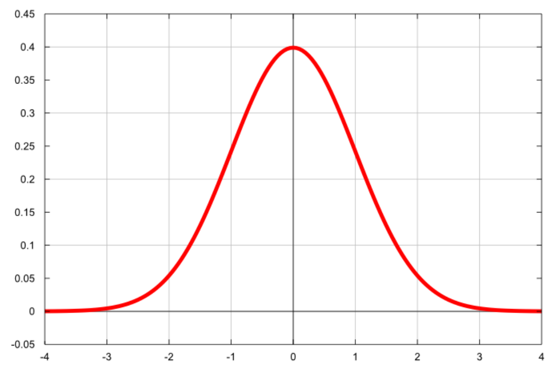
Continuous Random Variables

TF: Justin Zhu (justinzhu@college.harvard.edu), Credits to Timothy Kang

Continuous Random Variables

What is a Continuous Random Variable (CRV)? A continuous random variable can take on any possible value within a certain interval (for example, [0, 1]), whereas a discrete random variable can only take on variables in a list of countable values (for example, all the integers, or the values 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, etc.)

PMF's vs. PDF's Discrete R.V's have Probability *Mass* Functions, while continuous R.V.'s have Probability *Density* Functions. We visualize a PDF as a graph where the x axis is



Normal PDF

Intuitively, what do the y values represent? Well it doesn't make sense to say $P(X = x)$ for a continuous r.v. X because $P(X = x) = 0$ for all x. Think about the y value in the above graph as: **the relative frequency for getting a value within ϵ of the x value**

What is the Cumulative Density Function (CDF)? It is the following function of x.

$$F(x) = P(X \leq x)$$

With the following properties. 1) F is increasing. 2) F is right-continuous. 3) $F(x) \to 1$ as $x \to \infty$, $F(x) \to 0$ as $x \to -\infty$

What is the Probability Density Function (PDF)? The PDF, $f(x)$, is the derivative of the CDF.

$$F'(x) = f(x)$$

Or alternatively,

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Note that by the fundamental theorem of calculus,

$$F(b) - F(a) = \int_{a}^{b} f(x)dx$$

Thus to find the probability that a CRV takes on a value in an interval, you can integrate the PDF, thus finding the area under the density curve.

Two additional properties of a PDF: it must integrate to 1 (because the probability that a CRV falls in the interval $[-\infty, \infty]$ is 1, and the PDF must always be nonnegative.

$$\int_{-\infty}^{\infty} f(x)dx \qquad\qquad f(x) \geq 0$$

How do I find the expected value of a CRV? Where in discrete cases you sum over the probabilities, in continuous cases you integrate over the densities.

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Review: Expected value is *linear*. This means that for *any* random variables X and Y and any constants a, b, c, the following is true:

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

Discrete versus Continuous

	Discrete	Continuous
$P(X \leq x) =$	$F(x)$ (CDF)	$F(x)$ (CDF)
To find probabilities,	Add over PMF $P(X = x)$	Integrate over PDF $f(x) = F'(x)$
$E(X) =$	$\sum_x xP(X = x)$	$\int_{-\infty}^{\infty} xf(x)dx$
$E(g(X)) =$	$\sum_x g(x)P(X = x)$ (LOTUS Discrete)	$\int_{-\infty}^{\infty} g(x)f(x)dx$ (LOTUS Continuous)

Variance

What is Variance? Variance is the expected squared distance away from the mean. It is the square of the standard deviation.

$$\text{Var}(X) = E((X - EX)^2) = E(X^2) - E(X)$$

Oftentimes it's easier to calculate $E(X^2) - E(X)$ rather than $E(X - EX)$. You can find $E(X^2)$ with LotUS.

$$\text{Var}(aX + b) = a^2\text{Var}(X)$$

The following is true **only if X and Y are independent**:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

Universality of Uniform

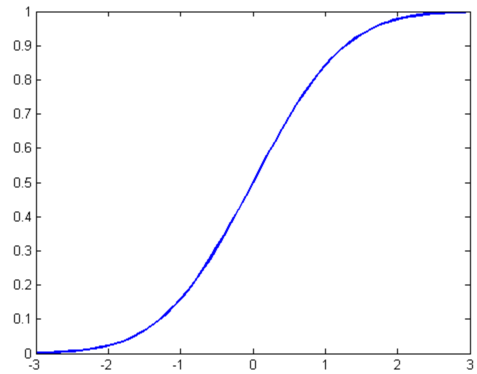
When you plug any random variable into its own CDF, you get a Uniform[0,1] random variable. When you put a Uniform[0,1] into an inverse CDF, you get the corresponding random variable. For example, let's say that a random variable X has a CDF

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

By the Universality of the the Uniform, if we plug in X into this function then we get a uniformly distributed random variable.

$$F(X) = 1 - e^{-X} \sim U$$

Similarly, since $F(X) \sim U$ then $X \sim F^{-1}(U)$. The key point is that *for any continuous random variable X, we can transform it into a uniform random variable and back by using its CDF.*



Normal CDF

In the CDF of the Normal(0,1) distribution above, if we choose a number between 0 and 1 uniformly at random on the y axis, we see that the distribution of x values is normal.

