13: Independent RVs

Lisa Yan
October 21, 2019
Probabilities from joint CDFs

Joint CDF: \( P(X \leq x, Y \leq y) = F_{X,Y}(x, y) \)

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Probabilities from joint CDFs

Joint CDF: \( P(X \leq x, Y \leq y) = F_{X,Y}(x, y) \)

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Probabilities from joint CDFs

Joint CDF: \( P(X \leq x, Y \leq y) = F_{X,Y}(x, y) \)

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) =
F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Probabilities from joint CDFs

Joint CDF: \( P(X \leq x, Y \leq y) = F_{X,Y}(x, y) \)

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \]

\[
F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Probabilities from joint CDFs

Joint CDF: \( P(X \leq x, Y \leq y) = F_{X,Y}(x,y) \)

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Probabilities from joint CDFs

Joint CDF: \( P(X \leq x, Y \leq y) = F_{X,Y}(x, y) \)

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Probabilities from joint CDFs

Joint CDF: \( P(X \leq x, Y \leq y) = F_{X,Y}(x, y) \)

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Probability with Instagram!

In image processing, a Gaussian blur is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise.

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)
\]
Gaussian blur

In a Gaussian blur, for every pixel:

- Weight each pixel by the probability that $X$ and $Y$ are both within the pixel bounds
- The weighting function is a Gaussian joint PDF with a standard deviation parameter $\sigma$.

### Gaussian blurring with $\sigma = 3$

**Joint PDF:**

$$f_{X,Y}(x, y) = \frac{1}{2\pi \cdot 3^2} e^{-(x^2+y^2)/2\cdot3^2}$$

**Joint CDF:**

$$F_{X,Y}(x, y) = \Phi \left( \frac{x}{3} \right) \Phi \left( \frac{y}{3} \right)$$

**Weight matrix:**

Center pixel: (0, 0)

Pixel bounds:

- $-0.5 < x \leq 0.5$
- $-0.5 < y \leq 0.5$
Gaussian blur

In a Gaussian blur:
• Weight each pixel by the probability that \( X \) and \( Y \) are both within the pixel bounds

What is the weight of the center pixel?

\[
P(-0.5 < X \leq 0.5, -0.5 < Y \leq 0.5) = F_{X,Y}(0.5,0.5) - F_{X,Y}(-0.5, 0.5) \\
- F_{X,Y}(0.5, -0.5) + F_{X,Y}(-0.5, -0.5) \\
= \Phi \left( \frac{0.5}{3} \right) \Phi \left( \frac{0.5}{3} \right) - 2 \cdot \Phi \left( \frac{-0.5}{3} \right) \Phi \left( \frac{0.5}{3} \right) \\
+ \Phi \left( \frac{-0.5}{3} \right) \Phi \left( \frac{-0.5}{3} \right) \\
\approx 0.5662^2 - 2 \cdot 0.5662 \cdot 0.4338 + 0.4338^2 \\
\approx 0.206
\]
Multiple events:

- Intersection:
  \[ P(E \cap F) = P(EF) \]

- Conditional probability:
  \[ P(E|F) = \frac{P(EF)}{P(F)} \]

- Independence:
  \[ P(EF) = P(E)P(F) \]

Joint \textbf{(Multivariate)} distributions

- Joint PMF/PDF:
  \[ p_{x,y}(x,y) \quad f_{x,y}(x,y) \]

- Conditional distributions:
  Yes! \textbf{(Wednesday)}

- Independent RVs:
  Yes! \textbf{(today)}

Model ALL the things!
Today’s plan (covered on midterm)

Independent RVs

Sum of independent RVs
- ✅ Binomial
- ✅ Convolution
- ✅ Poisson
- ✅ Normal
- ⚠️ Uniform

Expectation of sum of RVs (next class)
Independent discrete RVs

Recall the definition of independent events $E$ and $F$:

$$P(EF) = P(E)P(F)$$

Two discrete random variables $X$ and $Y$ are independent if:

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

Different notation, same idea:

$$p_{X,Y}(x,y) = p_X(x)p_Y(y)$$

Intuitively: knowing value of $X$ tells us nothing about the distribution of $Y$ (and vice versa)

If two variables are not independent, they are called dependent.
Dice (after all this time, still our friends)  

Let: \( D_1 \) and \( D_2 \) be the outcomes of two rolls  
\[ S = D_1 + D_2, \]  
the sum of two rolls  

• Each roll of a 6-sided die is an independent trial.  
• \( D_1 \) and \( D_2 \) are independent.  

