independence
Defn: Two events E and F are *independent* if
\[ P(EF) = P(E) \cdot P(F) \]

If \( P(F) > 0 \), this is equivalent to: \[ P(E|F) = P(E) \] (proof below)

Otherwise, they are called *dependent*
Roll two dice, yielding values $D_1$ and $D_2$

1) $E = \{ D_1 = 1 \}$
$F = \{ D_2 = 1 \}$
$P(E) = \frac{1}{6}, \ P(F) = \frac{1}{6}, \ P(EF) = \frac{1}{36}$

$P(EF) = P(E) \cdot P(F) \Rightarrow E$ and $F$ independent

*Intuitive; the two dice are not physically coupled*

2) $G = \{D_1 + D_2 = 5\} = \{(1,4),(2,3),(3,2),(4,1)\}$
$P(E) = \frac{1}{6}, \ P(G) = \frac{4}{36} = \frac{1}{9}, \ P(EG) = \frac{1}{36}$

*not independent!*

$E, \ G$ are dependent events

*The dice are still not physically coupled, but “$D_1 + D_2 = 5$” couples them mathematically: info about $D_1$ constrains $D_2$. (But dependence/independence not always intuitively obvious; “use the definition, Luke”).*
Two events $E$ and $F$ are independent if

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

If $P(F) > 0$, this is equivalent to: $P(E|F) = P(E)$

Otherwise, they are called dependent

Three events $E$, $F$, $G$ are independent if

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

$P(EG) = P(E) \cdot P(G)$  \hspace{4em} and \hspace{4em} $P(EFG) = P(E) \cdot P(F) \cdot P(G)$

$P(FG) = P(F) \cdot P(G)$

Example: Let $X, Y$ be each $\{-1, 1\}$ all outcomes equally likely

$E = \{X = 1\}$, $F = \{Y = 1\}$, $G = \{XY = 1\}$

$P(EF) = P(E)P(F)$, $P(EG) = P(E)P(G)$, $P(FG) = P(F)P(G)$

but $P(EFG) = 1/4$ !!
In general, events $E_1, E_2, \ldots, E_n$ are independent if for every subset $S$ of $\{1,2,\ldots, n\}$, we have

$$P \left( \bigcap_{i \in S} E_i \right) = \prod_{i \in S} P(E_i)$$

(Sometimes this property holds only for small subsets $S$. E.g., $E, F, G$ on the previous slide are pairwise independent, but not fully independent.)
Theorem: E, F independent \implies E, F^c independent

Proof: \[ P(EF^c) = P(E) - P(EF) \]
\[ = P(E) - P(E) P(F) \]
\[ = P(E) (1 - P(F)) \]
\[ = P(E) P(F^c) \]

Theorem: \( P(E)>0, P(F)>0 \)
E, F independent \iff P(E|F)=P(E) \iff P(F|E) = P(F)

Proof: Note \( P(EF) = P(E|F) P(F) \), regardless of in/dep.
Assume independent. Then

\[ P(E)P(F) = P(EF) = P(E|F) P(F) \implies P(E|F)=P(E) \quad (\div \text{ by } P(F)) \]

Conversely, \( P(E|F)=P(E) \implies P(E)P(F) = P(EF) \quad (\times \text{ by } P(F)) \)
Suppose a biased coin comes up heads with probability $p$, independent of other flips

$$P(n \text{ heads in } n \text{ flips}) = p^n$$

$$P(n \text{ tails in } n \text{ flips}) = (1-p)^n$$

$$P(\text{exactly } k \text{ heads in } n \text{ flips}) = \binom{n}{k} p^k (1-p)^{n-k}$$

Aside: note that the probability of some number of heads = \[ \sum_k \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1 \] as it should, by the binomial theorem.
Suppose a biased coin comes up heads with probability $p$, *independent* of other flips.

$$P(\text{exactly } k \text{ heads in } n \text{ flips}) = \binom{n}{k} p^k (1-p)^{n-k}$$

Note when $p=1/2$, this is the same result we would have gotten by considering $n$ flips in the “equally likely outcomes” scenario. But $p \neq 1/2$ makes that inapplicable. Instead, the *independence* assumption allows us to conveniently assign a probability to each of the $2^n$ outcomes, e.g.:

$$Pr(\text{HHTHTTT}) = p^2(1-p)p(1-p)^3 = p^{\#H}(1-p)^{\#T}$$
Consider the following parallel network

\[ P(\text{there is functional path}) = 1 - P(\text{all routers fail}) \]

\[ = 1 - p_1 p_2 \cdots p_n \]

\( n \) routers, \( i^{\text{th}} \) has probability \( p_i \) of failing, independently.
Contrast: a series network

\[ P(\text{there is functional path}) = P(\text{no routers fail}) = (1 - p_1)(1 - p_2) \cdots (1 - p_n) \]

n routers, \(i^{th}\) has probability \(p_i\) of failing, independently