Statistical Methods in Particle Physics / WS 13

Lecture II

Probability Density Functions

Niklaus Berger

Physics Institute, University of Heidelberg

Emmy Noether-Programm

Deutsche Forschungsgemeinschaft

DFG

Recap of Lecture I: Kolmogorov Axioms

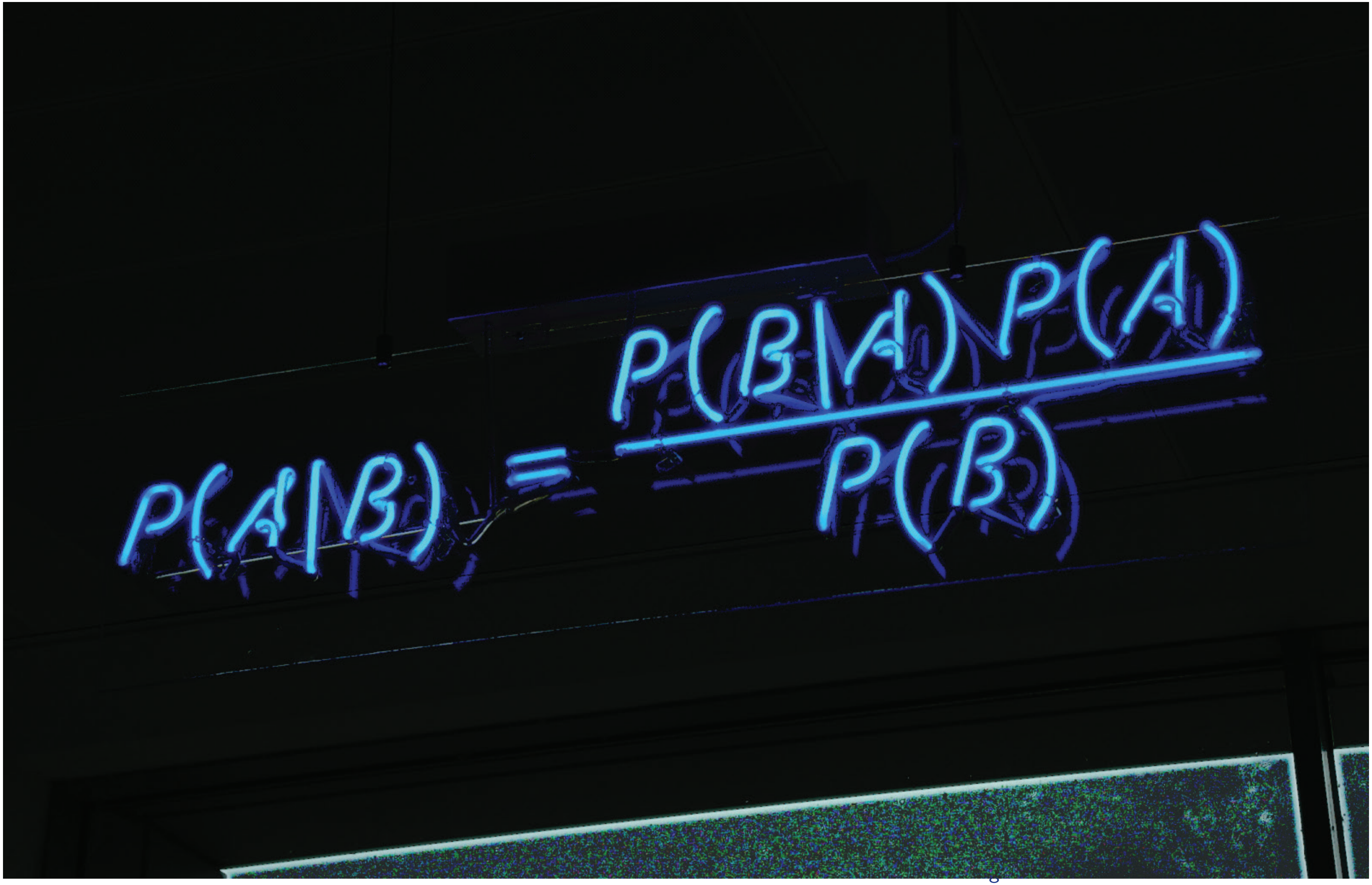
Ingredients:

- Set S of outcomes of the experiment (sample space)
- Subsets A, B... of S (technically need to form a Σ-Algebra)
- Mapping into the real numbers P(A) called the Probability

Kolmogorov Axioms:

- For all $A \subset S$, $P(A) \geq 0$
- $P(S) = 1$
- If $A \cap B = \emptyset$, $P(A \cup B) = P(A) + P(B)$

