Continuous Probability

CS70 Summer 2016 - Lecture 6A

David Dinh
25 July 2016
UC Berkeley

Logistics

Tutoring Sections - M/W 5-8PM in 540 Cory.

- Conceptual discussions of material
- No homework discussion (take that to OH/HW party, please)

Midterm is this Friday - 11:30-1:30, same rooms as last time.

- Covers material from MT1 to this Wednesday...
- ...but we will expect you to know everything we've covered from the start of class.
- One **double**-sided sheet of notes allowed (our advice: reuse sheet from MT1 and add MT2 topics to the other side).
- Students with time conflicts and DSP students should have been contacted by us - if you are one and you haven't heard from us, get in touch ASAP.

Today

- What is continuous probability?
- Expectation and variance in the continuous setting.
- Some common distributions.

Continuous Probability

Motivation I

Sometimes you can't model things discretely. Random real numbers. Points on a map. Time.

Probability space is **continuous**.

What is probability? Function mapping events to $[0, 1]$.

What is an event in continuous probability?

Motivation II

Class starts at 14:10. You take your seat at some "uniform" random time between 14:00 and 14:10.

What's an event here? Probability of coming in at exactly 14:03:47.32?

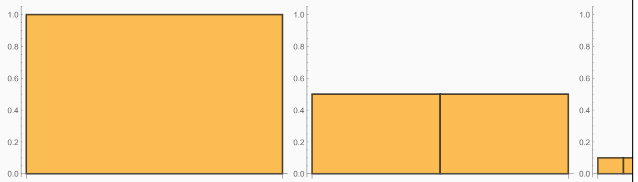
Sample space: all times between 14:00 and 14:10.

Size of sample space? How many numbers are there between 0 and 10? infinite

Chance of getting one event in an infinite sized uniform sample space? 0

Not so simple to define events in continuous probability!

Motivation III



Look at *intervals* instead of specific times.

Probability that you come in between 14:00 and 14:10? 1.

Probability that you come in between 14:00 and 14:05? 1/2.

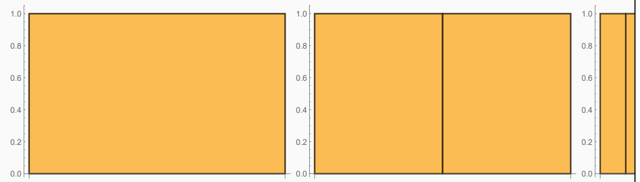
Probability that you come between 14:03 and 14:04? 1/10.

Probability that you come in some time interval of $10/k$ minutes? $1/k$.

PDF (no, not the file format)

What happens when you take $k \to \infty$? Probability goes to 0.

What do we do so that this doesn't disappear? If we split our sample space into k pieces - multiply each one by k.



The resulting curve as $k \to \infty$ is the **probability density function (PDF).**

Formally speaking...

PDF $f_X(t)$ of a random variable X is defined so that the probability of X taking on a value in $[t, t+\delta]$ is $\delta f(t)$ for infinitesimally small δ.

$$f_X(t) = \lim_{\delta \to 0} \frac{\Pr[X \in [t, t+\delta]]}{\delta}$$

Another way of looking at it:

$$\Pr[X \in [a,b]] = \int_a^b f_X(t)dt$$

f is nonnegative (negative probability doesn't make much sense).

Total probability is 1: $\int_{-\infty}^{\infty} f_X(t)dt = 1$

CDF

Cumulative distribution function (CDF): $F_X(t) = \Pr[X \le t]$.

Or, in terms of PDF...

$$F_X(t) = \int_{-\infty}^{t} F_X(z)dz$$

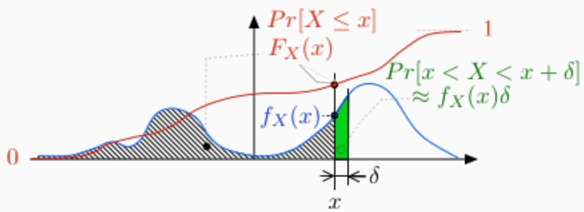
$$\begin{aligned} \Pr[X \in (a,b]] &= \Pr[X \le b] - \Pr[X \le a] \\ &= F_X(b) - F_X(a) \end{aligned}$$

$$F_X(t) \in [0,1]$$

$$\lim_{t \to -\infty} F_X(t) = 0$$

$$\lim_{t \to \infty} F_X(t) = 1$$

In Pictures



Expectation

Discrete case: $E[X] = \sum_{t=-\infty}^{\infty}(\Pr[X = t]t)$

Continuous case? Sum $\to$ integral.

