CONTINUOUS DISTRIBUTIONS

Jerzy Szulga

Department of Mathematics and Statistics
Auburn University

STAT 3600 SUMMER 2019

September 30, 2019

Outline

Continuous probability distributions

- cdf and pdf
- Additional properties

Special & named continuous distributions

- Exponential
- Uniform

Weibull, Pareto, etc.

cdf

A discrete random variable entails the pmf $f(x) = \mathrm{P}(X = x) > 0$ for an enumerable number of values x.

Many physical quantities are of continuous type, i.e., $f(x) = 0$ for all x, for example: time, length, area, mass, temperature, etc. In lieu of pmf, now useless, we will be using the **cumulative distribution function**, or cdf in short:

$$F(x) = \mathrm{P}(X \leq x)$$

Properties of cdf

We observe that

1. $F(x)$ is non-decreasing;
2. $\lim_{x \to -\infty} F(x) = 0$;
3. $\lim_{x \to \infty} F(x) = 1$;

   In addition

4. $F(x)$ possess the left limits $\lim_{h>0, h \to 0} F(x - h)$;
5. $F(x)$ is continuous from the right: $\lim_{h>0, h \to 0} F(x + h) = F(x)$.

Density or pdf

In most of the cases of interest the cdf $F(x)$ of a r.v. X is differentiable and it can be recovered from its derivative $f(x) = F'(x)$:

$$F(x) = \int_{-\infty}^{x} f(u)\, du.$$

Then we say that X has a **continuous probability distribution** and its derivative $f(x)$ is called a **probability density function**, or **pdf** in short.

Examples of pdf

The term "*density*" originated in physics.

Note the properties of a pdf:

1. $f(x) \geq 0$,
2. $\displaystyle\int_{-\infty}^{\infty} f(u)\, du = 1$.

Therefore any nonnegative integrable function on the real line can be turned into pdf:

$$a = \int_{-\infty}^{\infty} f(u)\, du < \infty \quad \mapsto \quad \text{replace} \quad f(x) := \frac{1}{a} f(x).$$

Support

By the support of a pdf (or a corresponding r.v. X) we understand the set

$$\{x : f(x) > 0\}$$

Recall that the **domain** is the set of x for which $f(x)$ is defined.

Example In the following case, $(-\infty, \infty)$ is the domain while $[0, \infty)$ is the support of the pdf:

$$f(x) = \begin{cases} e^{-x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

Support, notation

The expanded formula with curly brackets is often unwieldy, so we write crisply

$$f(x) = e^{-x},\ x \geq 0, \quad \text{or} \quad f(x) = e^{-x} \mathbb{1}_{\{x \geq 0\}} = e^{-x} \mathbb{1}_{[0,\infty)}(x)$$

The latter symbol is called an **indicator**. It is a function that takes two values: 1 (true) or 0 (false).

Again, the restriction that appears after the formula of a function describes the support, not the domain.

Continuous cdf - no worries about the end points

The cdf is well defined also in the discrete case. However, it is piecewise constant and increases only by jumps. It is quite unwieldy in the discrete case and we must carefully distinguish probabilities, e.g.,

$$\mathrm{P}(X \le x) \quad \text{versus} \quad \mathrm{P}(X < x)$$

In the continuous case, $\mathrm{P}(X = x)$ for all x. Hence

$$\mathrm{P}(a \le X \le b) = \mathrm{P}(a \le X < b) = \mathrm{P}(a < X \le b) = \mathrm{P}(a < X < b).$$

cdf - setting up probabilities

Disregarding the endpoints (whether included or excluded),

$$\mathrm{P}(a < X < b) = \int_a^b f(u)\,du, \quad \overline{F}(x) = \mathrm{P}(X > x) = \int_x^\infty f(u)\,du.$$

The latter function, called sometimes the (probability) **tail** or **excess function** (there are many other synonims) uniquely determines the probability distribution since

$$F(x) = 1 - \overline{F}(x), \quad \overline{F}(x) = 1 - F(x).$$

Mean and variance

Let X be of continuous type with pdf $f(x)$. Like before:

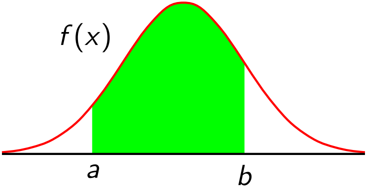
$$\begin{aligned} \mu &= \operatorname{E} X &&= \int_{-\infty}^{\infty} x\, f(x)\, dx \\ \sigma^2 &= \operatorname{Var}(X) &&= \int_{-\infty}^{\infty} (x-\mu)^2\, f(x)\, dx \end{aligned}$$

By the same token, the mean of a transformed r.v., e.g., the mgf:

$$\operatorname{E} g(X) = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx, \quad M(t) = \operatorname{E} e^{tX} = \int_{-\infty}^{\infty} e^{tx}\, f(x)\, dx.$$

Summary

A continuous random variable X takes a specific value with probability 0. The associated probabilities can be computed with the help of its probability density function $f(x)$



$$f(u) \geq 0, \quad \int_{-\infty}^{\infty} f(u)du = 1$$

$$P(a \leq X \leq b) = \int_a^b f(u)du$$

The mean $\mu = E(X)$ is a measure of center or location, and the variance $\sigma^2 = var(X)$ is a measure of variability or dispersion.

$$\mu = \int_{-\infty}^{\infty} xf(x)dx \qquad \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx.$$

Joint cdf and pdf

For two (or more) continuous r.vs., we put

$$F(x, y) = \mathrm{P}(X \le x,\ Y \le y)$$

$$F(x_1, x_2, \cdots, x_n) = \mathrm{P}(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n),$$

yielding the pdf:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x\, \partial y} \quad \text{(similarly for } n \text{ variables).}$$

Independence

$X_1, \ldots, X_n$ are independent if

$$F(x_1, x_2, \cdots, x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdots F_{X_n}(x_n).$$

or, equivalently

$$f(x_1, x_2, \cdots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdots f_{X_n}(x_n).$$

Smart formulas

Like in the discrete case:

$$\mathrm{Var}(X) = \mathrm{E}\,(X^2) - (\mathrm{E}\,X)^2$$

$$\mathrm{E}\,(c_1 X_1 + c_2 X_2 + \cdots) = c_1 \mathrm{E}\,X_1 + c_2 \mathrm{E}\,X_2 + \cdots$$

$$\mathrm{Var}(c_1 X_1 + c_2 X_2 + \cdots) = c_1^2\,\mathrm{Var}(X_1) + c_2^2\,\mathrm{Var}(X_2) + \cdots$$

(if independent)

$$\mathrm{E}\,(XY) = (\mathrm{E}\,X) \cdot (\mathrm{E}\,Y) \quad \text{(if independent)}$$

$$\mathrm{E}\,(X_1 \cdot X_2 \cdots) = (\mathrm{E}\,X_1) \cdot (\mathrm{E}\,X_2) \cdots \quad \text{(if independent)}$$

Mgf for sums of independent r.vs.

If X and Y are independent, then

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t),$$

and similarly for more independent factors.

In particular, when $X_1, \ldots X_n$ are iid with the common mgf $M(t)$, then

$$M_{X_1+\cdots+X_n}(t) = \Big(M(t)\Big)^n.$$

Tables

Discrete distributions appear on the first inner cover of the book.

Continuous distributions are on the last inner cover.

The lists contain the basic information:

- name
- pdf
- mean and variance
- Ross shows no mgf (it may be helpful)

Actually, Wikipedia under a given distribution's name has more information, including graphs.

