Conditional Probability & Independence

Conditional Probabilities

- **Question**: How should we modify $\mathbb{P}(E)$ if we learn that event F has occurred?
- **Definition**: the conditional probability of E given F is

$$\boxed{\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}, \quad \text{for} \quad \mathbb{P}(F) > 0}$$

Condition probabilities are useful because:

- Often want to calculate probabilities when some partial information about the result of the probabilistic experiment is available.
- Conditional probabilities are useful for computing ”regular” probabilities.

Example 1. 2 random cards are selected from a deck of cards.

(a)
- What is the probability that both cards are aces given that one of the cards is the ace of spaces?

(b)
- What is the probability that both cards are aces given that at least one of the cards is an ace?

$\Omega = \{\text{all unordered pairs of cards}\}$ uniform prob dist'n

(a) $$\Pr(AA \mid A\spadesuit) = \frac{\Pr(A\spadesuit \text{ and another ace})}{\Pr(A\spadesuit)}$$

$$= \frac{3}{51} \approx 0.059$$

(b) $$\Pr(AA \mid \geq 1\ A) = \frac{\Pr(AA)}{\Pr(\geq 1\ A)}$$

$$= \frac{\binom{4}{2}}{\underbrace{\binom{52}{2}}_{\text{all possibilities}} - \underbrace{\binom{48}{2}}_{\text{no aces}}} = \frac{4 \cdot 3}{52 \cdot 51 - 48 \cdot 47} \approx 0.03$$

Example 2. Deal a 5 card poker hand, and let

$E = \{\text{at least 2 aces}\}$, $F = \{\text{at least 1 ace}\}$,

$G = \{\text{hand contains ace of spades}\}$.

(a) Find $\mathbb{P}(E)$

$$Pr(E) = 1 - \frac{\binom{48}{5}}{\binom{52}{5}} - \frac{4\binom{48}{4}}{\binom{52}{5}}$$

(b) Find $\mathbb{P}(E \mid F)$

$$\frac{Pr(E \cap F)}{Pr(F)} = \frac{Pr(E)}{Pr(F)} \qquad Pr(F) = 1 - \frac{\binom{48}{5}}{\binom{52}{5}}$$

(c) Find $\mathbb{P}(E \mid G)$

$$\frac{Pr(E \cap G)}{Pr(G)} = \frac{Pr(\text{A of } \spadesuit + \text{at least 1 one other A})}{Pr(\text{contains A of } \spadesuit)}$$

$$= \frac{\binom{51}{4} - \binom{48}{4}}{\binom{51}{4}}$$

Best of 3 tournament between A & B

$Pr(\text{A wins first game}) = \frac{1}{2}$

after that $Pr(\text{A wins} \mid \text{any history in which won previous}) = \frac{2}{3}$

$\Rightarrow Pr\left(\text{A loses} \,\middle|\, \begin{array}{l}\text{any history where}\\ \text{won previous}\end{array}\right) = \frac{1}{3}$

$$\boxed{Pr(E \mid F) = 1 - Pr(E^c \mid F)}$$

Proof: $Pr(E \mid F) + Pr(E^c \mid F)$

$$= \frac{Pr(E \cap F) + Pr(E^c \cap F)}{Pr(F)}$$

$$= \frac{Pr(F)}{Pr(F)} = 1$$

Cond prob satisfies the usual prob axioms.

Suppose $(\mathbb{S}, \mathbb{P}(\cdot))$ is a probability space.

Then $(\mathbb{S}, \mathbb{P}(\cdot \,|\, F))$ is also a probability space (for $F \subset \mathbb{S}$ with $\mathbb{P}(F) > 0$).

- $0 \leq \mathbb{P}(\omega \,|\, F) \leq 1$
- $\sum_{\omega \in \mathbb{S}} \mathbb{P}(\omega \,|\, F) = 1$
- If $E_1, E_2, \ldots$ are disjoint, then

$$\mathbb{P}(\cup_{i=1}^{\infty} E_i \,|\, F) = \sum_{i=1}^{\infty} \mathbb{P}(E_i \,|\, F)$$

Thus all our previous propositions for probabilities give analogous results for conditional probabilities.

