Basic Methods of Probability

<u>Def:</u> A <u>probability space</u> (Ω, P) consists of a set of sample points or states Ω and a probability measure P which assigns probabilities in $[0,1]$ to events A, or subsets of Ω:

$$P : A \mapsto P(A) \in [0,1], \qquad A \subseteq \Omega.$$

<u>Example:</u> $\Omega = \{1,2,3,4,5,6\}$ = six outcomes of throwing dice

$P(\{i\}) = 1/6$ for $i = 1, \ldots, 6$

$A = \{2,4,6\}$ = event that even number is rolled

$P(A) = 1/2$

A <u>(vector-valued) random variable</u> $\tilde{X}$ is a map

$$\tilde{X} : \Omega \to \mathbb{R}^d \qquad (d = \text{dimension of } \tilde{X})$$

<u>Examples:</u> (i) Suppose that you bet a friend that you will get an even number in one roll of a dice, wagering \$10. Then, your earnings are a (scalar) random variable $\tilde{X}$ given by

$$\tilde{X}(2) = \tilde{X}(4) = \tilde{X}(6) = +10$$

$$\tilde{X}(1) = \tilde{X}(3) = \tilde{X}(5) = -10$$

Note that not every random variable has an ordinary function as its PDF! For example, in our example of the dice wager,

$$P_x(x) = \frac{1}{2}\delta(x+10) + \frac{1}{2}\delta(x-10).$$

The PDF is a generalized function (or distribution).
A more typical example of a PDF is the <u>Gaussian</u>

$$p(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$$

which is the PDF of a <u>$N(\mu,\sigma^2)$ - random variable</u> (or, simply, normal random variable).

Note that the PDF and CDF are related by

$$P_x(x) = \frac{\partial^d}{\partial x_1 \cdots \partial x_d} F_x(x)$$

and

$$F_x(x) = \int_{-\infty}^{x_1} du_1 \int_{-\infty}^{x_2} du_2 \cdots \int_{-\infty}^{x_d} du_d \, P_x(u).$$

The <u>expectation</u> $E(\tilde{x})$ of a random variable is defined by

$$E(\tilde{x}) = \int x \, P_x(x)\, d^d x \in \mathbb{R}^d$$

<u>Example:</u> In the dice wager,

$$E(X) = (-10)\frac{1}{2} + 10\frac{1}{2} = 0.$$

The distribution of $\tilde{X}$, P_X, is a probability measure on $\mathbb{R}^d$, defined by

$$P_X(A) = P(\{\tilde{X} \in A\}) = P(\tilde{X}^{-1}(A)), \quad A \subseteq \mathbb{R}^d.$$

Example: In previous example, $P_X(\{10\}) = P_X(\{-10\}) = \frac{1}{2}$ and $P_X(A) = 0$ if $10, -10 \notin A$.

Another way to specify the distribution of a random variable is by the cumulative distribution function (CDF) $F_{\tilde{X}} : \mathbb{R}^d \to [0,1]$ defined by

$$F_X(x) = P_X\left(\bigotimes_{i=1}^{d} (-\infty, x_i) \right).$$

It is an increasing function, seperately in each component of the vector variable $x = (x_1, \dots x_d)$.

Example: In previous example,

$$F_X(x) = \begin{cases} 0 & x < -10 \\ 1/2 & -10 \le x < 10 \\ 1 & 10 \le x \end{cases}$$

Still another way to specify the distribution of a random variable is by the probability density func. (PDF), $p_X : \mathbb{R}^d \to \mathbb{R}^+$ defined by

$$P_X(A) = \int_A p_X(x)\, d^d x.$$

Note that not every random variable has an expectation. For example, the Cauchy random variable with density

$$p(x;b) = \frac{b}{\pi} \frac{1}{x^2+b^2}$$

has $\int_{-\infty}^{\infty} x\, p(x;b)\, dx = -\infty + \infty =$ undefined.

The expectation $E(\tilde{x})$ may also be written as

$$E(\tilde{x}) = \int_{\Omega} \tilde{x}(\omega)\, dP(\omega)$$

as an integral w.r.t. P over Ω. It follows that

$$\mathbb{E}(\tilde{\mathbb{1}}_A) = \int_{\Omega} \mathbb{1}_A(\omega)\, dP(\omega) = \int_A dP(\omega) = \mathbb{P}(A)$$

where $\tilde{\mathbb{1}}_A$ is the indicator function

$$\tilde{\mathbb{1}}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}.$$

Generalizing the expectation, we may define the pth-order moments of $\tilde{x}$ by

$$\begin{aligned} \mu_{i_1 \dots i_p} &= \int x_{i_1} \cdots x_{i_p}\, P_x(x)\, d^d x \\ &= E(\tilde{x}_{i_1} \cdots \tilde{x}_{i_p}) \end{aligned}$$

for $(i_1, \dots i_p) \in \{1, 2, \dots d\}^p$. Of course, moments of all orders need not exist, if the PDF $P_x(x)$ has only power-law decay as $|x| \to \infty$.