Are \( S \) and \( D_1 \) independent?  

1. \( P(D_1 = 1, S = 7)? \)  
2. \( P(D_1 = 1, S = 5)? \)
Dice (after all this time, still our friends)

Let: \( D_1 \) and \( D_2 \) be the outcomes of two rolls
\[ S = D_1 + D_2, \] the sum of two rolls

- Each roll of a 6-sided die is an independent trial.
- \( D_1 \) and \( D_2 \) are independent.

Are \( S \) and \( D_1 \) independent?

1. \( P(D_1 = 1, S = 7) \)?
   
   Event \((S = 7)\): \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}
   
   \[
P(D_1 = 1)P(S = 7) = (1/6)(1/6) = 1/36 = P(D_1 = 1, S = 7)
   
   \boxed{\text{Independent events } (D_1 = 1),(S = 7) \}

2. \( P(D_1 = 1, S = 5) \)?
   
   Event \((S = 5)\): \{(1,4), (2,3), (3,2), (4,1)\}
   
   \[
P(D_1 = 1)P(S = 5) = (1/6)(4/36) \neq 1/36 = P(D_1 = 1, S = 5)
   
   \boxed{\text{Dependent events } (D_1 = 1),(S = 5) \}

All events \((X = x, Y = y)\) must be independent for \( X, Y \) to be independent random variables.
Coin flips

Flip a coin with probability \( p \) of “heads” a total of \( n + m \) times.

Let

\[ X = \text{number of heads in first } n \text{ flips. } X \sim \text{Bin}(n, p) \]
\[ Y = \text{number of heads in next } m \text{ flips. } Y \sim \text{Bin}(m, p) \]
\[ Z = \text{total number of heads in } n + m \text{ flips.} \]

1. Are \( X \) and \( Z \) independent? ❌

Counterexample: What if \( Z = 0 \)?
Coin flips

Flip a coin with probability $p$ of “heads” a total of $n + m$ times.

Let $X =$ number of heads in first $n$ flips. $X \sim \text{Bin}(n, p)$
$Y =$ number of heads in next $m$ flips. $Y \sim \text{Bin}(m, p)$
$Z =$ total number of heads in $n + m$ flips.

1. Are $X$ and $Z$ independent? X  
   Counterexample: What if $Z = 0$?

2. Are $X$ and $Y$ independent?

   Strategy:
   A. No, proof by counterexample
   B. Yes, proof by counting
   C. None/other
Coin flips

Flip a coin with probability $p$ of “heads” a total of $n + m$ times.

Let

- $X =$ number of heads in first $n$ flips. $X \sim \text{Bin}(n, p)$
- $Y =$ number of heads in next $m$ flips. $Y \sim \text{Bin}(m, p)$
- $Z =$ total number of heads in $n + m$ flips.

1. Are $X$ and $Z$ independent? $\times$

Counterexample: What if $Z = 0$?

2. Are $X$ and $Y$ independent? $✓$

Strategy:

- A. No, proof by counterexample
- B. Yes, proof by counting
- C. None/other
Coin flips

Flip a coin with probability $p$ of “heads” a total of $n + m$ times.

Let $X =$ number of heads in first $n$ flips. $X \sim \text{Bin}(n, p)$

$Y =$ number of heads in next $m$ flips. $Y \sim \text{Bin}(m, p)$

$Z =$ total number of heads in $n + m$ flips.

1. Are $X$ and $Z$ independent? $\times$

2. Are $X$ and $Y$ independent?

$P(X = x, Y = y) = P\left(\begin{array}{c}
\text{first } n \text{ flips have } x \text{ heads} \\
\text{and next } m \text{ flips have } y \text{ heads}
\end{array}\right)

= \binom{n}{x} p^x (1 - p)^{n-x} \binom{m}{y} p^y (1 - p)^{m-y}

= P(X = x)P(Y = y)$

Counterexample: What if $Z = 0$?
Independent continuous RVs

Two continuous random variables $X$ and $Y$ are independent if:

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

Equivalently:

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

More generally, $X$ and $Y$ are independent if joint density factors separately:

$$f_{X,Y}(x, y) = h(x)g(y), \text{ where } -\infty < x, y < \infty$$
Is the Gaussian blur distribution independent?

