17: Continuous Joint Distributions II

Jacob Mejia, Kanu Grover
February 16, 2024

Lecture Discussion on Ed
Convolution: Sum of independent Uniform RVs
Today’s lecture

Take what we’ve seen with discrete joint distributions...

...and generalize to continuous joint distributions!

For the most part, this is easy. For example:

Marginal distributions

\[ p_X(a) = \sum_y p_{X,Y}(a,y) \quad f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a,y) dy \]

Independent RVs

\[ p_{X,Y}(x,y) = p_X(x)p_Y(y) \quad f_{X,Y}(x,y) = f_X(x)f_Y(y) \]

But some concepts, while mathematically accessible, are difficult to implement in practice.

We’ll focus on some of these today.

Goal of CS109 continuous joint distributions unit: build mathematical maturity
Recall that for independent discrete random variables $X$ and $Y$:

$$P(X + Y = n) = \sum_{k} P(X = k)P(Y = n - k)$$

the convolution of $p_X$ and $p_Y$
Dance, Dance, Convolution

Recall that for independent discrete random variables $X$ and $Y$:

$$P(X + Y = n) = \sum_k P(X = k)P(Y = n - k)$$

the convolution of $p_X$ and $p_Y$

For independent continuous random variables $X$ and $Y$:

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x)f_Y(\alpha - x)dx$$

the convolution of $f_X$ and $f_Y$
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x) \, dx$$

Isn’t this just one??

Not so fast...
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x) \, dx$$

$f_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$

$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } 0 \leq \alpha - x \leq 1 \\ 0 & \text{otherwise} \end{cases}$

$\alpha$ is a constant in the integral w.r.t. $x$. 

Lisa Yan, Chris Piech, Mehran Sahami, and Jerry Cain, CS109, Winter 2024
Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0$

$X$ and $Y$ independent + continuous

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x) \, dx$$

$X$ and $Y$ independent + continuous

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \leq x \leq \alpha \\ 0 & \text{otherwise} \end{cases}$$
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0 \quad 0$

2. $\alpha = 1/2 \quad 1/2$

Integral = area under the curve
This curve = product of 2 functions of $x$
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0 \quad 0$

2. $\alpha = 1/2 \quad 1/2$

3. $\alpha = 1$

4. $\alpha = 3/2$

5. $\alpha \geq 2$

\[
\begin{align*}
X \text{ and } Y \text{ independent and continuous} \\
\int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x) \, dx
\end{align*}
\]

\[
f_X(x) = \begin{cases} 
1 & \text{if } 0 \leq x \leq 1 \\
0 & \text{otherwise} 
\end{cases}
\]

\[
f_Y(\alpha - x) = \begin{cases} 
1 & \text{if } \alpha - 1 \leq x \leq \alpha \\
0 & \text{otherwise} 
\end{cases}
\]
Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0 \quad 0$
2. $\alpha = 1/2 \quad 1/2$
3. $\alpha = 1 \quad 1$
4. $\alpha = 3/2$
5. $\alpha \geq 2$

\[ f_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \]

\[ f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \leq x \leq \alpha \\ 0 & \text{otherwise} \end{cases} \]
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0 \quad 0$

2. $\alpha = 1/2 \quad 1/2$

3. $\alpha = 1 \quad 1$

4. $\alpha = 3/2 \quad 1/2$

5. $\alpha \geq 2$

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0 \quad 0$

2. $\alpha = 1/2 \quad 1/2$

3. $\alpha = 1 \quad 1$

4. $\alpha = 3/2 \quad 1/2$

5. $\alpha \geq 2$
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0$  $0$
2. $\alpha = 1/2$  $1/2$
3. $\alpha = 1$  $1$
4. $\alpha = 3/2$  $1/2$
5. $\alpha \geq 2$  $0$
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \leq 0$  \hspace{1cm} 0
2. $\alpha = 1/2$  \hspace{1cm} 1/2
3. $\alpha = 1$  \hspace{1cm} 1
4. $\alpha = 3/2$  \hspace{1cm} 1/2
5. $\alpha \geq 2$  \hspace{1cm} 0