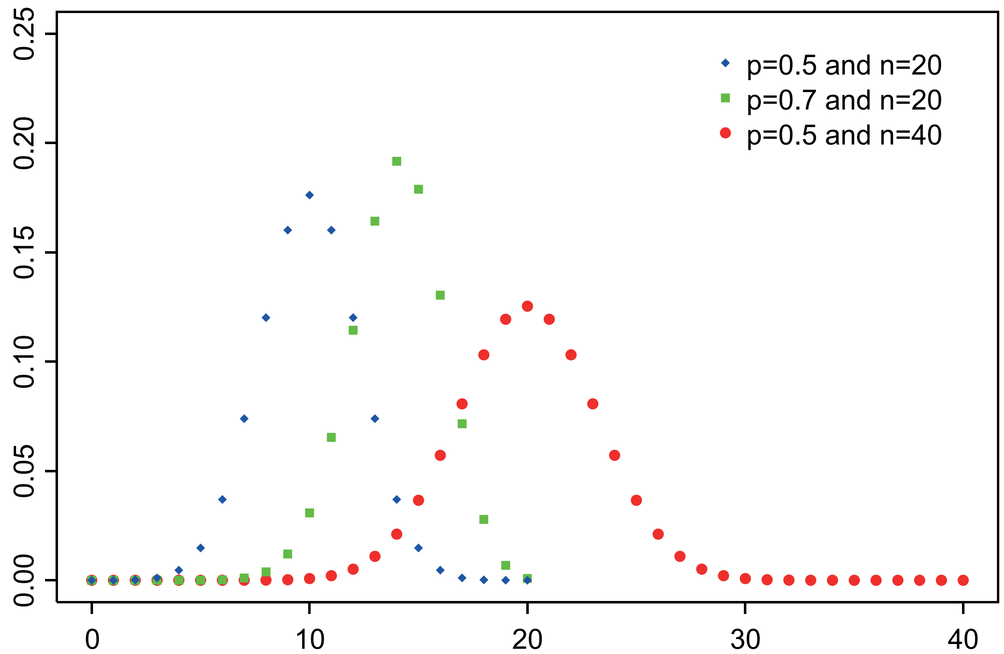
Recap of Lecture 1: Bayes' Theorem

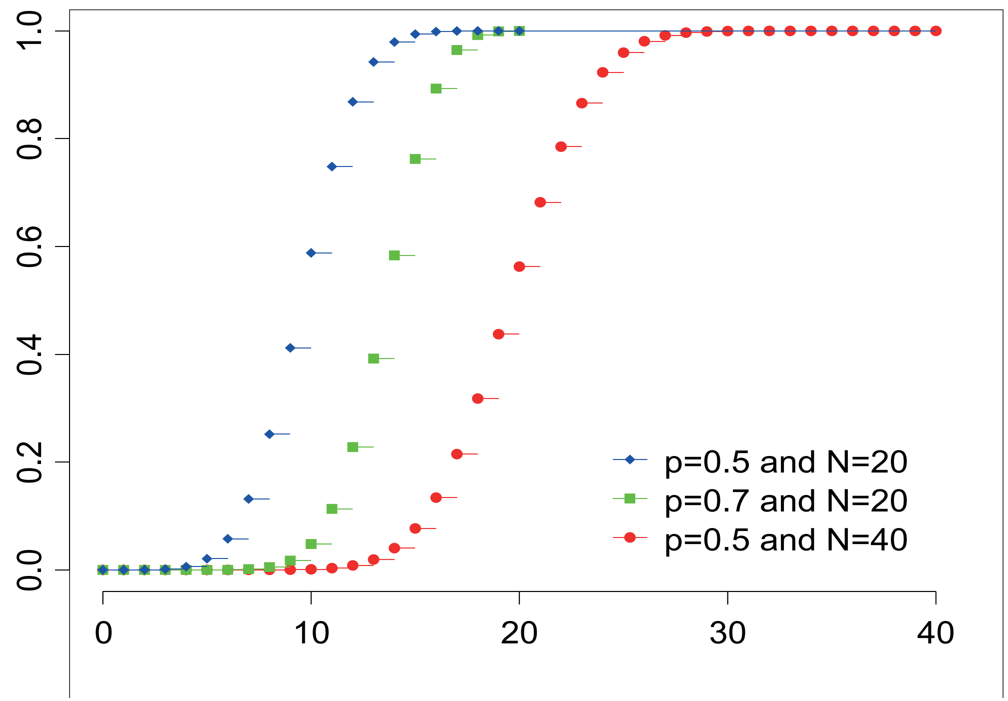


Part II:

Catalogue of Probability Density Functions

2.1. Binomial Distribution



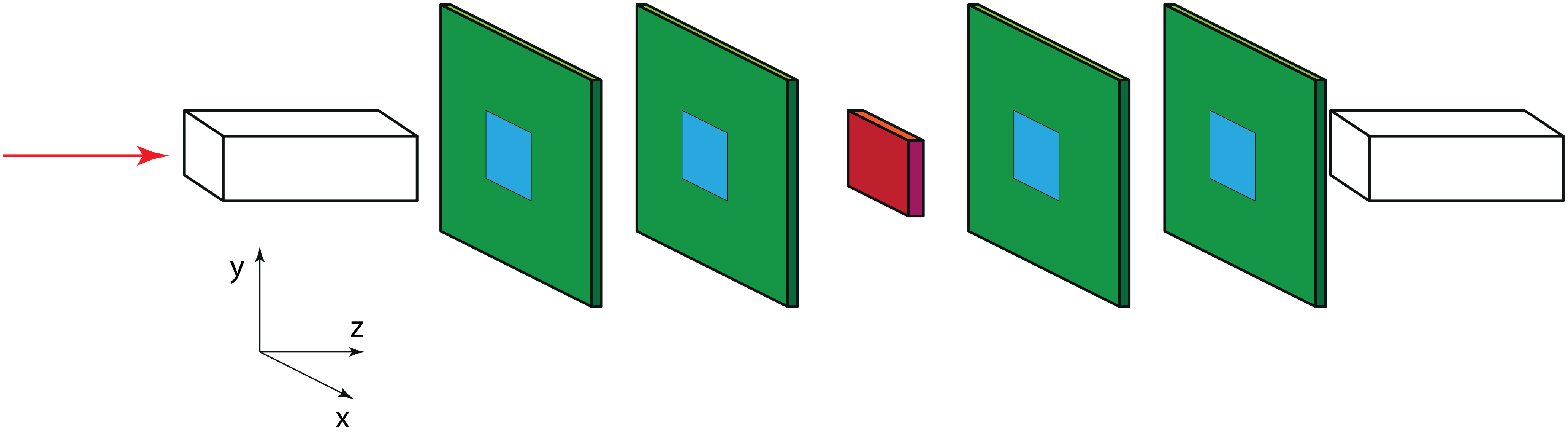


$$P_B(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

- Mean: $E[x] = np$
- Variance: $V[x] = np(1-p)$

Plot source: Wikipedia

Binomial Distribution: Application



Efficiency measurement:

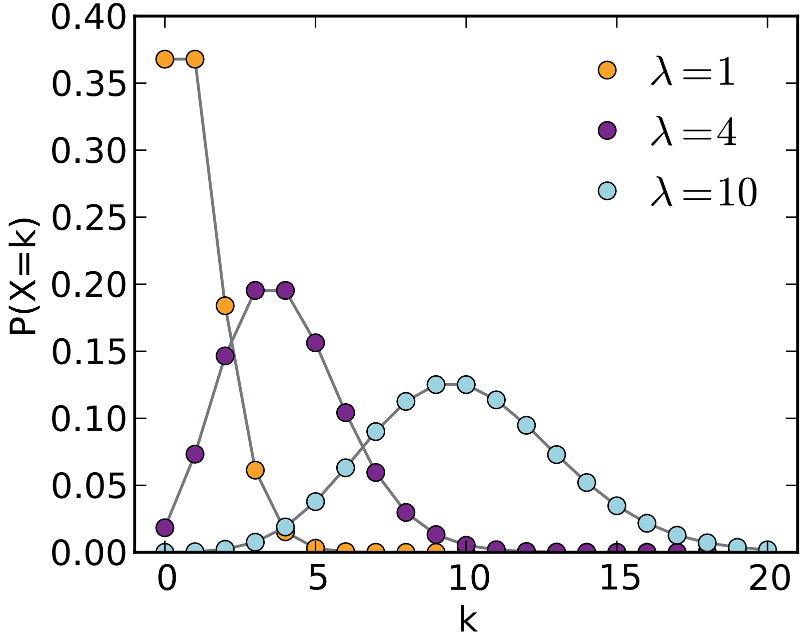
- **n** particles cross the Device Under Test (DUT)
- DUT has an efficiency to detect a particle of **p**
- Average number of hits seen in DUT **x** is **n p**
- Results of many experiments will follow a binomial distribution - use corresponding error bars
- Do NOT use counting errors on **n** and **x** and perform error propagation

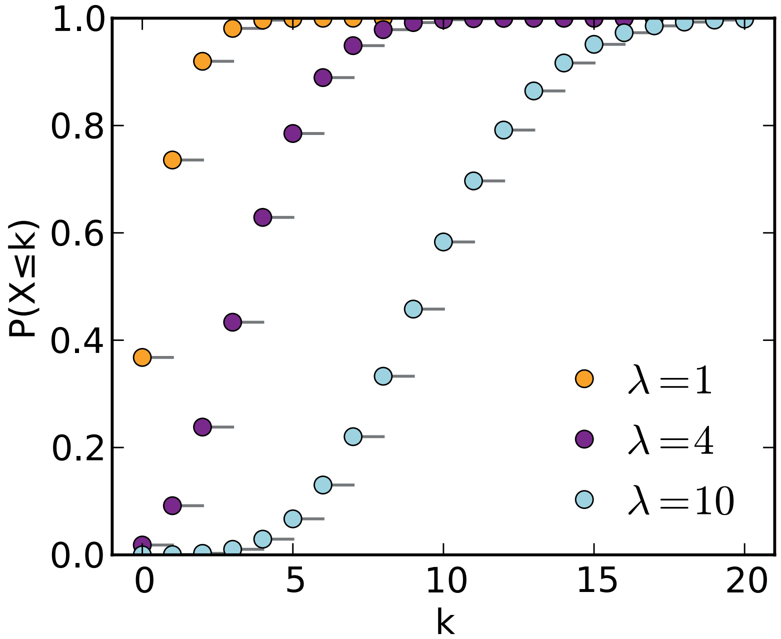
Rare events

- Take limit of binomial distribution for
  $n \to \infty$
  $p \to 0$
  with $n\,p$ fixed to ν