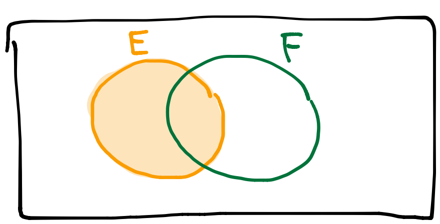
Examples

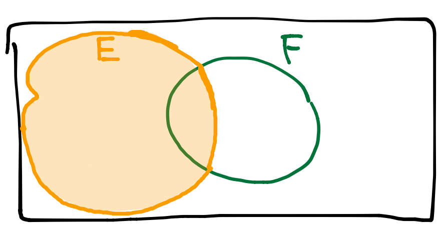
$\mathbb{P}(E^c \,|\, F) = 1 - \mathbb{P}(E \,|\, F)$

$\mathbb{P}(A \cup B \,|\, F) = \mathbb{P}(A \,|\, F) + \mathbb{P}(B \,|\, F) - \mathbb{P}(A \cap B \,|\, F)$

However,

$Pr(E|F) + Pr(E|F^c)$ not necessarily 1 !





The Multiplication Rule

- Re-arranging the conditional probability formula gives

$$\boxed{\mathbb{P}(E \cap F) = \mathbb{P}(F)\,\mathbb{P}(E \mid F)}$$

This is often useful in computing the probability of the intersection of events.

Example. Draw 2 balls at random without replacement from an urn with 8 red balls and 4 white balls. Find the chance that both are red.

8 R
4 W

draw 2 balls without replacement

$$Pr(\text{both } R) = Pr(\text{first } R)\, Pr(2^{nd} R \mid 1^{st} R)$$

$$= \frac{4}{12} \cdot \frac{7}{11}$$

The General Multiplication Rule

$$\boxed{\begin{aligned}&\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_n) = \\ &\quad \mathbb{P}(E_1) \times \mathbb{P}(E_2 \mid E_1) \times \mathbb{P}(E_3 \mid E_1 \cap E_2) \times \\ &\quad \cdots \times \mathbb{P}(E_n \mid E_1 \cap E_2 \cap \cdots \cap E_{n-1})\end{aligned}}$$

Example 1. Alice and Bob play a game as follows. A die is thrown, and each time it is thrown it is equally likely to show any of the 6 numbers. If it shows 5, A wins. If it shows 1, 2 or 6, B wins. Otherwise, they play a second round, and so on. Find $\mathbb{P}(A_n)$, for $A_n = \{\text{Alice wins on } n\text{th round}\}$.

N_i : event that nobody wins on i^{th} round

$$Pr(A_n) = Pr(N_1 \cap N_2 \cap \ldots \cap N_{n-1} \cap A_n)$$

$$= Pr(N_1)\,Pr(N_2 \mid N_1)\,Pr(N_3 \mid N_1 \cap N_2) \cdots Pr(N_{n-1} \mid N_1 \cdots N_{n-2})\,Pr(A_n \mid N_1 \cdots N_{n-1})$$