\[ f_{X+Y}(\alpha) = \begin{cases} \alpha & 0 \leq \alpha \leq 1 \\ 2 - \alpha & 1 \leq \alpha \leq 2 \\ 0 & \text{otherwise} \end{cases} \]
Dance, Dance, Convolution Extreme

\[ p(X = x) \]
\[ p(Y = y) \]
\[ + \]
\[ = \]

Independent \( X, Y \)

\[ f_X(x) \]
\[ + \]
\[ f_Y(y) \]
\[ = \]

Independent \( X, Y \)

\[ p(X + Y = n) \]

\[ f_{X+Y}(\alpha) \]
Ratio of PDFs
Relative probabilities of continuous random variables

Let $X = \text{time to finish problem set 4}$. Suppose $X \sim \mathcal{N}(10, 2)$.

How much more likely are you to complete in 10 hours than 5 hours?

$$\frac{P(X = 10)}{P(X = 5)} = \frac{f(10)}{f(5)}$$

A. $0/0 = \text{undefined}$
B. 2
C. $\frac{f(10)}{f(5)}$
D. $\frac{f(2)}{f(1)}$
Relative probabilities of continuous random variables

Let $X = \text{time to finish problem set 4}$. Suppose $X \sim \mathcal{N}(10, 2)$. How much more likely are you to complete in 10 hours than 5 hours?

\[
\frac{P(X = 10)}{P(X = 5)} =
\]

A. $0/0 = \text{undefined}$
B. 2
C. $\frac{f(10)}{f(5)}$
D. $\frac{f(2)}{f(1)}$
Relative probabilities of continuous random variables

Let $X = \text{time to finish problem set 4.}$
Suppose $X \sim \mathcal{N}(10,2)$.

How much more likely are you to complete in 10 hours than 5 hours?

$$
\frac{P(X = 10)}{P(X = 5)} = \frac{f(10)}{f(5)}
$$

Therefore

$$
P(X = a) = P \left( a - \frac{\varepsilon}{2} \leq X \leq a + \frac{\varepsilon}{2} \right) = \int_{a-\frac{\varepsilon}{2}}^{a+\frac{\varepsilon}{2}} f(x)dx \approx \varepsilon f(a)
$$

Ratios of PDFs are meaningful!

Stanford University
Continuous conditional distributions
Continuous conditional distributions

For continuous RVs $X$ and $Y$, the **conditional PDF** of $X$ given $Y$ is

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

where $f_Y(y) > 0$

Intuition: $P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \quad \leftrightarrow \quad f_{X|Y}(x|y) \varepsilon_x = \frac{f_{X,Y}(x,y) \varepsilon_x \varepsilon_y}{f_Y(y) \varepsilon_y}$

Note that conditional PDF $f_{X|Y}$ is a "true" density:

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y) dx = \int_{-\infty}^{\infty} \frac{f_{X,Y}(x,y)}{f_Y(y)} dx = \frac{f_Y(y)}{f_Y(y)} = 1$$
Why sums of random variables?

Sometimes modeling and understanding a complex RV, $X$, is difficult. But if we can decompose $X$ into the sum of independent simpler RVs,

- We can then compute distributions on $X$.
- We can then understand how $X$ changes as its parts change.

What can we model with a triangular PDF?

![Triangular PDF](image)

Sum of uniforms!

