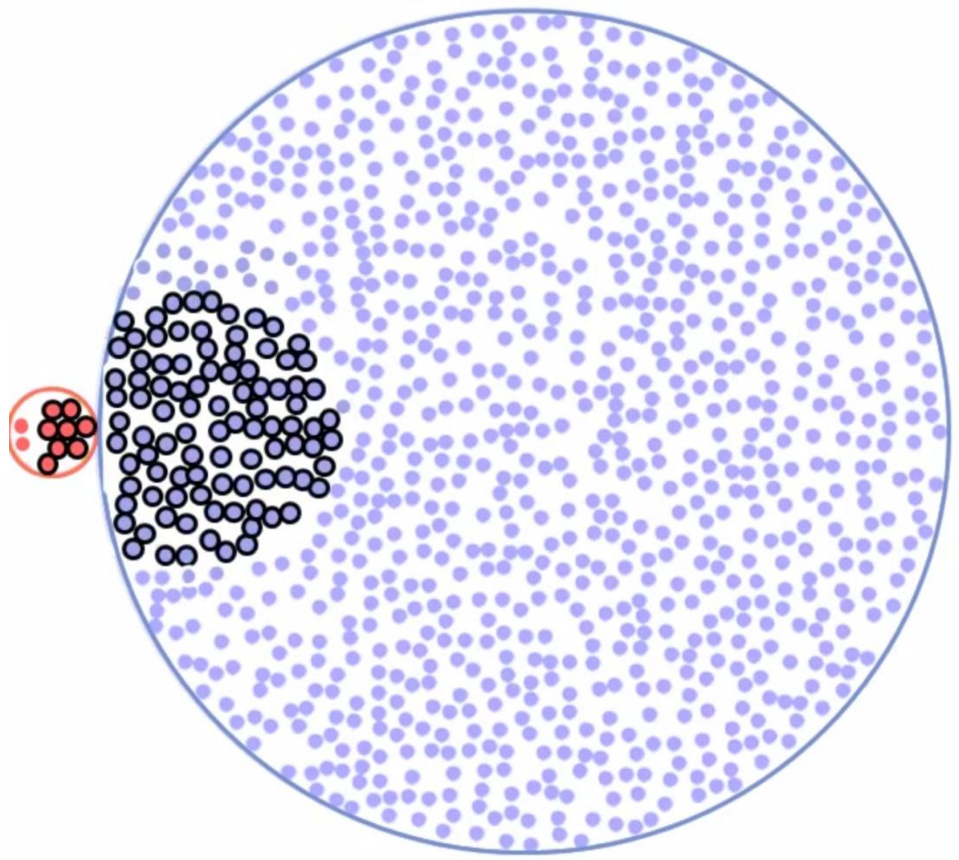
- In 80% of the sick women the test is *positive*



https://www.youtube.com/watch?v=D8VZqxcu0I0

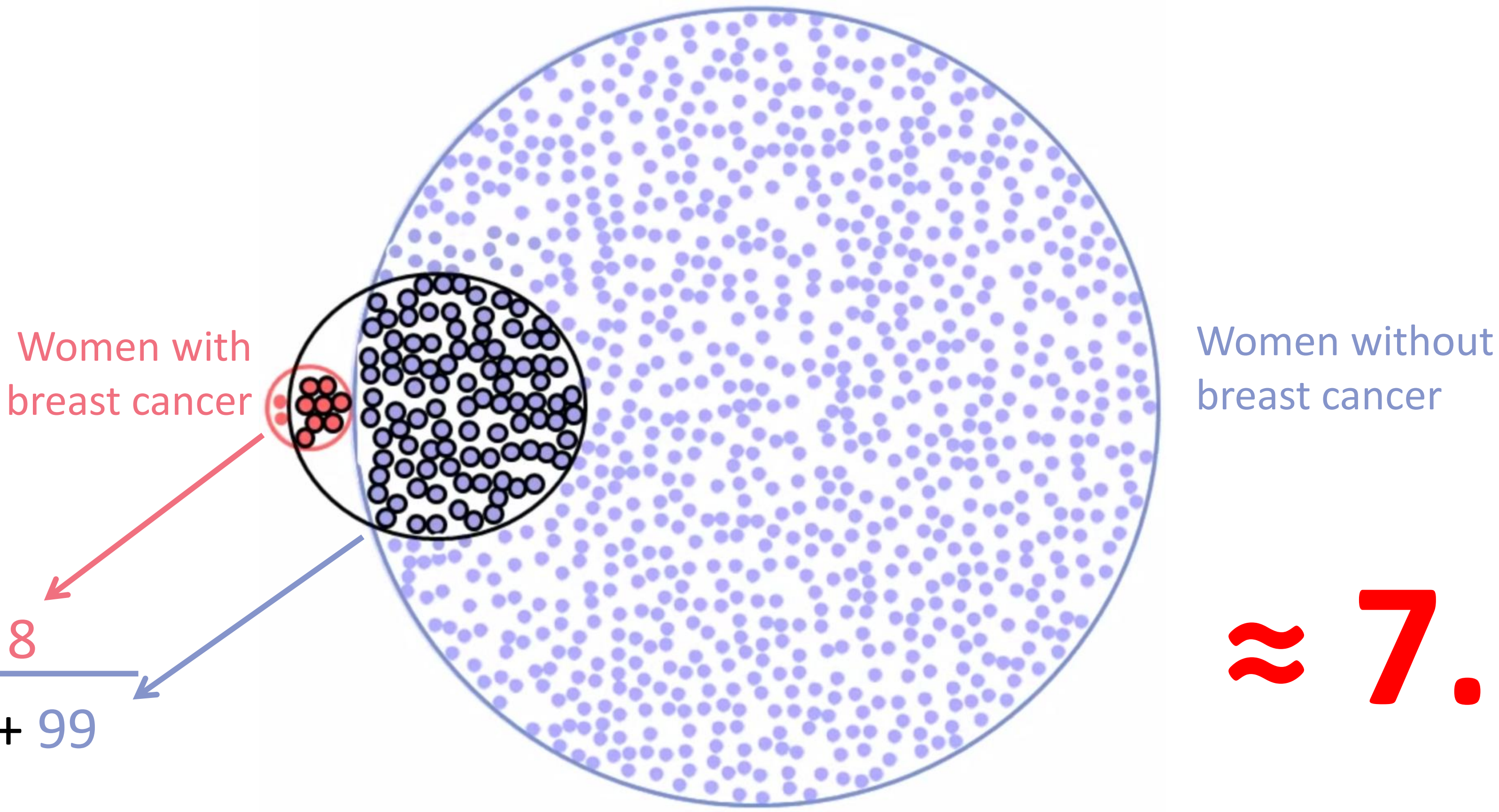
Breast cancer

- 10% of the mammographies in the healthy women are *positive*



Breast cancer

- You (40, ♀) have a *positive* test result. The probability for breast cancer is…



$$\frac{8}{8 + 99}$$

≈ 7.5%

Theorem of Bayes

- $P(A \mid B) := P(A \cap B)/P(B)$ "Probability of A given B"
- Relation between $P(A|B)$ and $P(B|A)$:

$$P(A|B) \cdot P(B) = P(A \cap B) = P(B \cap A) = P(B|A) \cdot P(A)$$

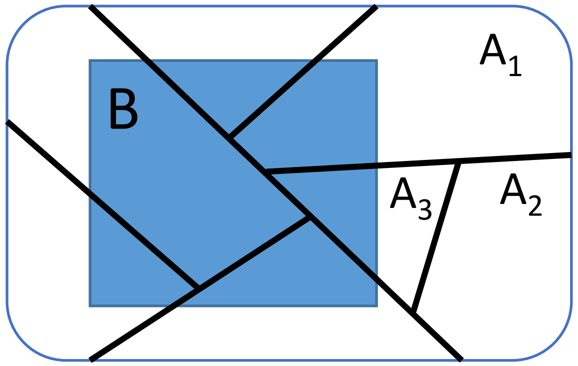
- From this the Theorem of Bayes follows:

$$P(A|B) := \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

Theorem of Total Probability

- Assume, we have a partition $A_i, i = 1, \dots, n$ of Ω, then follows:

$$P[B] = \sum_{i=1}^{n} P[A_i \cap B] = \sum_{i=1}^{n} P[B|A_i] P[A_i]$$



➢Compute the area of B without completing the puzzle.

Breast cancer – the calculation

- Event K: sick, Event T: test *positive*
  - given: $P(K) = 0.01, P(T|K) = 0.8, P(T|K^C) = 0.1$
  - wanted: $P(K|T)$

- Theorem of Total Probability:

$$P(T) = P(T|K)P(K) + P(T|K^C)P(K^C)$$

➢ $P(T) = 0.8 \cdot 0.01 + 0.1 \cdot 0.99 = 0.107$

- Theorem of Bayes:

$$P(K|T) = \frac{P(T|K)P(K)}{P(T)}$$

➢ $P(K|T) = \frac{0.8 \cdot 0.01}{0.107} \approx 0.0748$

$\approx$ 7.5%

LET'S MAKE A DEAL

Behind one of the doors there is a car to be won

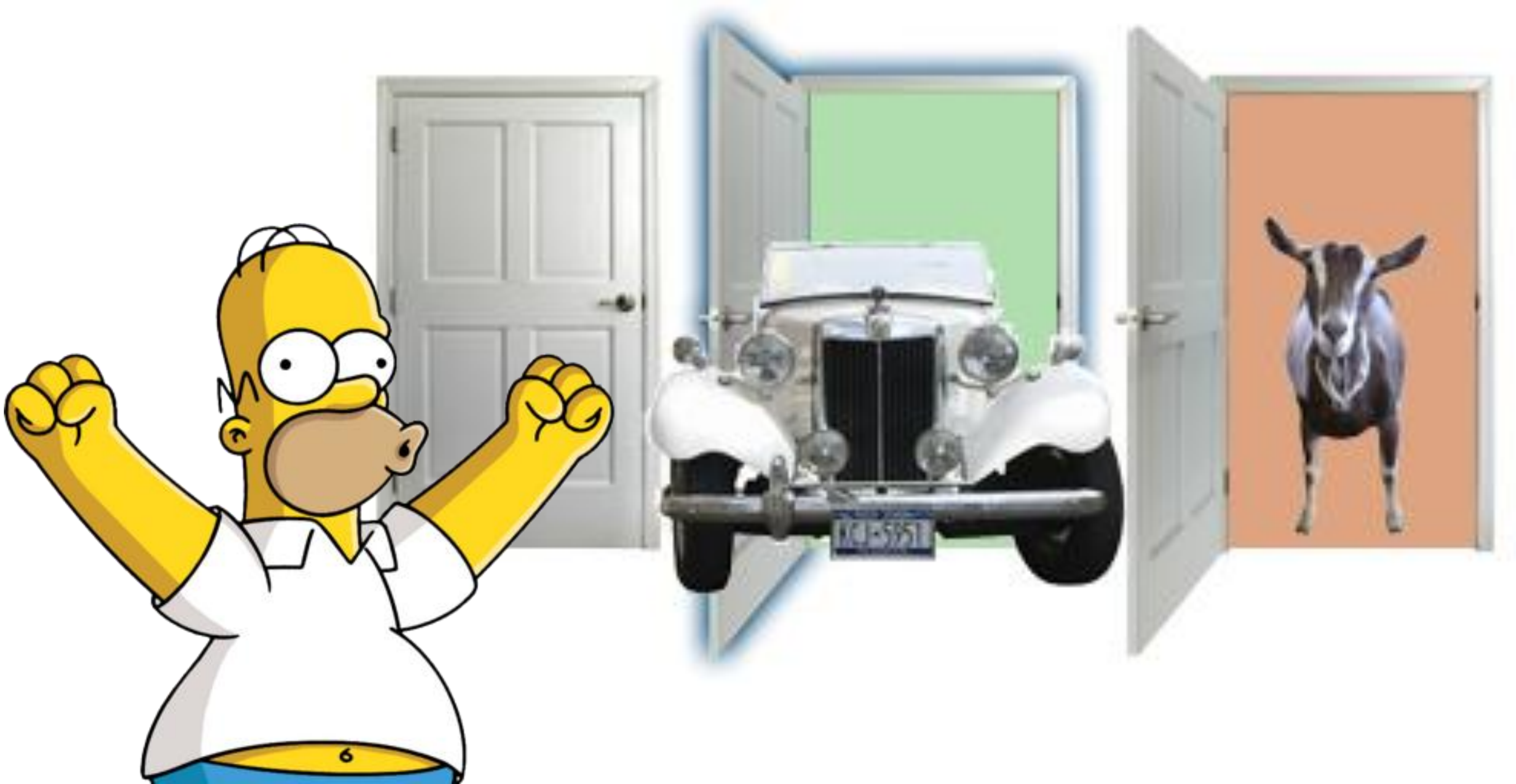
- The car has been placed randomly!