- "Rare events" with constant rate

2.2. Poisson Distribution





$$P_P(x;\nu) = \frac{\nu^x}{x!} e^{-\nu}$$

- Mean: $E[x] = \nu$
- Variance: $V[x] = \nu$
- Standard deviation: $\sigma = \sqrt{\nu}$

Plot source: Wikipedia

Poisson Distribution: Application

Any counting experiment with rare events will have results that are Poisson distributed:

- $\nu = \sigma \int L dt$ expected number of events given luminosity and cross-section
- Number of entries in a histogram bin
- Number of cosmic rays passing through you per second (what is ν?)

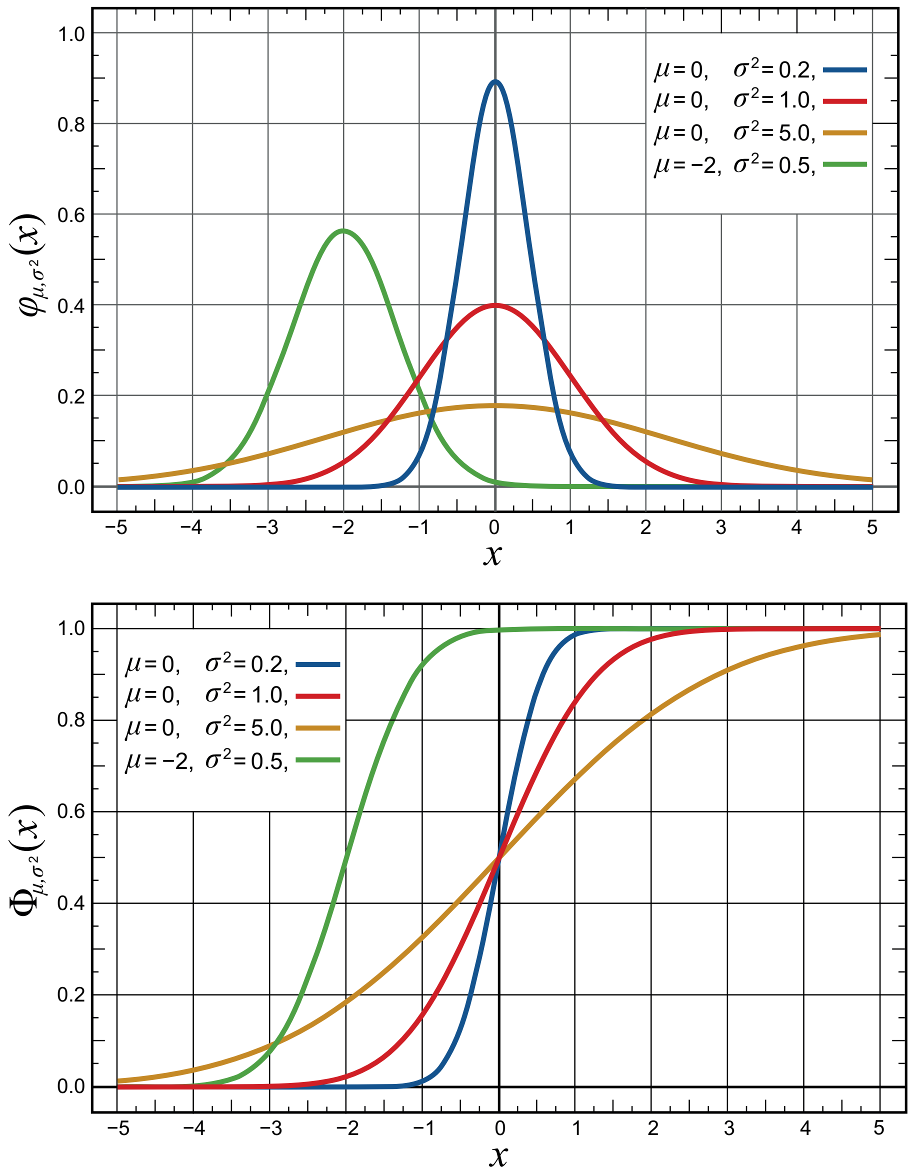
Many rare events

- Look at Poisson distribution for
  $\nu >> 1$

- Many "Rare events" with constant rate

- (and almost every other large N limit...)

