Skoltech
Skolkovo Institute of Science and Technology

Stochastic methods in Mathematical Modelling

Lecture 2. Random variables and distributions

Skoltech
Skolkovo Institute of Science and Technology

Random Variables

Probability space or a probability triple (Σ,F,P)

1. Σ is a *sample space*, i.e. the set of all possible outcomes.
2. F is an *event space*, which is a set of events F, an event being a set of outcomes in the sample space.
3. A *probability function* P, which assigns each event in the event space a probability, which is a number between 0 and 1.

The sample space { 1 , 2 , 3 , 4 , 5 , 6 }

The event space {1}, …, {6}, {1,1}, {1,2} …, {6,6}, ….

Probability maps the events to the number of outcomes:

P({4})=1/6 P({1,3,5})=1/2

Random Variables

Discrete, Continuous, Mixed

Distributions of RVs: How are the probabilities distributed across the random values?

State/phase/sample space (for events), **Σ**.

Example of discrete events: two states, Σ = {0, 1}, also called *Bernoulli random variable*. Probability of a state, *X*,

$$\forall\, X \in \Sigma : Prob(X) = P(X)$$

$$0 \leqslant P(x) \leqslant 1$$

$$\sum_{X \in \Sigma} P(X) = 1$$

For Bernoulli process, P (1) = β, P (0) = 1 − β

Q: Can you give an example of the Bernoulli distribution?

Random Variables

Another important discrete event distribution is the *Poisson distribution*. An event can occur k = 0, 1, 2, ... times in an interval. The average number of events in an interval is λ - called event rate. The probability of observing k events within the interval is

$$\forall\, k \in \mathbb{Z}^{*} = \{0\} \cup \mathbb{Z}: \qquad P(k) = \frac{\lambda^{k} e^{-\lambda}}{k!}$$

Q: Is it normalised?

Q: Are Bernoulli and Poisson distributions related? Can you ”design” Poisson from Bernoulli?

Skoltech
Skolkovo Institute of Science and Technology

Random Variables

Another important discrete event distribution is the *Poisson distribution*. An event can occur k = 0, 1, 2, ... times in an interval. The average number of events in an interval is λ - called event rate. The probability of observing k events within the interval is

$$\forall\, k \in \mathbb{Z}^{*} = \{0\} \cup \mathbb{Z}: \qquad P(k) = \frac{\lambda^{k} e^{-\lambda}}{k\,!}$$

Q: Is it normalised?

Q: Are Bernoulli and Poisson distributions related? Can you ”design” Poisson from Bernoulli?

Yes, you can. Consider repeated drawing of Bernoulli random numbers. You obtain a sequence of zeros and ones. Then you can check for ones only and record intervals with arrival of ones. If you study the probability distribution of *t* arrivals in *n* steps and go to the limit of $n \to \infty$ you get the Poisson distribution.

Distributions

The domain (support) can be continuous, bounded or unbounded.