You choose door 1 (so called «without loss of generality»; dt: OBdA: Ohne Beschränkung der Allgemeinheit)

I will show you, where the car is not ...

- Would you like to change the initially chosen door?

You changed to the 2nd door



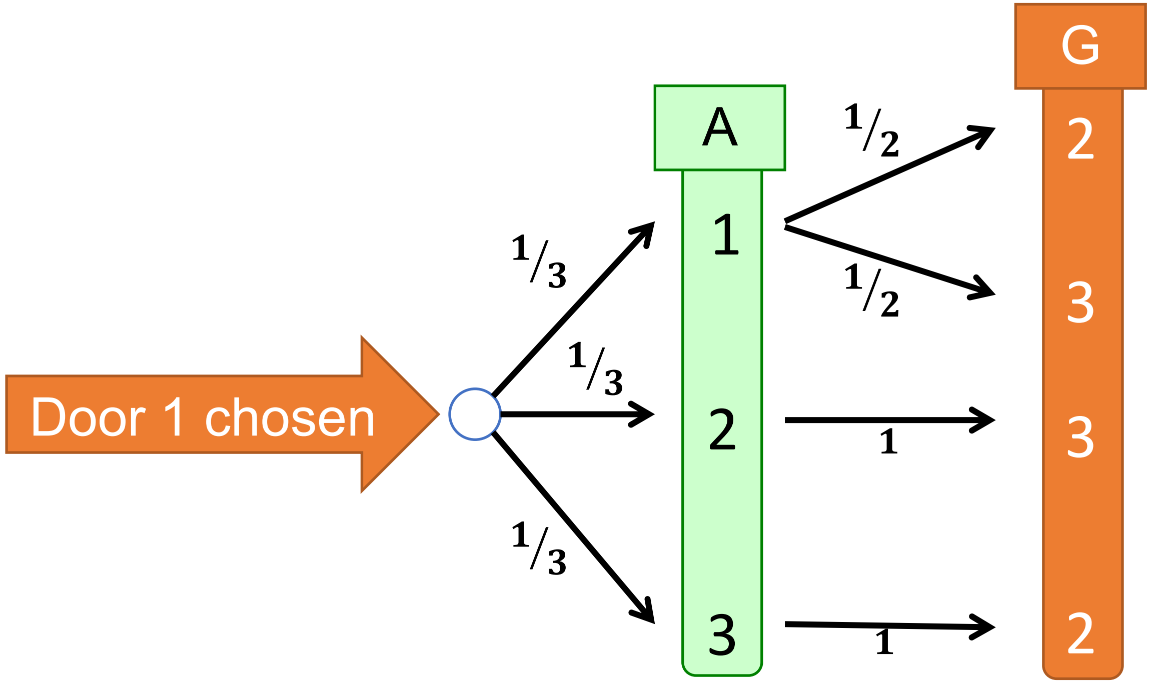
Conditional Probability

- Monty-Hall paradox
  - Is it useful to change the door?
- A = number of door with car
- G = number of opened door

- Is $P(A = 2|G = 3)$ larger than $P(A = 1|G = 3)$?
  - If yes, we should change
  - If not, we should stick with the initial choice
- $in\ General, we\ repeat$: $P(A|B) := \frac{P(A \cap B)}{P(B)}$
- $P(A = 2|G = 3) = \frac{P(A=2 \cap G=3)}{P(G=3)}$
- $P(A = 1|G = 3) = \frac{P(A=1 \cap G=3)}{P(G=3)}$

Conditional Probability (Numerator=Zähler)

- Monty-Hall paradox
  - Is it useful to change the door?
- A = number of door with car
- G = number of opened door



$$P(A=1 \cap G=2) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

$$P(A=1 \cap G=3) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

$$P(A=2 \cap G=3) = \frac{1}{3} \cdot 1 = \frac{1}{3}$$

$$P(A=3 \cap G=2) = \frac{1}{3} \cdot 1 = \frac{1}{3}$$

Conditional Probability (Denominator=Nenner)

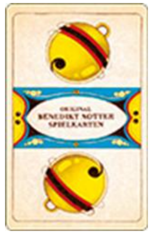
- What is missing to compute the conditional prob?
  - $P(G=3)\ =?$
- Theorem of Total Probability:
  - $P(G=3)\ = P(A=1 \cap G=3)\ + P(A=2 \cap G=3)\ + P(A=3 \cap G=3)$
    $= {}^{1}/_{6} + {}^{1}/_{3} + 0 = {}^{1}/_{2}$
- Definition of conditional probability:
  - $P(A|B) = \frac{P(A \cap B)}{P(B)}$
  - ➢ $P(A=2|G=3) = \frac{P(A=2 \cap G=3)}{P(G=3)} = \frac{1/3}{1/2} = \frac{2}{3}$
  - ➢ $P(A=1|G=3)\ = \frac{P(A=1 \cap G=3)}{P(G=3)} = \frac{1/6}{1/2} = \frac{1}{3}$

It's worth changing!

Random variables

- Example Jassen no Trumps (https://sonichits.com/video/Emil_Steinberger/Der_Jasser?track=1)

$\omega = Ace$



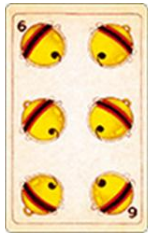
$\mapsto \quad X(\omega) = 11$

$\omega = King$

$\mapsto \quad X(\omega) = 4$

$\vdots \qquad \vdots \qquad \vdots$

$\omega = Sächsi$



$\mapsto \quad X(\omega) = 0$

X a function:

$$X: \ \Omega \to \mathbb{R}$$
$$\omega \mapsto X(\omega)$$

Probability for a score x

- $P(X = x) = P(\{\omega | X(\omega) = x\}) = \sum_{\omega; X(\omega)=x} P(\omega)$
- Exp. Jassen
  - Probability for 4 points, i.e. $P(X = 4)$
  - $= P(\{\omega; \omega = \text{any king}\})$
  - $= P(\text{oakling king}) + P(\text{rose king}) + P(\text{diamond king}) + P(\text{shield king})$

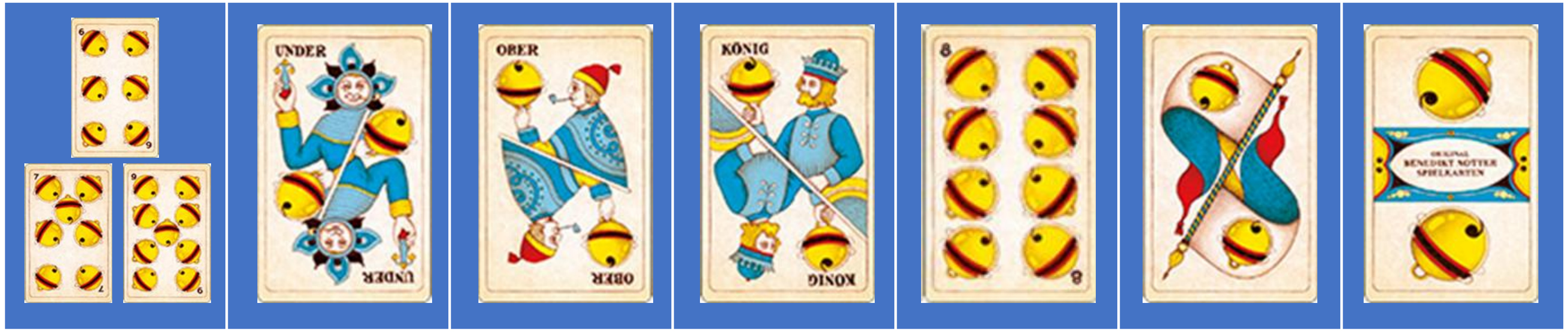
  - $= {}^{1}/_{36} + {}^{1}/_{36} + {}^{1}/_{36} + {}^{1}/_{36} = {}^{4}/_{36} = {}^{1}/_{9} \approx 11\%$

Probability function

- List all $P(X = x)$ for all possible values of x
- It holds that:

$$\sum_{all\ x} P(X = x) = 1$$

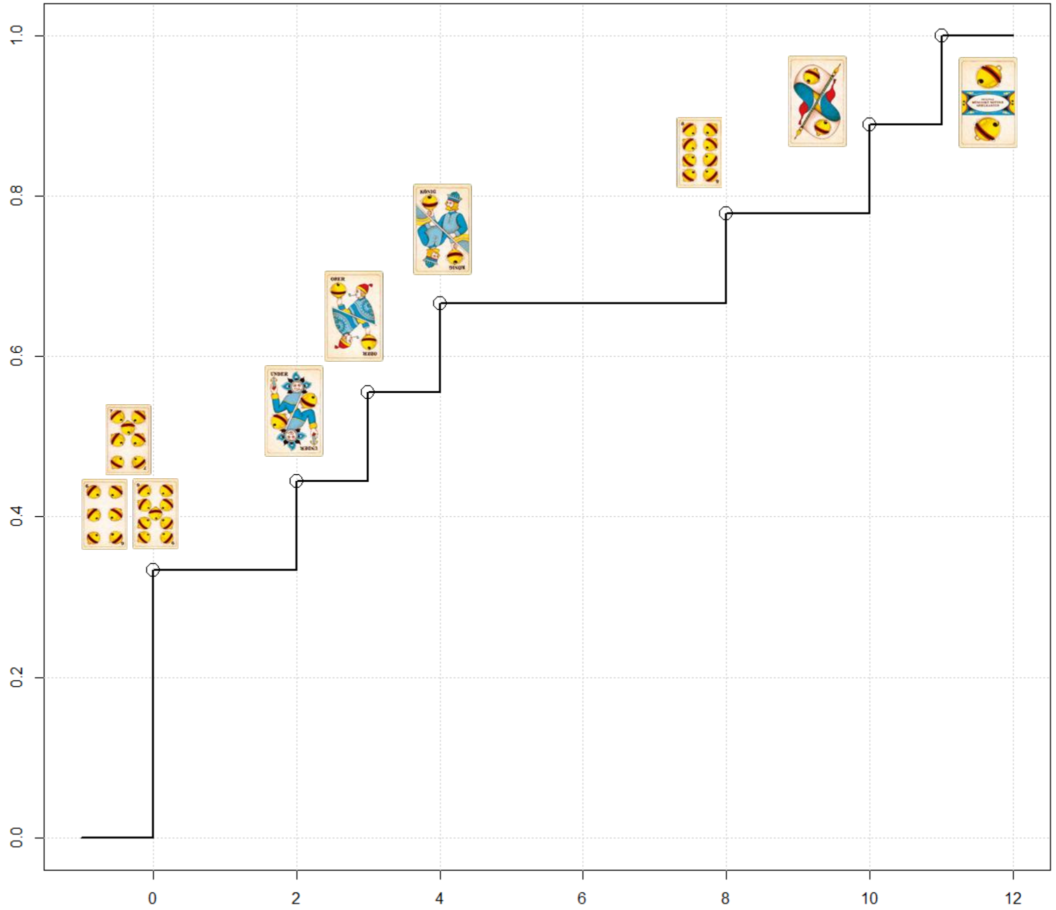
- Exp. Jassen (no trumps)



x	0	2	3	4	8	10	11
$\boldsymbol{P(X = x)}$	$\boldsymbol{12/36}$	$\boldsymbol{1/9}$	$\boldsymbol{1/9}$	$\boldsymbol{1/9}$	$\boldsymbol{1/9}$	$\boldsymbol{1/9}$	$\boldsymbol{1/9}$

Cumulative distribution function (c.d.f.)