$$= \left(\frac{2}{6}\right)^{n-1} \cdot \frac{1}{6}$$

Best of 3 tournament between A & B

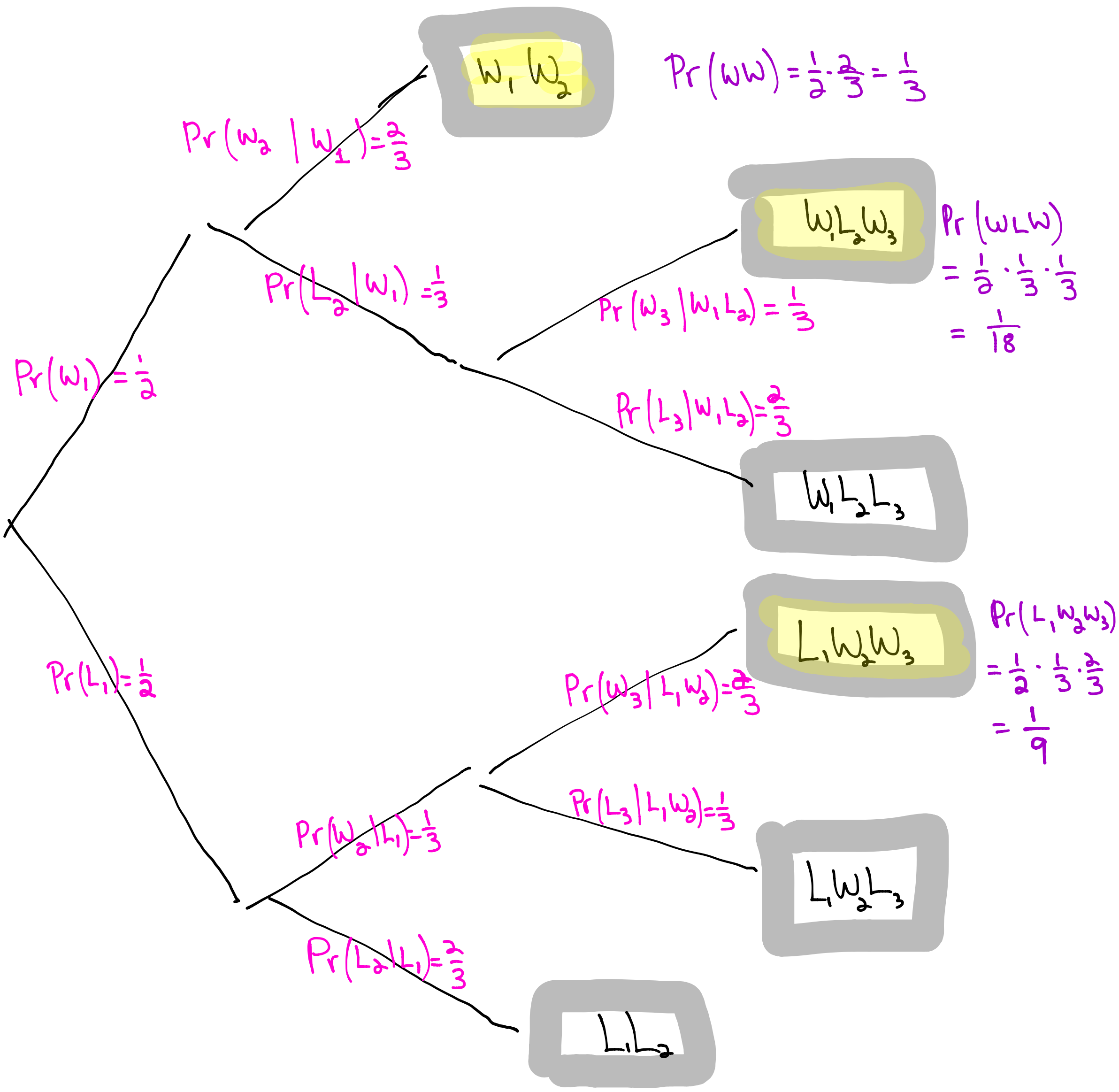
$\Pr(\text{A wins first game}) = \frac{1}{2}$

after that $\Pr(\text{A wins} \mid \text{any history in which won previous}) = \frac{2}{3}$

$\Pr(\text{A wins} \mid \text{any history in which lost previous}) = \frac{1}{3}$

Note: these 2 numbers do not need to add to 1

W_i : event that A wins i^{th} game



$$\Pr(\text{A wins tournament}) = \frac{1}{3} + \frac{1}{18} + \frac{1}{9} = \frac{1}{2}$$

Example 2. I have n keys, one of which opens a lock. Trying keys at random without replacement, find the chance that the kth try opens the lock.

E_i : event that i^{th} key opens lock

$$\Pr(E_k) = \Pr(\bar{E}_1 \bar{E}_2 \cdots \bar{E}_{k-1} E_k)$$

$$= \Pr(\bar{E}_1)\Pr(\bar{E}_2 | \bar{E}_1)\Pr(\bar{E}_3 | \bar{E}_1 \bar{E}_2) \cdots \Pr(\bar{E}_{k-1} | \bar{E}_1 \cdots \bar{E}_{k-2})\Pr(E_k | \bar{E}_1 \cdots \bar{E}_{k-1})$$

$$= \left(\frac{n-1}{n}\right)\left(\frac{n-2}{n-1}\right) \cdots \left(\frac{n-(k-1)}{n-k+2}\right)\left(\frac{1}{n-k+1}\right)$$

$$= \left(\frac{\cancel{n-1}}{n}\right)\left(\frac{\cancel{n-2}}{\cancel{n-1}}\right) \cancel{\cdots} \left(\frac{\cancel{n-(k-1)}}{\cancel{n-k+2}}\right)\left(\frac{1}{\cancel{n-k+1}}\right)$$

$$= \frac{1}{n}$$

The Law of Total Probability

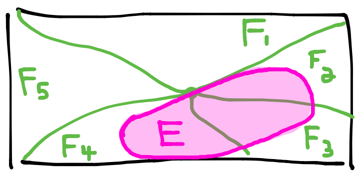
- We know that $\mathbb{P}(E) = \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^c)$. Using the definition of conditional probability,

  $$\boxed{\mathbb{P}(E) = \mathbb{P}(E \mid F)\,\mathbb{P}(F) + \mathbb{P}(E \mid F^c)\,\mathbb{P}(F^c)}$$

- This is **extremely useful**. It may be difficult to compute $\mathbb{P}(E)$ directly, but easy to compute it once we know whether or not F has occurred.