We’re covering the reverse direction for now; the forward direction will come on Friday.
**Everything**\(^*\) in probability is a sum or a product (or both)

\(^*\)except conditional probability (a ratio)

**Sum** of values that can be considered separately (possibly weighted by prob. of happening)

\[E[X] = \sum_x xp(x)\]

**Product** of values that can each be considered in sequence

\[P(E) = \sum_{i=1}^n P(E|F_i)P(F_i)\]

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

\[P(E) = \sum_{i=1}^n P(E_i)\]

Law of Total Probability

Axiom 3, \(E = E_1 \cup \cdots \cup E_n\)

\[P(E \cap F \cap G) = P(E)P(F|E)P(G|EF)\]

Chain Rule

\[f_{X,Y}(x,y) = f_X(x)f_Y(y)\]

Independent cont. RVs

\[P(X + Y = n) = \sum_k P(X = k)P(Y = n - k)\]

Sum of indep. discrete RVs (convolution)

Lisa Yan, Chris Piech, Mehran Sahami, and Jerry Cain, CS109, Winter 2024
Conditional probability and Bayes’ Theorem

Definition

\[ P(F|E) = \frac{P(E \cap F)}{P(E)} \]

Scaling to the correct sample space

Bayes’ Theorem

\[ P(F|E) = \frac{P(F)P(E|F)}{P(E)} \]

Prior: some prob. of event \( F \)

Likelihood

Posterior: prob. of \( F \) knowing that \( E \) happened

Scaling to the correct sample space

Independence

\( E, F \) independent

\[ P(F|E) = P(F) \]

Sample space doesn’t need to be scaled
Multiple Bayes’ Theorems

with events

\[ P(F|E) = \frac{P(F)P(E|F)}{P(E)} \]

with discrete RVs

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

You are given this value...

with continuous RVs

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

...so this is just a scalar

Really all the same idea!
Intense Exercise
You want to know the 2-D location of an object.

Your satellite ping gives you a noisy 1-D measurement of the distance of the object from the satellite (0,0).

Using the satellite measurement, where is the object?
Before measuring, we have some prior belief about the 2-D location of an object, $(X, Y)$.

We observe some noisy measurement $D = 4$, the Euclidean distance of the object to a satellite.

After the measurement, what is our updated (posterior) belief of the 2-D location of the object?
Tracking in 2-D space

- You have a **prior belief** about the 2-D location of an object, \((X, Y)\).
- You observe a **noisy distance measurement**, \(D = 4\).
- What is your **updated (posterior) belief** of the 2-D location of the object after observing the measurement?

Recall Bayes terminology:

\[
f_{X,Y|D}(x, y|d) = \frac{f_{D|X,Y}(d|x, y)f_{X,Y}(x, y)}{f_D(d)}
\]

- **Posterior belief**
- **Likelihood (of evidence)**
- **Prior belief**
- **Normalization constant**
1. Define prior

You have a prior belief about the 2-D location of an object, \((X, Y)\).

Let \((X, Y) = \text{object’s 2-D location, assuming satellite is at (0,0)}\)

Suppose the prior distribution is a symmetric bivariate normal distribution:

\[
    f_{X,Y}(x, y) = \frac{1}{2\pi 2^2} e^{-\frac{[(x-3)^2+(y-3)^2]}{2(2^2)}} = K_1 \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}
\]

normalizing constant
2. Define likelihood

You observe a noisy distance measurement, $D = 4$.

If you knew your actual location $(x, y)$, you could say how likely a measurement $D = 4$ is:

Let $D =$ distance from the satellite (radially).
Suppose you knew your actual position: $(x, y)$.
- $D$ is still noisy! Suppose noise is standard normal.
- On average, $D$ is your true Euclidean distance: $\sqrt{x^2 + y^2}$
2. Define likelihood

You observe a noisy distance measurement, \( D = 4 \).

