

Continuous Probability, RVs, Distributions

EECS 126 Fall 2019

September 17, 2019

Agenda

Announcements

Review

Continuous Probability Definitions
Cumulative Distribution Functions

Distributions

Uniform
Exponential
Gaussian

Analogs to Discrete Probability / RVs

Derived Distributions

Announcements

- ▶ HW3 AND Lab2 are due Friday (9/20).
- ▶ Feel free to come to Lab Party with HW questions on Thursday!
- ▶ HW4 will be optional to give you more time to study. We still recommend reading and attempting the problems.
- ▶ Midterm 1 is coming up quick on 9/26! You can find past exams on the Exams page of the website.

Probability Densities

In a continuous space, we describe distributions with probability density functions (PDFs) rather than assigned probability values.

A valid probability density of a continuous random variable X in R , $f_X(x)$, requires

- ▶ Non-negativity: $\forall x \in R \ f_X(x) \geq 0$
- ▶ Normalized: $\int_R f_X(x) dx = 1$

Continuous Probability Definitions

Getting probabilities from densities:

- ▶ $P(X \in B) = \int_B f_X(x) dx$
- ▶ $P(X \in [a, b]) = P(a \leq X \leq b) = \int_a^b f_X(x) dx$
(Note: $P(X = a) = 0$, so open and closed intervals do not matter here)

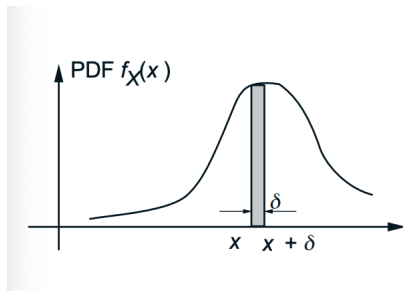


Figure: Geometric interpretation of the PDF

Questions

Suppose we uniformly sample a point in a ball of radius 1. What is the

- ▶ Probability of picking the origin?
- ▶ Probability density of picking the origin?
- ▶ Probability of picking a point on the surface?
- ▶ Probability of picking a point within a radius of $\frac{1}{2}$?

Answers

- ▶ Probability of picking the origin?
0.
- ▶ Probability density of picking the origin?
Volume of ball is $\frac{4}{3}\pi r^3 = \frac{4}{3}\pi$. Density is $\frac{3}{4\pi}$.
- ▶ Probability of picking a point on the surface?
0. A 2D surface has 0 volume in a 3D object.
- ▶ Probability of picking a point within a radius of $\frac{1}{2}$?
Since we're uniformly picking a point in the ball, we can just look at the ratio of the volumes. $\frac{\frac{4\pi}{3}(\frac{1}{2})^3}{\frac{4\pi}{3}} = \frac{1}{8}$.

Cumulative Distribution Functions (CDFs)

In both discrete and continuous distributions, the cumulative distribution is defined as $F_X(x) := P(X \leq x)$. However, they are computed slightly differently.

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

Consequently (by the Fundamental Theorem of Calculus),

$$f_X(x) = \frac{d}{dx} F_X(x)$$

More familiar definitions

Expectation:

- ▶ $E[X] := \int_{\mathcal{R}} xf_X(x)dx$
- ▶ $E[g(X)] := \int_{\mathcal{R}} g(x)f_X(x)dx$
- ▶ Linearity of expectation holds due to the linearity of integrals:
 $E[X + Y] = E[X] + E[Y]$

Variance stays the same

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

Questions

Let R be equal to the distance from the origin of a point randomly sampled on a unit ball. What is the

- ▶ CDF of R ?
- ▶ PDF of R ?
- ▶ Expectation of R ?

Answers

Let R be the distance from the origin of a point randomly sampled on a unit ball. What is the

- ▶ CDF of R ?

$$F_R(r) = \frac{3}{4\pi} \cdot \frac{4}{3}\pi r^3 = r^3.$$

- ▶ PDF of R ?

$$\frac{d}{dr} r^3 = 3r^2.$$

- ▶ Expectation of R ?

$$\int_0^1 r \cdot 3r^2 = \frac{3}{4}.$$

Uniform Distribution

The density is uniform across a bounded interval (a, b) . For $X \sim \text{Unif}(a, b)$

$$f_X(x) = \frac{1}{b-a}, \quad a < x < b$$

$$E[X] = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

Easy to work with distribution. Many problems can reduce to a uniform distribution!

