Random Variable. Moments. Characteristic Function

Vladimir Parfenyev

parfenius@gmail.com

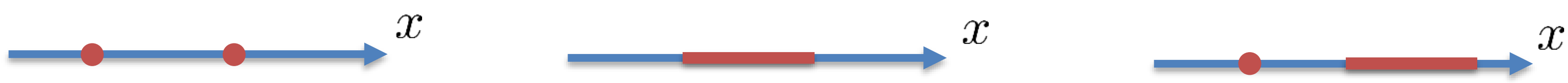
Skolkovo Institute of Science and Technology

03 April 2017

I. Random Variable

How to define Random Variable?

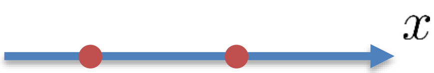
a) Set of Possible Values: $x \in \Omega$



discrete continuous mixed

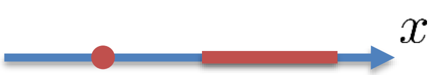
How to define Random Variable?

a) Set of Possible Values: $x \in \Omega$



discrete

continuous



mixed

b) Probability distribution: $p(x) \geq 0, \ \forall x \in \Omega$

$p(x)dx$ -- probability to find X in the interval [x, x+dx]

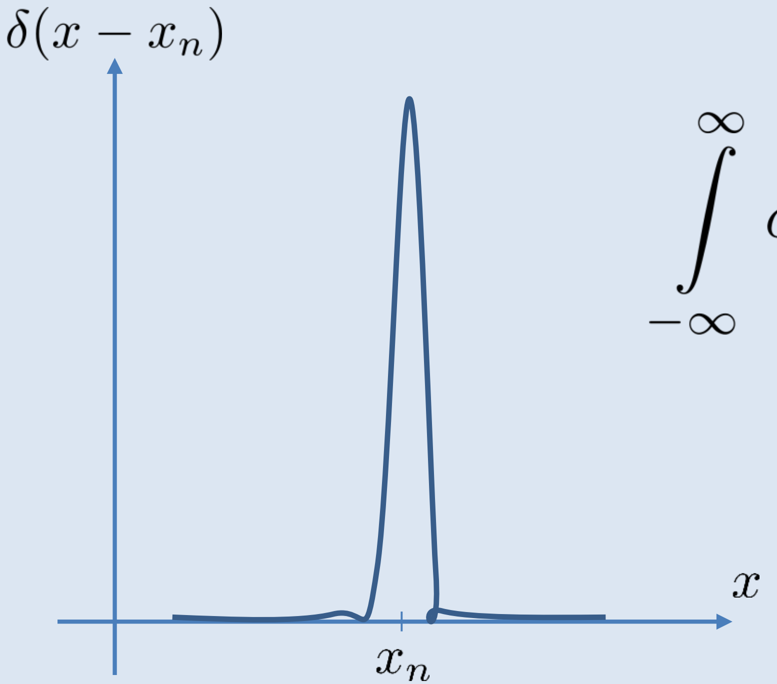
$\int\limits_{\Omega} p(x)dx = 1$ -- normalization

Mixed Set of Possible Values

$$p(x) = \sum_n p_n \delta(x - x_n) + \tilde{p}(x)$$

Mixed Set of Possible Values

$$p(x) = \sum_n p_n \delta(x - x_n) + \tilde{p}(x)$$



$$\int_{-\infty}^{\infty} \delta(x - x_n) f(x) = f(x_n)$$

$$\int_{-\infty}^{\infty} \delta(x - x_n) dx = 1$$

Cumulative Distribution Function

Total probability that the random variable X has a value less than x:

$$\mathcal{P}(X \leq x) = \int_{-\infty}^{x} p(x')dx'.$$

II. Moments

Average/Expectation

Random Variable $X \rightarrow f(X)$:

$$\mathbb{E}[f(x)] \equiv \langle f(X) \rangle = \int\limits_{\Omega} f(x)p(x)dx.$$

Average/Expectation

Random Variable $X \to f(X)$:

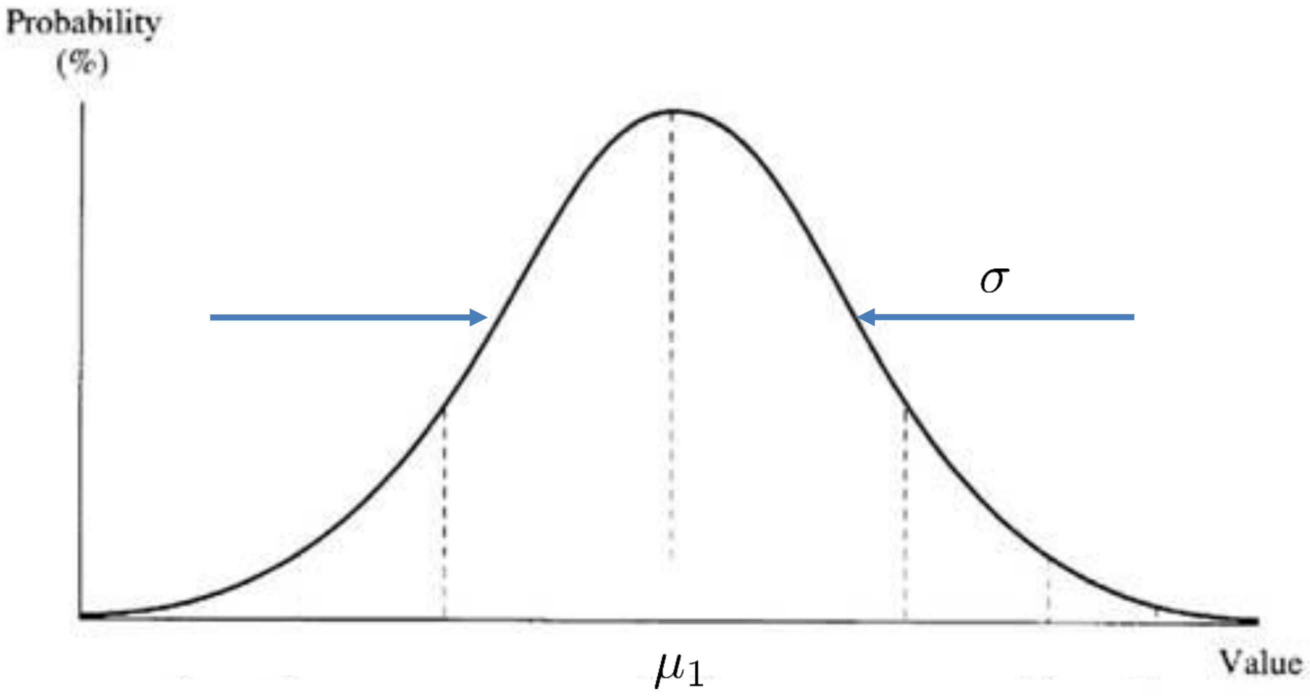
$$\mathbb{E}[f(x)] \equiv \langle f(X) \rangle = \int_{\Omega} f(x)p(x)dx.$$

In particular case, $f(X) = X^m$:

$$\mu_m \equiv \langle X^m \rangle \equiv \int_{\Omega} x^m p(x)dx$$ ← *m*-th moment

Mean and Dispersion

$$\mu_1 \equiv \langle X \rangle = \int\limits_{\Omega} x p(x) dx, \quad \sigma^2 = \langle (X - \langle X \rangle)^2 \rangle = \mu_2 - \mu_1^2$$

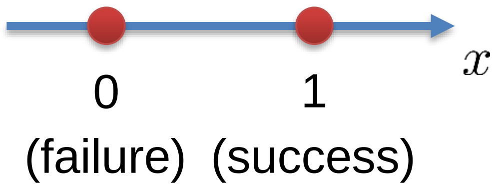


III. Important Distributions

Bernoulli Distribution

Represents (in particular) a coin toss:

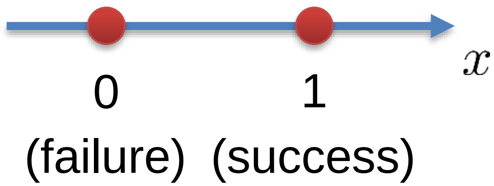
Ω :



Probability function: $p(x) = p\delta(x-1) + (1-p)\delta(x)$

Represents (in particular) a coin toss:

Ω :



