

12: Independent RVs

Jerry Cain
April 22, 2022

Table of Contents

2 Sums of Binomials
7 Convolutions and Poisson
15 Exercises



Sums of independent Binomial RVs

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:

for all x, y :

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

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**.

Sum of independent Binomials

$$\begin{array}{l} X \sim \text{Bin}(n_1, p) \\ Y \sim \text{Bin}(n_2, p) \\ X, Y \text{ independent} \end{array} \quad \Rightarrow \quad X + Y \sim \text{Bin}(n_1 + n_2, p)$$

Intuition:

- Each trial in X and Y is independent and has same success probability p
- Define $Z = \#$ successes in $n_1 + n_2$ 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:

$$\begin{array}{l} X_i \sim \text{Bin}(n_i, p) \\ X_i \text{ independent for } i = 1, \dots, n \end{array} \quad \Rightarrow \quad \sum_{i=1}^n X_i \sim \text{Bin}\left(\sum_{i=1}^n n_i, p\right)$$

If only it were
always so
simple...

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?
2. Are X and Y independent?



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? ❌

Counterexample: What if $Z = 0$?

2. Are X and Y independent? ✅

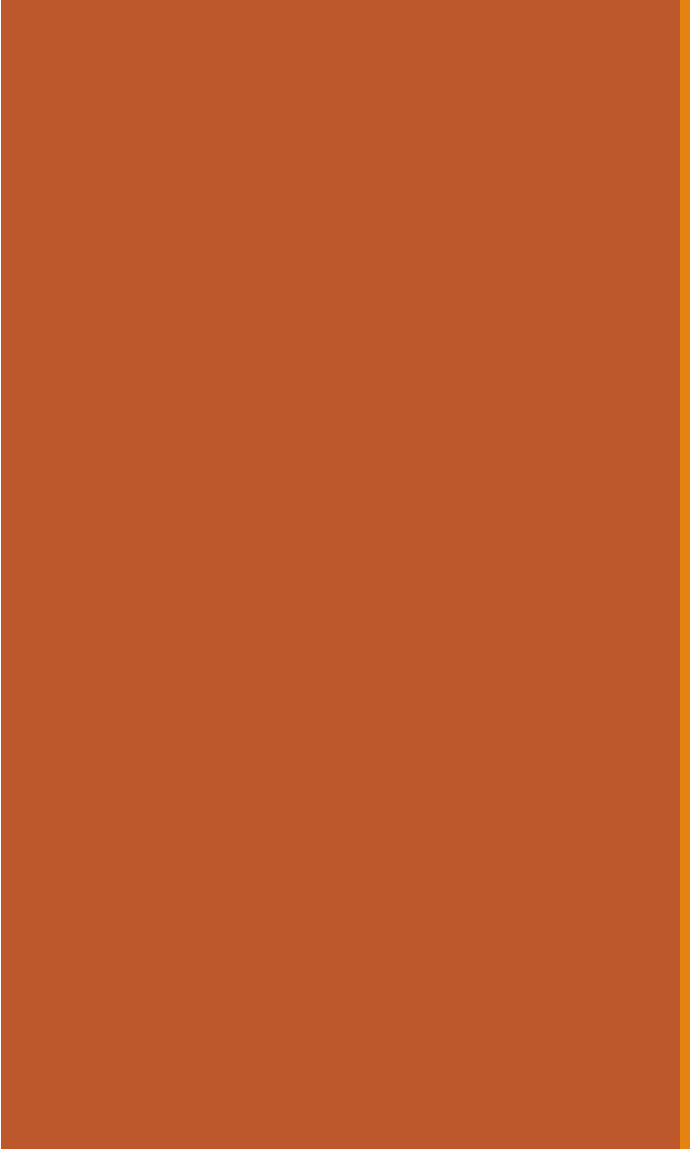
$$P(X = x, Y = y) = P(\text{first } n \text{ flips have } x \text{ heads} \\ \text{and next } m \text{ flips have } y \text{ heads})$$

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

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

of mutually exclusive outcomes in event : $\binom{n}{x} \binom{m}{y}$
 $P(\text{each outcome})$
 $= p^x (1-p)^{n-x} p^y (1-p)^{m-y}$

This probability (found through counting) is the product of the marginal PMFs.