- To generalize, say events $F_1, \dots, F_n$ form a **partition** if they are disjoint and $\bigcup_{i=1}^{n} F_i = \mathbb{S}$.

- Since $E \cap F_1, E \cap F_2, \dots E \cap F_n$ are a disjoint partition of E. $\mathbb{P}(E) = \sum_{i=1}^{n} \mathbb{P}(E \cap F_i)$.



$$Pr(E) = \cancel{Pr(E \cap F_1)} + Pr(E \cap F_2) + Pr(E \cap F_3) + Pr(E \cap F_4) + \cancel{Pr(E \cap F_5)}$$

- Apply conditional probability to give the **law of total probability**,

  $$\boxed{\mathbb{P}(E) = \sum_{i=1}^{n} \mathbb{P}(E \mid F_i)\,\mathbb{P}(F_i)}$$

Example 1. Eric's girlfriend comes round on a given evening with probability 0.4. If she does not come round, the chance Eric watches *The Wire* is 0.8. If she does, this chance drops to 0.3. Find the probability that Eric gets to watch *The Wire*.

$$\Pr(\text{watches}) = \Pr(\text{watches} \mid GF)\Pr(GF) + \Pr(\text{watches} \mid \overline{GF})\Pr(\overline{GF})$$

$$= 0.3 \cdot 0.4 + 0.6 \cdot 0.6$$

$$= 0.48$$

Bayes Formula

- Sometimes $\mathbb{P}(E \mid F)$ may be specified and we would like to find $\mathbb{P}(F \mid E)$.

Example 2. I call Eric and he says he is watching *The Wire*. What is the chance his girlfriend is around?

Pr(girlfriend comes by) = 0.4
Pr(watches | GF) = 0.3
Pr (watches | $\overline{GF}$) = 0.6

- A simple manipulation gives **Bayes' formula**,

$$\boxed{\mathbb{P}(F \mid E) = \frac{\mathbb{P}(E \mid F)\,\mathbb{P}(F)}{\mathbb{P}(E)}}$$

since $Pr(E|F)Pr(F) = Pr(E \cap F)$

- Combining this with the law of total probability,

$$\boxed{\mathbb{P}(F \mid E) = \frac{\mathbb{P}(E \mid F)\,\mathbb{P}(F)}{\mathbb{P}(E \mid F)\,\mathbb{P}(F) + \mathbb{P}(E \mid F^c)\,\mathbb{P}(F^c)}}$$

since denominator = Pr(E) by law of total probability

$$Pr(GF \mid Watching) = \frac{Pr(Watches \mid GF)\,Pr(GF)}{Pr(watches)}$$

$$= \frac{0.3 \cdot 0.4}{0.48} = \frac{1}{4}$$

• Sometimes conditional probability calculations can give quite unintuitive results.

Example 3. I have three cards. One is red on both sides, another is red on one side and black on the other, the third is black on both sides. I shuffle the cards and put one on the table, so you can see that the upper side is red. What is the chance that the other side is black?

• is it $1/2$, or $> 1/2$ or $< 1/2$?

Solution

prob model 1: pick random card

put R side up if it has a red side

$$\Pr(RB \mid \text{see } R) = \frac{\Pr(RB \cap \text{see } R)}{\Pr(\text{see } R)} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

prob model 2: pick random card

pick random side to show

$$\Pr(RB \mid \text{see } R) = \frac{\Pr(RB \cap \text{see } R)}{\Pr(\text{see } R)}$$

$$= \frac{\Pr(\text{see } R \mid RB)\Pr(RB)}{\Pr(\text{see } R \mid RB)\Pr(RB) + \Pr(\text{see } R \mid RR)\Pr(RR) + \Pr(\text{see } R \mid BB)\Pr(BB)}$$

$$= \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3} + 1\cdot\frac{1}{3}} = \frac{1}{3}$$

Example: Spam Filtering

- 60% of email is spam.
- 10% of spam has the word "Viagra".
- 1% of non-spam has the word "Viagra".
- Let V be the event that a message contains the word "Viagra".
- Let J be the event that the message is spam.

