Continuous random variables

Gabriel T. Landi

Additional reading:

- Ross, chapter 5
- Reif, chapter 1
- Tomé, de Oliveira, chapter 1
- Salinas, chapter 1.

PDF and CDF

When X is a random variable that varies continually, instead of probabilities we work with the probability density function (PDF)

$$p(x) \quad \text{or} \quad p_X(x) = \text{PDF} \tag{1}$$

The probability that X is in some interval $[a, b]$ is then

$$P(a \leqslant X \leqslant b) = \int_a^b p(x)\,dx \tag{2}$$

The probability that you find X anywhere is 1 so

$$\int_{-\infty}^{\infty} p(x)\,dx = 1 \tag{3}$$

which is the normalization condition for PDFs.

If you look for the prob. of finding X in a small interval $[x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2}]$ we get

$$P\left(x - \frac{\Delta x}{2} \leqslant X \leqslant x + \frac{\Delta x}{2}\right) = \int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} p(x)\,dx \approx p(x)\,\Delta x \tag{4}$$

Note that $p(x)$ is a density, so if x has units of length, then $p(x)$ will have units of 1/length. This is most evident from (2) or (3) since probabilities have no units. Note also that $p(x) \geqslant 0$, but it does not have to be smaller than 1.

The computation of moments is analogous to the discrete case

$$\langle X \rangle = \int_{-\infty}^{\infty} x\, p(x)\, dx \qquad (5)$$

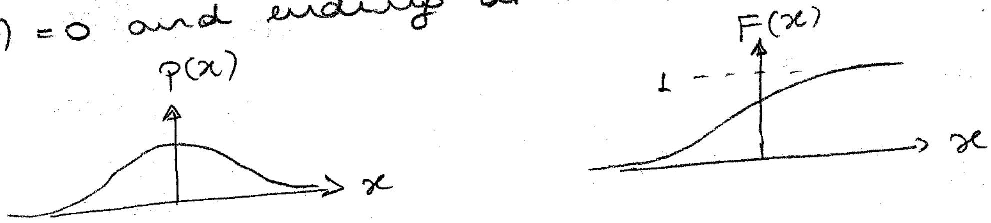
Or, if $f(x)$ is an arbitrary function of x, then

$$\langle f(X) \rangle = \int_{-\infty}^{\infty} f(x) p(x)\, dx \qquad (6)$$

Another important quantity, is the cumulative distribution function, CDF. It is defined as

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} p(y)\, dy \qquad (7)$$

The CDF is an actual prob.; ie, it is restricted to $[0, 1]$. It is also a monotonically increasing function of x, starting at $F(-\infty) = 0$ and ending at $F(\infty) = 1$



From (7) and the fundamental theorem of calculus we have that

$$p(x) = \frac{dF}{dx} \qquad (8)$$

Moreover, again from (7), we have

$$P(a \leqslant X \leqslant b) = F(b) - F(a) \qquad (9)$$

The uniform distribution

The uniform distribution Unif (a, b) is defined as

$$\boxed{X \sim \text{Unif}(a,b) \quad : \quad p(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}} \tag{10}$$

Of particular importance is Unif $(0,1)$, also called the standard uniform.

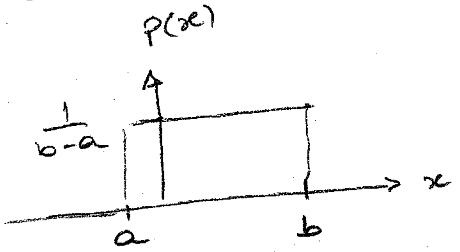
The CDF is found from the definition (7)

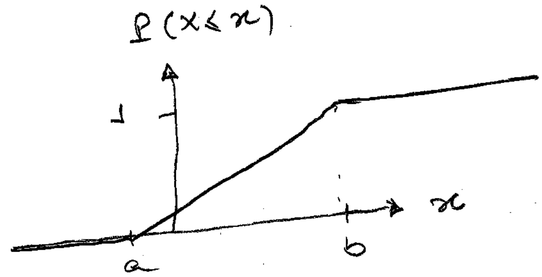
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(y)\,dy = \int_{a}^{x} \frac{1}{b-a}\,dy$$