If you knew your actual location \((x, y)\), you could say how likely a measurement \( D = 4 \) is:

Let \( D \) = distance from the satellite (radially).
Suppose you knew your actual position: \((x, y)\).
- \( D \) is still noisy! Suppose noise is standard normal.
- On average, \( D \) is your true Euclidean distance: \( \sqrt{x^2 + y^2} \)

\[
D|X,Y \sim N(\mu = (A), \sigma^2 = (B))
\]

\[
f_{D|X,Y}(D = d|X = x, Y = y) = \frac{1}{(C) \sqrt{2\pi}} e^{-\frac{(D)}{2}}
\]
2. Define likelihood

You observe a noisy distance measurement, $D = 4$. If you knew your actual location $(x, y)$, you could say how likely a measurement $D = 4$ is:

Let $D$ = distance from the satellite (radially). Suppose you knew your actual position: $(x, y)$.
- $D$ is still noisy! Suppose noise is standard normal.
- On average, $D$ is your true Euclidean distance: $\sqrt{x^2 + y^2}$

$$D|X, Y \sim N \left( \mu = \sqrt{x^2 + y^2}, \sigma^2 = 1 \right)$$

$$f_{D|X,Y}(D = d|X = x, Y = y) = \frac{1}{\sqrt{2\pi}} e^{-\left(d-\sqrt{x^2+y^2}\right)^2}$$

$$= K_2 \cdot e^{-\left(d-\sqrt{x^2+y^2}\right)^2}$$

normalizing constant
3. Compute posterior

What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

Compute:

Posterior belief

\[
f_{X,Y|D}(x, y|4) = f_{X,Y|D}(X = x, Y = y|D = 4)
\]
3. Compute posterior

What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

Compute:

\[ f_{X,Y|D}(x,y|4) = f_{X,Y|D}(X = x, Y = y|D = 4) \]

Know:

Prior belief \[ f_{X,Y}(x,y) = K_1 \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}} \]

Observation likelihood \[ f_{D|X,Y}(d|x,y) = K_2 \cdot e^{-\frac{(d-\sqrt{x^2+y^2})^2}{2}} \]

Tips
- Use Bayes’ Theorem!
- \( f_D(4) \) is just a scaling constant. Why?
- How can we approximate the final scaling constant with a computer?
Tracking in 2-D space

What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

\[
f_{X,Y|D}(X = x, Y = y|D = 4) = \frac{f_{D|X,Y}(D = 4|X = x, Y = y)f_{X,Y}(x, y)}{f(D = 4)}
\]

Key: Once we know the part dependent on \(x, y\), we can computationally approximate \(K_4\) so that \(f_{X,Y|D}\) is a valid PDF.
Tracking in 2-D space

With this continuous version of Bayes’ theorem, we can explore new domains.

• Before measuring, we have some prior belief about the 2-D location of an object, \((X, Y)\).

• We observe some noisy measurement of the distance of the object to a satellite.

• After the measurement, what is our updated (posterior) belief of the 2-D location of the object?
Tracking in 2-D space: Posterior belief

\[ f_{X,Y}(x, y) = K_1 \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}} \]

\[ f_{X,Y|D}(x, y|4) = K_4 \cdot e^{-\left[\frac{(4-\sqrt{x^2 + y^2})^2}{2} + \frac{(x-3)^2 + (y-3)^2}{8}\right]} \]
How’d you compute that $K_4$?

To be a valid conditional PDF, \[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y|D}(x, y|4) \, dx \, dy = 1
\]

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_4 \cdot e^{-\left[\frac{(4-\sqrt{x^2+y^2})^2}{2} + \frac{(x-3)^2+(y-3)^2}{8}\right]} \, dx \, dy = 1
\]

\[
\frac{1}{K_4} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\left[\frac{(4-\sqrt{x^2+y^2})^2}{2} + \frac{(x-3)^2+(y-3)^2}{8}\right]} \, dx \, dy
\]

(pull out $K_4$, divide)

Approximate:

\[
\frac{1}{K_4} \approx \sum_{y} \sum_{x} e^{-\left[\frac{(4-\sqrt{x^2+y^2})^2}{2} + \frac{(x-3)^2+(y-3)^2}{8}\right]} \Delta x \Delta y
\]

Use a computer!