Probability function: $p(x) = p\delta(x-1) + (1-p)\delta(x)$

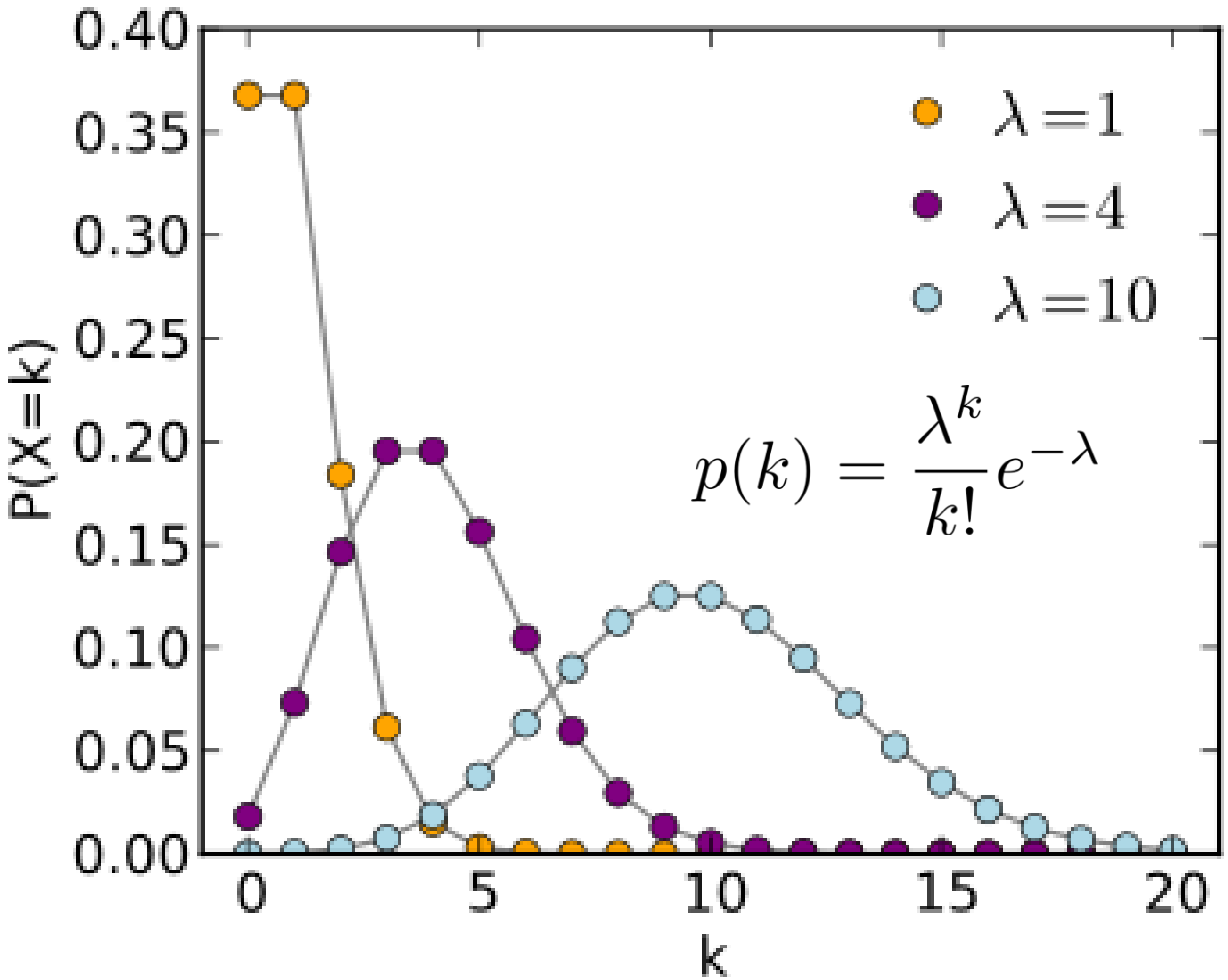
Exercise 1: Calculate *n*-th moment and dispersion.

Poisson Distribution

expresses the probability to observe *k* events within a fixed interval of time, if these events occur independently of time.