- $P(X \leq x) \in [0,1]$
- $P(X \leq x) = \sum_{z \leq x} P(X = z)$
- monotonously increasing from 0 to 1



x	0	2	3	4	8	10	11
$\boldsymbol{P(X = x)}$	$\boldsymbol{^{12}/_{36}}$	$\boldsymbol{^{1}/_{9}}$	$\boldsymbol{^{1}/_{9}}$	$\boldsymbol{^{1}/_{9}}$	$\boldsymbol{^{1}/_{9}}$	$\boldsymbol{^{1}/_{9}}$	$\boldsymbol{^{1}/_{9}}$

Parameters of a discrete distribution

- Mean of X (Erwartungswert/Mittelwert)

$$E(X) = \sum_x x \cdot P(X = x)$$

- Standard deviation of X (Standardabweichung) – **not: standard error/standardfehler**

$$\sigma(X) = \sqrt{Var(X)}$$

- ...where variance of X (Varianz)

$$Var(X) = \sum_x (x - E(X))^2 \cdot P(X = x)$$

- Exp.: Jassen (no trumps)
  - $E(X) = 0 \cdot \frac{3}{9} + 2 \cdot \frac{1}{9} + 3 \cdot \frac{1}{9} + 4 \cdot \frac{1}{9} + 8 \cdot \frac{1}{9} + 10 \cdot \frac{1}{9} + 11 \cdot \frac{1}{9} \approx 4.2$ (Center of Gravity)
  - $Var(X) = \frac{3}{9}(0 - 4.2)^2 + \frac{1}{9}(2 - 4.2)^2 + \cdots + \frac{1}{9}(11 - 4.2)^2 \approx 17.1 \;\rightarrow \sigma(X) \approx 4.1$

How to understand continous random variables

(Probability-)Density (for continuous values)

- $F(x) \coloneqq P(X \leq x)$ cumulative distribution function (c.d.f.)

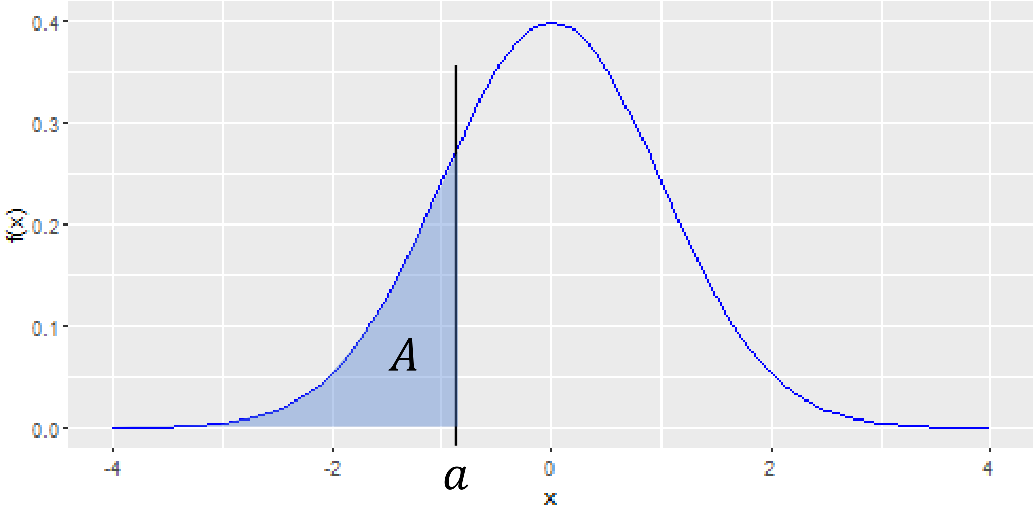
$$\Rightarrow F(x) = \int_{-\infty}^{x} f(y)dy$$

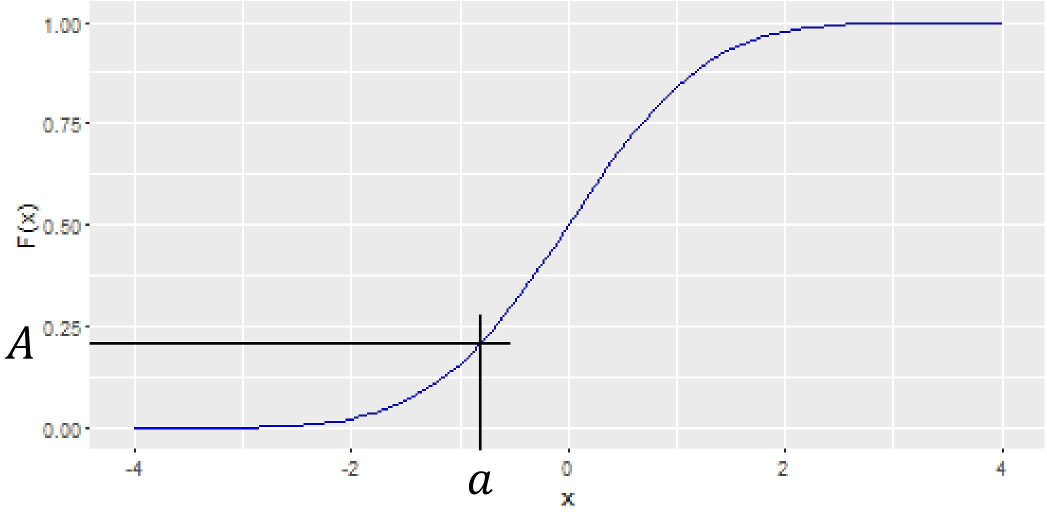
$f(x) = \frac{d}{dx}F(x)$, so f = F'

- Properties of $f(x)$
  - $f(x) \geq 0$ for all x
  - $P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x)dx$
  - $\int_{-\infty}^{\infty} f(x)dx = 1$

(Probability-)Density

$F(x) \coloneqq P(X \leq x)$ cumulative distribution function (c.d.f.)





$P(X \leq a) = A$

$$P(X \leq a) = \int_{-\infty}^{a} f(y)dy = A$$

Parameters of a continuous distribution

- Mean

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

- Variance

$$\mathrm{Var}(X) = \mathrm{E}\left(\left(\mathrm{X} - \mathrm{E}(\mathrm{X}) \right)^2 \right) = \int_{-\infty}^{\infty} \left(x - E(X) \right)^2 \cdot f(x) dx$$

Comparison between discrete and continuous

- Mean of discrete X

$$E(X) = \sum_x x \cdot P(X = x)$$

- Variance of discrete X

$$Var(X) = \sum_x (x - E(X))^2 \cdot P(X = x)$$

- Mean of continuous X

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

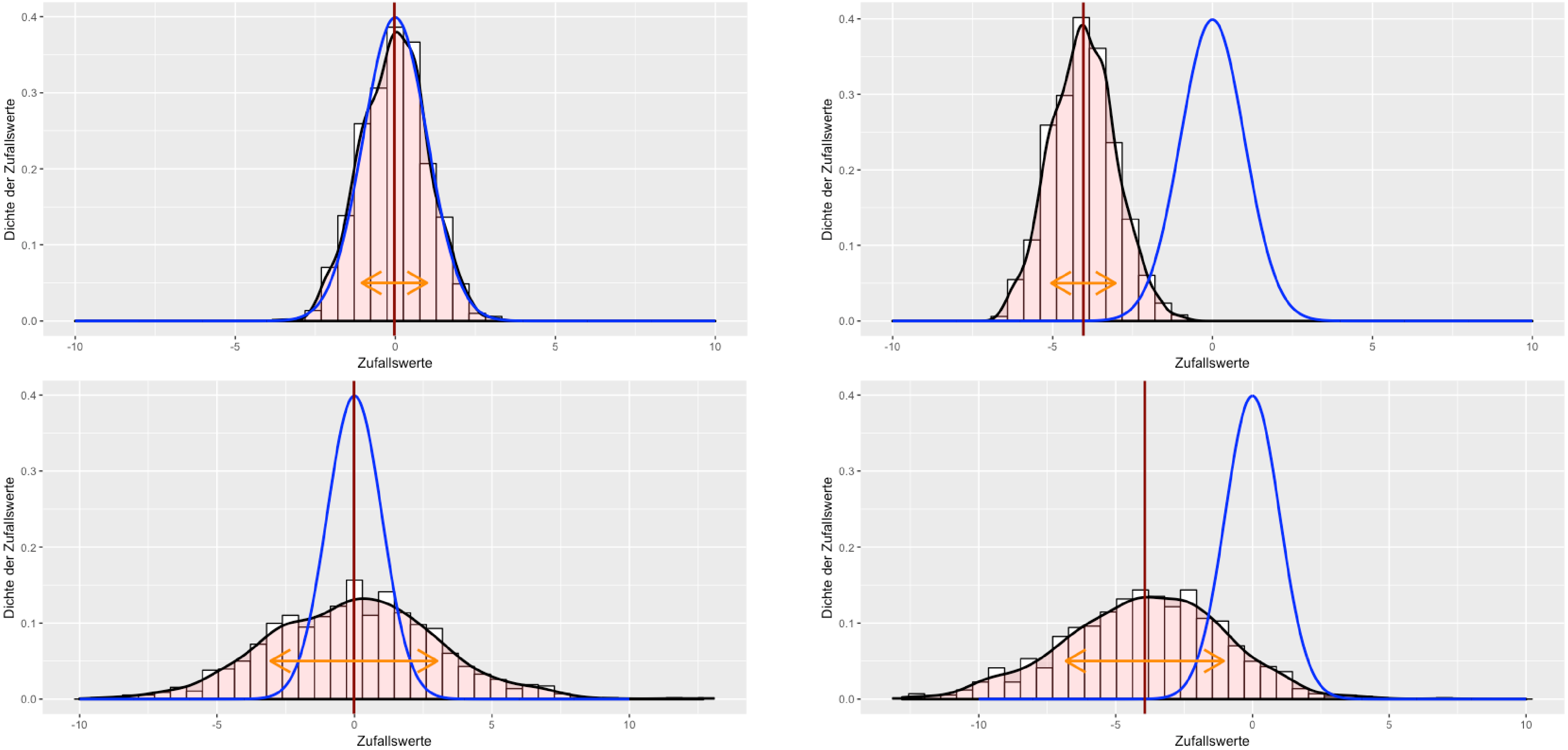
- Variance of continuous X

$$\mathrm{Var}(X) = \mathrm{E}\left(\left(\mathrm{X} - E(\mathrm{X})\right)^2\right)$$

$$= \int_{-\infty}^{\infty} \left(x - E(X)\right)^2 \cdot f(x)dx$$

https://stekhoven.shinyapps.io/mean_and_sd/

Mean and standard deviation



Binomial distribution - $\boldsymbol{Bin(n,p)}$

- Situation
  - Buy n tickets in a Tombola lottery
  - All tickets have the same chance of winning
  - Tickets are independent
- Random variable X: number of winnings in n tickets