$$E[X] = \int_{-\infty}^{\infty} tf_X(t)dt$$

Expectation of a function?

$$E[g(X)] = \int_{-\infty}^{\infty} g(t)f_X(t)dt$$

Linearity of expectation:

$$E[aX + bY] = aE[X] + bE[Y]$$

Proof: similar to discrete case.

If X, Y, Z are mutually independent, then $E[XYZ] = E[X]E[Y]E[Z]$.

Proof: also similar to discrete case.

Exercise: try proving these yourself.

Variance

Variance is defined exactly like it is for the discrete case.

$$\begin{aligned} Var(X) &= E[(X - E[X])^2] \\ &= E[X^2] - E[X]^2 \end{aligned}$$

The standard properties for variance hold in the continuous case as well.

$$Var(aX) = a^2 Var(X)$$

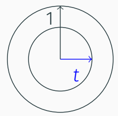
For independent r.v. X, Y:

$$Var(X + Y) = Var(X) + Var(Y)$$

.

Target shooting

Suppose an archer always hits a circular target with 1 meter radius, and the exact point that he hits is distributed uniformly across the target. What is the distribution the distance between his arrow and the center (call this r.v. X)?



Probability that arrow is closer than t to the center?

$$\begin{aligned} Pr[X \le t] &= \frac{\text{area of small circle}}{\text{area of dartboard}} \\ &= \frac{\pi t^2}{\pi} = t^2. \end{aligned}$$

Target shooting II

CDF:

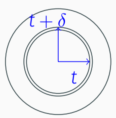
$$F_Y(t) = Pr[Y \le t] = \begin{cases} 0 & \text{for } t < 0 \\ t^2 & \text{for } 0 \le t \le 1 \\ 1 & \text{for } t > 1 \end{cases}$$

PDF?

$$f_Y(t) = F_Y(t)' = \begin{cases} 2t & \text{for } 0 \le t \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Target shooting III

Another way of attacking the same problem: what's the probability of hitting some ring with inner radius t and outer radius $t + \delta$ for small δ?



Area of circle: π

Area of ring:

$$\pi((t+\delta)^2 - t^2) = \pi(t^2 + 2t\delta + \delta^2 - t^2) = \pi(2t\delta + \delta^2) \approx \pi 2t\delta$$

Probability of hitting the ring: $2t\delta$.

PDF for $t \le 1$: $2t$

Shifting & Scaling

Let $f_X(x)$ be the pdf of X and $Y = a + bX$ where $b > 0$. Then

$$\begin{aligned} Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] \\ &= f_X(\frac{y-a}{b})\frac{\delta}{b}. \end{aligned}$$

Left-hand side is $f_Y(y)\delta$. Hence,

$$f_Y(y) = \frac{1}{b} f_X(\frac{y-a}{b}).$$

Continuous Distributions

Uniform Distribution: CDF and PDF

PDF is constant over some interval $[a, b]$, zero outside the interval.

What's the value of the constant in the interval?

$$\int_{-\infty}^{\infty} k dt = \int_a^b k dt = b - a = 1$$

so PDF is $1/(b-a)$ in $[a, b]$ and 0 otherwise.

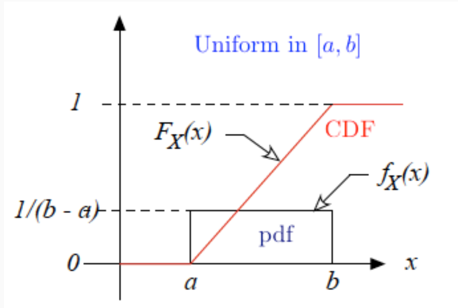
CDF?

$$\int_{-\infty}^{t} 1/(b-a) dz$$

0 for $t < a$, $(t-a)/(b-a)$ for $a < t < b$, and 1 for $t > b$.