Convolution: Sum of independent Poisson RVs

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 :

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

the **convolution** of p_X and p_Y

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

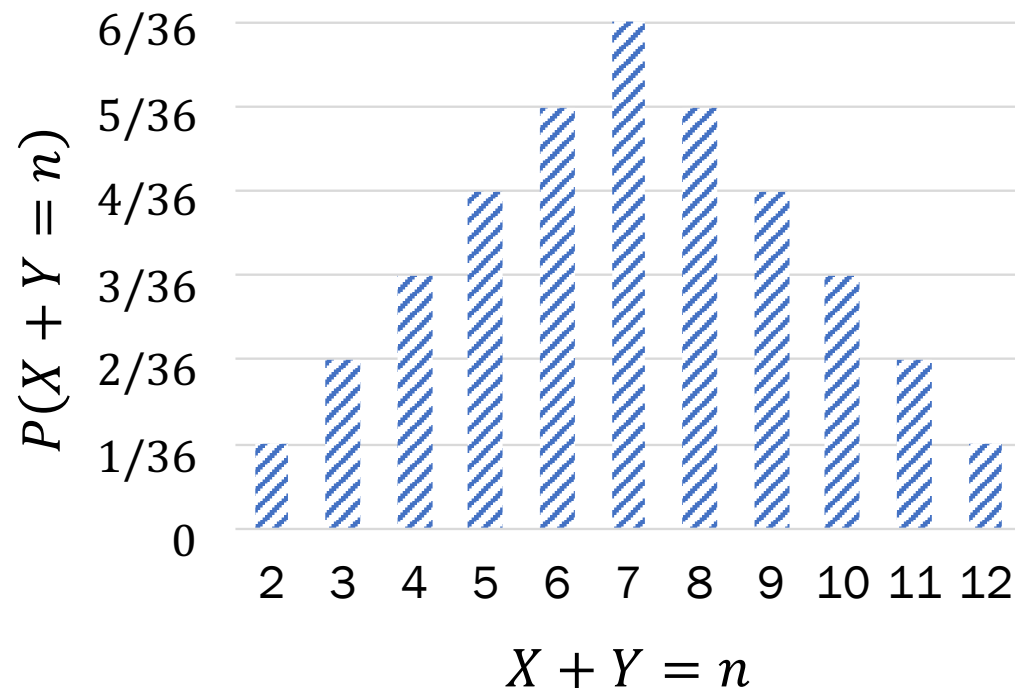
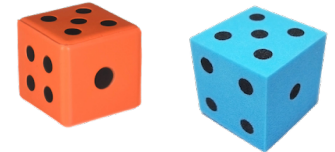
		X						
		0	1	2	...	n	$n + 1$...
Y	0					✓		
			
	$n - 2$			✓				
	$n - 1$		✓					
	n	✓						
	$n + 1$							
	...							

- ✓: event where $X + Y = n$
- Each event has probability:

$$P(X = k, Y = n - k)$$

$$= P(X = k)P(Y = n - k)$$
 (because X, Y are independent)
- $P(X + Y = n) =$ sum of mutually exclusive events

