

# MATH 180A: INTRO TO PROBABILITY (FOR DATA SCIENCE)

[www.math.ucsd.edu/~tkemp/180A](http://www.math.ucsd.edu/~tkemp/180A)

Today: § 2.4-2.5, 4.4

Next: § 3.3

Lab 2.2 due **TONIGHT**  
(maybe?)

HW3 due Friday.

Midterm next wed evening

\* Regrade requests for Lab1: window **Thursday 10/17**

**8am - 11pm**

↳ separate request for each problem

↳ detailed, polite responses please.

\* The material slotted for today's lecture is the **cutoff**  
for **Midterm 1**.

# Independent Random Variables

2.3

A collection  $X_1, X_2, X_3, \dots, X_n$  of random variables defined on the same sample space are independent if

for any  $B_1, B_2, \dots, B_n \subseteq \mathbb{R}$ , the events

E.g.  $\{X_1 \in B_1\}, \{X_2 \in B_2\}, \dots, \{X_n \in B_n\}$  are independent.

$$\mathbb{P}(\{X_1 \in B_1\} \cap \{X_2 \in B_2\} \cap \dots \cap \{X_n \in B_n\}) = \mathbb{P}(X_1 \in B_1) \mathbb{P}(X_2 \in B_2) \dots \mathbb{P}(X_n \in B_n)$$

Special Case: if the  $X_j$  are discrete random variables, it suffices to check the simpler condition

for any real numbers  $t_1, t_2, \dots, t_n$

$$\mathbb{P}(X_1 = t_1, X_2 = t_2, \dots, X_n = t_n) = \mathbb{P}(X_1 = t_1) \mathbb{P}(X_2 = t_2) \dots \mathbb{P}(X_n = t_n)$$

E.g. Let  $X_1, X_2, \dots, X_n$  be fair coin tosses. Denote  $H \sim 1, T \sim 0$ .

$$\mathbb{P}(X_1 = t_1, X_2 = t_2, \dots, X_n = t_n) = \frac{1}{2^n} = \underbrace{\frac{1}{2} \cdot \frac{1}{2} \dots \frac{1}{2}}_n = \mathbb{P}(X_1 = t_1) \mathbb{P}(X_2 = t_2) \dots \mathbb{P}(X_n = t_n)$$

# Independent Trials

2.4

Experiments can have numerical observables, but sometimes you only observe whether there is **success** or **failure**

We model this with a random variable  $X$  taking value 1 with some probability  $p$ , and value 0 with probability  $1-p$ .

$$X \sim \text{Ber}(p) \quad (\text{Bernoulli})$$

In practice, we usually repeat the experiment many times, making sure to use the same setup each trial. The previous trials do not influence the future ones.

$X_1, X_2, X_3, \dots, X_n$  independent  $\text{Ber}(p)$  r.v.'s.  $(1-p)^4 p^2$

$$\begin{aligned} & \mathbb{P}(X_1=0, X_2=1, X_3=0, X_4=0, X_5=1, X_6=0) \quad (1-p) p (1-p) (1-p) p (1-p) \\ & = \mathbb{P}(X_1=0) \mathbb{P}(X_2=1) \mathbb{P}(X_3=0) \mathbb{P}(X_4=0) \mathbb{P}(X_5=1) \mathbb{P}(X_6=0) \end{aligned}$$

## How many successful trials?

Run  $n$  independent trials, each with success probability  $p$ .

$$X_1, X_2, \dots, X_n \sim \text{Ber}(p).$$

Let  $S_n = \#$  successful trials

$$= X_1 + X_2 + X_3 + \dots + X_n.$$

What is the distribution of  $S_n$ ?  $S_n \in \{0, 1, 2, \dots, n\}$

$$\begin{aligned} \mathbb{P}(S_n = k) &= \mathbb{P}(\{\text{exactly } k \text{ of the } n \text{ trials are successful}\}) \\ 0 \leq k \leq n & \quad = \binom{n}{k} p^k (1-p)^{n-k} \end{aligned}$$

$p = \frac{1}{2}, 1-p = \frac{1}{2}$   
 $p^k (1-p)^{n-k} = \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \frac{1}{2^n}$

Binomial Distribution Bin( $n, p$ )

$\text{Ber}(p) = \text{Bin}(1, p)$  | #heads in  $n$  <sup>fair</sup> coin tosses  $\approx \text{Bin}(n, \frac{1}{2})$

Eg. Roll a fair die 10 times. What is the probability that success  $\rightarrow$  (6) comes up at least 3 times?