Uniform Distribution: CDF and PDF, Graphically

$$f_X(t) = \begin{cases} 1/(b-a) & a < t < b \\ 0 & \text{otherwise} \end{cases} \qquad F_X(t) = \begin{cases} 0 & t < a \\ (t-a)/(b-a) & a < t < b \\ 1 & \text{b<t} \end{cases}$$



Uniform Distribution: Expectation and Variance

Expectation?

$$E[X] = \int_a^b \frac{t}{b-a} dt = \frac{1}{2}\frac{b^2 - a^2}{b-a} = \frac{b+a}{2}$$

Variance?

$$\begin{aligned} Var[X] &= E[X^2] - E[X]^2 \\ &= \int_a^b \frac{t^2}{b-a} dt - \left(\frac{b+a}{2}\right)^2 \\ &= \frac{t^3}{3(b-a)}\Big|_a^b - \left(\frac{b+a}{2}\right)^2 \\ &= \frac{(a-b)^2}{12} \end{aligned}$$

Exponential Distribution: Motivation

Continuous-time analogue of the geometric distribution.

How long until a server fails? How long does it take you to run into pokemon?

Can't "continuously flip a coin". What do we do?

Look at geometric distributions representing processes with higher and higher granularity.

Exponential Distribution: Motivation II

Suppose a server fails with probability λ every day.

Probability that server fails on the same day as time t:

$$(1-\lambda)^{\lceil t \rceil - 1}\lambda$$

More precision! What's the probability that it fails in a 12-hour period? $\lambda/2$ if we assume that there is no time that it's more likely to fail than another.

Generally: server fails with probability λ/n during any $1/n$-day time period.

Probability that server fails on the same $1/n$-day time period as t:

$$\left(1 - \frac{\lambda}{n}\right)^{\lceil tn \rceil - 1} \frac{\lambda}{n}$$

Exponential Distribution: Motivation III

$$\left(1 - \frac{\lambda}{n}\right)^{\lceil tn \rceil - 1} \frac{\lambda}{n}$$

What happens when we try to take n to ∞?

Probability goes to zero...but we can make a PDF out of this!

Remove the width of the interval $(1/n)$ and take the limit as $n \to \infty$ to get:

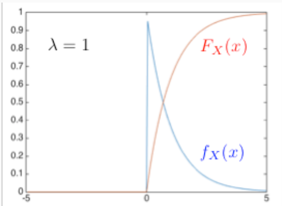
$$\begin{aligned} \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{\lceil tn \rceil - 1} \lambda &= \lambda \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{tn-1} \\ &= \lambda e^{-\lambda t} \end{aligned}$$

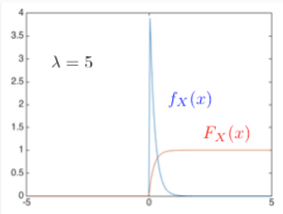
This is the PDF of the **exponential distribution**!

Exponential Distribution: PDF and CDF

The exponential distribution with parameter $\lambda > 0$ is defined by

$$f_X(t) = \begin{cases} 0, & \text{if } t < 0 \\ \lambda e^{-\lambda t}, & \text{if } t \geq 0. \end{cases} \qquad F_X(t) = \begin{cases} 0, & \text{if } t < 0 \\ 1 - e^{-\lambda t}, & \text{if } t \geq 0. \end{cases}$$





Note that $Pr[X > t] = e^{-\lambda t}$ for $t > 0$.

Expectation & Variance of the Exponential Distribution

$X = Expo(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x \lambda e^{-\lambda x} dx = -\int_0^\infty x de^{-\lambda x}.$$

Integration by parts:

$$\begin{aligned} \int_0^\infty x de^{-\lambda x} &= [xe^{-\lambda x}]_0^\infty - \int_0^\infty e^{-\lambda x} dx \\ &= 0 - 0 + \frac{1}{\lambda} \int_0^\infty de^{-\lambda x} = -\frac{1}{\lambda}. \end{aligned}$$

So: expectation is $E[X] = \frac{1}{\lambda}$.

Variance: $1/\lambda^2$

Properties of the Exponential Distribution: Memorylessness

Similar to memorylessness for geometric distributions.

"If your server doesn't fail today, it's in the same state as it was before today."

