M378K: August 31^{st}, 2026.

3. Independent events

What if knowing that an event happened in fact does **not** give any information about the probability of another event?

Definition 3.1. *We say that events E and F on Ω are* independent *if*

$$\mathbb{P}[E \cap F] = \mathbb{P}[E]\mathbb{P}[F].$$

In the case when E or F have a positive probability, it's possible to rewrite the above condition in a different (illustrative!) way. *How?*

Assume $\mathbb{P}[E] > 0$.

$$\mathbb{P}[F \mid E] = \frac{\mathbb{P}[E \cap F]}{\mathbb{P}[E]} = \frac{\cancel{\mathbb{P}[E]} \cdot \mathbb{P}[F]}{\cancel{\mathbb{P}[E]}} = \mathbb{P}[F]$$

Now that we know the notion of **independence**, we can construct random variables in many creative ways.

Example 3.2. *A fair coin is tossed repeatedly and* **independently** *until the first* Heads. *Let the random variable Y represent the total number of* Tails *observed by the end of the procedure.*

What is the support of the random variable Y?

$$S_Y = \{0, 1, 2, \dots\} =: \mathbb{N}_0$$

What is the **probability mass function** *of the random variable Y?*

for all $y \in S_Y$:

$$p_Y(y) = \mathbb{P}[Y = y] = \left(\frac{1}{2}\right)^y \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{y+1}$$

Moreover, now that we remember the definition of **conditional probability**, we can solve interesting problems such as this one:

Problem 3.1. *The number of pieces of gossip that break out in a particular high school in a week is modeled by a random variable Y with the following probability mass function:*

$$p_n = p_Y(n) = \frac{1}{(n+1)(n+2)} \quad \textit{for all } n \in \mathbb{N}_0.$$

(i) Is the above a well-defined probability mass function? ✓

(ii) $\boxed{\textit{Calculate the probability that at least one piece of gossip occurred in a week } \textbf{given} \textit{ that at most four pieces of gossip occurred.}}$

(i)

- $p_n \geq 0$ ✓
- $\sum_{n=0}^{\infty} p_n \overset{?}{=} 1$ ✓

Let $N \in \mathbb{N}_0$:

$$\sum_{n=0}^{N} p_n = \sum_{n=0}^{N} \frac{1}{(n+1)(n+2)}$$

$$= \sum_{n=0}^{N} \left(\frac{1}{n+1} - \frac{1}{n+2} \right)$$

$$= \frac{1}{0+1} - \cancel{\frac{1}{0+2}} + \cancel{\frac{1}{1+1}} - \cancel{\frac{1}{1+2}} + \dots$$

$$\dots + \cancel{\frac{1}{N+1}} - \frac{1}{N+2}$$

$$= 1 - \frac{1}{N+2} \xrightarrow[N \to \infty]{} 1$$

(ii) $$\mathbb{P}[Y \geq 1 \mid Y \leq 4] = \frac{\mathbb{P}[1 \leq Y \leq 4]}{\mathbb{P}[Y \leq 4]}$$

$$= \frac{p_1 + p_2 + p_3 + p_4}{p_0 + p_1 + p_2 + p_3 + p_4}$$

$$= \frac{\frac{1}{2} - \cancel{\frac{1}{3}} + \cdots + \cancel{\frac{1}{5}} - \frac{1}{6}}{1 - \cancel{\frac{1}{2}} + \cancel{\frac{1}{2}} - \cancel{\frac{1}{3}} + \cdots + \cancel{\frac{1}{5}} - \frac{1}{6}}$$

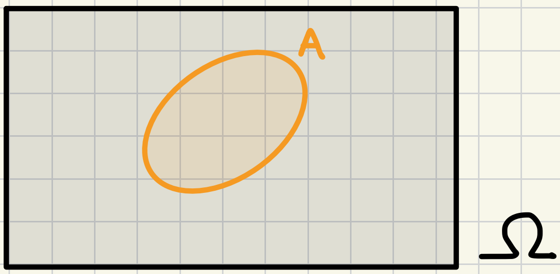
$$= \frac{\frac{1}{2} - \frac{1}{6}}{1 - \frac{1}{6}} = \frac{\frac{2}{6}}{\frac{5}{6}} = \frac{2}{5} \quad \square$$

Named Discrete Distributions.

Def'n. **Bernoulli Trials** have two possible outcomes.

Usually, the outcomes are interpreted as
$$\begin{cases} 1 \text{ for "success"} \\ 0 \text{ for "failure"} \end{cases}$$

They are also known as **indicators** or **indicator random variables.**



$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

$$\mathbb{I}_A = \begin{cases} 1 & \text{if } A \text{ happened} \\ 0 & \text{if } A \text{ did not happen} \end{cases}$$

Example. Y_i, $i = 1, 2$, are the results of a throw of two regular dice.

We win if

- the number on the 1^{st} die is even

and

- the number on the 2^{nd} die is prime.

Define: $I_1 = \mathbb{I}_{\{Y_1 \in \{2,4,6\}\}}$

$I_2 = \mathbb{I}_{\{Y_2 \in \{2,3,5\}\}}$

The indicator of a win is $\boxed{I_1 \cdot I_2}$

Practically.
- Insurance. An indicator of whether the deductible is met.
- Quality control. ... whether a component is defective or not.

Bernoulli dist'n. $Y \sim B(p)$ w/ $p \in (0,1)$

y	0	1
$p_Y(y)$	q	p

w/ $q := 1 - p$

Q: Say we repeat independently n Bernoulli trials w/ the same success probability p.

We denote by Y the number of successes.

$$S_Y = \cancel{?}\ \{0, 1, \dots, n\}$$

$$p_Y = ?$$

for $k = 0, \dots, n$

$$p_Y(k) = \mathbb{P}[Y=k] = \boxed{\binom{n}{k}} p^k (1-p)^{n-k}$$

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$$

BINOMIAL COEFFICIENT

The dist'n of Y is called the binomial dist'n.

We write $\boxed{b(n,p)}$