➢ $X \sim Bin(n,p)$

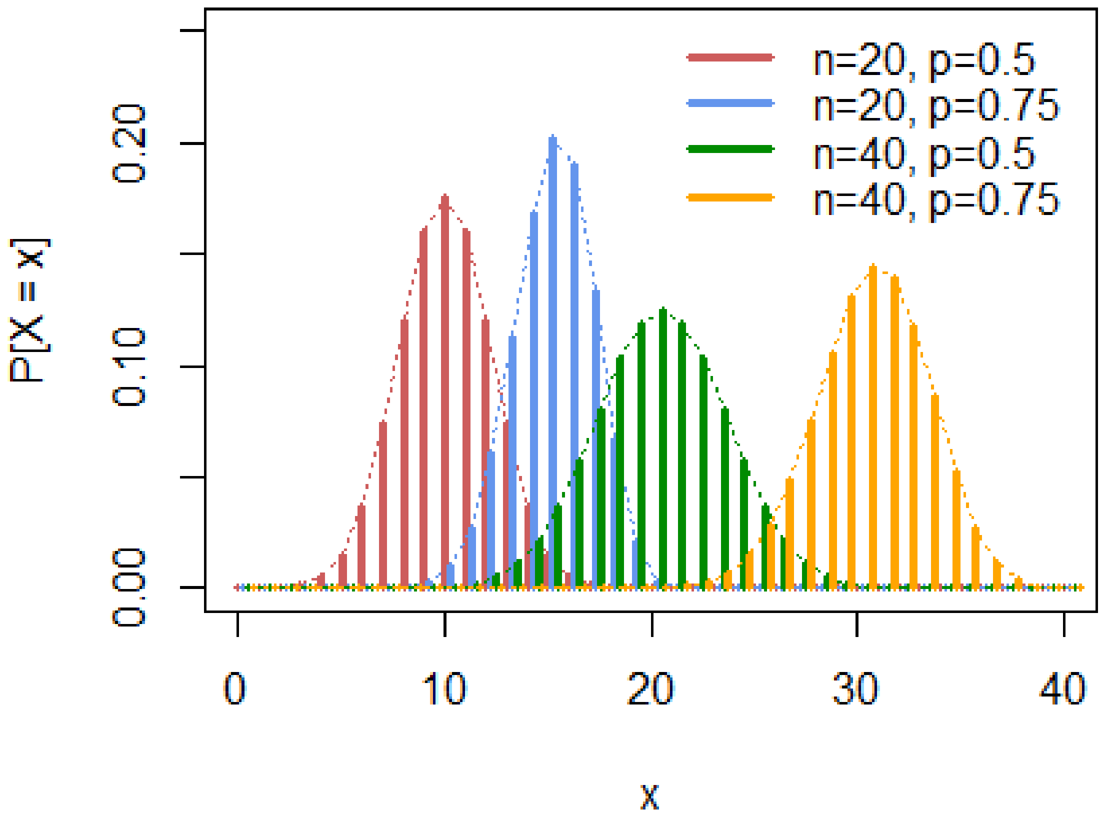
"X is binomially distributed with parameters n and p"

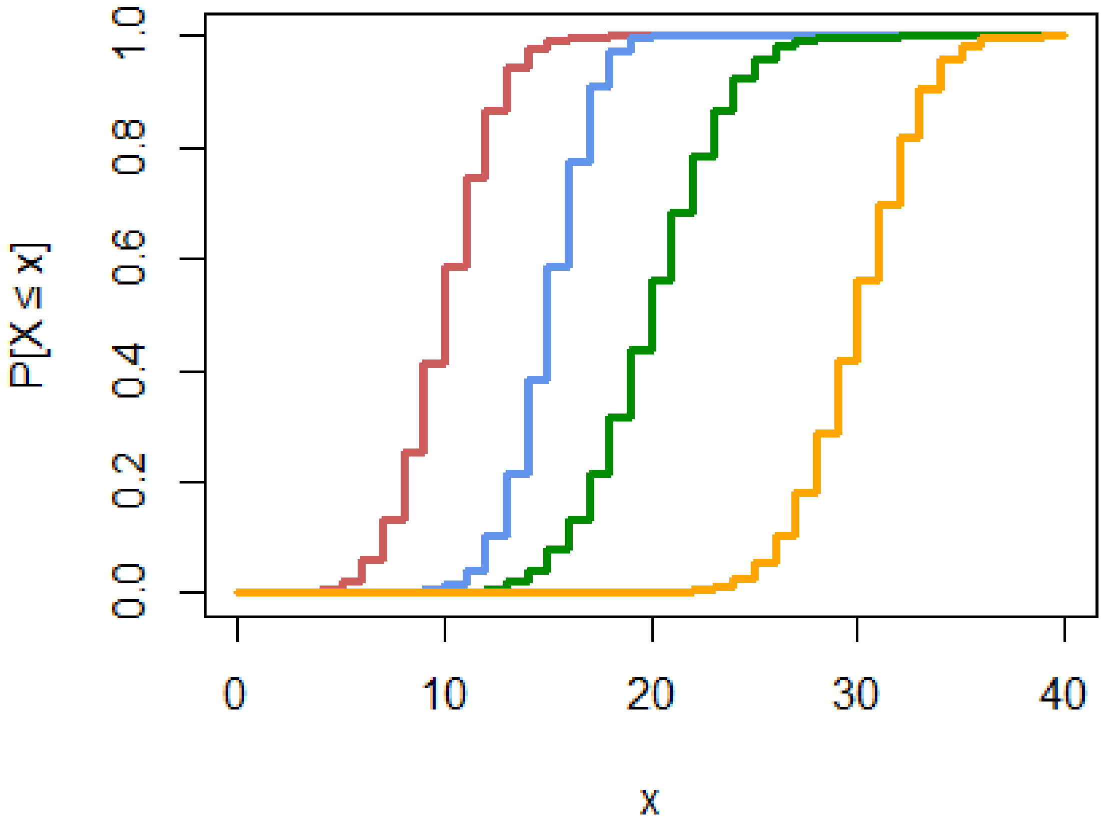
- Binomial coefficient: $\binom{n}{x} = \frac{n!}{x!(n-x)!}$

"How many possibilities to choose k objects from n?"

- $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{0, 1, 2, \dots, n\}$
- $E(X) = n \cdot p$
- $Var(X) = n \cdot p \cdot (1-p)$

Binomial distribution - $\boldsymbol{Bin(n,p)}$

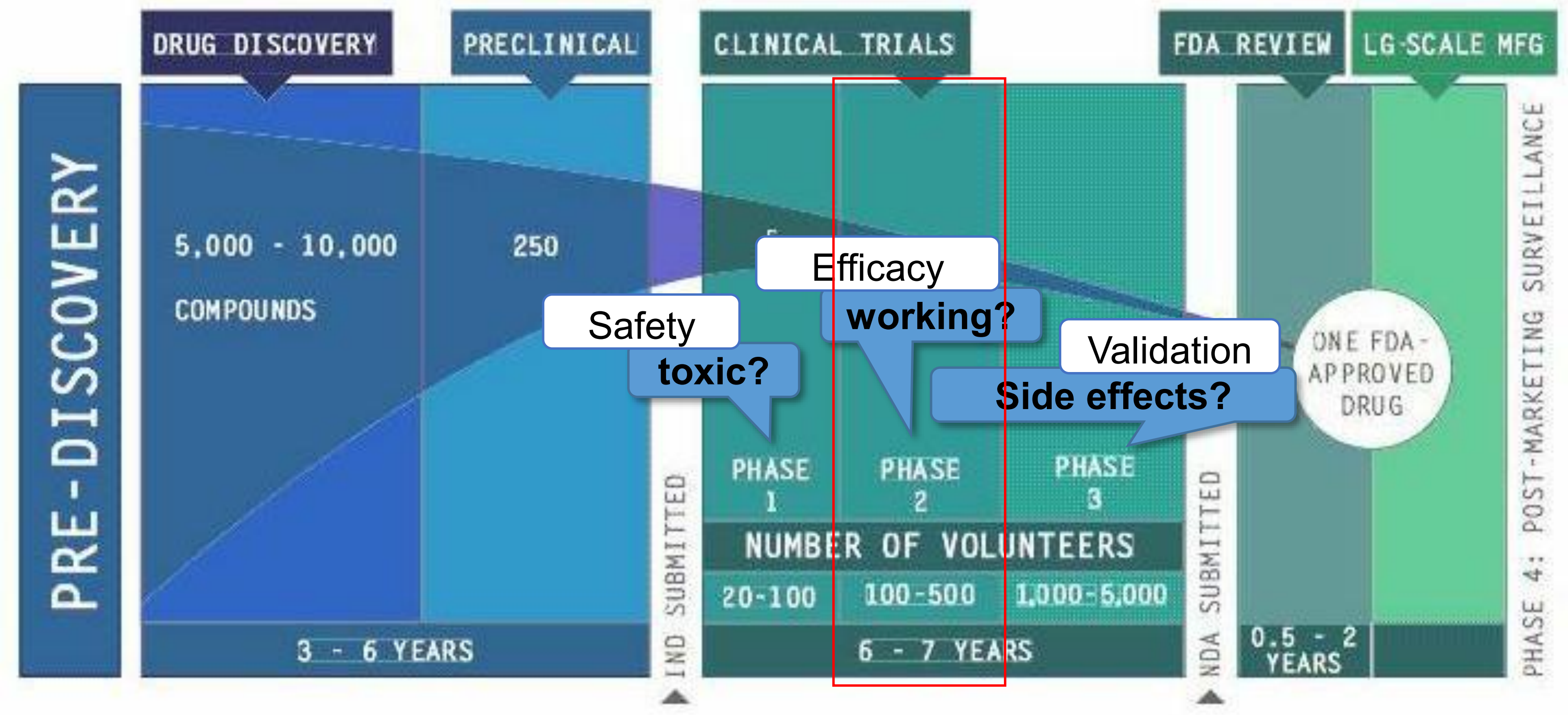




- Example: **LD10**, $X \sim Bin(20, 0.1)$, $\quad P(X = 2) = \binom{20}{2} \cdot 0.1^2 \cdot 0.9^{18} \approx 0.285$

$$E(X) = n \cdot p = 20 \cdot 0.1 = 2$$

Example «Clinical Trial»



- **Tickets**: all possible patients
- n **tickets bought**: participants in the trial
- **Winning**: participant is healed / goes into remission
- p: proportion of all possible patients healed / going into remission
- **IND** = Investigatin New Drug ; **NDA** = New Drug Application

Example – Phase 2 clinical trial (efficacy)

- Producer claims: drug heals in 80% of the cases
- Phase 2 trial: of 100 participants 73 are healed
  - Is this plausible, given a healing probability of 80%?
- X: no. of healed participants
- if producer is correct:

$$X \sim Bin(n = 100, p = 0.8)$$

- We test «$p = 0.8$» versus «$p < 0.8$» … why?

