Continuous RVs Continued: Independence, Conditioning, Gaussians, CLT

CS 70, Summer 2019

Lecture 25, 8/6/19

Not Too Different From Discrete...

Discrete RV:
X and Y are independent iff for all a, b:

$$\mathbb{P}[X = a, Y = b] = \mathbb{P}[X = a] \cdot \mathbb{P}[Y = b]$$

Continuous RV:
X and Y are independent iff for all $a \leq b$, $c \leq d$:

$$\mathbb{P}[a \leq X \leq b, c \leq Y \leq d] =$$

A Note on Independence

For continuous RVs, what is weird about the following?

$$\mathbb{P}[X = a, Y = b] = \mathbb{P}[X = a] \cdot \mathbb{P}[Y = b]$$

What we **can** do: consider a interval of length dx around a and b!

Independence, Continued

If X, Y are independent, their joint density is the product of their individual densities:

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

Example: If X, Y are independent exponential RVs with parameter λ:

Example: Max of Two Exponentials

Let $X \sim \text{Expo}(\lambda)$ and $Y \sim \text{Expo}(\mu)$.
X and Y are **independent**.
Compute $\mathbb{P}[\max(X, Y) \geq t]$.

Use this to compute $\mathbb{E}[\max(X, Y)]$.

Min of n Uniforms

Let $X_1, \ldots, X_n$ be **i.i.d.** and uniform over $[0, 1]$.

What is $\mathbb{P}[\min(X_1, \ldots, X_n) \leq x]$?

Use this to compute $\mathbb{E}[\min(X_1, \ldots, X_n)]$.

Min of n Uniforms

What is the CDF of $\min(X_1, \ldots, X_n)$?

What is the PDF of $\min(X_1, \ldots, X_n)$?

Memorylessness of Exponential

We can't talk about independence without talking about **conditional probability**!

Let $X \sim \text{Expo}(\lambda)$. X is **memoryless**, i.e.

$$\mathbb{P}[X \geq s + t | X > t] = \mathbb{P}[X \geq s]$$

Conditional Density

What happens if we condition on events like $X = a$? These have 0 probability!

The same story as discrete, except we now need to define a conditional **density**:

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

Think of $f(y|x)$ as

$$\mathbb{P}\left[Y \in [y, y + dy] | X \in [x, x + dx]\right]$$

Conditional Density, Continued

Given a conditional density $f_{Y|X}$, compute

$$\mathbb{P}[Y \leq y | X = x] =$$

If we know $\mathbb{P}[Y \leq y | X = x]$, compute

$$\mathbb{P}[Y \leq y] =$$

Go with your gut! What worked for discrete also works for continuous.

Example: Sum of Two Exponentials

Let X_1, X_2 be **i.i.d** $\text{Expo}(\lambda)$ RVs.
Let $Y = X_1 + X_2$.

What is $\mathbb{P}[Y < y | X_1 = x]$?

What is $\mathbb{P}[Y < y]$?

Example: Total Probability Rule

What is the CDF of Y?

What is the PDF of Y?

Break

If you could immediately gain one new skill, what would it be?

The Normal (Gaussian) Distribution

X is a **normal** or **Gaussian** RV if:

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

Parameters:

Notation: $X \sim$

$\mathbb{E}[X] =$ $\quad$ $\mathrm{Var}(X) =$

Standard Normal:

Gaussian Tail Bound

Let $X \sim \mathcal{N}(0, 1)$.
Easy upper bound on $\mathbb{P}[|X| \geq \alpha]$, for $\alpha \geq 1$?
(Something we've seen before...)

Gaussian Tail Bound, Continued

Turns out we can do better than Chebyshev.

Idea: Use $\int_\alpha^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \leq$

Shifting and Scaling Gaussians

Let $X \sim \mathcal{N}(\mu, \sigma)$ and $Y = \frac{X-\mu}{\sigma}$. Then:

$Y \sim$

Proof: Compute $\mathbb{P}[a \leq Y \leq b]$.

Change of variables: $x = \sigma y + \mu$.

Shifting and Scaling Gaussians

Can also go the other direction:

If $X \sim \mathcal{N}(0, 1)$, and $Y = \mu + \sigma X$:
Y **is still Gaussian!**

$\mathbb{E}[Y] =$

$\mathrm{Var}(Y) =$

Sum of Independent Gaussians

Let X, Y be **independent** standard Gaussians.

Let $Z = [aX + c] + [bY + d]$.
Then, Z is **also Gaussian!** (Proof optional.)

$\mathbb{E}[Z] =$

$\text{Var}(Z) =$

Example: Height

Consider a family of a two parents and twins with the same height. The parents' heights are independently drawn from a $\mathcal{N}(65, 5)$ distribution. The twins' height are independent of the parents', and from a $\mathcal{N}(40, 10)$ distribution.

Let H be the sum of the heights in the family.
Define relevant RVs:

Example: Height

$\mathbb{E}[H] =$

$\text{Var}[H] =$

Sample Mean

We sample a RV X **independently** n times.
X has mean μ, variance σ^2.

Denote the **sample mean** by $A_n = \frac{X_1+X_2+\ldots+X_n}{n}$

$\mathbb{E}[X] =$

$\text{Var}(X) =$

The Central Limit Theorem (CLT)

Let $X_1, X_2, \ldots, X_n$ be **i.i.d.** RVs with mean μ, variance σ^2. (Assume mean, variance, are finite.)

Sample mean, as before: $A_n = \frac{X_1+X_2+\ldots+X_n}{n}$
Recall: $\mathbb{E}[A_n] =$
$\text{Var}(A_n) =$

Normalize the sample mean:

$$A'_n =$$

Then, as $n \to \infty$, $\mathbb{P}[A'_n] \to$

Example: Chebyshev vs. CLT

Let $X_1, X_2, \ldots$ be **i.i.d** RVs with $\mathbb{E}[X_i] = 1$ and $\text{Var}(X_i) = \frac{1}{2}$. Let $A_n = \frac{X_1+X_2+\ldots+X_n}{n}$.

$\mathbb{E}[A_n] =$

$\text{Var}(A_n) =$

Normalize to get A'_n:

Example: Chebyshev vs. CLT

Upper bound $\mathbb{P}[A'_n \geq 2]$ for **any** n.
(We don't know if A'_n is **non-neg** or **symmetric**.)

If we take $n \to \infty$, upper bound on $\mathbb{P}[A'_n \geq 2]$?

Summary

- Independence and conditioning also generalize from the **discrete** RV case.
- The Gaussian is a very important continuous RV. It has several nice properties, including the fact that adding independent Gaussians gets you another Gaussian
- The CLT tells us that if we take a **sample average** of a RV, the distribution of this average will approach a **standard normal**.