Thus

$$P(X \leq x) = \frac{x-a}{b-a}, \quad x \in [a,b] \tag{11}$$

This looks like this





Consider now Unif $(0,1)$ and let $0 \leq c \leq d \leq 1$. Then

$$P(c \leq X \leq d) = d - c \tag{12}$$

Please remember this:

$$\boxed{\text{Uniform: Prob} \propto \text{length of the interval}} \tag{13}$$

Next we compute the average:

$$\langle x \rangle = \int_a^b \frac{1}{b-a}\, x\, dx = \frac{1}{b-a}\left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{1}{2}\,\frac{(b-a)(b+a)}{b-a}$$

Thus

$$\boxed{\langle x \rangle = \frac{b+a}{2}} \qquad (14)$$

And the 2nd moment

$$\langle x^2 \rangle = \frac{1}{b-a}\int_a^b x^2\, dx = \frac{1}{(b-a)}\,\frac{(b^3-a^3)}{3} = \frac{1}{3}\,\frac{(b-a)(a^2+ab+b^2)}{b-a}$$

Thus

$$\langle x^2 \rangle = \frac{a^2+ab+b^2}{3} \qquad (15)$$

The variance is then

$$\mathrm{Var}(x) = \langle x^2 \rangle - \langle x \rangle^2 = \frac{a^2+ab+b^2}{3} - \frac{(a^2+2ab+b^2)}{4}$$

$$= \frac{1}{12}\left(a^2+b^2-2ab\right)$$

Thus

$$\boxed{\mathrm{var}(x) = \frac{(a-b)^2}{12}} \qquad (16)$$

The exponential distribution

The exponential distribution Expo (λ) is defined by

$$\boxed{X \sim \text{Expo}(\lambda), \qquad p(x) = \lambda\, e^{-\lambda x}, \quad x > 0} \tag{17}$$

Let's check the normalization:

$$\int_{-\infty}^{\infty} p(x)\,dx = \lambda \int_0^{\infty} e^{-\lambda x}\,dx = \frac{\lambda}{-\lambda}\, e^{-\lambda x}\Big|_0^{\infty} = -1\,(0-1) = 1$$

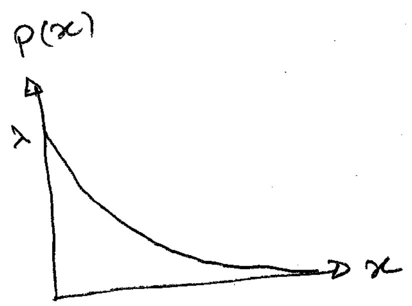
Yay!

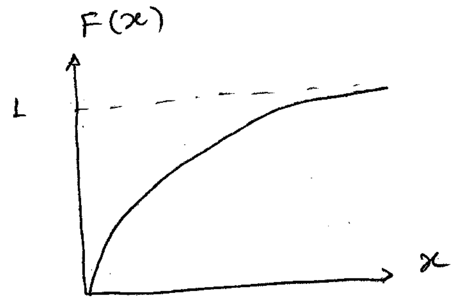
The CDF is

$$F(x) = \int_{-\infty}^{x} p(x')\,dx' = \lambda \int_0^{x} e^{-\lambda x'}\,dx' = \frac{\lambda}{-\lambda}\, e^{-\lambda x'}\Big|_0^{x}$$

whence

$$F(x) = 1 - e^{-\lambda x} \tag{18}$$





The mean is

$$\langle x \rangle = \lambda \int_0^{\infty} x\, e^{-\lambda x}\,dx$$

There is a neat trick to evaluate this integral. We write it as

$$\langle x \rangle = -\lambda \frac{\partial}{\partial \lambda} \int_0^\infty e^{-\lambda x} dx = -\lambda \frac{\partial}{\partial \lambda}\left(\frac{1}{\lambda}\right) = -\lambda\left(-\frac{1}{\lambda^2}\right) = \frac{1}{\lambda}$$

$$\therefore \quad \langle x \rangle = \frac{1}{\lambda} \tag{19}$$

We use the same trick to compute the second moment

$$\langle x^2 \rangle = \lambda \int_0^\infty x^2 e^{-\lambda x} dx = \lambda \frac{\partial^2}{\partial \lambda^2} \int_0^\infty e^{-\lambda x} dx$$