Sum of 2 dice rolls

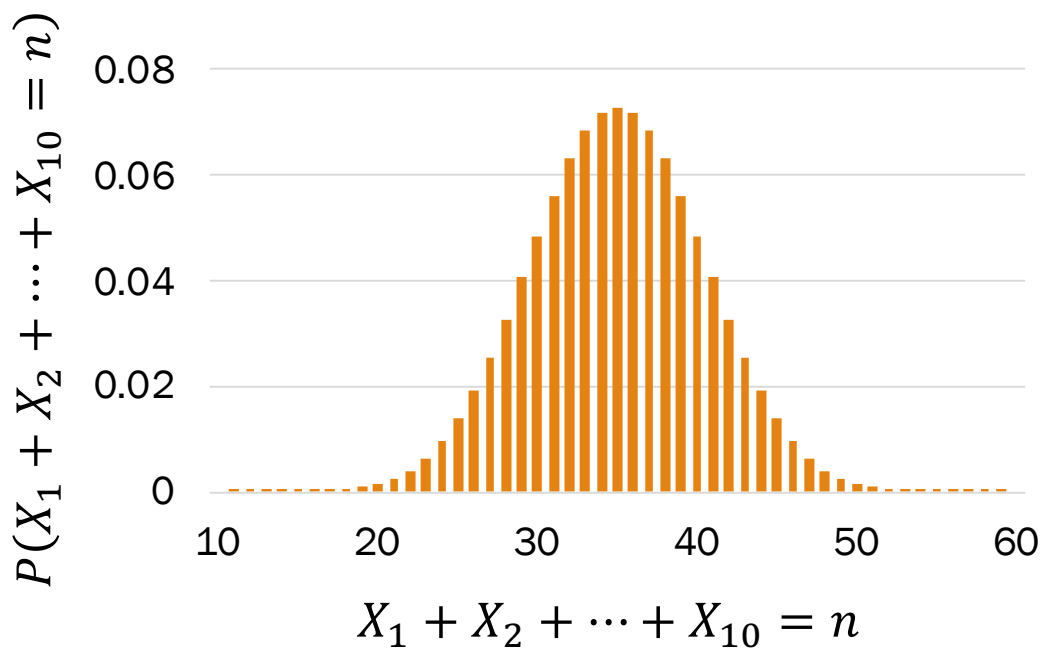


The distribution of a sum of 2 dice rolls is a convolution of 2 PMFs.

Example:

$$\begin{aligned} P(X + Y = 4) = & P(X = 1)P(Y = 3) \\ & + P(X = 2)P(Y = 2) \\ & + P(X = 3)P(Y = 1) \end{aligned}$$

Sum of 10 dice rolls (fun preview)



The distribution of a sum of 10 dice rolls is a convolution 10 PMFs.

Looks kinda Normal...???
(more on this in Week 7)

Sum of independent Poissons

$X \sim \text{Poi}(\lambda_1), Y \sim \text{Poi}(\lambda_2)$
 X, Y independent



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

Proof (just for reference):

$$\begin{aligned} P(X + Y = n) &= \sum_k P(X = k)P(Y = n - k) \\ &= \sum_{k=0}^n e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} = e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^n \frac{\lambda_1^k \lambda_2^{n-k}}{k! (n-k)!} \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{k=0}^n \frac{n!}{k! (n-k)!} \lambda_1^k \lambda_2^{n-k} = \underbrace{\frac{e^{-(\lambda_1 + \lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n}_{\text{Poi}(\lambda_1 + \lambda_2)} \end{aligned}$$

X and Y independent,
convolution

PMF of Poisson RVs

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 independent for $i = 1, \dots, n$



$$\sum_{i=1}^n X_i \sim \text{Poi}\left(\sum_{i=1}^n \lambda_i\right)$$



Sum of independent Poissons

$$\begin{array}{l} X \sim \text{Poi}(\lambda_1), Y \sim \text{Poi}(\lambda_2) \\ X, Y \text{ independent} \end{array} \quad \Rightarrow \quad X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$$

- n servers with independent number of requests/minute
- Server i 's requests each minute can be modeled as $X_i \sim \text{Poi}(\lambda_i)$

What is the probability that the total number of web requests received at all servers in the next minute exceeds 10?