2.3. Normal (Gaussian) Distribution



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

- Mean: $E[x] = \mu$
- Variance: $V[x] = \sigma^2$
- CDF is called "error function"

Plot source: Wikipedia

Central limit theorem

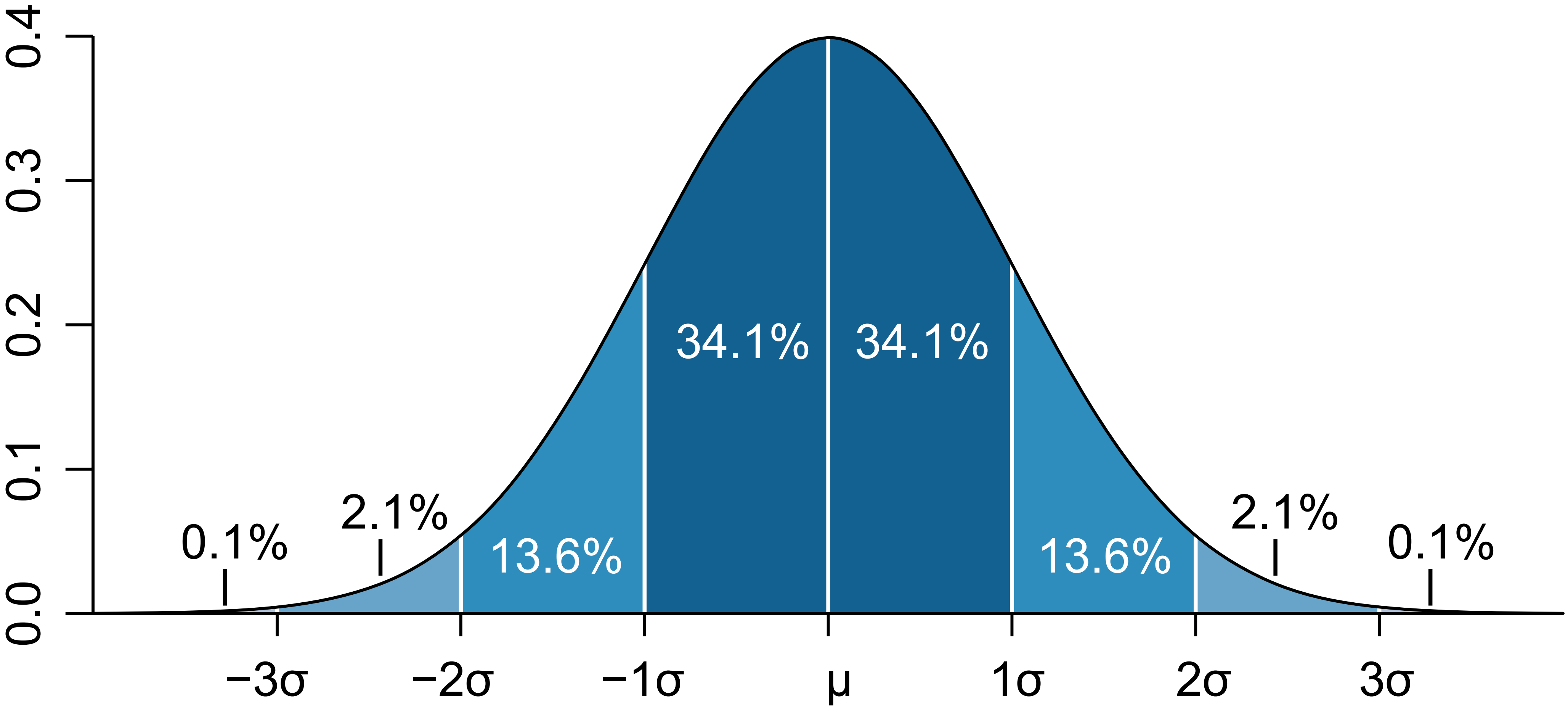
The arithmetic mean of a sufficiently large number of uncorrelated random variables drawn from a distribution with well defined mean and well defined and finite variance will be approximately normal distributed

- This is independent of the underlying distribution
- If your measurement is sufficiently messy (i.e. influenced by "enough" independent effects) it will be normal distributed
- Usually leads to the assumption, that "everything" is normal distributed, which is wrong...

Examples for the normal distribution

- Good example: Size distribution among students of a particular sex in a class
- Not so good example: Distribution of scattering angles of particles passing through a slab of material
  (There are large angle scatters (Rutherford!) that produce non-Gaussian tails)
- Bad example: Energy loss of particles passing through a slab of material
  (Dominated by few large losses - Landau distribution)