Gaussian blurring with $\sigma = 3$

Joint PDF:

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

Joint CDF:

$$F_{X,Y}(x, y) = \Phi\left(\frac{x}{3}\right) \Phi\left(\frac{y}{3}\right)$$

Weight matrix:

Center pixel: (0, 0)

Pixel bounds:

$$-0.5 < x \leq 0.5$$

$$-0.5 < y \leq 0.5$$
Independent continuous RVs

Two continuous random variables $X$ and $Y$ are independent if:

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

Equivalently:

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

More generally, $X$ and $Y$ are independent if joint density factors separately:

$$f_{X,Y}(x, y) = g(x)h(y), \text{ where } -\infty < x, y < \infty$$
Pop quiz! (just kidding)

Are $X$ and $Y$ independent in the following cases?

1. $f_{X,Y}(x, y) = 6e^{-3x}e^{-2y}$
   where $0 < x, y < \infty$

2. $f_{X,Y}(x, y) = 4xy$
   where $0 < x, y < 1$

3. $f_{X,Y}(x, y) = 4xy$
   where $0 < x + y < 1$
Pop quiz! (just kidding)

Are $X$ and $Y$ independent in the following cases?

1. $f_{X,Y}(x, y) = 6e^{-3x}e^{-2y}$
   where $0 < x, y < \infty$
   Separable functions: $g(x) = 3e^{-3x}$
   $h(y) = 2e^{-2y}$

2. $f_{X,Y}(x, y) = 4xy$
   where $0 < x, y < 1$
   Separable functions: $g(x) = 2x$
   $h(y) = 2y$

3. $f_{X,Y}(x, y) = 4xy$
   where $0 < x + y < 1$
   Cannot capture constraint on $x + y$
   into factorization!

👉 If you can factor densities over all of the support, you have independence.
Break for jokes/announcements
Announcements

Midterm exam
When: Tuesday, October 29th, 7:00pm-9:00pm
Where: Hewlett 200
Covers: Up to (and including) week 4 + Lecture Notes #13
Practice: http://web.stanford.edu/class/cs109/exams/midterm.html
Review session: Saturday, 10am-12pm, Shiram 104

Problem Set 4
Out: later today
Due: Wednesday 11/6
Midterm coverage: First half (marked)

Concept checks
Week 4’s: Tuesday 10/22 1pm
Week 5’s: Wednesday 10/31 1pm
Today’s plan

Independent RVs

- Binomial
- Convolution
- Poisson
- Normal
- Uniform

Sum of independent RVs

Expectation of sum of RVs (next class)
Sum of independent Binomials

Intuition:

• Each trial in $X$ and $Y$ is independent and has same success probability $p$
• Define $Z = n_1 + n_2$ independent trials, each with success probability $p$
  $Z \sim \text{Bin}(n_1 + n_2, p)$, and also $Z = X + Y$

Holds in general case:

$X_i \sim \text{Bin}(n_i, p)$
$X_i$ independent for $i = 1, \ldots, n$

$\sum_{i=1}^{n} X_i \sim \text{Bin}(\sum_{i=1}^{n} n_i, p)$
Convolution: Sum of independent random variables

For any discrete random variables $X$ and $Y$:

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

In particular, 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 $p_Y$
Insight into convolution

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

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

Suppose $X$ and $Y$ are independent, both with support \{0, 1, ...\}:

<table>
<thead>
<tr>
<th>$X = k$</th>
<th>$Y = n - k$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$n$</td>
<td>$P(X = 0)P(Y = n)$</td>
</tr>
<tr>
<td>1</td>
<td>$n - 1$</td>
<td>$P(X = 1)P(Y = n - 1)$</td>
</tr>
<tr>
<td>2</td>
<td>$n - 2$</td>
<td>$P(X = 2)P(Y = n - 2)$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$n$</td>
<td>0</td>
<td>$P(X = n)P(Y = 0)$</td>
</tr>
<tr>
<td>$n + 1$</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

Sum of mutually exclusive events
The distribution of a sum of dice rolls is a convolution.