Uniform Distribution (Continuous)

Let us say that U is distributed $\text{Unif}(a, b)$. We know the following:

Properties of the Uniform For a uniform distribution, the probability of an draw from any interval on the uniform is proportion to the length of the uniform. The PDF of a Uniform is just a constant, so when you integrate over the PDF, you will get an area proportional to the length of the interval.

PDF and CDF

$$\text{Unif}(a, b) \qquad f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases} \qquad F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x > b \end{cases}$$

Normal Distribution (Continuous) (aka Gaussian)

Let us say that X is distributed $\mathcal{N}(\mu, \sigma^2)$. We know the following:

Central Limit Theorem The Normal distribution is ubiquitous because of the central limit theorem, which states that averages of independent identically-distributed variables will approach a normal distribution regardless of the initial distribution.

Transformable Every time we stretch or scale the normal distribution, we change it to another normal distribution. If we add c to a normally distributed random variable, then its mean increases additively by c. If we multiply a normally distributed random variable by c, then its variance increases multiplicatively by c^2. Note that for every normally distributed random variable $X \sim \mathcal{N}(\mu, \sigma^2)$, we can transform it to the standard $\mathcal{N}(0, 1)$ by the following transformation:

$$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Example Heights are normal. Measurement error is normal. By the central limit theorem, the sampling average from a population is also normal.

Standard Normal - The Standard Normal, denoted Z, is $Z \sim \mathcal{N}(0, 1)$

PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

CDF - It's too difficult to write this one out, so we express it as the function $\Phi(x)$

Continuous Distributions

Distribution	PDF and Support	Expected Value	Variance
Uniform Unif(a, b)	$f(x) = \frac{1}{b-a}, x \in [a, b]$ $f(x) = 0, x \notin [a, b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal $\mathcal{N}(\mu, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Exponential Expo(λ)	$f(x) = \lambda e^{-\lambda y}, x \in [0, \infty)$ $f(x) = 0, x \notin [0, \infty)$	$1/\lambda$	$1/\lambda^2$

Exponential Distribution (Continuous)

Let us say that X is distributed Expo(λ). We know the following:

Story You're sitting on an open meadow right before the break of dawn, wishing that airplanes in the night sky were shooting stars, because you could really use a wish right now. You know that shooting stars come on average every 15 minutes, but it's never true that a shooting star is ever "due" to come because you've waited so long. Your waiting time is memorylessness, which means that the time until the next shooting star comes does not depend on how long you've waited already.

Example The waiting time until the next shooting star is distributed Expo(4). The 4 here is λ, or the rate parameter, or how many shooting stars we expect to see in a unit of time. The expected time until the next shooting star is $\frac{1}{\lambda}$, or $\frac{1}{4}$ of an hour. You can expect to wait 15 minutes until the next shooting star.

All Exponentials are Scaled Versions of Each Other

$$Y \sim \text{Expo}(\lambda) \rightarrow X = \lambda Y \sim \text{Expo}(1)$$

PDF and CDF The PDF and CDF of a Exponential is:

$$f(x) = \lambda e^{-\lambda x}, \quad x \in [0, \infty) \qquad\qquad F(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad x \in [0, \infty)$$

Memorylessness The Exponential Distribution is the sole continuous memoryless distribution. This means that it's always "as good as new", which means that the probability of it failing in the next infinitesimal time period is the same as any infinitesimal time period. This means that for an exponentially distributed X and any real numbers t and s,

$$P(X > s + t | X > s) = P(X > t)$$

Given that you've waited already at least s minutes, the probability of having to wait an additional t minutes is the same as the probability that you have to wait more than t minutes to begin with. Here's another formulation.

$$X - a | X > a \sim \text{Expo}(\lambda)$$

Practice Problems

Example 1. Uniform Power.
Let $U \sim \text{Unif}(-1, 1)$

(a) Compute $E(U), \text{Var}(U), E(U^4)$.

(b) Find the PDF and CDF of U^2. Is it also a uniform distribution?

Example 2. Normal Squared.
Let $Z \sim N(0, 1)$ with CDF Φ. The PDF of Z^2 is the function given by:

$$g(w) = \frac{1}{\sqrt{2\pi w}} e^{-w/2}$$

with a support of $w \geq 0$.

(a) Find expressions for $E(Z^4)$ as integrals in two different ways, one based on the PDF of Z and the other based on the PDF of Z^2.

(b) Find $E(Z^2 + Z + \Phi(Z))$.

(c) Find the CDF of Z^2 in terms of Φ; do not find the PDF of g.

Example 3. Universality of the Uniform.
Let $U \sim \text{Unif}(0, 1)$, and let $X = -(\log(1 - U))^{1/3}$. Find the CDF and PDF of X.

Example 4. Continuous RV manipulation.
Let X be a continuous r.v. with CDF F and PDF f.

(a) Find the conditional CDF X given that $X > a$ (where a is a constant with $P(X > a) \neq 0$).

(b) Find the conditional PDF of X given $X > a$ (this is the derivative of the conditional CDF).

(c) Check that the conditional PDF from (b) is a valid PDF, by showing directly that it is nonnegative and integrates to 1.