Uniform distribution

$$\forall x \in [0,1]: \quad p(x) = 1,$$
$$\int_0^1 dx p(x) = 1,$$

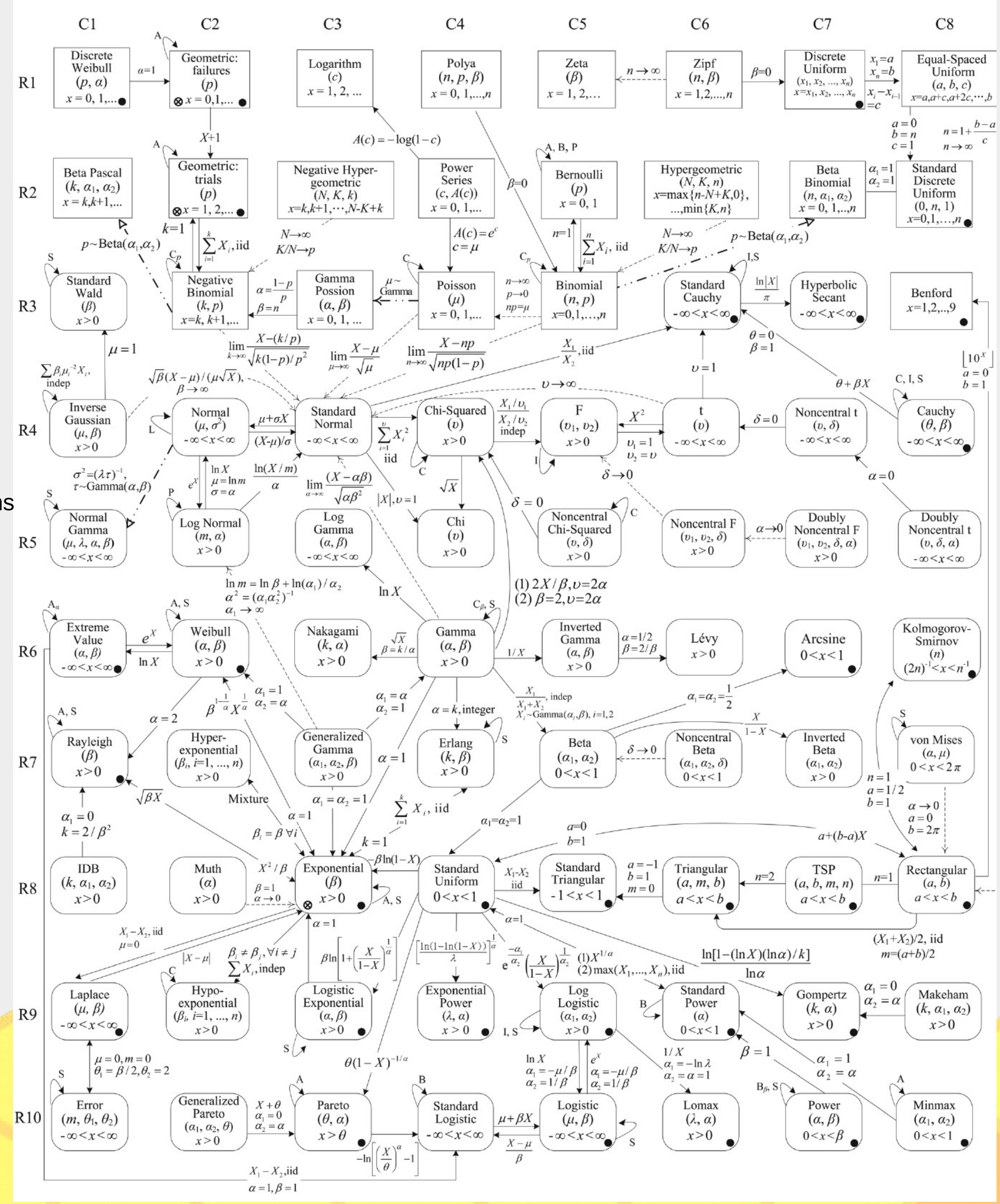
Gaussian distribution

$$\forall x \in \mathbb{Z}: \quad p(x|\sigma,\mu) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$
$$-\infty < x < \infty$$

Skolkovo Institute of Science and Technology

Distributions. The zoo

W. T. Song and Y. Chen, "Eighty Univariate Distributions and Their Relationships Displayed in a Matrix Format," in IEEE Transactions on Automatic Control, vol. 56, no. 8, pp. 1979-1984, Aug. 2011



Skoltech

Skolkovo Institute of Science and Technology

Distributions. Characterisation

Raw moments

$$\langle X^m \rangle = \int_{-\infty}^{\infty} x^m f(x)\,dx$$

Central moments

$$\langle (X-\mu)^m \rangle = \int_{-\infty}^{\infty} (x-\mu)^m f(x)\,dx$$

Example of a *central* moment

The important moments are

1st. *Mean (expectation value)*

$E[X]=\mu$

Example of a raw moment

2nd. *Variance*

$\sigma^2=\mathrm{Var}[X]=E(X-E[X])^2$

σ is *standard deviation*

3rd. *Skewness*

$$\tilde{\mu}_3 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

4th. *Kurtosis*

$$\mathrm{Kurt}[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = \frac{\mu_4}{\sigma^4}$$