Uniform Variance Proof

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

$$\begin{aligned} E[X] &= \int_a^b x \frac{1}{b-a} dx \\ &= \frac{x^2}{2(b-a)} \Big|_a^b \\ &= \frac{a+b}{2} \end{aligned}$$

$$\begin{aligned} E[X^2] &= \int_a^b x^2 \frac{1}{b-a} dx \\ &= \frac{x^3}{3(b-a)} \Big|_a^b \\ &= \frac{b^3 - a^3}{3(b-a)} \end{aligned}$$

$$\text{Var}(X) = \frac{b^3 - a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

Exponential Distribution

The exponential distribution PDF:

$$f_X(x) = \lambda e^{-\lambda x}, x > 0$$

The exponential distribution CDF:

$$F_X(x) = 1 - e^{-\lambda x}, x > 0$$

$$E[X] = \frac{1}{\lambda}, \text{Var}(X) = \frac{1}{\lambda^2}$$

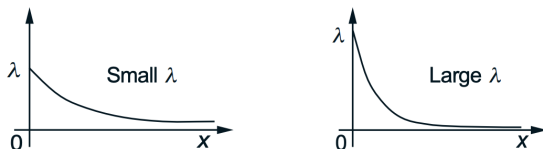


Figure: Exponential distribution for varying λ

Memoryless Property

The defining characteristic of the exponential is the memoryless property. Recall the memoryless property is:

$$P(X > x + a | X > x) = P(X > a)$$

Think about banging your head on the wall.

What distribution does this remind you of?

Connection to Geometric

One can think of the exponential distribution as the continuous analog to the geometric distribution.

Remark: These are the only distributions in discrete and continuous spaces respectively with the memoryless property!

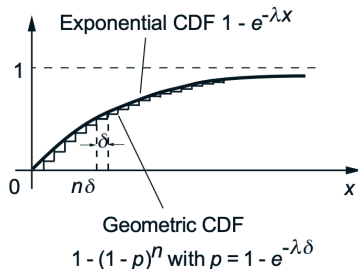


Figure: Relating the Exponential dist. to the Geometric dist.

Connection to Geometric cont.

Intuition that the geometric distribution approaches the exponential distribution as trials per second approaches infinity.

Let $X \sim \text{Geo}(p)$, $Y \sim \text{Expo}(\lambda)$. Recall the CDF of the geometric distribution

$$F_X(n) = 1 - (1 - p)^n$$

If we let $\delta = \frac{-\ln(1-p)}{\lambda}$, we have $e^{-\lambda\delta} = 1 - p$. Thus, $F_X(n) = F_Y(n\delta)$. If we drive δ down, we can interpret this as a geometric r.v. holding infinitely many trials per second while making sure that the expected number of trials passed stays the same. As $\delta \rightarrow 0$, we approach a continuous exponential distribution.

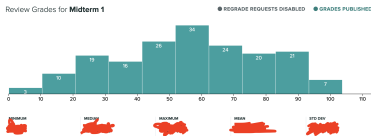
Normal / Gaussian Distribution

The Gaussian is seen abundantly in nature (e.g. exam scores). This can be explained by the Central Limit Theorem (CLT), which we will go over later in the course.

Gaussian PDF and CDF for mean μ and variance σ^2 :

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

$F_X(x) = \Phi(x)$, (cannot be expressed in elementary functions)



Properties of the Gaussian

- ▶ The sum of two independent Gaussians is Gaussian. If $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, and $Z = X + Y$, then

$$Z \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

- ▶ The sum of two dependent Gaussians isn't always Gaussian. Consider the following example.

$$X \sim N(0, 1)$$

$$Y = \begin{cases} X & \text{w.p. } \frac{1}{2} \\ -X & \text{w.p. } \frac{1}{2} \end{cases}$$

They are both Gaussian but $X + Y$ is not Gaussian.