$$= \lambda \frac{\partial^2}{\partial \lambda^2}\left(\frac{1}{\lambda}\right)$$

whence

$$\langle x^2 \rangle = \frac{2}{\lambda^2} \tag{20}$$

The variance is then

$$\boxed{\mathrm{Var}(x) = \frac{1}{\lambda^2}} \tag{21}$$

Connection between the exponential and the Poisson

Now I want to show you an interesting interpretation of the exponential distribution. Let N_t be the number of buses which go through a bus stop during a time t. It is reasonable to assume that

$$N_t \sim \mathrm{Pois}(\lambda t) \tag{22}$$

where λ = "average number of buses per hour" (assuming t is measured in hours).

Now suppose you just arrived at the bus stop. Let T denote the time until the first bus arrives. I will show now that

$$T \sim \mathrm{Expo}(\lambda) \tag{23}$$

The basis of this calculation is the statement

$$P(T > t) = P(N_t = 0) \tag{24}$$

The LHS is the prob. that the first bus only arrives after t. The RHS is the prob that the number of buses which passed up to time t is zero. These are just two ways of saying the same thing.

Now, $N_t \sim \mathrm{Pois}(\lambda t)$ so

$$P(N_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

and therefore

$$P(T > t) = P(N_t = 0) = e^{-\lambda t} \tag{25}$$

The quantity $P(T > t)$ is related to the CDF $P(T \leq t)$ by

$$P(T > t) + P(T \leq t) = 1$$

Thus

$$P(T \leq t) = 1 - P(T > t) = 1 - e^{-\lambda t} \qquad (26)$$

This is the CDF of Expo (λ).

Moral of the story: don't do drugs.

Moral of the story: if $N_t \sim$ Pois (λt) counts the "arrival" or occurrence of certain events, then the time before the first arrival is Expo (λ).

The <u>memoryless property</u>: you may have noticed something strange in our previous calculation. We did not have to specify when we first arrived at the bus stop. We could have arrived immediately after a bus passed or it could have been 30 min since the last bus passed. This is actually a special property of the exponential distribution: it is memoryless. No matter how long you wait, you always start fresh. In symbols

$$P(T \geq t+s \mid T \geq s) = P(T \geq t) \qquad (26')$$

The proof is simple. From the definition of conditional prob

$$P(T \geq t+s \mid T \geq s) = \frac{P(T \geq t+s, \overbrace{T \geq s}^{\text{Redundant}})}{P(T \geq s)} = \frac{P(T \geq t+s)}{P(T \geq s)}$$

But for Expo (λ), using (18), $P(T \geq t) = 1 - P(T \leq t) = e^{-\lambda t}$

Thus

$$P(T \geq t+s \mid T \geq s) = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = P(T \geq t)$$

which is (26'). The memoryless property is very special. There are only two distributions which satisfy it: Expo(λ) and Geom(p).

Random number generators

Generating random numbers in a computer is extremely important in physics in general and statistical mechanics in particular. They are used specially in a method called Monte-Carlo simulations, which we will discuss later on.

Generating random numbers is also an art. And one which has been evolving quickly in recent years. If you ever encounter this type of problem I suggest you never try to implement a random number generator by yourself. Use an implementation made by a professional. Not only will it be faster than yours, it will also be more reliable: a crappy random number generator will give wrong physical answers to your simulations.

The important distributions like $\mathrm{Unif}(0,1)$ and $N(0,1)$ have dedicated generators. If you want to generate some other distribution, then you first generate $\mathrm{Unif}(0,1)$ and then use the following transformation method.

Let $U \sim \mathrm{Unif}(0,1)$ be a uniform r.v. Also let X be some r.v. you want to simulate, with $F(x)$ being its corresponding CDF. Then the following result is something worth remembering:

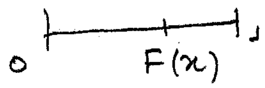
$$\boxed{\text{If } U \sim \mathrm{Unif}(0,1) \text{ then } X = F^{-1}(U) \sim F}$$

In words, find the inverse function and apply to it U.

