Important Continuous Distributions for Statistics and Data Science

For continuous distributions, the probability density function (pdf) $f(x)$ for a random variable X does not equal the probability that X equals x. Continous random variables have so many possible values that the probability of observing any single one of them is zero. Instead, the pdf is a curve under which area represents probability.

Uniform

Consider the continuous "uniform" distribution on the interval (a, b) given by the pdf

$$f(x) = \frac{1}{b-a} \cdot I_{(a,b)}(x).$$

We write $X \sim unif(a, b)$.

Notice that if you graph this pdf, it is a flat line over the interval (a, b). Since the total area under the pdf must be 1, this forces the height of the line to be $1/(b-a)$.

The uniform distribution is the continuous version of "equally likely outcomes". Consider the probability that X is in the interval (c, d) where $a < c < d < b$. One can "slide" this interval around, and, as long as it remains fully contained in (a, b), the probability that X falls in the interval remains the same!

Exponential

Let X be a continuous random variable with pdf

$$\begin{aligned} f(x) &= \begin{cases} \lambda e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad x < 0. \end{cases} \\ \\ &= \lambda e^{-\lambda x} \, I_{(0,\infty)}(x). \end{aligned}$$

Suppose that you are standing near the door of a grocery store watching customers arrive.

Suppose further that

- the arrival rate is a constant 15.2 people per minute, and
- the number of arrivals in non-overlapping periods of time are independent.

Let

$$X = \text{the time (in minutes) between any two consecutive arrivals.}$$

One can show that X has the exponential pdf given above with $\lambda = 15.2$. We will write $X \sim exp(rate = \lambda)$.

Note that some people write the exponential pdf as $f(x) = \frac{1}{\lambda} e^{-x/\lambda} I_{(0,\infty)}(x)$. In this case, λ is known a a "mean" paramter for reasons which will become apparent in this course. We will write $X \sim exp(mean = \lambda)$.

Be advised that most people and textbooks simply write $X \sim exp(\lambda)$.

Normal

Let X be a continuous random variable with pdf

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{for} \quad -\infty < x < \infty.$$

Then X is said to have a normal distribution with mean μ and variance σ^2. We write $X \sim N(\mu, \sigma^2)$.

We will not include an indicator on this pdf since it would be equal to 1 for all x and won't be "zeroing out" anywhere.

The graph of the $N(\mu, \sigma^2)$ pdf is the infamous "bell curve" in statistics. It is centered at μ and the value of σ^2 controls how wide and spread out it is.

Gamma

Let X have a "**gamma distribution** with parameters α and β" This means that X is a continuous random variable with pdf

$$f(x) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\beta x} I_{(0,\infty)}(x)$$

for some parameters $\alpha > 0$ and $\beta > 0$.

We write $X \sim \Gamma(\alpha, \beta)$.

Notes:

1. Just as with the exponential distribution, some people/books, write $X \sim \Gamma(\alpha, \beta)$ to mean that X has pdf

   $$f_X(x) = \frac{1}{\Gamma(\alpha)} (1/\beta)^\alpha x^{\alpha-1} e^{-x/\beta} I_{(0,\infty)}(x).$$

   Here, α and β are known as the "shape" and "scale" parameters, respectively.

   For our form of the gamma pdf, β is known as the "inverse scale parameter".

2. The pdf involves the "gamma function", $\Gamma(\alpha)$ which we define below. It is just the constant that ensures tha the pdf integrates to 1. The constant $\Gamma(\alpha)$ should not be confused with $\Gamma(\alpha,\beta)$ (two arguments) which is the name of a distribution.

An Aside: The Gamma Function

The pdf for the gamma distribution was defined using the **gamma function** which is denoted by $\Gamma(\cdot)$.

The gamma function, is defined, for $\alpha > 0$, as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\,dx.$$

Note that, for any $\beta > 0$,

$$\int_0^\infty \beta^\alpha x^{\alpha-1} e^{-\beta x}\,dx = \int_0^\infty (\beta x)^{\alpha-1} e^{-\beta x}\,\beta\,dx \stackrel{u=\beta x}{=} \int_0^\infty u^{\alpha-1} e^{-u}\,du = \Gamma(\alpha)$$

(Here we have used the fact that $du = \beta\,dx$ and that if x goes from 0 to ∞, then $u = \beta x$ also goes from 0 to ∞ since $\beta > 0$.)

Now,

$$\int_0^\infty \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\beta x}\,dx = \frac{1}{\Gamma(\alpha)} \int_0^\infty \beta^\alpha x^{\alpha-1} e^{-\beta x}\,dx = \frac{1}{\Gamma(\alpha)} \cdot \Gamma(\alpha) = 1,$$

so basically $1/\Gamma(\alpha)$ is the constant that makes $\beta^\alpha x^{\alpha-1} e^{-\beta x}$ into a proper pdf over $x \geq 0$!

Properties of the Gamma Function

1. $\Gamma(1) = 1$

   Proof: $\Gamma(1) = \int_0^\infty x^{1-1} e^{-x}\,dx = \int_0^\infty e^{-x}\,dx = 1.$

2. For $\alpha > 1$,

$$\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1).$$

   Proof: $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\,dx$ Using integration by parts

$$\int u\,dv = uv - \int v\,du$$

   with $u = x^\alpha$ and $dv = e^{-x}\,dx$ (So $du = (\alpha-1)x^{\alpha-2}\,dx$ and $v = \int e^{-x}\,dx = -e^{-x}$.), we have

$$\begin{aligned} \Gamma(\alpha - 1) &= -x^{\alpha-1} e^{-x}\Big|_0^\infty + \textstyle\int_0^\infty (\alpha-1) x^{\alpha-2} e^{-x}\,dx \\ &= 0 + (\alpha - 1) \textstyle\int_0^\infty x^{\alpha-2} e^{-x}\,dx = (\alpha-1)\cdot\Gamma(\alpha-1) \ \checkmark \end{aligned}$$

3. If $n \geq 1$ is an integer,

$$\Gamma(n) = (n-1)!.$$

Proof: By repeated application of property 2,

$$\begin{aligned}\Gamma(n) &= (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) \\ &= \cdots = (n-1)(n-2)\cdots(1)\underbrace{\Gamma(1)}_{1} = (n-1)!\end{aligned}$$