$X_1, X_2, \dots, X_{10}$  indep.  $\text{Ber}(\frac{1}{6})$

$\therefore S_{10} = \text{Bin}(10, \frac{1}{6})$

$$P(S_{10} \geq 3) = \sum_{k=3}^{10} P(S_{10} = k)$$

$$\frac{566299}{2519424} \approx 22.51\%$$

$$\begin{aligned} 1 - P(S_{10} < 3) &= 1 - (P(S_{10} = 0) + P(S_{10} = 1) + P(S_{10} = 2)) \quad || \\ &= 1 - \left( \binom{10}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{10} + \binom{10}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^9 + \binom{10}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^8 \right) \end{aligned}$$

Eg. What is the probability that no 6 is rolled in the 10 rolls?

$$P(S_{10} = 0) = \binom{10}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{10} = \left(\frac{5}{6}\right)^{10} \approx 16\%$$

Now, keep rolling. Let  $N$  denote the first roll where a 6 appears.  $N$  is a random variable. What is its distribution?

## First Success Time

$N$  = first success in repeated independent trials (success rate  $p$ ).

Model trials with (unlimited number) of independent  $\text{Ber}(p)$ 's:

$$X_1, X_2, X_3, X_4, \dots \quad N \in \{1, 2, 3, \dots\}$$

$$\{N=k\} = \{X_1=0, X_2=0, X_3=0, \dots, X_{k-1}=0, X_k=1\}$$

$$\begin{aligned} P(N=k) &= P(\text{"}) = P(X_1=0) P(X_2=0) \dots P(X_{k-1}=0) P(X_k=1) \\ &= (1-p) \cdot (1-p) \dots (1-p) \cdot p = (1-p)^{k-1} p \end{aligned}$$

Geometric Distribution  $\text{Geom}(p)$  on  $\{0, 1, 2, 3, \dots\} = N$ .

$$\sum_{k=1}^{\infty} P(N=k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=1}^{\infty} (1-p)^{k-1} = p \sum_{k-1=l}^{\infty} (1-p)^l$$

geometric series  $\left[ = p \frac{1}{1-(1-p)} \right] = \frac{p}{p} = 1$ .

# Rare Events

4.4

If  $S_n = S_{n,p} \sim \text{Bin}(n, p)$ ,  $S_n$  is the number of successes in  $n$  independent trials each with success probability  $p$ .

What if  $p$  is very small, but  $n$  is very large?

One way to handle this mathematically is a **scaling limit**

↳ For each  $n$ , take  $p \propto \frac{1}{n}$ .  $p = \frac{\lambda}{n}$  for some  $\lambda > 0$ .

$$\begin{aligned} \mathbb{P}(S_{n,p} = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

$$\lim_{n \rightarrow \infty} \left( \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \right) = ?$$

What happens as  $n \rightarrow \infty$ ?

$$P(S_{n,p}=k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$n \rightarrow \infty$   
 $k$  is fixed

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \dots \frac{n-k+1}{n} \cdot \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$\begin{matrix} \downarrow & \downarrow & \downarrow & \dots & \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 1 & 1 & \dots & 1 & \frac{\lambda^k}{k!} & e^{-\lambda} & 1 \end{matrix}$$

$$\left(1 + \frac{\lambda}{n}\right)^n \rightarrow e^\lambda$$

$$\therefore \lim_{n \rightarrow \infty} P(S_{n,\lambda/n} = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$



# Poisson Distribution

A random variable  $X$  has the  $\text{Poisson}(\lambda)$  distribution if

$$P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k=0,1,2,\dots$$

Check:  $\sum_{k=0}^{\infty} P(X=k) = 1.$

# Summary

Sampling independent trials, the most important (discrete) probability distributions are:

- $\text{Ber}(p)$ :  $P(X=1)=p, P(X=0)=1-p$   $0 \leq p \leq 1$   
(single trial with success probability  $p$ )
- $\text{Bin}(n,p)$ :  $P(S_n=k) = \binom{n}{k} p^k (1-p)^{n-k}$   $0 \leq k \leq n$   
(number of successes in  $n$  independent trials with rate  $p$ )
- $\text{Geom}(p)$   $P(N=k) = (1-p)^{k-1} p$   $k=0,1,2,\dots$   
(first successful trial in repeated independent trials with rate  $p$ )
- $\text{Poisson}(\lambda)$   $P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$   $k=0,1,2,\dots$   $\lambda > 0$ .  
(Approximates  $\text{Bin}(n, \lambda/n)$ ; number of rare events in many trials)