__Proof__: $\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x))$

But for Unif $(0,1)$, prob. = length so

$$\mathbb{P}(U \leq F(x)) = F(x)$$



qed.

__Example__: generating Expo (λ).

The CDF of Expo (λ) is [Eq (17)]

$$F(x) = 1 - e^{-\lambda x}$$

we now find the inverse : $F(x) = u \rightsquigarrow x = F^{-1}(u)$

$$e^{-\lambda x} = 1 - u$$

$$-\lambda x = \ln(1-u$$

$$x = -\frac{1}{\lambda} \ln(1-u)$$

Thus

$$\boxed{F^{-1}(u) = -\frac{1}{\lambda} \ln(1-u)}$$

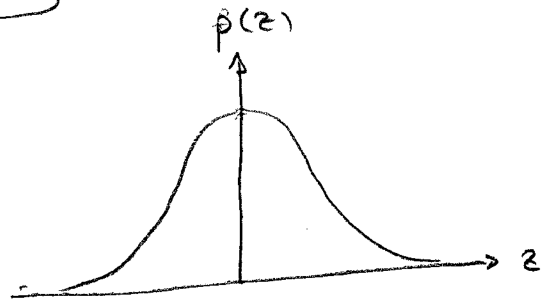
Generate $U \sim$ Unif $(0,1)$. Then $F^{-1}(U) \sim$ Expo (λ)

The Normal/Gaussian distribution

The normal distribution is the most important distribution in all of physics. This is due to a result known as the central limit theorem, which states that the distribution of the sum of a bunch of iid r.v.s will be normal.

The simplest normal, called $N(0,1)$, is defined by the distribution

$$\boxed{z \sim N(0,1) : \quad p(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}} \tag{27}$$



The normalization is ensured by the famous Gaussian integral

$$\boxed{\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}} \tag{28}$$

Setting $a = 1/2$ gives $\sqrt{2\pi}$.

The mean is

$$\langle z \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z\, e^{-z^2/2}\, dz = 0 \tag{29}$$

I know it is zero because of parity: this is the product of an even with an odd function

The 2^{nd} moment is

$$\langle z^2 \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz$$

I like to compute this integral using a trick similar to (18). First we compute the more general result

$$\int_{-\infty}^{\infty} z^2 e^{-az^2} dz = -\frac{\partial}{\partial a} \int_{-\infty}^{\infty} e^{-z^2} dz = -\frac{\partial}{\partial a} \sqrt{\frac{\pi}{a}} = \frac{1}{2} \frac{\sqrt{\pi}}{a^{3/2}}$$

Now I set $a = 1/2$, which gives

$$\langle z^2 \rangle = \frac{1}{\sqrt{2\pi}} \frac{\pi}{2} 2^{3/2} = 1$$

Thus, the variance is

$$\boxed{Var(z) = 1} \tag{30}$$

The numbers appearing in $N(0,1)$ are therefore the mean and the variance.

The CDF of $N(0,1)$ is

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2} dy := \frac{1}{2} erfc\left(-\frac{z}{\sqrt{2}}\right) \tag{31}$$

This integral cannot be expressed in terms of elementary functions. It is written in terms of the complex error function $erfc(x)$.

The normal $N(\mu, \sigma^2)$

The more general normal is constructed from the $N(0,1)$ as

$$X = \mu + \sigma Z \qquad Z \sim N(0,1) \tag{32}$$

where μ, σ are two parameters. The mean is

$$\langle X \rangle = \mu + \sigma \langle Z \rangle = \mu$$

and the variance is

$$Var(X) = \sigma^2 \, Var(Z) = \sigma^2$$

Thus:

$$\boxed{X \sim N(\mu, \sigma^2) : \quad \begin{array}{l} \langle X \rangle = \mu \\ Var(X) = \sigma^2 \end{array}} \tag{33}$$

To find the distribution it is easier to start with

$$F(x) = P(X \leqslant x) = P(\mu + \sigma Z \leqslant x) = P\left(Z \leqslant \frac{x-\mu}{\sigma}\right)$$