- $P(X \leq 73) = 0.056$

Uniform distribution – discrete – $Unif(n)$

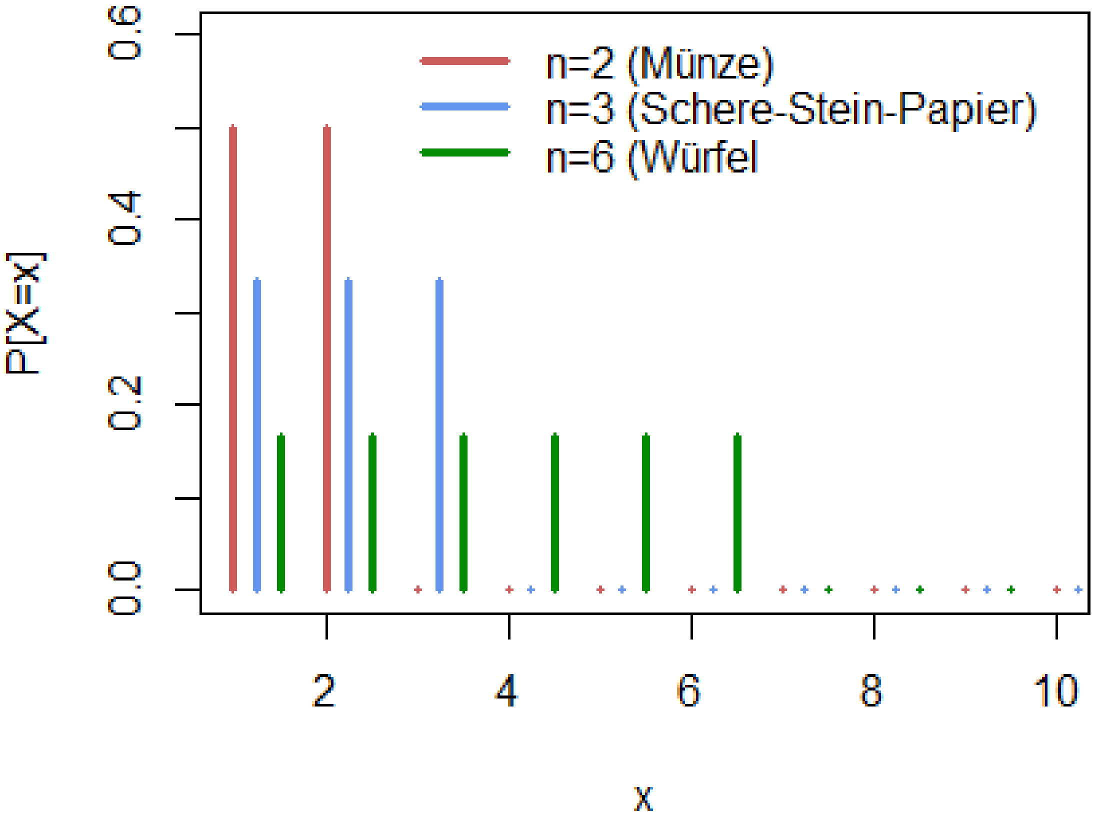
- Situation
  - Draw a number from {1, 2, 3, …, n}
  - All numbers have the same probability
- Random variable X: number drawn

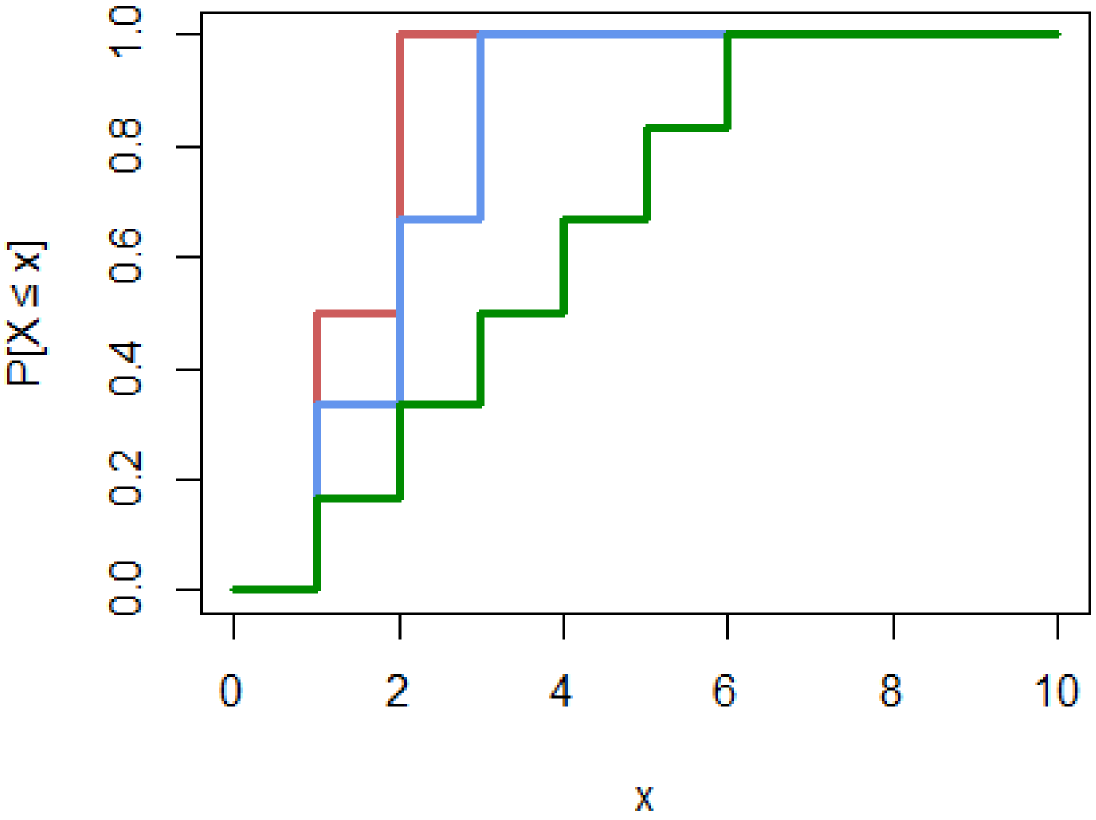
➢ $X \sim Unif(n)$

"X is uniformly distributed on the numbers 1 to n"

- $P(X = x) = \frac{1}{n}, \quad x \in \{1, 2, \dots, n\}$
- $E(X) = \frac{n+1}{2}$
- $Var(X) = \frac{(n+1)(n+2)}{12}$

Uniform distribution – discrete – $Unif(n)$



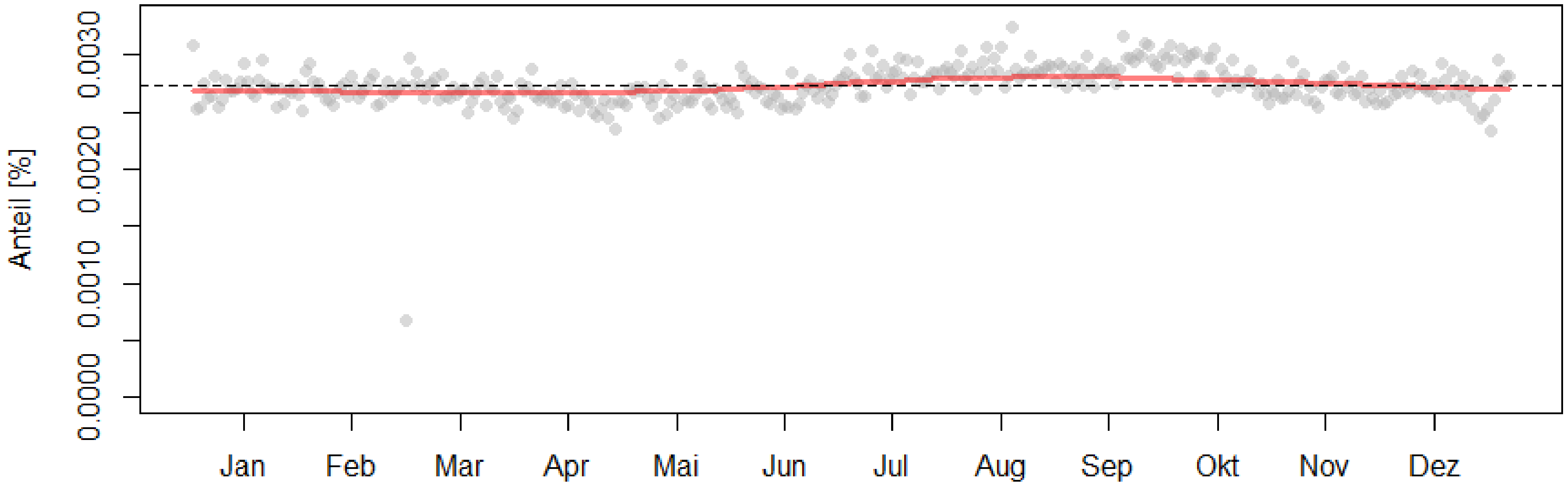


- Example: **Dice**, $X \sim Unif(6)$, $\mathrm{P}(X = x) = \frac{1}{6}$

$$E(X) = \frac{6+1}{2} = 3.5$$

Are birthdays uniformly distributed?

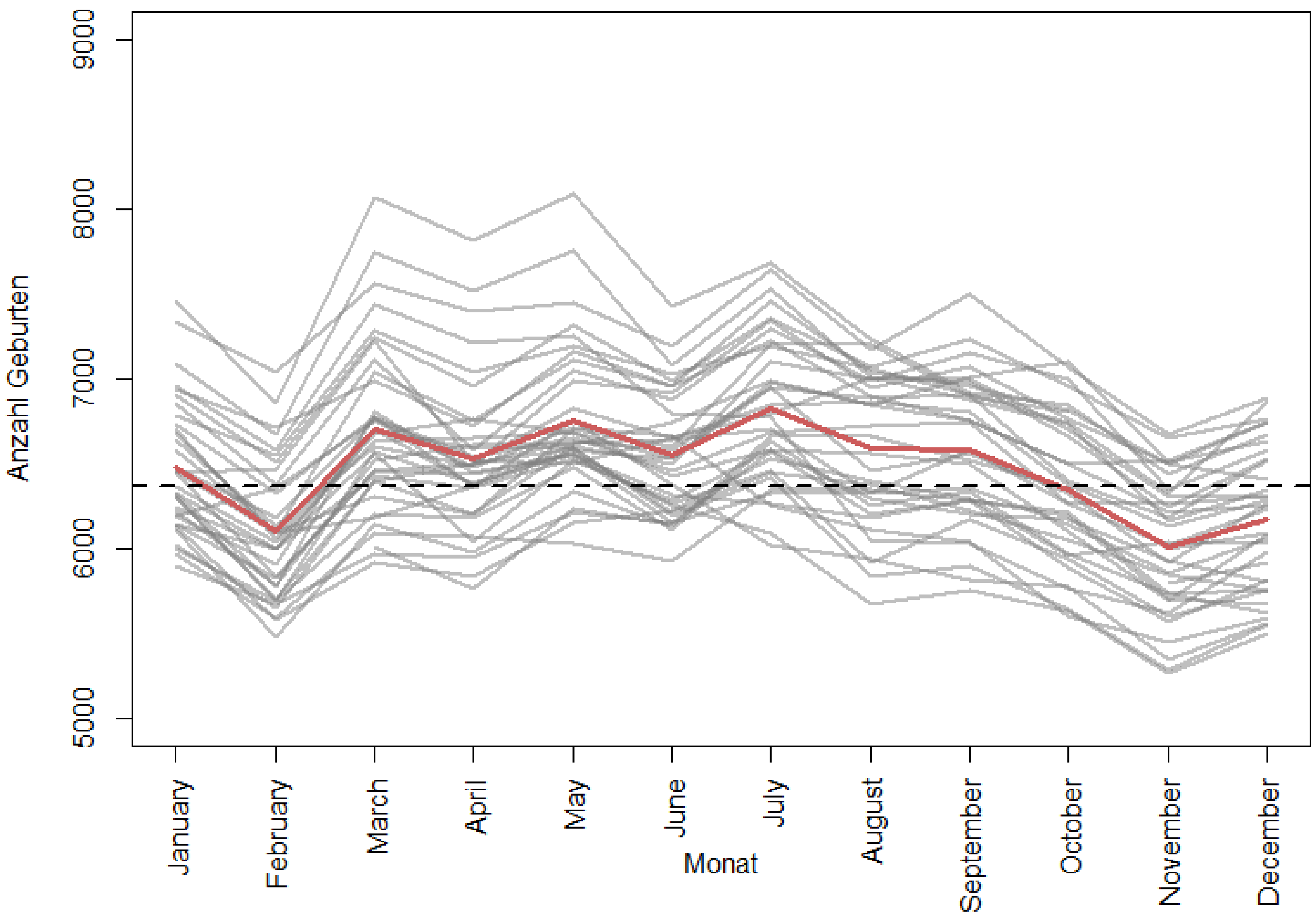
- Birthdays from a life insurance 1981 – 1994



- It seems on average they are!

