Introduction to probability theory and stochastic processes

Cameron Ray Smith[1], Aviv Bergman[1,2,3,4]

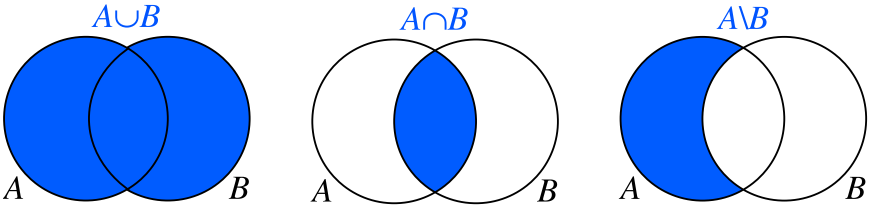
[1]Department of Systems and Computational Biology
[2]Dominick P. Purpura Department of Neuroscience
[3]Department of Pathology
Albert Einstein College of Medicine
[4]Santa Fe Institute

EINSTEIN
Albert Einstein College of Medicine

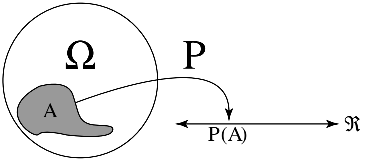
August 17, 2012

Outline

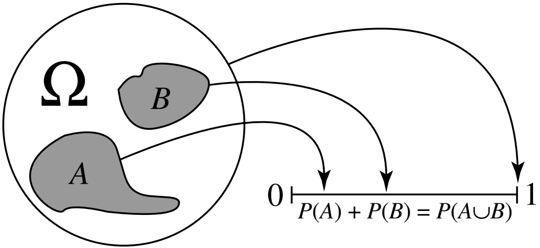
sets



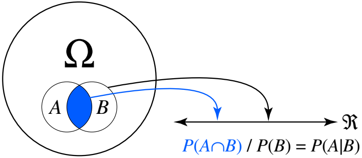
probability measure



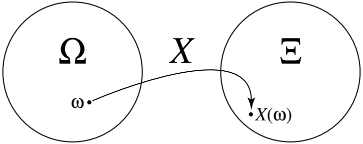
probability axioms



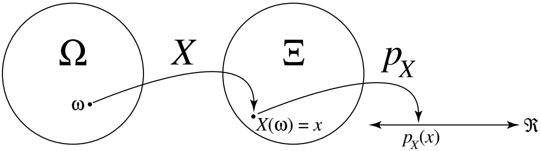
conditional probability



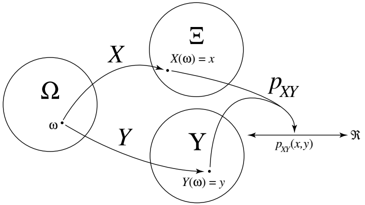
random variables



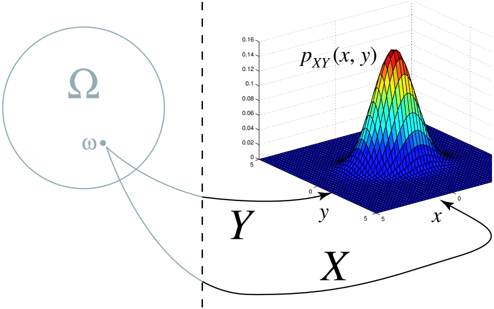
probability densities



joint probability densities



the reality



- For the general theory of measure spaces, we first need a *measurable space* (Ω, Σ), that is a set equipped with a collection Σ of **measurable sets** complete under certain operations. Then this becomes a measure space (Ω, Σ, μ) by throwing in a function μ from Σ to a space of values (such as the real line) that gets along with the set-theoretic operations that Σ has. If E is a measurable set, then $\mu(E)$ is called the **measure** of E with respect to μ. [1, 2]

1. Given a set Ω,
2. a σ-**algebra** is a collection of subsets of Ω that is closed under complementation, countable unions, and countable intersections.
3. A **measurable space**, by the usual modern definition, is a set Ω equipped with a σ-algebra Σ.
4. The elements of Σ are called the **measurable sets** of Ω (or more properly, the measurable subsets of (Ω, Σ)).