Examples of *standardised* moments

Distributions. Characterisation

1st. The *mean* (*expectation value*)

$$E[X] = \mu$$

2nd. The *variance*

$$\sigma^2 = Var[X] = E(X - E[X])^2$$

σ is *standard deviation*

3rd. The *skewness*

$$\tilde{\mu}_3 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

4th. The *kurtosis*

$$Kurt[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$$

Mean

Variance σ^2

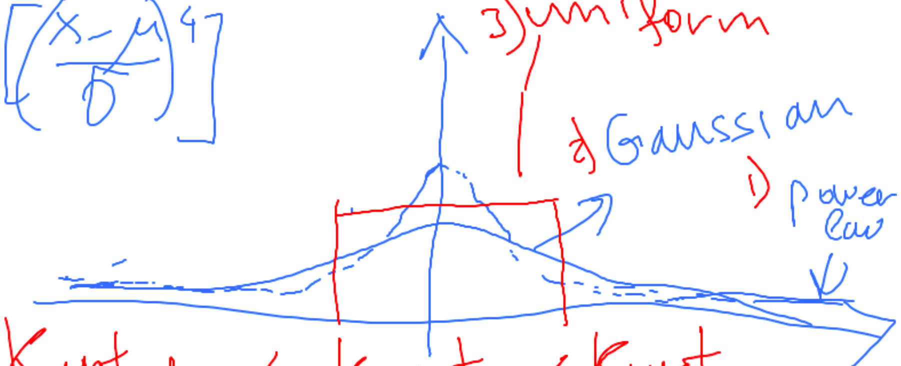
Skewness

positive

negative

Kurtosis

$$E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right]$$



$$\mathrm{Kurt}_1 < \mathrm{Kurt}_2 < \mathrm{Kurt}_3$$

$$\boxed{\mathrm{Kurt}_{Gauss} = 3}$$

Skoltech
Skolkovo Institute of Science and Technology

Distributions. Characterisation

Characteristic function

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

$$G(0) = 1, \quad |G(k)| \le 1$$

$$G(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \langle X^m \rangle$$

hence

$$\langle X^m \rangle = \frac{1}{i^m} \frac{\partial^m}{\partial k^m} G(k) \bigg|_{k=0}$$

Distributions. Characterisation

Cumulants κ_m: An alternative to moments

$$\ln G(k) = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \kappa_m$$

Moments are linked to the cumulants in the sense that any two probability distributions whose moments are identical will have identical cumulants as well, and similarly the cumulants determine the moments

The reason to use cumulants is that for some problems it is more convenient

Distributions. Characterisation

Skolkovo Institute of Science and Technology

Qs

Compute mean, variance, skewness, kurtosis, the first 2 cumulants and the characteristic functions for the following distributions:

$$\tilde{\mu}_3 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

$$Kurt[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$$

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

$$G(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \langle X^m \rangle$$

$$\ln G(k) = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \kappa_m$$

1) Poisson distribution $P(k) = \frac{\lambda^k}{k!} e^{-\lambda}$

2) Bernoulli distribution

3) Gaussian distribution $P(x) \sim e^{\frac{-x^2}{2}}$

4) Exponential distribution $P(x) \sim e^{-x}$

5) Cauchy-Lorentz distribution $P(x) = \frac{1}{\pi(x^2+1)}$

Example: Bernoulli distribution

Prob $p \to 1$ $\qquad p + q = 1$

Prob $q \to 0$ $\qquad p(x) = p\,\delta(x-1) + q\,\delta(x)$

Char. function

$$G(k) = \int_{-\infty}^{+\infty} e^{ikx}\left(p\,\delta(x-1) + q\,\delta(x)\right)dx = p\,e^{ik} + 1 - p$$

$$\mu_m = \frac{\partial^m}{\partial (ik)^m}\left(1 - p + p\,e^{ik}\right)\Big|_{k=0} = p, \ \forall m$$

$$\ln G(k) = \ln\left(1 - p + p\,e^{ik}\right) \simeq 0 + \underbrace{ik\,p}_{\text{cumulants}} + \underbrace{\frac{(ik)^2}{2!}\,p(1-p)}_{\text{cumulants}} + \dots$$