Exercises

Independent questions

1. Let $X \sim \text{Bin}(30, 0.01)$ and $Y \sim \text{Bin}(50, 0.02)$ be independent RVs.
 - How do we compute $P(X + Y = 2)$ using a Poisson approximation?
 - How do we compute $P(X + Y = 2)$ exactly?
2. Let $N = \#$ of requests to a web server per day. Suppose $N \sim \text{Poi}(\lambda)$.
 - Each request independently comes from a human (prob. p), or bot ($1 - p$).
 - Let X be $\#$ of human requests/day, and Y be $\#$ of bot requests/day.

Are X and Y independent? What are their marginal PMFs?

1. Approximating the sum of independent Binomial RVs

Let $X \sim \text{Bin}(30, 0.01)$ and $Y \sim \text{Bin}(50, 0.02)$ be independent RVs.

- How do we compute $P(X + Y = 2)$ using a Poisson approximation?

- How do we compute $P(X + Y = 2)$ exactly?

$$\begin{aligned} P(X + Y = 2) &= \sum_{k=0}^2 P(X = k)P(Y = 2 - k) \\ &= \sum_{k=0}^2 \binom{30}{k} 0.01^k (0.99)^{30-k} \binom{50}{2-k} 0.02^{2-k} 0.98^{50-(2-k)} \approx \mathbf{0.2327} \end{aligned}$$

2. Web server requests

Let $N = \#$ of requests to a web server per day. Suppose $N \sim \text{Poi}(\lambda)$.

- Each request independently comes from a human (prob. p), or bot ($1 - p$).
- Let X be $\#$ of human requests/day, and Y be $\#$ of bot requests/day.

Are X and Y independent? What are their marginal PMFs?

$$P(X = x, Y = y) = P(X = x, Y = y | N = x + y)P(N = x + y) + P(X = x, Y = y | N \neq x + y)P(N \neq x + y) \quad \text{Law of Total Probability}$$

$$= P(X = x | N = x + y)P(Y = y | X = x, N = x + y)P(N = x + y) \quad \text{Chain Rule}$$

$$= \binom{x + y}{x} p^x (1 - p)^y \cdot 1 \cdot e^{-\lambda} \frac{\lambda^{x+y}}{(x + y)!} \quad \text{Given } N = x + y \text{ indep. trials, } X | N = x + y \sim \text{Bin}(x + y, p)$$

$$= \frac{(x + y)!}{x! y!} e^{-\lambda} \frac{(\lambda p)^x (\lambda(1 - p))^y}{(x + y)!} = e^{-\lambda p} \frac{(\lambda p)^x}{x!} \cdot e^{-\lambda(1-p)} \frac{(\lambda(1 - p))^y}{y!}$$

$$= P(X = x)P(Y = y) \quad \text{where } X \sim \text{Poi}(\lambda p), Y \sim \text{Poi}(\lambda(1 - p))$$

Yes, X and Y are independent!

Independence of multiple random variables

Recall independence of n events E_1, E_2, \dots, E_n :

for $r = 1, \dots, n$:

for every subset E_1, E_2, \dots, E_r :

$$P(E_1, E_2, \dots, E_r) = P(E_1)P(E_2) \cdots P(E_r)$$

We have independence of n discrete random variables X_1, X_2, \dots, X_n if for all x_1, x_2, \dots, x_n :

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i)$$

Independence is symmetric

If X and Y are independent random variables, then X is independent of Y , and Y is independent of X

Well, yeah....

Captain
Obvious



Let N be the number of times you roll 2 dice repeatedly until a 4 is rolled (the player wins), or a 7 is rolled (the player loses).

Let X be the value (4 or 7) of the final throw.

- Is N independent of X ?
 $P(N = n|X = 7) = P(N = n)?$
 $P(N = n|X = 4) = P(N = n)?$
 - Is X independent of N ?
 $P(X = 4|N = n) = P(X = 4)?$
 $P(X = 7|N = n) = P(X = 7)?$
- } (yes, easier to intuit)

Redux: Independence is not always intuitive, but it is **always** symmetric.