Looking at (26) we will then have

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-y^2/2} \, dy$$

The PDF is then found by differentiation [Eq (5)]:

$$p(x) = \frac{d}{dx} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

thus:

$$\boxed{X \sim N(0,1): \quad p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}} \tag{34}$$

The Dirac δ function as a prob. distribution

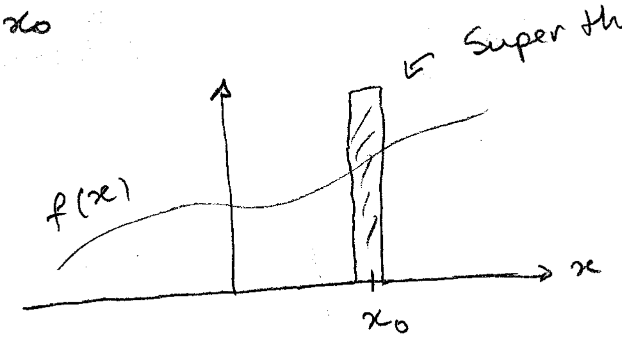
The Dirac δ function is defined as

$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases} \quad \text{but} \quad \int_{-\infty}^{\infty} \delta(x)\, dx = 1 \qquad (35)$$

If $f(x)$ is an arbitrary function, it then follows that

$$\int_{-\infty}^{\infty} f(x)\, \delta(x - x_0) = f(x_0) \qquad (36)$$

So $\delta(x-x_0)$ functions as a "window", which takes only the part of $f(x)$ around x_0

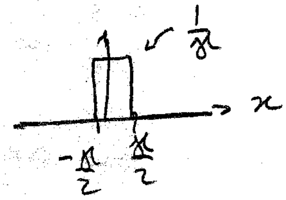


The fact that $\delta(x)$ integrates to 1 allows us to interpret it as a probability distribution. It defines a distribution that is infinitely sharp, meaning that x is simply a deterministic variable.

You may also think about $\delta(x)$ as a limiting case of other distributions, with $\gamma \to 0$

$$X \sim \text{Unif}\left(-\frac{\gamma}{2}, \frac{\gamma}{2}\right) : \quad \delta(x) \simeq \begin{cases} \dfrac{1}{\gamma} & -\dfrac{\gamma}{2} \leqslant x \leqslant \dfrac{\gamma}{2} \\ 0 & \text{otherwise} \end{cases} \qquad (37)$$

" Very sharp uniform"



or

$$X \sim N(0, \gamma) : \quad \delta(x) \simeq \frac{1}{\sqrt{2\pi}\,\gamma} e^{-x^2/2\gamma^2}$$

" very sharp Gaussian" (38)

or

$$X \sim C(0, \gamma) : \quad \delta(x) \simeq \frac{1}{\pi} \frac{\gamma}{\gamma^2 + x^2}$$

Cauchy distribution. We will learn about it later. (39)

Integral representation of the δ-function

There is a way to represent the δ function which is really, really really useful in practice:

REMEMBER ME!

$$\boxed{\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\,dk} \qquad (40)$$

This formula is worth remembering. It could save your life one day.

Before we prove it, let us try to understand it: if $x = 0$ we get

$$\int_{-\infty}^{\infty} 1\,dk = \infty$$

which agrees with (35). When $x \neq 0$ the integral can be split into sines and cosines. These oscillate indefinitely, which leads to destructive interference. Hence when $x \neq 0$ we get 0.

To prove Eq (38) consider first the following integral ($\gamma > 0$)

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx - \gamma|k|}\,dk = \frac{1}{2\pi}\int_{0}^{\infty} e^{(ix-\gamma)k}\,dk + \frac{1}{2\pi}\int_{-\infty}^{0} e^{(ix+\gamma)k}\,dk$$

$$= \frac{1}{2\pi}\left.\frac{e^{(ix-\gamma)k}}{ix-\gamma}\right|_0^{\infty} + \frac{1}{2\pi}\left.\frac{e^{(ix+\gamma)k}}{ix+\gamma}\right|_{-\infty}^{0}$$

$$= \frac{1}{2\pi}\left[\frac{1}{\gamma - ix} + \frac{1}{\gamma + ix}\right]$$

$$= \frac{1}{2\pi}\left[\frac{\gamma + ix + \gamma - ix}{\gamma^2 + x^2}\right]$$

$$= \frac{1}{2\pi}\,\frac{2\gamma}{\gamma^2 + x^2}$$

Thus we conclude that

$$\boxed{\frac{1}{\pi}\frac{\varepsilon}{x^2+\varepsilon^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx-\varepsilon|k|}\,dk} \qquad (41)$$

This is an exact result. Now we use Eq (39) to note that, when $\varepsilon\to 0$ the LHS becomes $\delta(x)$. In this way we arrive at Eq (40).