Are birthdays uniformly distributed?

- What if just one country (Switzerland, 1973-2013)



Uniform distribution – continuous – $\boldsymbol{U(a,b)}$

- **Situation**:
  Each value inside $[a, b]$ has the same probability
- Random variable X: a value in $[a, b]$
- $X \sim Unif(a, b)$
- Density:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{else} \end{cases}$$

- C.d.f.:

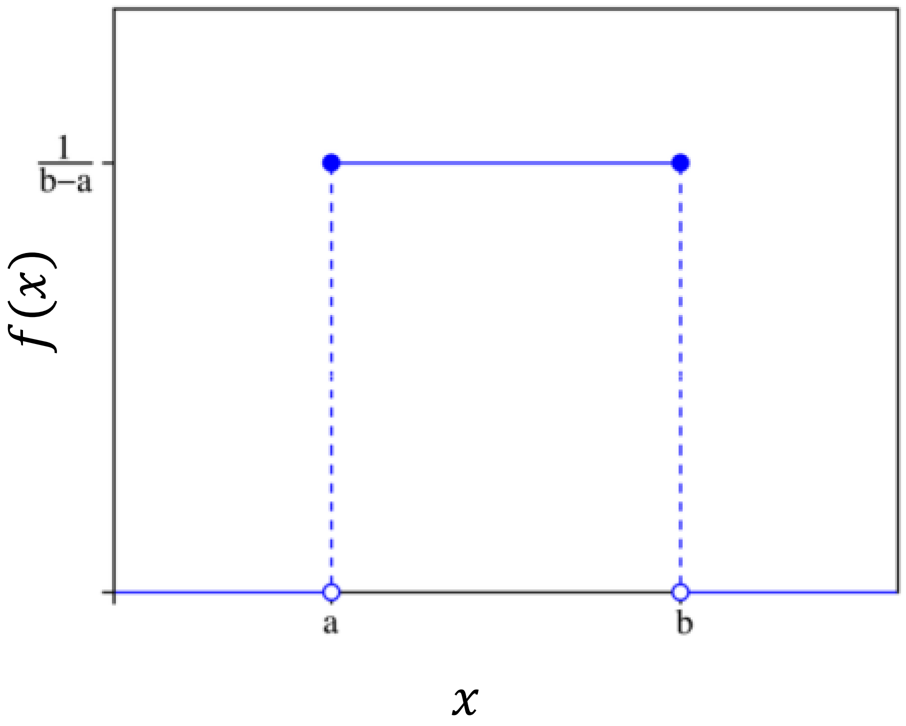
$$F(x) = \begin{cases} 0 & if\ x < a \\ \dfrac{x-a}{b-a} & if\ a \leq x < b \\ 1 & if\ x \geq b \end{cases}$$

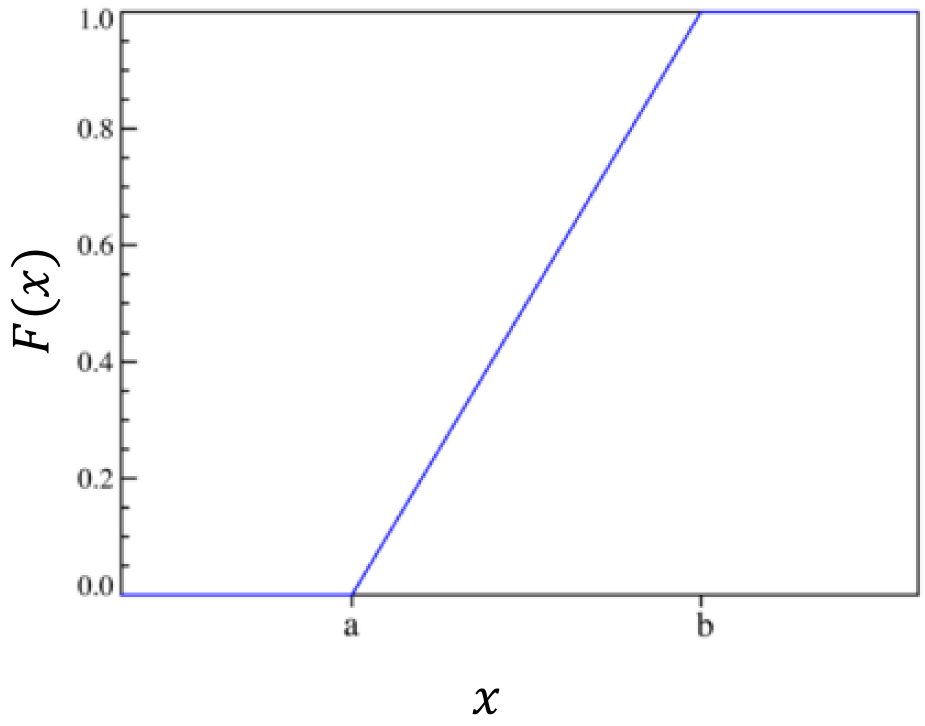
"X is uniformly distributed on the interval [a, b]"

$$\boldsymbol{E(X) = \frac{b+a}{2}}$$

$$\boldsymbol{Var(X) = \frac{(b-a)^2}{12}}$$

Uniform distribution – continuous – $\boldsymbol{U(a, b)}$





Uniform distribution – continuous – $\boldsymbol{U(a,b)}$

- Trams come every 7 minutes in Zurich.
- You arrive by chance at a stop.
- What is the probability that you only have to wait for at most 1 minute?

- X: waiting time in minutes

➢ $X \sim Unif(0,7)$

- $P(X \leq 1) = F(1) = \frac{1-0}{7-0} = \frac{1}{7}$

Poisson distribution – $Pois(\lambda)$

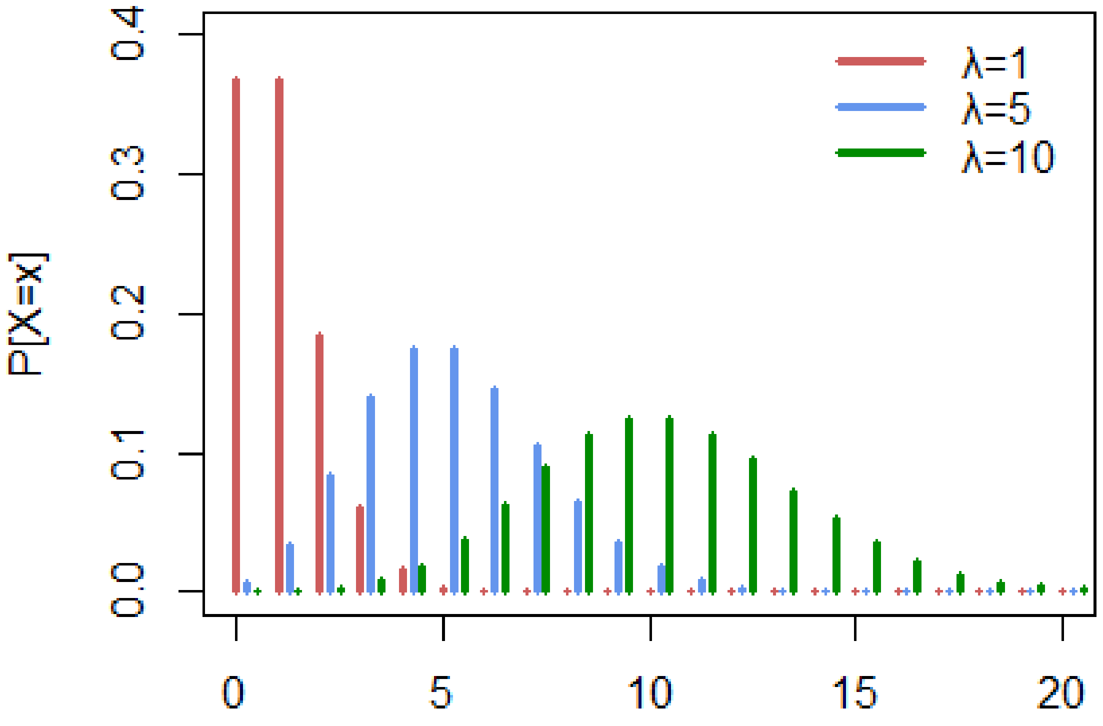
- Situation
  - Number of (rare) events are counted within a fixed window of time
- Random variable X: number of rare events

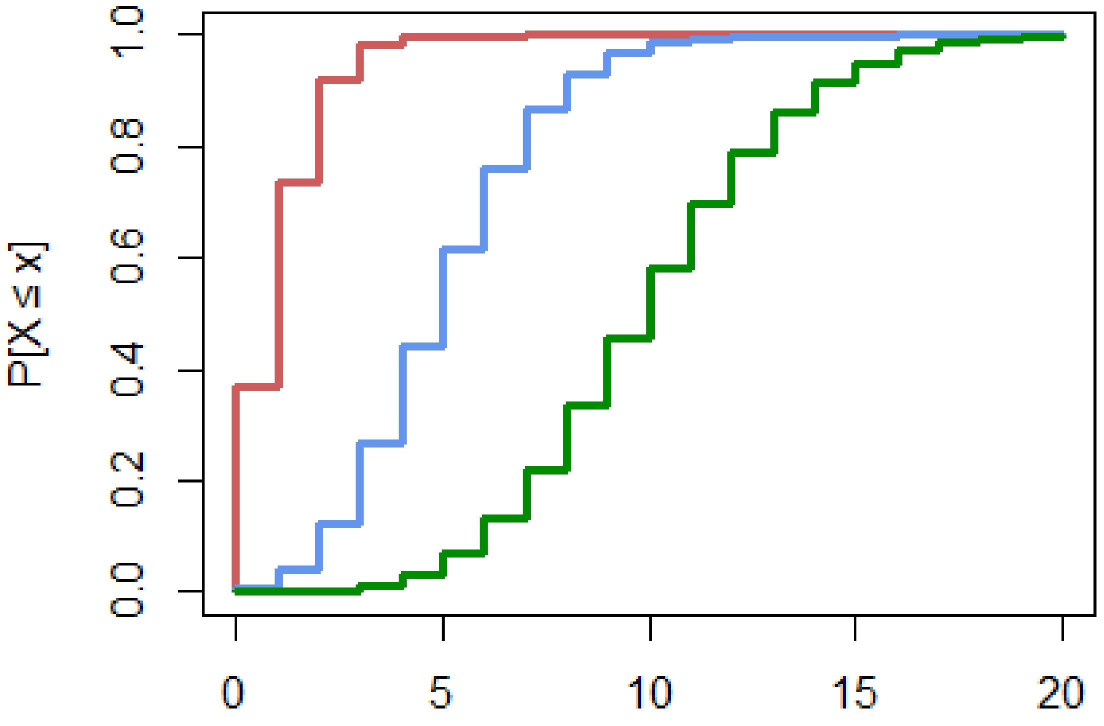
➢ $X \sim Pois(\lambda)$

“X is Poisson distributed with a rate of λ”

- $P(X = x) = \frac{\lambda^x}{x!}\exp(-\lambda), \quad x \in \{0,1,\dots,\infty\}$
- $E(X) = \lambda$
- $Var(X) = \lambda$

Poisson distribution – $Pois(\lambda)$





- Example: **Cesium-137**, $X \sim Pois\left(\frac{\ln 2}{27}\right) \approx Pois(0.026)$
  - ...rate per year, i.e. $8.2 \times 10^{-10}\,\frac{1}{s}$

$$P(X=1) = \frac{0.026^1}{1!} e^{-0.026} \approx 0.025$$
$$E(X) = \lambda = 0.026$$