Note for $k, n - k$ in the support,

$$P(X = k, Y = n - k) = P(X = k)P(Y = n - k) = 1/36$$
Today’s plan

Independent RVs

- Binomial
- Convolution
- Poisson
- Normal
- Uniform

Sum of independent RVs

- ✅ Binomial
- 🚫 Poisson
- 🚫 Normal
- 🚫 Uniform

Expectation of sum of RVs (next class)
Sum of independent Poissons

\[ X \sim \text{Poi}(\lambda_1), \ Y \sim \text{Poi}(\lambda_2) \]

\[ X, Y \text{ independent} \]

\[ X + Y \sim \text{Poi}(\lambda_1 + \lambda_2) \]

Proof (just for reference):

\[
P(X + Y = n) = \sum_{k=0}^{n} P(X = k)P(Y = n - k)
\]

\[
= \sum_{k=0}^{n} e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} = e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^{n} \frac{\lambda_1^k \lambda_2^{n-k}}{k!(n-k)!}
\]

\[
= e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{\lambda_1^k \lambda_2^{n-k}}{n!} = \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n
\]

\[ \text{PMF of Poisson RVs} \]

\[ X \text{ and } Y \text{ independent, convolution} \]

\[ \text{Binomial Theorem:} \]

\[ (a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \]
General sum of independent Poissons

Holds in general case:

\[ X_i \sim \text{Poi}(\lambda_i) \]

\[ X_i \text{ independent for } i = 1, \ldots, n \]

\[ \sum_{i=1}^{n} X_i \sim \text{Poi} \left( \sum_{i=1}^{n} \lambda_i \right) \]
Sum of independent Gaussians

\[ X \sim \mathcal{N}(\mu_1, \sigma_1^2), \]
\[ Y \sim \mathcal{N}(\mu_2, \sigma_2^2) \]
\[ X, Y \text{ independent} \quad \Rightarrow \quad X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \]

(proof left to Wikipedia)

Holds in general case:

\[ X_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \]
\[ X_i \text{ independent for } i = 1, \ldots, n \]
\[ \sum_{i=1}^{n} X_i \sim \mathcal{N} \left( \sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2 \right) \]
Virus infections

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected } \geq 55)$?

1. Define RVs & state goal

Let $A = \#$ infected in G1.

$A \sim \text{Bin}(200, 0.1)$

$B = \#$ infected in G2.

$B \sim \text{Bin}(100, 0.4)$

Want: $P(A + B \geq 55)$

Strategy:

A. Dance, Dance, Convolution
B. Sum of indep. Binomials
C. (approximate) Sum of indep. Poissons
D. (approximate) Sum of indep. Normals
E. None/other
Virus infections

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected} \geq 55)$?

1. Define RVs & state goal

Let $A = \# \text{ infected in G1.}$

$A \sim \text{Bin}(200, 0.1)$

$B = \# \text{ infected in G2.}$

$B \sim \text{Bin}(100, 0.4)$

Want: $P(A + B \geq 55)$

Strategy:

A. Dance, Dance, Convolution
B. Sum of indep. Binomials
C. (approximate) Sum of indep. Poissons
D. (approximate) Sum of indep. Normals
E. None/other
Virus infections

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected } \geq 55)$?

1. Define RVs & state goal

   Let $A = \#$ infected in G1.
   
   $A \sim \text{Bin}(200, 0.1)$
   
   $B = \#$ infected in G2.
   
   $B \sim \text{Bin}(100, 0.4)$

   Want: $P(A + B \geq 55)$

2. Approximate as sum of Normals

   $A \approx X \sim \mathcal{N}(20, 18)$
   
   $B \approx Y \sim \mathcal{N}(40, 24)$

   $P(A + B \geq 55) \approx P(X + Y \geq 54.5)$

   continuity correction

3. Solve
Virus infections

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected } \geq 55)$?

1. Define RVs & state goal

   Let $A = \#\text{ infected in G1.}$
   $A \sim \text{Bin}(200,0.1)$
   $B = \#\text{ infected in G2.}$
   $B \sim \text{Bin}(100,0.4)$

   Want: $P(A + B \geq 55)$

2. Approximate as sum of Normals

   $A \approx X \sim \mathcal{N}(20,18)$
   $B \approx Y \sim \mathcal{N}(40,24)$

   $P(A + B \geq 55) \approx P(X + Y \geq 54.5)$

   continuity correction

3. Solve

   Let $W = X + Y \sim \mathcal{N}(20 + 40 = 60, 18 + 24 = 42)$

   $P(W \geq 54.5) = 1 - \Phi \left( \frac{54.5 - 60}{\sqrt{42}} \right) \approx 1 - \Phi(-0.85)$

   $\approx 0.8023$
Linear transforms vs. independence

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = X + X$. What is the distribution of $Y$?

- Are both approaches valid?

### Independent RVs approach

Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ be independent.

Then $Y = X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

### Linear transform approach

Let $X \sim \mathcal{N}(\mu, \sigma^2)$.

If $Y = aX + b$,
then $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$. 