Affine transformation

The most popular transformation is done with the help of an **affine function** $g(x) = ax + b$:

$$Y = g(X) = aX + b.$$

By simple algebra and elementary calculus:

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right), \quad f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right).$$

For the mean, variance, mgf

$$\mathrm{E}\, Y = a\mathrm{E}\, X + b, \quad \mathrm{Var}(Y) = a^2 \mathrm{Var}(X), \quad M_Y(t) = e^{tb}\, M_X\left(at\right).$$

Normalization

On the other hand, with $\mu = \mathsf{E}\, X$ and $\sigma^2 = \mathrm{Var}(X)$ we **normalize** or **standardize**:

$$Y = \frac{X - \mu}{\sigma}, \quad \text{so } \mathsf{E}\, Z = 0,\ \mathrm{Var}(Z) = 1.$$

In particular, even if X has a physical dimension, the normalized Y has none, it is a pure scalar.

When needed, we can "un-normalize":

$$X = \sigma Y + \mu.$$

The most user friendly distribution

The **standard exponential pdf** has a clear-cut mgf

$$f(x) = e^{-x},\ x \geq 0 \quad \mapsto \quad M(t) = \frac{1}{1-t}, \quad (t < 1).$$

Whence by differentiation all moments follow quickly

$$\mathrm{E}\, X^n = M^{(n)}(0) = n!, \quad n = 0, 1, 2, \dots,$$

In particular $\mathrm{E}\, X = 1$, $\mathrm{Var}(X) = 1$. Scaling $Y = \theta\, X$ yields the **general exponential distribution with parameter $\theta > 0$**:

$$f_Y(y) = \frac{1}{\theta}\, e^{x/\theta},\ x > 0, \quad \mathrm{E}\, Y = \theta, \quad \mathrm{Var}(Y) = \theta^2.$$

Lifetime model and Lack of Memory

Like any nonnegative continuous r.v., an exponential X may serve as a model of a lifetime of an organism, gadget, or system. Then the tail may be called the **survival function**:

$$\mathrm{P}(X > x),$$

the probability of surviving past x units of time.

The continuous exponential distribution shares with the discrete geometric distribution the **lack of memory**.

$$\mathrm{P}(X > x + h) | X > x) = \mathrm{P}(X > h).$$

Standard uniform pdf and cdf

The simplest is U[0,1]:

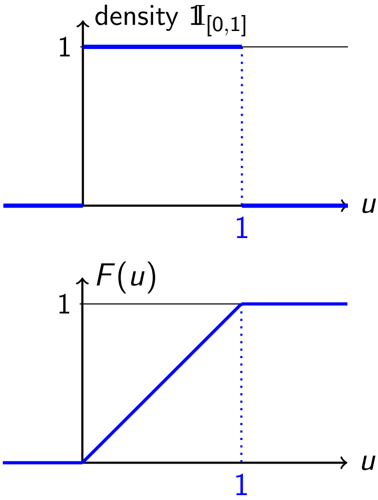
$$f = \mathbb{1}_{[0,1]}, \quad F(x) = \begin{cases} 0 & \text{if } x < 0, \\ x & \text{if } 0 \le x < le1, \\ 1 & \text{if } x > 1. \end{cases}$$

By geometry and physics, $\mathrm{E}\, X = \dfrac{1}{2}$. The moments:

$$\mathrm{E}\, X^p = \int_0^1 u^p \, dx = \frac{1}{p+1}, \quad p > -1.$$

Hence, $\mathrm{E}\,(X^2) = 1/3$, so $\mathrm{Var}(X) = 1/3 - (1/2)^2 = 1/12$.

Standard uniform - graphs



The general uniform on $[a, b]$

The unit interval $[0, 1]$ can be transformed to an arbitrary interval $[a, b]$ by the affine transformation

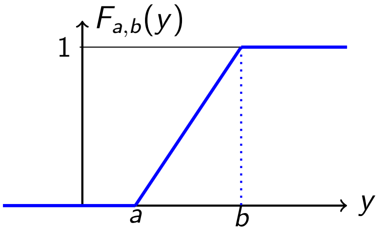
$$g(x) = a + (b - a)x, \quad 0 \le x \le 1.$$