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

Poisson Distribution



Poisson Distribution

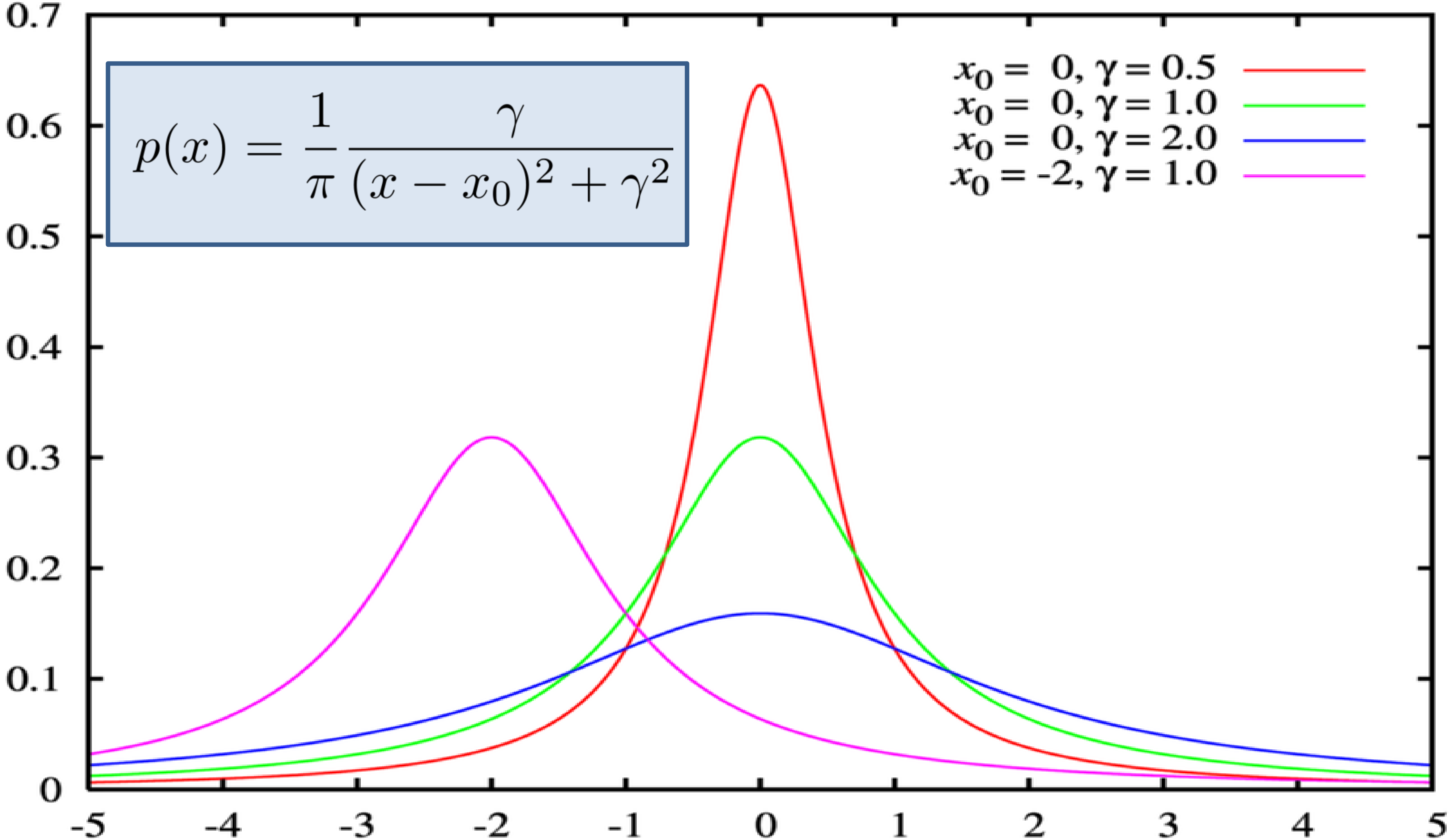
expresses the probability to observe *k* events within a fixed interval of time, if these events occur independently of time.

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

Exercise 2: Check the normalization, calculate 1-st and 2-nd moments and dispersion.

Lorentz or Cauchy distribution

describes resonance behavior (e.g. the form of laser linewidth)



Lorentz or Cauchy distribution

$$p(x) = \frac{1}{\pi} \frac{\gamma}{(x-a)^2 + \gamma^2}, \qquad -\infty < x < +\infty$$

First and second moments:

$$\mu_1 = \frac{\gamma}{\pi} \int\limits_{-\infty}^{+\infty} \frac{x dx}{(x-a)^2 + \gamma^2} = a,$$

$$\mu_2 = \frac{\gamma}{\pi} \int\limits_{-\infty}^{+\infty} \frac{x^2 dx}{(x-a)^2 + \gamma^2} = \infty.$$

IV. Probabilistic Inequalities

Markov Inequality

Intuitively one would say that it is rare for an observation to deviate greatly from the expected value.