— // —

Since we are talking about $\delta(x)$, I want to mention a formula which we will need soon. Let $f(x)$ be an arbitrary function which is zero at certain values x_i (that is, $f(x_i)=0$). Then

$$\boxed{\delta(f(x)) = \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|}} \qquad (41')$$

To see why this works, suppose $f(x)$ only has one root at x_0. Then

$$\int_{-\infty}^{\infty} g(x)\,\delta(f(x))\,dx = \int_{-\infty}^{\infty} g(x(y))\frac{\delta(y)}{|f'(x)|}\,dy = \frac{g(x_0)}{|f'(x_0)|}$$

$y=f(x)$

$dy = f'(x)dx$

I wrote $|f'(x)|$ because if $f'(x_0)<0$ then the limits of integration are also inverted. The result $g(x_0)/|f'(x_0)|$ is the same we would get from

$$\int_{-\infty}^{\infty} g(x)\frac{\delta(x-x_0)}{|f'(x_0)|}\,dx = \frac{g(x_0)}{|f'(x_0)|}$$

To extend the proof to multiple roots, split the integral into several parts, each centered around x_i. Then apply the same procedure for each integral.

Expressing discrete r.v.s using a PDF

Let X be a r.v. with probabilities $P(X=k)$. We may also interpret X as a continuous random variable, with PDF

$$\boxed{p(x) = \sum_k P(X=k)\,\delta(x-k)} \tag{42}$$

This allows us to treat discrete and continuous r.v.s using a unified formalism. Normalization now translates as

$$1 = \int_{-\infty}^{\infty} p(x)\,dx = \int_{-\infty}^{\infty} \sum_k P(X=k)\,\delta(x-k)\,dx$$

$$= \sum_k P(X=k) \underbrace{\int_{-\infty}^{\infty} \delta(x-k)\,dx}_{=1}$$

Thus

$$1 = \int_{-\infty}^{\infty} p(x)\,dx = \sum_k P(X=k) \tag{43}$$

In the description (42) a discrete r.v. is therefore interpreted as if it varied continuously, but with a PDF that has extremely sharp peaks.

This type of description is convenient for a variety of reasons. First, it allows for a unified framework. Second, it may be used to encompass those situations where the r.v. should be discrete but, due to experimental uncertainties, there is some blur around it.

Finally, we may also use this type of description when we have r.v.s that take on a mix of discrete and continuous states. For instance, the energy of an electron in the presence of a proton is discrete if $E<0$ (the energy levels of the hydrogen atom) and continuous if $E>0$ (in this case the electron is not bound to the p^+, but just passing by).

Transformation of variables

Let X be a r.v. with PDF $P_X(x)$ and let $Y = f(X)$, where f is an arbitrary function. What is $P_Y(y)$? This type of problem appears frequently in prob. theory and statistical mechanics. For instance, if the velocity v of a particle is a r.v., what is the PDF of the kinetic energy?

Note that, if you only want expectation values then the problem is much easier. For instance

$$\langle Y \rangle = \langle f(X) \rangle = \int f(x)\, P_X(x)\, dx$$

But if you want to know the full distribution $P_Y(y)$, then the problem is usually harder.

I will give you two formulas for transforming variables. The first is simple to use but only holds when $f(x)$ is a one-to-one function [e^x is one-to-one but x^2 is not. A one-to-one function crosses each point of the vertical axis only once]. Then I will give you a formula which is more abstract, but holds for any situation.

The proofs of these formulas will be given in the next set of notes, when we discuss the characteristic function.