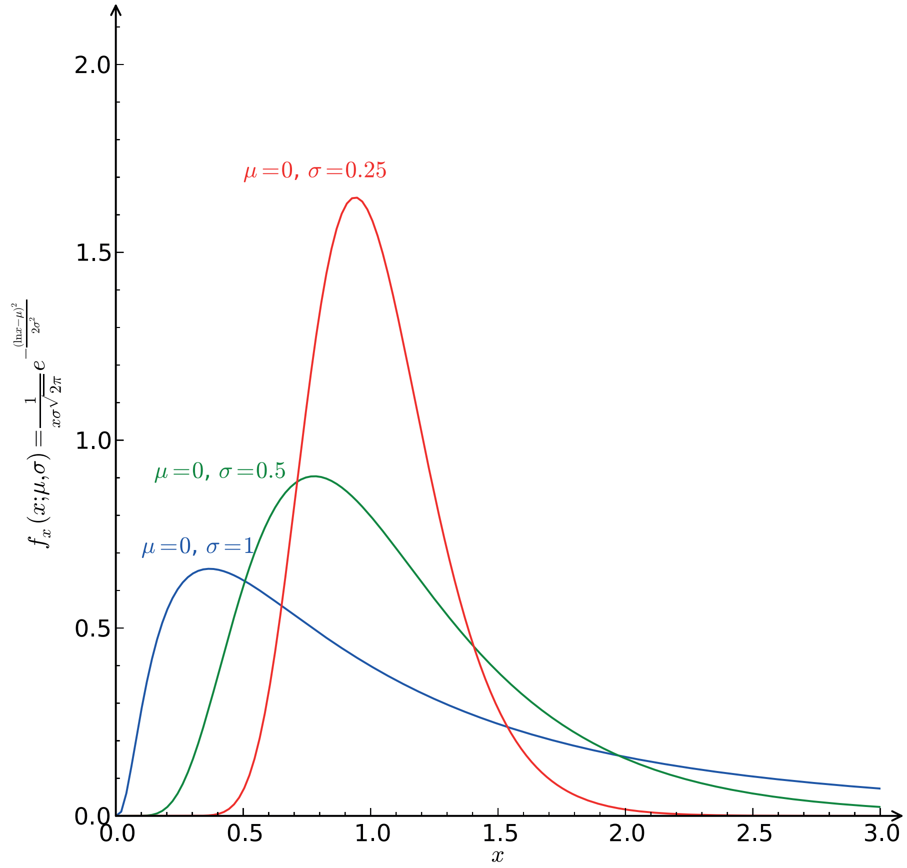
Quantiles of the Normal Distribution: Standard Deviations



Plot source: Wikipedia

2.4. Log Normal Distribution

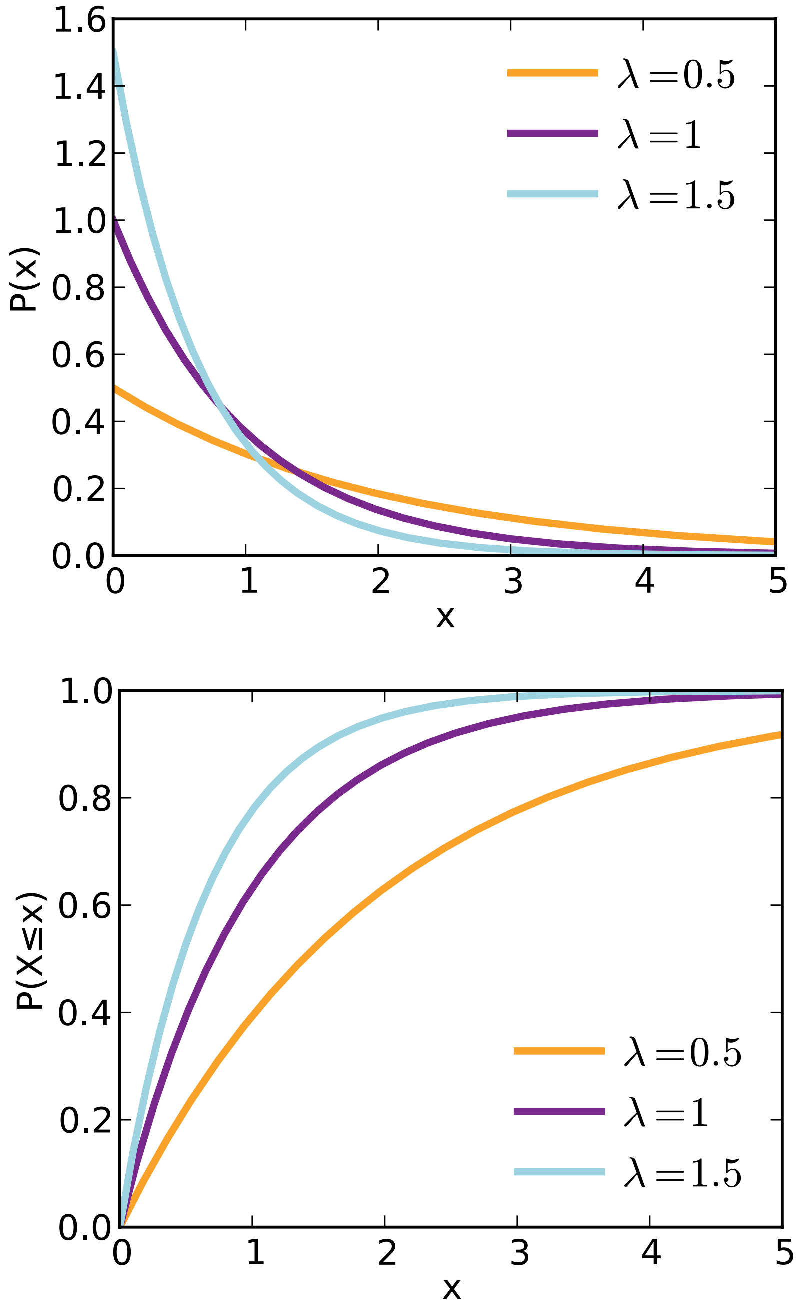
$$P_G(x;\mu,\sigma) = \frac{1}{x\sqrt{2\pi}\sigma} e^{-(\ln x-\mu)^2/2\sigma^2}$$