➢ $1\mu g$ ^{137}Cs contains 10^{15} nuclei, i.e. $\mu = N \cdot p = 8.2 \times 10^5$ decays per second

A speciality of Poisson

- Assume:
  - $X \sim Pois(\lambda_1), Y \sim Pois(\lambda_2)$
  - X, Y are independent
- Form a new random variable $Z = X + Y$

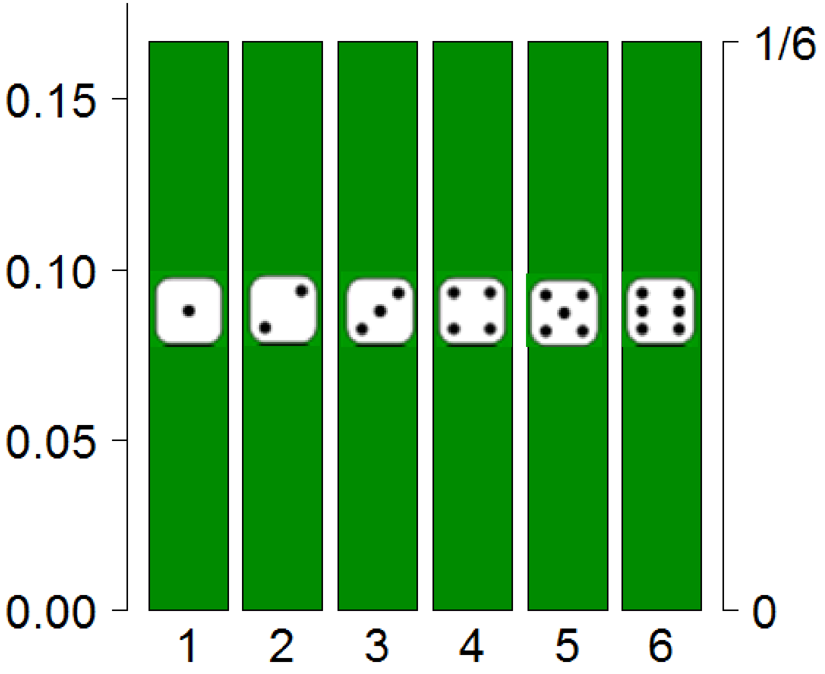
$$\Rightarrow Z \sim Pois(\lambda_1 + \lambda_2)$$

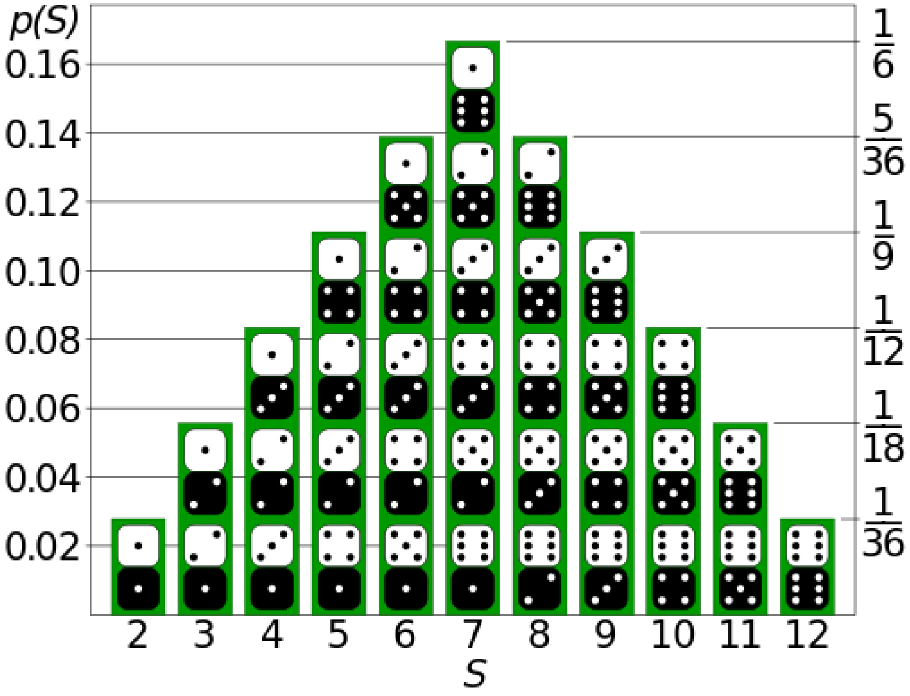
Such results are not necessarily true with other distributions!

Sums of RVs tend to give something else…

Exp: X and Y are two independent, fair dice

$S = X + Y$ is not uniformly distributes, 2 is rare, 7 is common





Hypergeometric distribution – $Hyper(N, n, m)$

- Situation
  - Urn with N balls, m white and $N - m$ black
  - Draw n balls (without replacements)
  - How many drawn balls are white?
- Random variable X: number of white balls drawn

➢ $X \sim Hyper(N, n, m)$

"X is hypergeometrically distributed with parameters N, n and m"

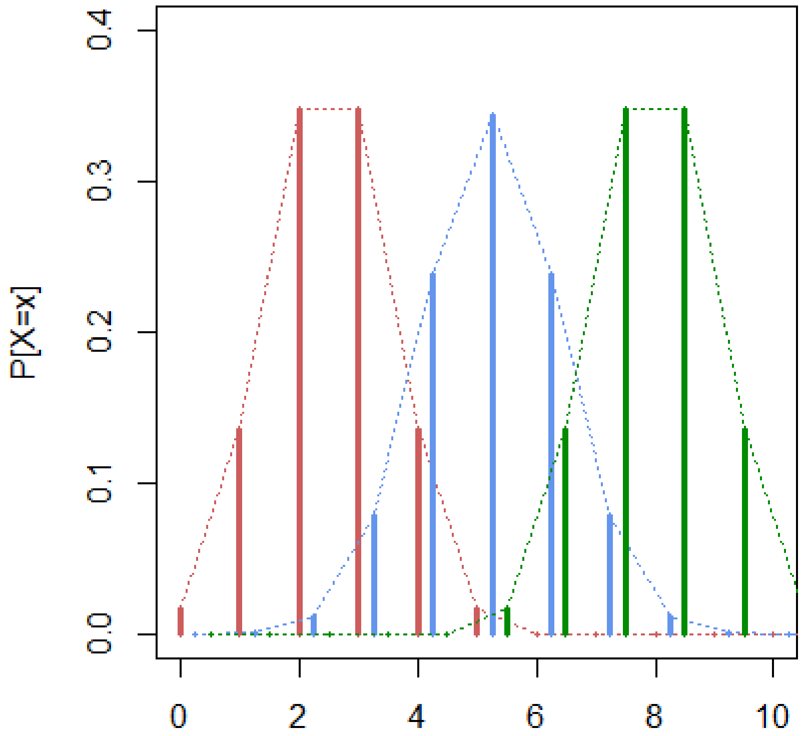
"favorable"

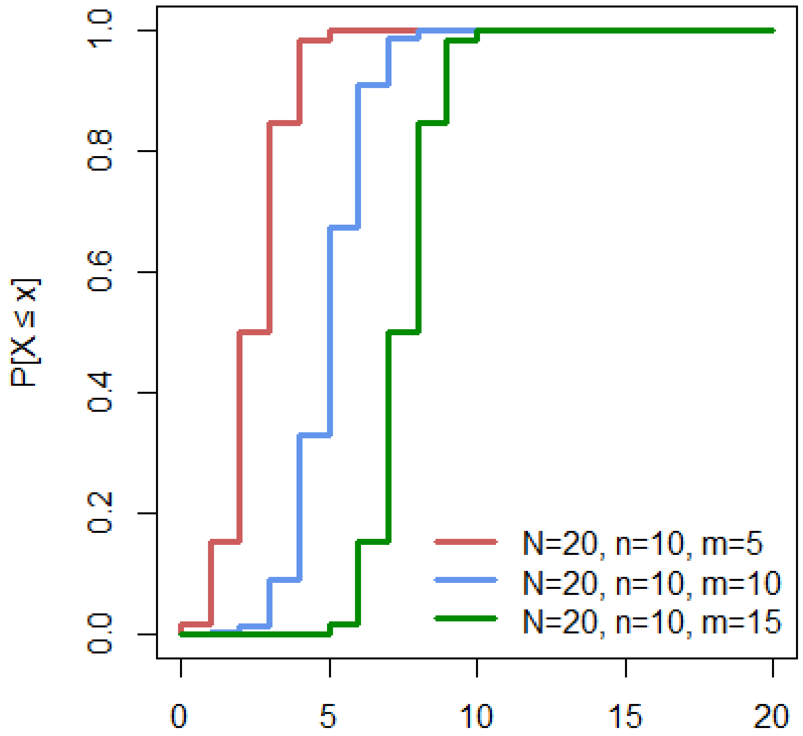
- $P(X = x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, \quad x \in \{\max(0, n + m - N), 1, \dots, \min(m, n)\}$

"possible"

- $E(X) = \frac{(n \cdot m)}{N}$, $Var(X)$ totally weird, look it up yourself on WIKIPEDIA

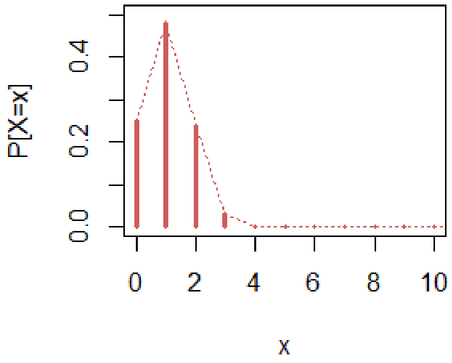
Hypergeometric distribution – $Hyper(N, n, m)$





- Example: **Urn**, $X \sim Hyper(N = 20, n = 3, m = 7)$
  - 20 balls, 7 white, 3 drawn without replacements

$$P(X = 1) = \frac{\binom{7}{1}\binom{13}{2}}{\binom{20}{3}} \approx 0.48, \qquad E(X) = \frac{3 \cdot 7}{20} = 1.05$$



Example– Phase 2 clinical trial

$$P(X = x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

- Double-blind, randomized trial

"drawn white balls" = x

"white balls" = m

	treatment	control	Total
Healed	15	9	24
Not healed	10	11	21
Total	**25**	**20**	**45**

"balls drawn" = n

"total balls= N

- If the drug is not working: 24 participants are healed irrespective of their association to a treatment group

Example– Phase 2 clinical trial

$$P(X = x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

- RV X: number of participants healed in treatment group
- Under the $\mathcal{H}_0$ (no efficacy): $X \sim Hyper(N = 45, n = 25, m = 24)$

	treatment	control	Total
Healed	15	9	24
Not healed	10	11	21
Total	**25**	**20**	**45**

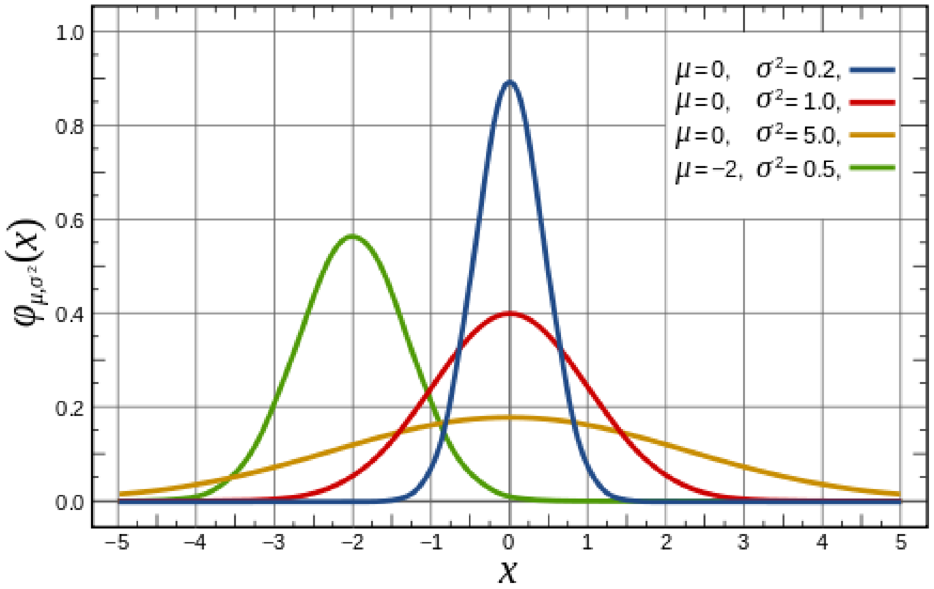
- Is it plausible to see 15 healed participants in the treatment group?