A **measure space** is a **measurable space** equipped with a **measure**. There are many different types of measures parametrized by the type of their codomains. Let (Ω, Σ) be a measurable space. A **probability measure** on Ω (due to Kolmogorov) is a function μ from the collection Σ of measurable sets to the unit interval $[0, 1]$ such that:

1. The measure of the empty set is zero: $\mu(\emptyset) = 0$;
2. The measure of the entire space is one: $\mu(\Omega) = 1$;
3. Countable additivity: $\mu(\bigcup_{i=1}^{\infty} S_i) = \sum_{i=1}^{\infty} \mu(S_i)$ whenever the S_i are mutually disjoint sets—disjoint. (Part of the latter condition is the requirement that the sum on the right-hand side must converge.)

It is sometimes stated (but in fact follows from the previous) that:

- Finitary additivity: $\mu(S \cup T) = \mu(S) + \mu(T)$ whenever S and T are disjoint.
- μ is increasing: $\mu(A) \leq \mu(B)$ if $A \subseteq B$.

Measures can be thought of in terms of integrals and densities are defined in terms of measures. Let A be, for example, one of the measurable sets from the collection of measurable sets, Σ, of our sample space Ω.

- $\mu(A) = \int_A dx$ or $\mu(A) = \int_A p(x)dx$
- $\mu(A)$ represents the mass of A which can be interpreted geometrically as an *abstract volume* or probabilistically as *the probability mass of the event "random variable X takes a value within A"*
- A **density** can then be defined intuitively as a function that transforms some measure μ_1 into a measure μ_2 by pointwise reweighting on the sample space Ω. Thus, densities are always relative measures.
- $d\mu_2(x) = f(x)d\mu_1(x)$ or $\frac{d\mu_2}{d\mu_2}(x) = f(x)$

Does a density always exist?

- A density function f is thus a function that is integrated to obtain information in terms of measure μ_2 from information in terms of measure μ_1.
- $\mu_2(A) = \int_A d\mu_2(x) = \int_A f(x) d\mu_1(x)$ is not defined if $\mu_1(A) = 0$ and $\mu_2(A) \neq 0$.
- If this is never the case for all $A \in \Sigma$, then μ_2 is referred to as *absolutely continuous* with respect to μ_1 and this relationship is written $\mu_2 \ll \mu_1$.
- This conclusion is formalized in the **Radon-Nikodym theorem** which states that μ_2 has a density with respect to μ_1 if and only if $\mu_2 \ll \mu_1$.

... so the answer is ... no, which is the reason for going through all this abstract stuff

Outline

Intuitively, **stochastic processes** are ∞-dimensional probability distributions.

- In most applications, stochastic processes model systems that evolve randomly in time, which is likely the origin of the word *process* in the name.
- The order of this evolution can be described through the use of an index and an index set respectively $t \in T$.

Definition

Consider a random experiment with sample space X, a σ-algebra Σ, a base probability measure $\mu : \Sigma \to [0, 1]$, and a collection of random variables S_t indexed by a set T. A **stochastic process** *is then defined by the set $\{S_t, t \in T\}$.*

- This definition can be specialized to the case of discrete or continuous stochastic processes by taking the index set to be $T \in \mathbb{N}$ or $T \in \mathbb{R}_+$ respectively.

Outline

Theoretical background

Stochastic processes

References

[1] David Insua, Fabrizio Ruggeri, and Mike Wiper.
Bayesian Analysis of Stochastic Process Models.
Wiley, 2012.

[2] NLab.
Measure space.
http://ncatlab.org/nlab/show/measure+space.