Lisa Yan, CS109, 2019
Linear transforms vs. independence

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = X + X$. What is the distribution of $Y$?

- Are both approaches valid?

Independent RVs approach \[ \times \]

Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ be independent.

Then $Y = X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

$Y = X + X$

$X + X \sim \mathcal{N}(\mu + \mu, \sigma^2 + \sigma^2)$ \[ \text{is NOT independent of } X! \]

$Y \sim \mathcal{N}(2\mu, 2\sigma^2)$?

Linear transform approach \[ \checkmark \]

Let $X \sim \mathcal{N}(\mu, \sigma^2)$.

If $Y = aX + b$,

then $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

$x = 2X$

$Y \sim \mathcal{N}(2\mu, 4\sigma^2)$
Motivating idea: Zero sum games

Want: \( P(\text{Warriors win}) = P(A_W > A_B) = P(A_W - A_B > 0) \)

Assume \( A_W, A_B \) are independent. Let \( D = A_W - A_B \).

What is the distribution of \( D \)?

A. \( D \sim \mathcal{N}(1657 - 1470, 200^2 - 200^2) \)
B. \( D \sim \mathcal{N}(1657 - 1470, 200^2 + 200^2) \)
C. \( D \sim \mathcal{N}(1657 + 1470, 200^2 + 200^2) \)
D. Dance, Dance, Convolution
E. None/other
Motivating idea: Zero sum games

Want: \( P(\text{Warriors win}) = P(A_W > A_B) \)
\[= P(A_W - A_B > 0) \]

Assume \( A_W, A_B \) are independent.
Let \( D = A_W - A_B \).

What is the distribution of \( D \)?

A. \( D \sim \mathcal{N}(1657 - 1470, 200^2 - 200^2) \)
B. \( D \sim \mathcal{N}(1657 - 1470, 200^2 + 200^2) \)
C. \( D \sim \mathcal{N}(1657 + 1470, 200^2 + 200^2) \)
D. Dance, Dance, Convolution
E. None/other

If \( X \sim \mathcal{N}(\mu_1, \sigma^2) \),
then \((-X) \sim \mathcal{N}(-\mu, (-1)^2 \sigma^2 = \sigma^2) \)
Motivating idea: Zero sum games

Want: \( P(\text{Warriors win}) = P(A_W > A_B) \)
\[
= P(A_W - A_B > 0)
\]

Assume \( A_W, A_B \) are independent.
Let \( D = A_W - A_B \).
\[
D \sim \mathcal{N}(1657 - 1470, 200^2 + 200^2)
\sim \mathcal{N}(187, 2 \cdot 200^2) \quad \sigma \approx 283
\]

\[
P(D > 0) = 1 - F_D(0) = 1 - \Phi\left(\frac{0 - 187}{283}\right)
\approx 0.7454
\]

Compare with 0.7488, calculated by sampling!
Today’s plan

Independent RVs

Sum of independent RVs

- ✔️ Binomial
- ✔️ Convolution
- ✔️ Poisson
- ✔️ Normal
- ⚠️ Uniform

Expectation of sum of RVs (next class)
Dance, Dance, Convolution Extreme

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 random variables. What is the distribution of \( X + Y \), \( f_{X+Y} \)?

\[
f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(k) f_Y(\alpha - k) dk
\]

\( f_X(k) f_Y(\alpha - k) = 1 \) when: (select one)

A. between 0 and 1
B. \( 0 \leq k \leq 1 \)
C. \( 0 \leq \alpha - k \leq 1 \)
D. \( 0 \leq \alpha \leq 2 \)
E. Other
Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent random variables.

What is the distribution of $X + Y$, $f_{X+Y}$?

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

$$f_X(k) f_Y(\alpha - k) = 1:$$

- $0 \leq \alpha \leq 2$
- $0 \leq k \leq 1$
- $0 \leq \alpha - k \leq 1$
- $\alpha - 1 \leq k \leq \alpha$

The precise integration bounds on $k$ depend on $\alpha$.

What are the bounds on $k$ when:

1. $\alpha = 1/2$?
   
   $0 \leq k \leq \alpha$
   
   $\int_{k=0}^{\alpha} 1 dk = \alpha = 1/2$

2. $\alpha = 3/2$?
   
   $\alpha - 1 \leq k \leq 1$
   
   $\int_{k=\alpha-1}^{1} 1 dk = \alpha - 1 = 1/2$

3. $\alpha = 1$?
   