- ▶ A Gaussian multiplied by a constant is Gaussian. If $X \sim N(\mu, \sigma^2)$ and $Y = aX$, then

$$Y \sim N(a \cdot \mu, a^2 \cdot \sigma^2)$$

Scaling to the Standard Gaussian

- ▶ The properties on the previous slide allow us to convert any Gaussian into the standard Gaussian.
- ▶ If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X - \mu}{\sigma}$$

is distributed with $Z \sim N(0, 1)$.

- ▶ Intuition: I got 1 SD on midterm 1.

Joint PDFs

Just how multiple discrete RVs have a joint PMF, multiple continuous RVs have a joint PDF.

- ▶ **Discrete**

$$p_{X,Y}(x,y)$$

- ▶ **Continuous**

$$f_{X,Y}(x,y)$$

- ▶ Still needs to be non-negative.
- ▶ Still needs to integrate to 1.

Joint CDFs

- ▶ **Single RV**

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

- ▶ **Multiple RVs**

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

- ▶ **Single RV**

$$\frac{d}{dx} F_X(x) = f_X(x)$$

- ▶ **Multiple RV**

$$\frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y) = f_{X,Y}(x, y)$$

Marginal Probability Density

- ▶ **Discrete**

$$p_X(x) = \sum_{y \in \mathcal{Y}} p_{X,Y}(x, y)$$

- ▶ **Continuous**

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

- ▶ $f_X(x)$ is still a density, not a probability.

Conditional Probability Density

- ▶ **Discrete**

$$p_{X|Y}(x | y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

- ▶ **Continuous**

$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

- ▶ By definition, Multiplication Rule still holds.

Independence

Similar to discrete, 3 equivalent definitions.

- ▶ For all x and y ,

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

- ▶ For all x and y ,

$$f_{X|Y}(x | y) = f_X(x)$$

- ▶ For all x and y ,

$$f_{Y|X}(y | x) = f_Y(y)$$

Bayes Rule

- ▶ **Discrete** (simple form)

$$p_{X|Y}(x | y) = \frac{p_{Y|X}(y | x)p_X(x)}{p_Y(y)}$$

- ▶ **Discrete** (extended form)

$$p_{X|Y}(x | y) = \frac{p_{Y|X}(y | x)p_X(x)}{\sum_{x' \in \mathcal{X}} p_{Y|X}(y | x')p_X(x')}$$

- ▶ **Continuous** (simple form)

$$f_{X|Y}(x | y) = \frac{f_{Y|X}(y | x)f_X(x)}{f_Y(y)}$$

- ▶ **Continuous** (extended form)

$$f_{X|Y}(x | y) = \frac{f_{Y|X}(y | x)f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y | t)p_X(t) dt}$$

Conditional Expectation

► **Discrete**

$$E[Y | X = x] = \sum_{y \in \mathcal{Y}} y \cdot p_{Y|X}(y | x)$$

► **Continuous**

$$E[Y | X = x] = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y | x) dy$$

Combining Discrete and Continuous RVs

- ▶ You can also have discrete and continuous RVs defined jointly.
- ▶ Ex. let X be the outcome of a dice roll and Y be $Exp(X)$.

$$p_X(x) = \frac{1}{6}$$
$$f_{Y|X}(y | x) = xe^{-xy}$$

Change of Variables / Derived Distributions

- ▶ Let $X \sim U[0, 1]$, and $Y = 2X$. Then is it true that

$$f_Y(y) = P(Y = y) = P(2X = y) = P(X = \frac{y}{2}) = f_X(\frac{y}{2})$$

- ▶ No, this won't integrate to 1.
- ▶ You have to use the CDF.

$$F_Y(y) = P(Y \leq y) = P(2X \leq y) = P(X \leq \frac{y}{2}) = F_X(\frac{y}{2})$$

- ▶

$$f_Y(y) = \frac{d}{dy} F_X(\frac{y}{2}) = f_X(\frac{y}{2}) \cdot \frac{1}{2}$$

References

Introduction to probability. DP Bertsekas, JN Tsitsiklis - 2002