1) when $y = f(x)$ is one-to-one

$$\boxed{p_Y(y) = p_X(x)\left|\frac{dx}{dy}\right| \quad \text{with} \quad x = f^{-1}(y)} \tag{44}$$

<u>Example</u>: $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$.

$$\left|\frac{dx}{dy}\right| = \frac{1}{|a|} \quad \text{and} \quad x = \frac{y-b}{a}$$

Thus

$$p_Y(y) = p_X(x)\frac{1}{|a|} = \frac{1}{|a|}\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{\sqrt{2\pi}\,|a|\sigma} \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{y-b}{a} - \mu\right)^2\right\}$$

$$= \frac{1}{\sqrt{2\pi}\,|a|\sigma} \exp\left\{-\frac{1}{2(a\sigma)^2}(y - b - a\mu)^2\right\}$$

this is the PDF of a Gaussian again. thus we conclude that

$$\boxed{\text{If } X \sim N(\mu, \sigma^2) \text{ then } aX + b \sim N(a\mu + b, |a|^2\sigma^2)} \tag{45}$$

As a particular case, if $Z \sim N(0,1)$ then $X = \sigma Z + \mu \sim N(\mu, \sigma^2)$ which is (32), (33).

Another interesting consequence is

$$X \sim N(\mu, \sigma^2) \quad \text{then} \quad -X \sim N(-\mu, \sigma^2) \tag{46}$$

2) General formula

$$\boxed{P_Y(y) = \int_{-\infty}^{\infty} \delta(y - f(x))\, P_X(x)\, dx} \qquad (47)$$

To use this formula we need Eq. (41').

$$\delta(y - f(x)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|} \qquad (48)$$

where x_i are the roots of $y = f(x)$. For instance, if $f(x)$ is one-to-one we get $\delta(y - f(x)) = \frac{\delta(x - x_0)}{|f'(x_0)|}$, where $x_0 = f^{-1}(y)$.

then (47) becomes

$$P_Y(y) = \int_{-\infty}^{\infty} \frac{\delta(x - x_0)}{|f'(x_0)|} P_X(x)\, dx = \frac{P_X(x_0)}{|f'(x_0)|}$$

This is exactly Eq (44) since $|f'(x_0)| = |dy/dx|$.

Example ; chi-squared distribution

Let $Z \sim N(0,1)$ and $X = Z^2$. What is $P_X(x)$?

Well, in this case $x = f(z)$ with $f(z) = z^2$. Thus, the equation $x = z^2$ has two roots at $z = \pm\sqrt{x}$. We then get

$$\delta(y - z^2) = \frac{\delta(z - \sqrt{x})}{2\sqrt{x}} + \frac{\delta(z + \sqrt{x})}{2\sqrt{x}}$$

Applying Eq (47) we then get

$$P_X(x) = \frac{1}{2\sqrt{x}} \int_{-\infty}^{\infty} \delta(z - \sqrt{x})\, P_Z(z)\,dz + \frac{1}{2\sqrt{x}} \int_{-\infty}^{\infty} \delta(z + \sqrt{x})\, P_Z(z)\,dz$$

$$= \frac{1}{2\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-x/2} + \frac{1}{2\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-x/2}$$

Whence

$$P_X(x) = \frac{1}{\sqrt{2\pi x}} e^{-x/2} \qquad x > 0 \tag{49}$$

This is called the $\chi(1)$ distribution. In general if $X = Z_1 + \ldots + Z_n$ we will obtain the $\chi(n)$ distribution. We will deal with this case later.

Note : if $y = x$ then, according to Eq (47) we may write

$$P_X(x) = \int_{-\infty}^{\infty} \delta(x - x')\, P_X(x')\,dx'$$

Thus, we see that

$$\boxed{P_X(x) = \langle \delta(x - x') \rangle}$$

The PDF is the average of the δ function.

The lognormal distribution

When you produce micro or nanoparticles, for instance in chemical or farmaceutical processes, each particle will usually have a different diameter. Thus, the diameter of a nanoparticle is a random variable. Usually the diameter follows a lognormal distribution.