Let $X = Expo(\lambda)$. Then, for $s, t > 0$,

$$\begin{aligned} Pr[X > t + s \mid X > s] &= \frac{Pr[X > t + s]}{Pr[X > s]} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} \\ &= Pr[X > t]. \end{aligned}$$

Properties of the Exponential Distribution: Scaling

Let $X = Expo(\lambda)$ and $Y = aX$ for some $a > 0$. Then

$$\begin{aligned} Pr[Y > t] &= Pr[aX > t] = Pr[X > t/a] \\ &= e^{-\lambda(t/a)} = e^{-(\lambda/a)t} = Pr[Z > t] \text{ for } Z = Expo(\lambda/a). \end{aligned}$$

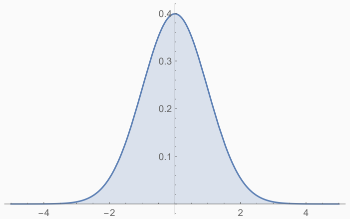
Thus, $a \times Expo(\lambda) = Expo(\lambda/a)$. Also, $Expo(\lambda) = \frac{1}{\lambda} Expo(1)$.

Normal Distribution

Continuous counterpart to Binomial dist. (more on this later)

Normal (or Gaussian) distribution with parameters μ, σ^2, denoted $\mathcal{N}(\mu, \sigma^2)$:

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



Sometimes called a "bell curve". Above: $\mathcal{N}(0, 1)$.

Normal Distribution: Properties

PDF: $f_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$

CDF: involves an integral with no nice closed form (often expressed in terms of "erf", the error function). Won't discuss it here.

Expectation: μ (notice that PDF is symmetric around μ).

Variance: σ^2 (fairly straightforward integration)

Scaling/Shifting: if $X \sim \mathcal{N}(0, 1)$ and $Y = \mu + \sigma X$, then $Y \sim \mathcal{N}(\mu, \sigma^2)$.

Central Limit Theorem

Basically: if you take a lot of i.i.d random variables from any* distribution with zero mean and the same variance and sum them up, the sum will converge to a random Gaussian with the same mean and variance.

Suppose $X_1, X_2, \ldots$ are i.i.d. random variables with expectation μ and variance σ^2. Let

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{(\sum_i X_i) - n\mu}{\sigma/\sqrt{n}}$$

Then S_n tends towards $\mathcal{N}(0,1)$ as $n \to \infty$.

Or:

$$\Pr[S_n \leq a] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx$$

Proof: EE126

Sum of Bernoullis (binomial) tends towards normal!

Summary

Continuous probability: translation of discrete probability to a continuous sample space with an infinite number of events.

Concepts of variance, expectation, etc. translate to continuous too.

Geometric distribution $\to$ exponential distribution.

Binomial distribution $\to$ normal distribution.

Central limit theorem: everything converges to normal if we take enough samples

Today's Gig: Cauchy Distribution

Cauchy

Augustin-Louis Cauchy (1789-1857)

Practically invented complex analysis. Made fundamental contributions to calculus and group theory.

"More concepts and theorems have been named for Cauchy than for any other mathematician."

Was also a baron because he tutored a duke... who ended up hating math.

Definition

Actually first written about by Poisson in 1824. Cauchy became associated with it in 1853!

Suppose I have a wall on the x-axis. Stand at (0,1) and point a laser at a uniform random angle such that the laser hits the wall.

What is the distribution of the point on the wall?

$$\tan\theta = t$$

$$\theta = \tan^{-1} t$$

$$d\theta = \frac{1}{1+t^2} dt$$

$$\frac{d\theta}{\pi} = \frac{1}{1+t^2} \frac{dt}{\pi}$$

Properties

PDF:

$$\frac{1}{\pi(1+t^2)}$$

Expectation?

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{t}{\pi(1+t^2)} dt &= \lim_{a\to\infty} \int_{-a}^{a} \tfrac{t}{\pi(1+t^2)} dt = 0 \\ &= \lim_{a\to\infty} \int_{-a}^{2a} \tfrac{t}{\pi(1+t^2)} dt \neq 0 \end{aligned}$$

Expectation doesn't exist!

If you try to estimate the expectation by sampling points and averaging, you'll get crazy results.

Variance doesn't exist either.

Main takeaway: there are some really badly-behaved distributions out there.

Questions?