Probability Refresher École Polytechnique-HEC 2017-18

Lectures : Lucas Gerin `lucas.gerin@polytechnique.edu`
Tutorials : Raphaël Forien `raphael.forien@polytechnique.edu` / Thibaut Mastrolia `thibaut.mastrolia@polytechnique.edu`

1. Probability spaces

Sample spaces and probability measures definitions, first examples
Independence and conditioning

2. Random variables and expectation

RV and their laws examples of *usual* laws, cdf, densities of r.v.
Abstract expectation $\mathbb{E}[X]$, $\mathbb{E}[\phi(X)]$
Inequalities Jensen, Markov and Bienaymé-Tchebychev
L^2 **spaces**, definitions, orthogonal projection.

3. Random vectors

Random vectors joint laws/densities/cdf. Fubini's theorems
Transformations applications of Jacobian formula
(In)dependence covariance/correlation
Sums of independent random variables

4. Gaussian vectors

Definitions. Covariance, Linear transformations of gaussian vectors.

5. Conditioning

conditional expectation discrete/continuous case.
Conditioning = orthogonal projections
conditional distributions

6. More on random variables

Concentration inequalities
Order statistics
Mixtures

7. Convergences of random variables

Convergences of r.v. $\neq$ types of convergences.
Law(s) of Large Numbers.

8. Convergences of distributions

Definitions and criteria Characteristic functions, cdf. The discrete case (Binomial $\rightarrow$ Poisson).
The Central Limit Theorem Confidence intervals : asymptotic vs non-asymptotic

---

References :

- G.Grimmett, D.Stirzaker. *Probability and Random Processes*, Oxford Univ.Press.
  A very nice and easy-to-read introduction to Theoretical Probability, with many examples.
- A.DasGupta. *Probability for Statistics and Machine Learning*, Springer.
  A modern and very comprehensive book, with a special emphasis on what is needed for Machine learning.

Usual distributions

1 Discrete distributions

- **Uniform** distribution in $\{1,2,\dots,n\}$, $\mathbb{P}(X=i)=1/n$ for each n.

$$\mathbb{E}[X]=\frac{n+1}{2},\quad \mathrm{Var}(X)=\frac{n^2-1}{12}.$$

- **Bernoulli** law X with parameter $p\in[0,1]$,
$\mathbb{P}(X=1)=p$, $\mathbb{P}(X=0)=1-p$.

$$\mathbb{E}[X]=p,\quad \mathrm{Var}(X)=p(1-p).$$

- **Binomial** law $\mathrm{Bin}(n,p)$ with parameters $n\geq 1, p\in[0,1]$
(= number of successes in n Bernoulli trials)
$\mathrm{Bin}(n,p)\in\{0,1,\dots,n\}$ and

$$\mathbb{P}(\mathrm{Bin}(n,p)=k)=\binom{n}{k}p^k(1-p)^{n-k}.$$

$$\mathbb{E}[X]=np,\quad \mathrm{Var}(X)=np(1-p).$$

- **Geometric** law $\mathcal{G}$ with parameter $p\in[0,1]$
(= first success in Bernoulli trials)
$\mathcal{G}\in\{1,2,\dots\}$ and

$$\mathbb{P}(\mathcal{G}=k)=p(1-p)^{k-1}.$$

$$\mathbb{E}[X]=1/p,\quad \mathrm{Var}(X)=(1-p)/p^2.$$

- **Poisson** law $\mathcal{P}$ with parameter $\lambda>0$.
$\mathcal{P}\in\{0,1,2,\dots\}$ and

$$\mathbb{P}(\mathcal{P}=k)=e^{-\lambda}\frac{\lambda^k}{k!}.$$

$$\mathbb{E}[X]=\lambda,\quad \mathrm{Var}(X)=\lambda.$$

2 Continuous distributions

- **Uniform** distribution in $[a,b]$:

$$f(x)=\frac{1}{b-a}\text{ for }a\leq x\leq b.$$

$$\mathbb{E}[X]=(a+b)/2,\quad \mathrm{Var}(X)=(b-a)^2/12.$$

- **Exponential** distribution with parameter $\lambda>0$:

$$f(x)=\lambda\exp(-\lambda x)\text{ for }x\geq 0.$$

$$\mathbb{E}[X]=1/\lambda,\quad \mathrm{Var}(X)=1/\lambda^2.$$

- **Normal** (or **gaussian**) distribution $\mathcal{N}(\mu,\sigma^2)$ with parameters μ,σ^2 :

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\text{ for all }x.$$

$$\mathbb{E}[X]=\mu,\quad \mathrm{Var}(X)=\sigma^2.$$

- $\boldsymbol{\chi^2}$ (**Khi-square**) distribution with d degrees of freedom :

$$X=X_1^2+\dots+X_d^2,$$

where X_i's are i.i.d. $\mathcal{N}(0,1)$.

$$\mathbb{E}[X]=d,\quad \mathrm{Var}(X)=2d.$$