$$\sigma^2 = \int p(x)\,(x-p)^2\,dx = pq$$

Skewness

$$\tilde{\mu}_3 = \int_{-\infty}^{+\infty}\left(p\,\delta(x-1) + q\,\delta(x)\right)\frac{(x-p)^3}{(pq)^{3/2}}\,dx = \frac{(1-p)^3}{(pq)^{3/2}}\,p - \frac{\overset{=1-p}{q}\,p^3}{(pq)^{3/2}} = \frac{\overset{=(q+p)(q-p),\ q+p=1}{q^2 - p^2}}{\sqrt{qp}} = \frac{q-p}{\sqrt{qp}}$$

$$\text{Kurt} = \int_{-\infty}^{+\infty} \frac{(x-p)^4}{(pq)^2} p(x)\,dx = \frac{1-3qp}{pq}$$

Excessive Kurtosis $\quad \text{Kurt} - 3 = \dfrac{1-6qp}{pq}$

Poisson distribution

$$G(p) = \sum_k \frac{(\lambda e^{ip})^k}{k!} e^{-\lambda} = e^{\lambda(e^{ip}-1)}$$

$$\int_{-\infty}^{+\infty} \sum_k e^{ipx} \delta(x-k) \frac{\lambda^k}{k!} e^{-\lambda} dx$$

$$\ln G(p) = \lambda(e^{ip}-1)$$

$$\left.\frac{\partial^n \ln G(p)}{\partial (ip)^n}\right|_{ip=0} = \lambda = \varkappa_n \quad \forall n$$

$$\langle X^m \rangle = \left.\frac{\partial^m G(p)}{\partial (ip)^m}\right|_{p=0}$$

$\mu_1 = \lambda$

$\mu_2 = \lambda^2 + \lambda$

$$\underbrace{\text{Var}\,P = \lambda}$$

Cauchy distribution

$$P(x) = \frac{1}{\pi(x^2+1)}.$$

Mean

$$\mu_1 = \int_{-\infty}^{+\infty} \frac{x}{\pi(1+x^2)} dx = \infty$$

$$\mu_2 = \int_{-\infty}^{+\infty} \frac{x^2}{\pi(1+x^2)} dx = \infty$$

Characteristic function

$$\varphi(k) = \int_{-\infty}^{+\infty} \frac{e^{ikx}}{\pi(1+x^2)} dx = e^{-|k|}$$

Skoltech
Skolkovo Institute of Science and Technology

Distributions. Characterisation

Probabilistic inequalities

Markov's inequality

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

X≥0,C>0

Chebyshev's inequality

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

Chernoff bound

$$P(x \geq a) = P(e^{tx} \geq e^{ta}) \leq \frac{\mathbb{E}[e^{tx}]}{e^{ta}}$$

Markov inequality

$$P(X \geq c) \leq \frac{E[X]}{c}, \quad x \geq 0, c > 0$$

$$E[X] = \sum_{x, p(x) > 0} x\, p(x) \geq \sum_{x \geq c} x\, p(x) \geq \sum_{x \geq c} c\, p(x) = c\, P(X \geq c)$$

$$P(X \geq c) \leq \frac{E[X]}{c}$$

$$Y = |X - \mu|$$

$$P(Y^2 \geq c^2) \leq \frac{\boxed{E[Y^2]}}{c^2}, \quad E[Y^2] = \sigma^2$$

$$P(Y \geq c) \leq \frac{\sigma^2}{c^2} \rightarrow \text{Chebyshev inequality}$$

Distributions. Characterisation

Inequalities

Qs

Prove Chernoff bound

Coupon collector’s problem

Assume that 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?

Coupon collector's problem

Prob that we do not get a particular coupon after 1 step $1-\frac{1}{n}$

$\left(1-\frac{1}{n}\right)^t$ after t steps

$n\left(1-\frac{1}{n}\right)^t$ coupons missing after t steps

$P($ ~~some are~~ number of missing $\geq 1) \leq \frac{n\left(1-\frac{1}{n}\right)^t}{1} = n\left(1-\frac{1}{n}\right)^t \leq n e^{-t/n}$