- Mean: $E[x] = e^{\mu + \sigma^2/2}$
- Variance: $V[x] = (e^{\sigma^2}-1)\, e^{2\mu+\sigma^2}$
- Order of magnitude is normally distributed: e.g. growth of droplets

Plot source: Wikipedia

2.5. Exponential Distribution



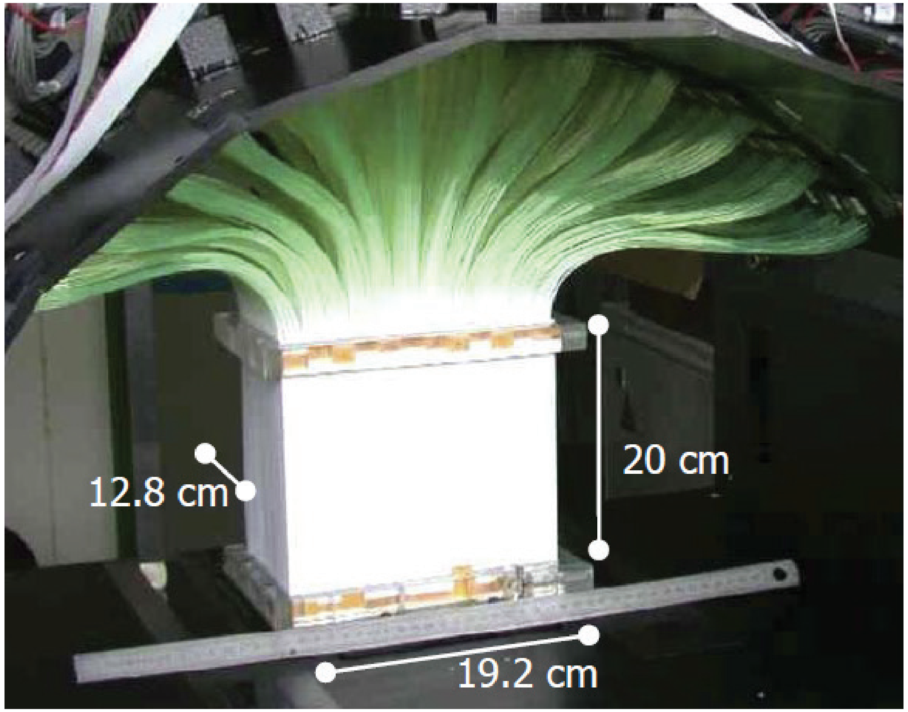
$$P_E(x;\tau) = \frac{1}{\tau}\, e^{-x/\tau} \quad \text{for } x \geq 0$$

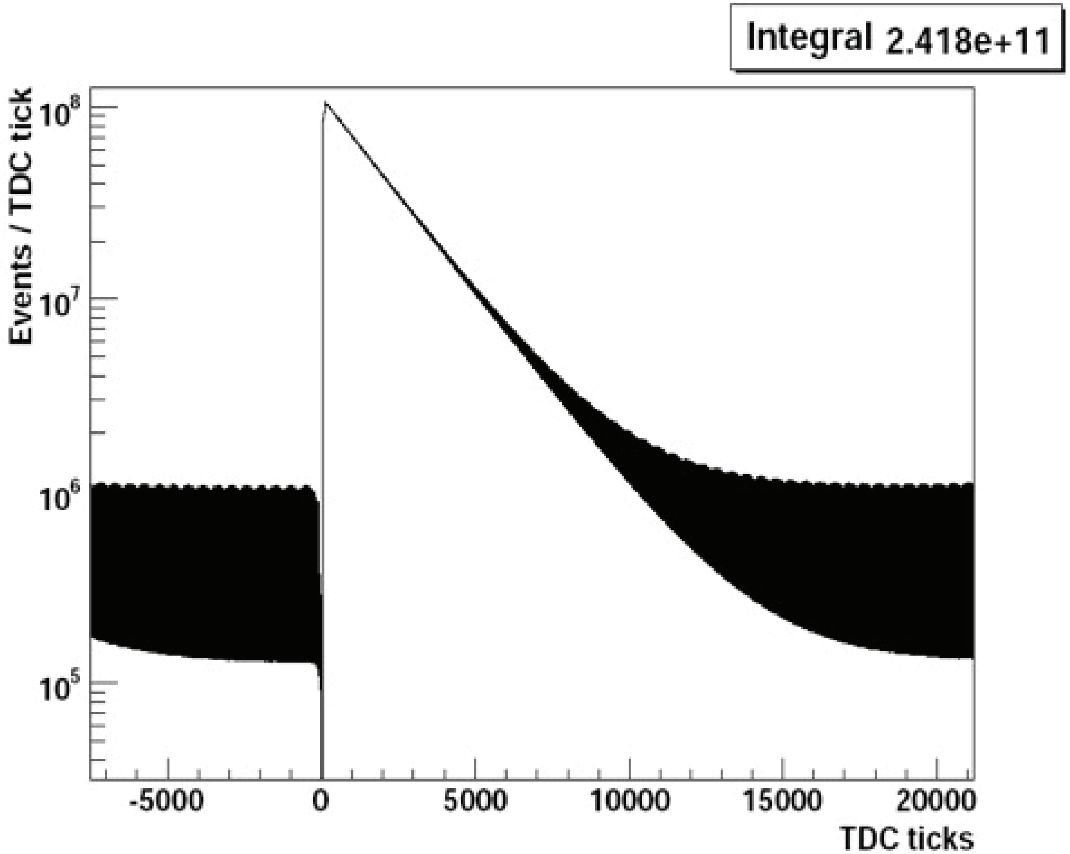
- Mean: $E[x] = \tau$
- Variance: $V[x] = \tau^2$