   $0 \leq k \leq \alpha$
   
   $\int_{k=0}^{\alpha} 1 dk = \alpha = 1$
   
   (the other bound works too)
Sum of independent Uniforms

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

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

$f_X(k) f_Y(\alpha - k) = 1$ when:

- $0 \leq \alpha \leq 2$
- $0 \leq k \leq 1$
- $0 \leq \alpha - k \leq 1$
- $\alpha - 1 \leq k \leq \alpha$

The precise integration bounds on $k$ depend on $\alpha$.

$$f_{X+Y}(\alpha) = \begin{cases} 
    a & 0 \leq a \leq 1 \\
    2 - a & 1 \leq a \leq 2 \\
    0 & \text{otherwise}
\end{cases}$$
Today’s plan

Independent RVs

Sum of independent RVs
• ✅ Binomial
• Convolution
• ✅ Poisson
• ✅ Normal
• ⚠ Uniform

Expectation of sum of RVs (next class)
Properties of Expectation, extended to two RVs

1. Linearity:

\[ E[aX + bY + c] = aE[X] + bE[Y] + c \]

2. Expectation of a sum = sum of expectation:

\[ E[X + Y] = E[X] + E[Y] \]

(we’ve seen this; we’ll prove this next)

3. Unconscious statistician:

\[ E[g(X, Y)] = \sum_x \sum_y g(x, y)p_{X,Y}(x, y) \]
\[ E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f_{X,Y}(x, y) \, dx \, dy \]
Proof of expectation of a sum of RVs

\[ E[X + Y] = E[g(X, Y)] = \sum_x \sum_y g(x, y)p_{X,Y}(x, y) = \sum_x \sum_y (x + y)p_{X,Y}(x, y) \]

\[ E[X + Y] = \sum_x x \sum_y p_{X,Y}(x, y) + \sum_y y \sum_x p_{X,Y}(x, y) \]

\[ E[X + Y] = \sum_x xp_X(x) + \sum_y yp_Y(y) \]

\[ = E[X] + E[Y] \]

Even if the joint distribution is unknown, you can calculate the expectation of sum as sum of expectations.

Example: \( E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i] \) despite dependent trials \( X_i \).
Expectations of common RVs

\[ X \sim \text{Bin}(n, p) \quad E[X] = np \]

\[ X = \sum_{i=1}^{n} X_i \quad \text{Let } X_i = \text{ith trial is heads} \]

\[ X_i \sim \text{Ber}(p), E[X_i] = p \]

\[ E[X] = E \left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np \]
Expectations of common RVs

**$X \sim \text{Bin}(n, p)$**  \hspace{1cm} $E[X] = np$

\[ X = \sum_{i=1}^{n} X_i \text{ Let } X_i = \text{ith trial is heads} \]
\[ X_i \sim \text{Ber}(p), E[X_i] = p \]

\[ E[X] = E \left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np \]

**$Y \sim \text{NegBin}(r, p)$**  \hspace{1cm} $E[Y] = \frac{r}{p}$

Suppose:

\[ Y = \sum_{i=1}^{?} Y_i \]

How should we define $Y_i$?

A. $Y_i = \text{ith trial is heads. } Y_i \sim \text{Ber}(p), i = 1, \ldots, n$

B. $Y_i = \# \text{ trials to get ith success (after (i - 1)th success)}$
\[ Y_i \sim \text{Geo}(p), i = 1, \ldots, r \]

C. $Y_i = \# \text{ successes in } n \text{ trials}$
\[ Y_i \sim \text{Bin}(n, p), i = 1, \ldots, r, \text{ we look for } P(Y_i = 1) \]
Expectations of common RVs

\( X \sim \text{Bin}(n, p) \quad E[X] = np \)

\[ X = \sum_{i=1}^{n} X_i \quad \text{Let } X_i = \text{ith trial is heads} \quad X_i \sim \text{Ber}(p), E[X_i] = p \]

\[ E[X] = E \left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np \]

\( Y \sim \text{NegBin}(r, p) \quad E[Y] = \frac{r}{p} \)

\[ Y = \sum_{i=1}^{r} Y_i \quad \text{Let } Y_i = \text{# trials to get ith success (after} \quad (i - 1)\text{th success)} \quad Y_i \sim \text{Geo}(p), E[Y_i] = \frac{1}{p} \]

\[ E[Y] = E \left[ \sum_{i=1}^{r} Y_i \right] = \sum_{i=1}^{r} E[Y_i] = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p} \]