Let $X \sim U[0, 1]$. Put $Y = a + (b - a)X$. Then

$$f_Y = \frac{1}{b-a}\mathbb{1}_{[a,b]}, \quad \mathrm{E}\, Y = \frac{a+b}{2}, \quad \mathrm{Var}(Y) = \frac{(b-a)^2}{12}$$

General uniform - cdf and graph

$$F_{a,b}(y) = P(a + (b-a)X \le y) = \begin{cases} 0, & y < a; \\ \dfrac{y-a}{b-a}, & a \le y \le b \\ 1, & y > b. \end{cases}$$



Standard uniform - the mother of all distributions

Let $F(x)$ be an arbitrary cdf. For simplicity assume that it is strictly increasing. Then the inverse function $G = F^{-1}$ exist. Let $X = G(U)$, where $U \sim U[0,1]$. Then

$$\mathrm{P}(X \leq x) = \mathrm{P}(F^{-1}(U) \leq x) = \mathrm{P}(U \leq F(x)) = F(x).$$

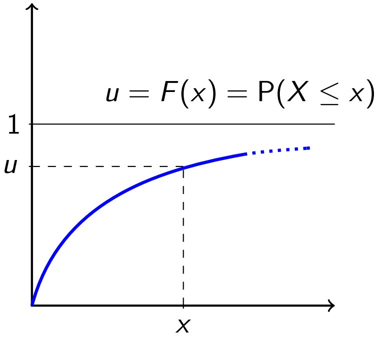
In other words, any cdf $F(x)$ is the cdf of a transformed standard uniform r.v.

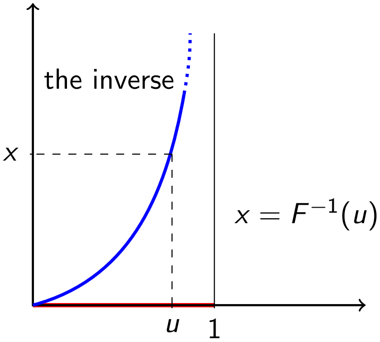
If $F(x)$ has constant segments (e.g. the cdf of a discrete distribution), it is still possible to define a generalized inverse.

MOAD - illustration

In particular, the unit interval $[0,1]$ with the usual length as the probability serves as the universal probability space.

A cdf $F(x)$ is the cdf of the transformed uniform $X = F^{-1}(U)$.





Transformed Uniform - Examples

$V = -\log U$ has the standard exponential distribution.

Indeed, for $v > 0$

$$P(V \leq v) = P(-\ln U \leq v) = \mathrm{P}(U \geq e^{-v}) = 1 - e^{-v}.$$

Hence $\theta\, V = -\ln U^{\theta}$ will be the general exponential with mean θ.

Weibull - power of exponential

Let V be exponential with mean θ.

Consider $W = V^{1/\beta}$, where $\beta > 0$. The tail:

$$\mathrm{P}(W > w) = \mathrm{P}(V^{1/beta} > w) = \mathrm{P}(V > w^{\beta}) = e^{-w^{\beta}/\theta}.$$

The pdf (differentiating): $f(w) = \dfrac{\beta}{\theta}\, w^{\beta-1}\, e^{-w^{\beta}/\theta}$.

This is a **Weibull distribution**, also serving as a model of lifetime.

The right tail is **light**.

$\beta < 1$: a vertical asymptote at 0;

$\beta > 1$: a flat at 0 followed by a hump.

Pareto - negative power of uniform

Let U be U[0,1]. Put $X = U^{-1/p}$. Then, for $x \geq 1$,

$$F(x) = \mathrm{P}(X \leq x) = \mathrm{P}(U > x^{-p}) = 1 - x^{-p}, \quad f(x) = F'(x) = \frac{p}{x^{p+1}}$$

(support $[1, \infty)$). The right light is **heavy** (a.k.a. **fat**) for small p.

Not just this one, but any probability density with a power right tail is called a **Pareto distribution**. The left tails (near 0) may vary.

Note that it has infinite mean for $p \leq 1$. Its mgf is not a mgf.
When $1 < p \leq 2$ is has a finite mean but infinite variance.