Markov's Inequality: For a nonnegative random variable X, and for any positive real number C>0

$$P(X \geq C) \leq \frac{\mathbb{E}[X]}{C}.$$

Chebyshev Inequality

Chebyshev's Inequality: For a random variable X, and for any positive real number C>0

$$P(|X - \mathbb{E}[X]| \geq C) \leq \frac{\sigma^2}{C^2}.$$

Hint: apply Makov's inequality to the random variable $Y = (X - \mathbb{E}[X])^2$.

Coupon Collector's Problem

Exercise 3: There are n different coupons and you want to collect all of them. At every step you can get only one random coupon. What is the probability that you still do not have all coupons after t steps?

Coupon Collector's Problem

Exercise 3: There is *n* different coupons and you want to collect all of them. At every step you can get only one random coupon. What is the probability that you still do not have all coupons after *t* steps?

Hint: X – number of coupons still missing after *t* steps. Estimate $P(X \geq 1)$ using Markov's inequality.

Coupon Collector's Problem

Exercise 3: There is *n* different coupons and you want to collect all of them. At every step you can get only one random coupon. What is the probability that you still do not have all coupons after *t* steps?

Hint: X – number of coupons still missing after *t* steps. Estimate $P(X \geq 1)$ using Markov's inequality.

Hint2: Probability that you haven't particular coupon is $(1 - 1/n)^t$.

Coupon Collector's Problem

Exercise 3: There is *n* different coupons and you want to collect all of them. At every step you can get only one random coupon. What is the probability that you still do not have all coupons after *t* steps?

Hint: X – number of coupons still missing after *t* steps. Estimate $P(X \geq 1)$ using Markov's inequality.

Hint2: Probability that you haven't particular coupon is $(1 - 1/n)^t$.

Solution: $P(X \geq 1) \leq n(1 - 1/n)^t \leq ne^{-t/n}$.

V. Characteristic Function

Characteristic Function

$$G(k) = \langle e^{ikX} \rangle = \int\limits_{-\infty}^{+\infty} e^{ikx} p(x) dx$$ -- Fourier Transform

Properties: $G(0) = 1, \ |G(k)| \leq 1.$

Characteristic Function

$$G(k) = \langle e^{ikX} \rangle = \int\limits_{-\infty}^{+\infty} e^{ikx} p(x) dx$$ -- Fourier Transform

Properties: $G(0) = 1, \ |G(k)| \leq 1.$

Contains information about all moments:

$$G(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \mu_m \quad \Rightarrow \quad \mu_m = \frac{1}{i^m} \frac{\partial^m}{\partial k^m} G(k) \Big|_{k=0}.$$

Characteristic Function

$$G(k) = \langle e^{ikX} \rangle = \int\limits_{-\infty}^{+\infty} e^{ikx} p(x) dx$$ -- Fourier Transform

Properties: $G(0) = 1, \ |G(k)| \leq 1.$

Contains information about all moments:

$$G(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \mu_m \quad \Rightarrow \quad \mu_m = \frac{1}{i^m} \frac{\partial^m}{\partial k^m} G(k) \Big|_{k=0}.$$

Exercise 4: Calculate *G(k)* for Bernoulli distribution and find its *m*-th moment.

Cumulants

$$\ln G(k) = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \kappa_m \quad \Rightarrow \quad \kappa_m = \frac{\partial \ln G(k)}{\partial (ik)^m} \Big|_{k=0}$$

Taylor series starts from *m=1* since *G(0)=1*.

Cumulants

$$\ln G(k) = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \kappa_m \quad \Rightarrow \quad \kappa_m = \frac{\partial \ln G(k)}{\partial (ik)^m}\bigg|_{k=0}$$

Taylor series starts from *m=1* since *G(0)=1*.

Moments determine cumulants and vice versa:

$$\kappa_1 = \mu_1, \; \kappa_2 = \mu_2 - \mu_1^2 = \sigma^2, \ldots$$

Cumulants

$$\ln G(k) = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \kappa_m \quad \Rightarrow \quad \kappa_m = \frac{\partial \ln G(k)}{\partial (ik)^m}\Big|_{k=0}$$

Taylor series starts from *m=1* since *G(0)=1*.

Moments determine cumulants and vice versa:

$$\kappa_1 = \mu_1,\ \kappa_2 = \mu_2 - \mu_1^2 = \sigma^2, \ldots$$

Exercise 5: Calculate *G(k)* for Poisson distribution and find its *m*-th cumulant.