$$P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.76 = 0.24$$

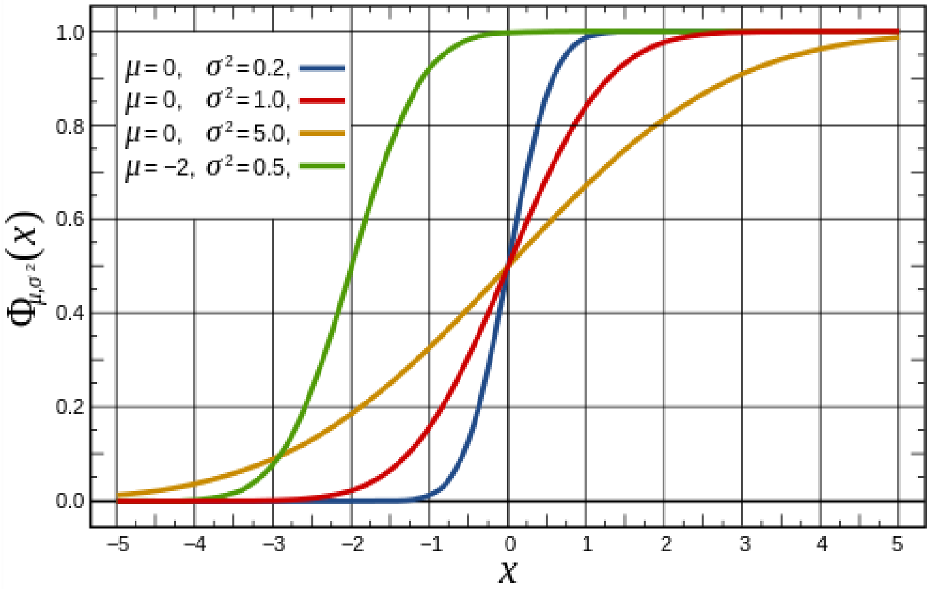
- If there is **no** efficacy, it seems quite possible to see 15 or more participants…

Normal distribution – $\mathcal{N}(\mu, \sigma^2)$

Density

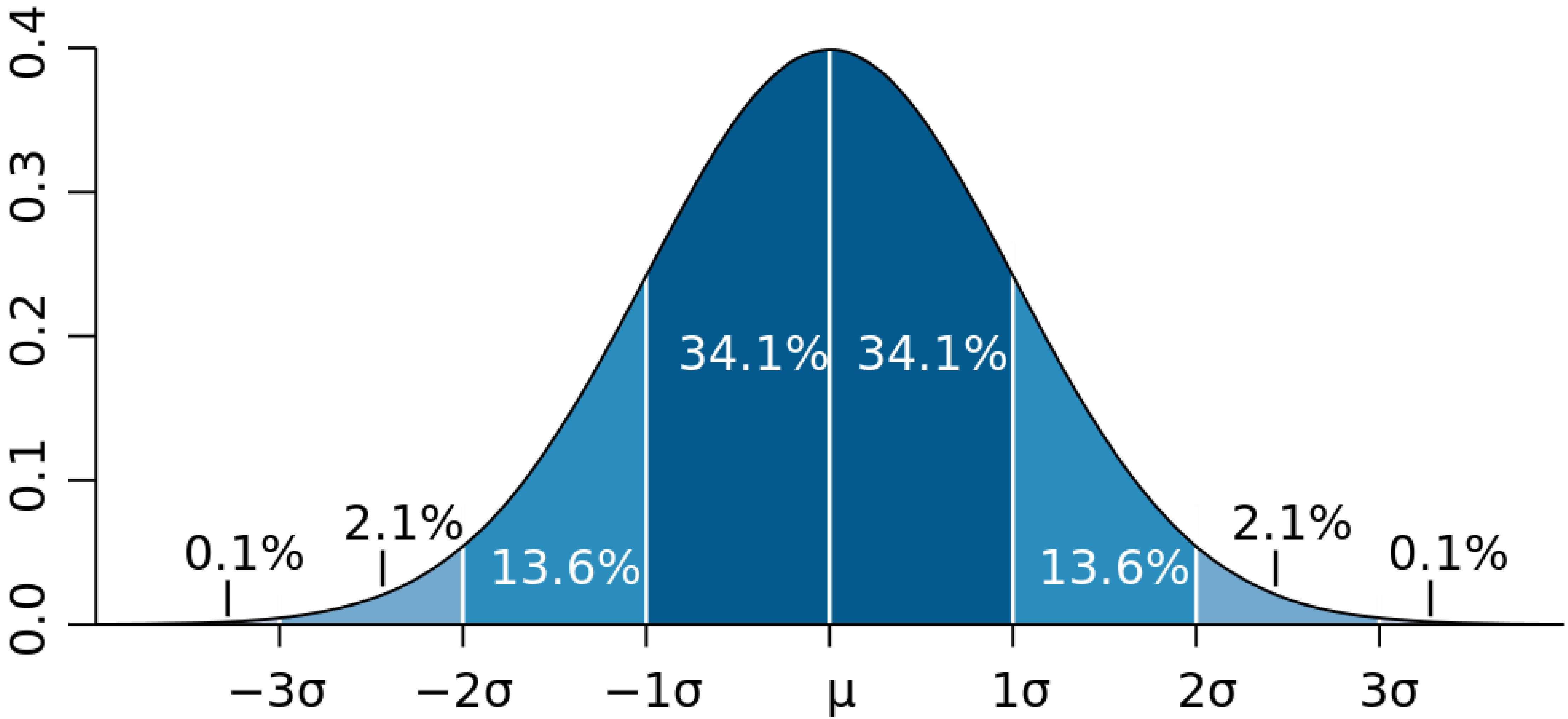


c.d.f



Speciality: sum of normal is normal again

Normal distribution – $\mathcal{N}(\mu, \sigma^2)$



Standard normal distribution

- $Z \sim \mathcal{N}(0,1)$
- Density denoted with φ : $\varphi(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right)$
- CDF denoted with ϕ : $\phi(x) = \int_{-\infty}^{x} \varphi(y)dy$

analytically not solvable

- Exp.: 95%-quantile

$$P(Z < 1.64) = \phi(1.64) = 0.9465$$

Speciality: normal distribution

- **Z-Transform**, let $X \sim \mathcal{N}(\mu, \sigma^2)$

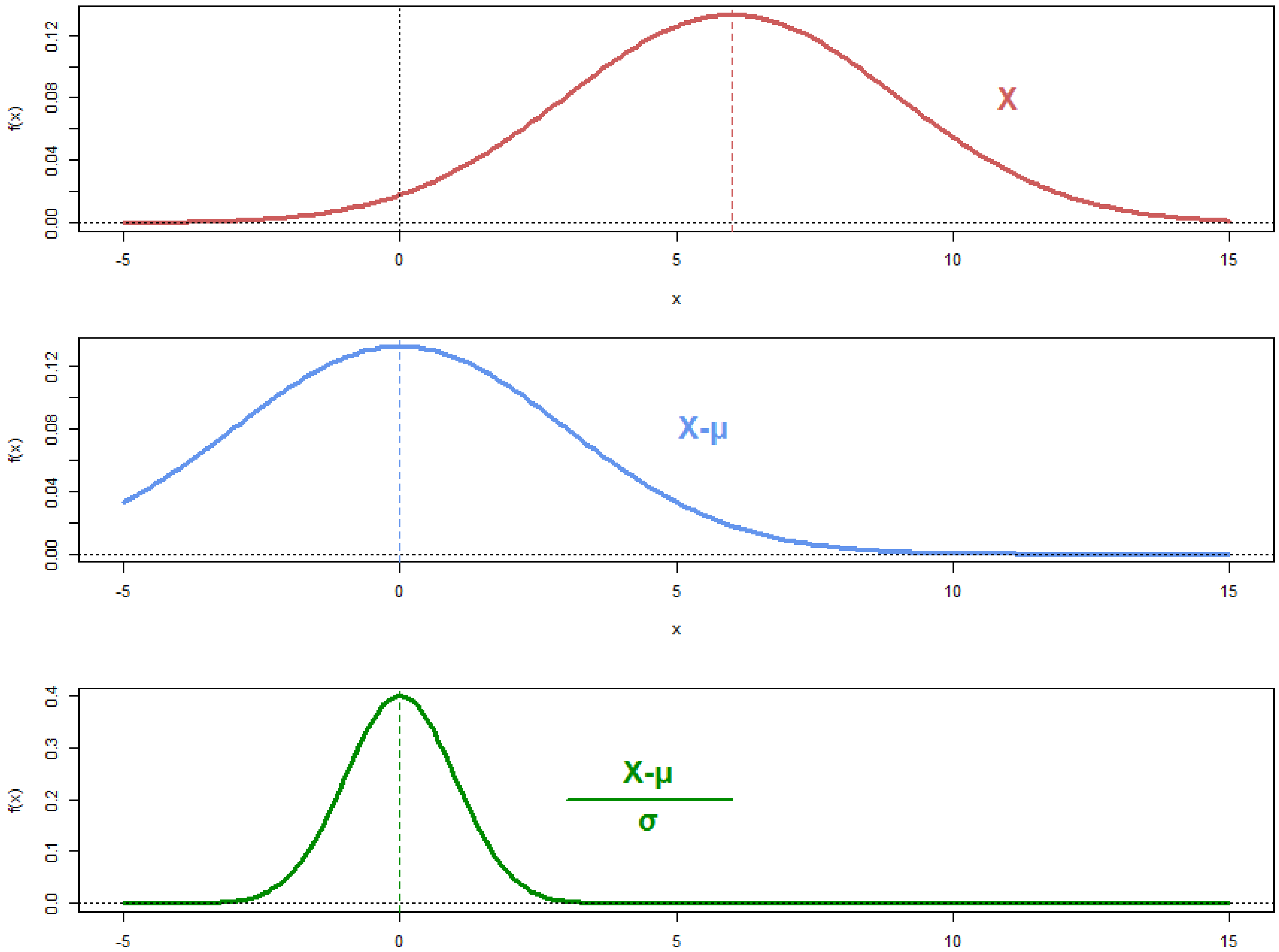
$$Z := \frac{X - \mu}{\sigma} = -\frac{\mu}{\sigma} + \frac{1}{\sigma} \cdot X = a + b \cdot X$$

$$\Rightarrow E(Z) = -\frac{\mu}{\sigma} + \frac{1}{\sigma} \cdot \mu = 0$$

$$\Rightarrow Var(Z) = \frac{1}{\sigma^2} \cdot \sigma^2 = 1$$

- ...then $Z \sim \mathcal{N}(0,1)$ → easy to handle!

Standardize (Z-transform), $\boldsymbol{X} \sim \boldsymbol{\mathcal{N}}(\mathbf{6}, \mathbf{9})$



Speciality: normal distribution

Example Z-Transform

- $X \sim \mathcal{N}(2, 2^2)$, what's $P(X \leq 5)$?

$$P(X \leq 5) =$$

Summary

- Conditional probability
- Random variables and their distributions
- Parameters of distributions: mean and variance
- Binomial, Uniform, Poisson, Hypergeometric, Normal