What is the probability of J given V?

Solution.

$$\Pr(\text{spam} \mid V) = \frac{\Pr(V \mid \text{spam}) \Pr(\text{spam})}{\Pr(V \mid \text{spam}) \Pr(\text{spam}) + \Pr(V \mid \text{not spam}) \Pr(\text{not spam})}$$

$$= \frac{0.1 \cdot 0.6}{0.1 \cdot 0.6 + 0.01 \cdot 0.4}$$

Discussion problem. Suppose 99% of people with HIV test positive, 95% of people without HIV test negative, and 0.1% of people have HIV. What is the chance that someone testing positive has HIV?

$$\Pr(HIV^+ \mid test^+) =$$

$$= \frac{\Pr(test^+ \mid HIV^+)\Pr(HIV^+)}{\Pr(test^+ \mid HIV^+)\Pr(HIV^+) + \Pr(test^+ \mid HIV^-)\Pr(HIV^-)}$$

$$= \frac{.99 \cdot .001}{.99 \cdot .001 + .05 \cdot .999} = 0.019$$

$\sim 2\%$!!

what if people who get tested have 10% chance of being HIV^+

$\sim 68\%$

50% chance of being HIV^+

$\sim 95\%$

Example: Statistical inference via Bayes' formula

Alice and Bob play a game where A tosses a coin, and wins \$1 if it lands on H or loses \$1 on T. B is surprised to find that he loses the first ten times they play. If B's **prior belief** is that the chance of A having a two headed coin is 0.01, what is his **posterior belief**?

Note. Prior and posterior beliefs are assessments of probability before and after seeing an outcome. The outcome is called **data** or **evidence**.

Solution.

$$\Pr(\text{2-headed} \mid \text{loses } 10\times) = \frac{\Pr(\text{2 headed \& loses } 10\times)}{\Pr(\text{loses } 10\times)}$$

$$= \frac{\Pr(\text{loses } 10\times \mid \text{2 headed}) \Pr(\text{2-headed})}{\Pr(\text{loses } 10\times \mid \text{2 headed}) \Pr(\text{2 headed}) + \Pr(\text{loses } 10\times \mid \text{regular}) \Pr(\text{reg})}$$

$$= \frac{0.01}{0.01 + \left(\frac{1}{2}\right)^{10} 0.99}$$

Example: A plane is missing, and it is equally likely to have gone down in any of three possible regions. Let α_i be the probability that the plane will be found in region i given that it is actually there. What is the conditional probability that the plane is in the second region, given that a search of the first region is unsuccessful?

$$\Pr(\text{in } 2^{nd} \mid \text{search of first failed})$$

$$= \frac{\Pr(\text{in } 2^{nd} \text{ \& search of } 1^{st} \text{ failed})}{\Pr(\text{search of } 1^{st} \text{ failed})}$$

$$= \frac{\frac{1}{3}\Pr(\text{search of } 1^{st} \text{ failed} \mid \text{in } 2^{nd})}{\frac{1}{3}\cdot\alpha_1 + \frac{2}{3}}$$

$$= \frac{\frac{1}{3}}{\frac{1}{3}\alpha_1 + \frac{2}{3}}$$

Independence

- Intuitively, E is independent of F if the chance of E occurring is not affected by whether F occurs. Formally,

$$\mathbb{P}(E \mid F) = \mathbb{P}(E) \tag{1}$$

- We say that E and F are **independent** if

$$\boxed{\mathbb{P}(E \cap F) = \mathbb{P}(E)\,\mathbb{P}(F)} \tag{2}$$

Note. (2) and (1) are equivalent.

Note 1. It is clear from (2) that independence is a symmetric relationship. Also, (2) is properly defined when $\mathbb{P}(F) = 0$.

Note 2. (1) gives a useful way to think about independence; (2) is usually better to do the math.

Proposition. If E and F are independent, then so are E and F^c.

Proof.

$$\Pr(F \mid E) = \Pr(F)$$

$$\begin{aligned} \Pr(F^c \mid E) &= 1 - \Pr(F \mid E) \\ &= 1 - \Pr(F) \\ &= \Pr(F^c) \end{aligned}$$

Example 1: Independence can be obvious

Draw a card from a shuffled deck of 52 cards. Let E = card is a spade and F = card is an ace. Are E and F independent?

Solution

$$\Pr(E \cap F) = 0$$

So no...