<u>Remark:</u> In physics, expectation $E(\tilde{X})$ is often denoted as $\langle \tilde{X} \rangle$ and, thus, moments $\mu_{i_1 \cdots i_p}$ as $\langle \tilde{X}_{i_1} \cdots \tilde{X}_{i_p} \rangle$!

Random vectors may be infinite-dimensional, taking values in function spaces, e.g. An interesting example of the latter is the <u>fine-grained PDF</u> of a random vector $\tilde{X}$, defined by

$$\tilde{P}_X(x) = \delta^d(x - \tilde{X})$$

whose value is a generalized function! It has the property

$$P_X(x) = \mathbb{E}\left(\tilde{P}_X(x)\right)$$

as may be seen from

$$\begin{aligned} \mathbb{E}\left(\tilde{P}_X(x)\right) &= \int \delta^d(x - \tilde{X}(\omega))\, dP(\omega) \\ &= \int \delta^d(x - x')\, P_X(x')\, d^d x' \\ &= P_X(x). \end{aligned}$$

In the same way, we define a <u>fine-grained CDF</u>

$$\tilde{F}_X(x) = \prod_{i=1}^{d} \theta(x_i - \tilde{X}_i(\omega))$$

where $\theta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x = 0 \end{cases}$ is the <u>Heaviside step-function.</u>

It has the property that

$$\mathbb{E}\left(\tilde{F}_X(x)\right) = F_X(x).$$

Stochastic dynamics

We now consider the more particular case of a random process $\tilde{X}(t,\omega)$, $t \in \mathbb{R}$

which may be regarded as a continuous sequence of random vectors indexed by time $t \in \mathbb{R}$.

Alternatively, this may be regarded as an infinite-dimensional random vector, taking a value in a function space such as

$$\mathcal{X} = C(\mathbb{R}, \mathbb{R}^d) \quad \text{or} \quad C([t_0, t_f], \mathbb{R}^d)$$

which is called path space.

E.g. the process may be the solution of a stochastic differential equation (SDE) or Langevin equation, such as

$$d\tilde{X}(t) = f(\tilde{X},t)\,dt + g(\tilde{X},t)\,dW_t$$

which is a special case called a (nonlinear) diffusion process.

All of the concepts introduced for a general random vector can be applied to random processes. For example, the distribution $P_{\tilde{X}}$ of the random process $\tilde{X}(t)$ is the measure on path-space defined by

$$P_{\tilde{X}}(A) = P(\tilde{X} \in A), \quad A \subseteq \mathcal{X}$$

One may introduce other statistical measures such as ...

the <u>single-time PDF</u>

$$P_{X}(x,t) = \mathbb{E}\left(\delta^d(x-\tilde{X}(t))\right)$$

or the <u>multi-time PDF's</u>

$$P_X(x_1,t_1,\dots x_p,t_p) = \mathbb{E}\left(\prod_{i=1}^{p} \delta^d(x_i - \tilde{X}(t_i))\right)$$

The times t may be regarded just as additional "components" of the infinite-dimensional random vector.

One can also introduce <u>expectations</u>

$$\mathbb{E}(\tilde{X}(t)) = \int_\Omega \tilde{X}(t,\omega)\, d\mathbb{P}(\omega) = \int_{\mathbb{R}^d} x\, P_X(x,t)\, d^d x$$

and, similarly, multi-time moments

$$\mathbb{E}\left(\tilde{X}_{i_1}(t_1)\cdots\tilde{X}_{i_p}(t_p)\right) = \int_\Omega \tilde{X}_{i_1}(t_1,\omega)\cdots\tilde{X}_{i_p}(t_p,\omega)\, d\mathbb{P}(\omega)$$

$$= \int_{\mathbb{R}^d} d^d x_1 \cdots \int_{\mathbb{R}^d} d^d x_p \; x_{1,i_1}\cdots x_{p,i_p}\, P_X(x_1,t_1,\dots x_p,t_p)$$

Just as with any random vector, these quantities are not guaranteed to exist, if the PDF's decay too slowly at large values of $|x_i|$, $i=1,\dots,p$.