The word "lognormal" means that "the log is normal". That is, if Y is lognormal then $X = \ln Y$ will be normal. We may find the PDF of Y starting with $X \sim N(\mu, \sigma^2)$ and $Y = e^X$. The exponential is one-to-one so we may use (44). We have

$$\frac{dy}{dx} = e^x = y$$

so

$$P_Y(y) = P_X(x)\left|\frac{1}{y}\right| = \frac{1}{\sqrt{2\pi}\,y\sigma}\exp\left[-\frac{1}{2\sigma^2}(\ln y - \mu)^2\right] \qquad (50)$$

$y>0$ so $|y| = y$

Multivariate distributions

If we have two random variables X and Y, they may be described together using the joint PDF

$$p(x,y) \quad \text{or} \quad p_{X,Y}(x,y) \tag{51}$$

It is normalized as

$$\int p(x,y)\,dx\,dy = 1 \tag{52}$$

If we eventually get tired of one of them, we may compute the marginal distribution by integrating over one of them:

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dy$$
$$p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx \tag{53}$$

Notice that this is a one way process: from the joint you get the marginal, but from the marginal there is no way of getting back the joint.

To see how everything fits together, consider

$$\langle x \rangle = \int x \, p_{X,Y}(x,y)\, dx\, dy = \int dx \; x \underbrace{\int dy \; p_{X,Y}(x,y)}_{p_X(x)}$$

$$= \int dx \; x \, p_X(x)$$

Statistical independence may now be defined as

$$\boxed{X, Y \text{ are independent if } \; p_{X,Y}(x,y) = p_X(x)\, p_Y(y)} \qquad (54)$$

Remember this

$$\boxed{\text{Independent} = \text{prob. is a product}} \qquad (55)$$

Finally, we define conditional probability as

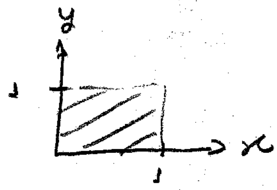
$$\boxed{\begin{aligned} p_{X,Y}(x,y) &= p_{X|Y}(x|y)\, p_Y(y) \\ &= p_{Y|X}(y|x)\, p_X(x) \end{aligned}} \qquad (56)$$

The interpretation is the same as in the discrete case

Ex: the 2D uniform

1) (X, Y) are uniform on a square

$$x, y \in [0,1]$$



The joint PDF is

$$p_{X,Y}(x,y) = \begin{cases} 1 & x, y \in [0,1] \\ 0 & \text{otherwise} \end{cases} \tag{54}$$

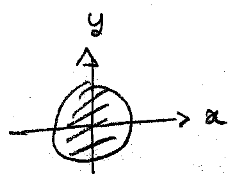
The "1" there is the area of the square. In analogy with the 1D case

$$\text{Prob} = \text{area} \tag{46}$$

In this case X and Y are independent Unif $(0,1)$.

2) (X, Y) are uniform on a circle:

$$x^2 + y^2 = 1$$



The area of the circle is π so the joint dist. is

$$p_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 = 1 \\ 0 & \text{otherwise} \end{cases} \tag{55}$$

But now X and Y are not independent. For instance

$$X = 0 \quad \Rightarrow \quad -1 \leq Y \leq 1$$

$$X = 1 \quad \Rightarrow \quad Y = 0$$

Let us marginalize on X:

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x,y)\, dy = \frac{1}{\pi}\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}} dy = \frac{2}{\pi}\sqrt{1-x^2}$$

Thus

$$p_X(x) = \frac{2}{\pi}\sqrt{1-x^2} \qquad x \in [-1,1] \qquad (18)$$

The PDF is larger in the middle, where X may take on more values. Clearly, X is not Uniform.

By symmetry $p_Y(y) = \frac{2}{\pi}\sqrt{1-y^2}$.

Now let us find the conditional prob. From (44)

$$p_{Y|X}(y|x) = \frac{p_{X,Y}(x,y)}{p_X(x)} = \frac{1}{2\sqrt{1-x^2}}$$

This will hold for $-\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}$. Thus, if we know that $X=x$, Y will be uniform

$$Y|X=x \sim \text{Unif}\left(-\sqrt{1-x^2}, \sqrt{1-x^2}\right) \qquad (19)$$

—//—

We will discuss more about multivariate distributions later on in the course.