Example 2: Independence can be surprising

Toss a coin 3 times. Define

A={at most one T}=$\{HHH, HHT, HTH, THH\}$

B={both H and T occur}=$\{HHH, TTT\}^c$.

Are A and B independent?

Solution

$$\Pr(A) = \frac{4}{8} = \frac{1}{2}$$

$$\Pr(B) = \frac{6}{8} = \frac{3}{4}$$

$$\Pr(A \cap B) = \Pr(\{HHT, HTH, THH\})$$

$$= \frac{3}{8} = \Pr(A) \cdot \Pr(B)$$

Independence as an Assumption

- It is often convenient to suppose independence. People sometimes assume it without noticing.

Example. A sky diver has two chutes. Let

$E = \{\text{main chute opens}\}, \qquad \mathbb{P}(E) = 0.98;$

$F = \{\text{backup opens}\}, \qquad \mathbb{P}(F) = 0.90.$

Find the chance that at least one opens, **making any necessary assumption clear**.

$$1 - Pr(\bar{E} \cap \bar{F}) = 1 - 0.02 \cdot 0.1$$

$$= 0.998$$

Note. Assuming independence does not justify the assumption! Both chutes could fail because of the same rare event, such as freezing rain.

Independence of Several Events

- Three events E, F, G are **independent** if

$$\begin{aligned}
\mathbb{P}(E \cap F) &= \mathbb{P}(E) \cdot \mathbb{P}(F) \\
\mathbb{P}(F \cap G) &= \mathbb{P}(F) \cdot \mathbb{P}(G) \\
\mathbb{P}(E \cap G) &= \mathbb{P}(E) \cdot \mathbb{P}(G) \\
\mathbb{P}(E \cap F \cap G) &= \mathbb{P}(E) \cdot \mathbb{P}(F) \cdot \mathbb{P}(G)
\end{aligned}$$

- If E, F, G are independent, then E will be independent of any event formed from F and G.

Example. Show that E is independent of $F \cup G$.

Proof.

$$\Pr(F \cup G \mid E) = \Pr(F \mid E) + \Pr(G \mid E) - \Pr(F \cap G \mid E)$$

$$= \Pr(F) + \Pr(G) - \Pr(F \cap G)$$

$$= \Pr(F \cup G)$$

Pairwise Independence

- E, F and G are **pairwise independent** if E is independent of F, F is independent of G, and E is independent of G.

Example. Toss a coin twice. Set $E = \{HH, HT\}$, $F = \{TH, HH\}$ and $G = \{HH, TT\}$.

(a) Show that E, F and G are pairwise independent.

$\Pr(E \cap F) = \frac{1}{4} = \Pr(E)\Pr(F)$ etc...

(b) By considering $\mathbb{P}(E \cap F \cap G)$, show that E, F and G are NOT independent.

$\Pr(E \cap F \cap G) = \frac{1}{4} \neq \left(\frac{1}{2}\right)^3$

Note. Another way to see the dependence is that $\mathbb{P}(E \mid F \cap G) = 1 \neq \mathbb{P}(E)$.

Example: Insurance policies

Insurance companies categorize people into two groups: accident prone (30%) or not. An accident prone person will have an accident within one year with probability 0.4; otherwise, 0.2. What is the conditional probability that a new policyholder will have an accident in his second year, given that the policyholder has had an accident in the first year?

$$\Pr(\text{acc in } 2^{nd} \text{ year} \mid \text{acc in first})$$

$$= \frac{\Pr(\text{acc in both})}{\Pr(\text{acc in first})}$$

$$= \frac{\Pr(\text{acc in both} \mid AP)\Pr(AP) + \Pr(\text{acc in both} \mid \overline{AP})\Pr(\overline{AP})}{\Pr(\text{acc in first} \mid AP)\Pr(AP) + \Pr(\text{acc in both} \mid \overline{AP})\Pr(\overline{AP})}$$

$$= \frac{(0.4)^2 \cdot .3 + (0.2)^2 \cdot 0.7}{0.4 \cdot .3 + 0.2 \cdot 0.7}$$

Note that we are assuming that the event of a person having an accident this year is indep of the event of having an accident the following year

Note: We can study a probabilistic model and determine if certain events are independent or we can define our probabilistic model via independence.

Example: Supposed a biased coin comes up heads with probability p, independent of other flips

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

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

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

$\mathbb{P}(\text{HHTHTTT}) = p^2(1-p)p(1-p)^3 = p^{\sharp \text{H}}(1-p)^{\sharp \text{T}}$