Plot source: Wikipedia

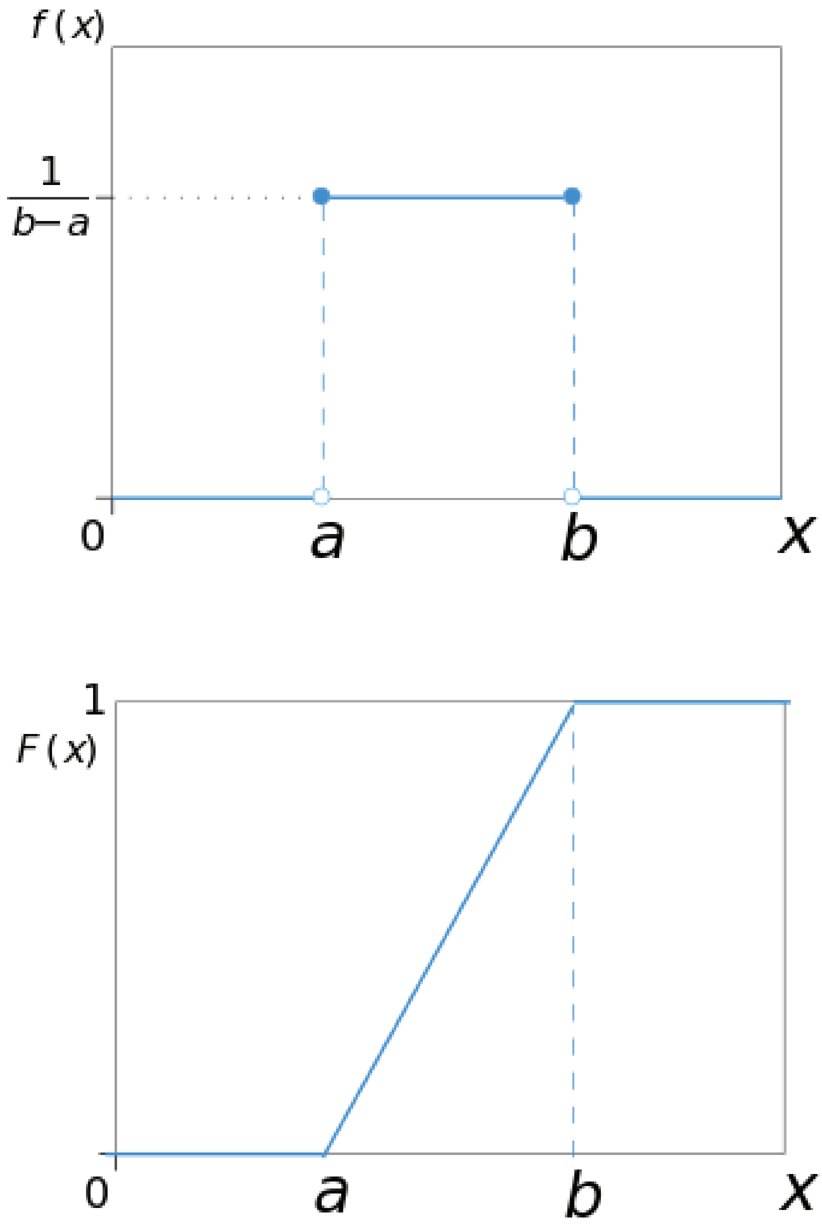
Exponential Distribution: Application

- Lifetime of a particle
- Has no memory: $f(t - t_0 | t > t_0) = f(t)$





2.6. Uniform Distribution



$$P_U(x; a, b) = \frac{1}{a\text{-}b} \text{ for } a \leq x \leq b$$

- Mean: $E[x] = 1/2\,(a+b)$
- Variance: $V[x] = 1/12\,(a\text{-}b)^2$

Plot source: Wikipedia