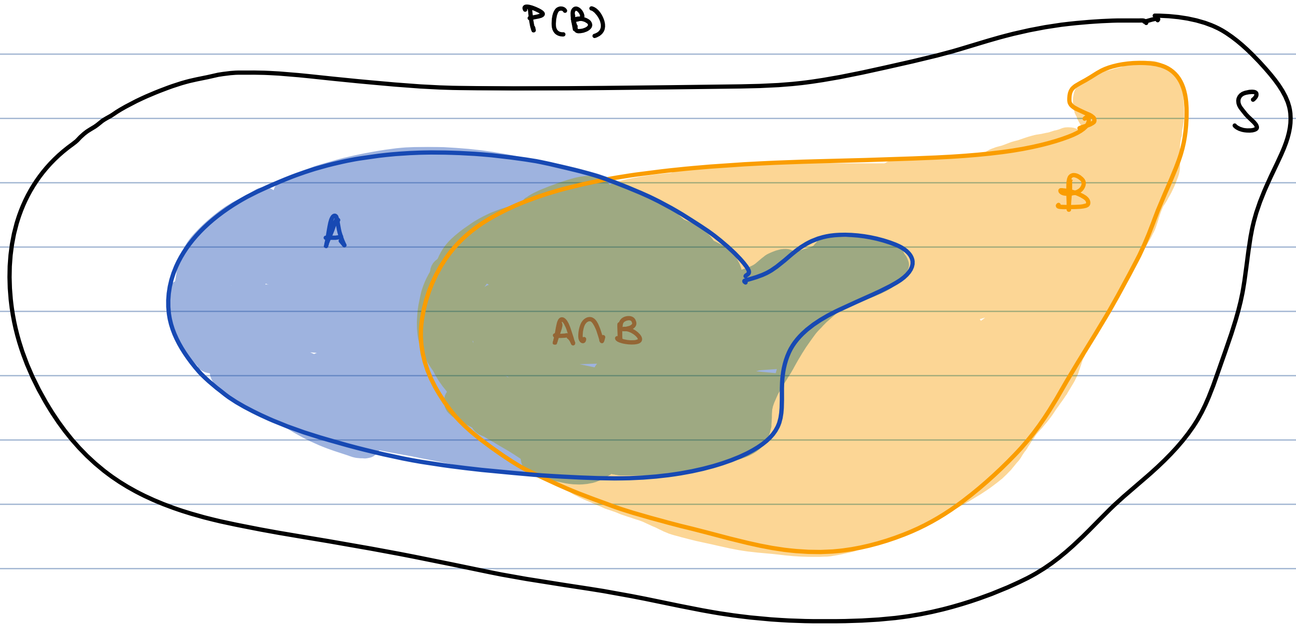
Conditional Probability

Definition: Let (S, P) be a probability space and let B be an event for which $P(B) > 0$. Then for any event $A \subset S$ we define the probability of A, given B, denoted $P(A|B)$ as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



Example: A bit string of length four is generated at random so that each of the 16 bit strings of length four is equally likely. What is the probability that it contains at least two consecutive 0's, given that its bit is a 0?

Solution:

The sample space S is:

$$S = \{x_1x_2x_3x_4 \mid x_i \in \{0,1\},\ i = 1,2,3,4\}$$

Let

$$B = \{abcd \in S \mid a = 0\},$$

and

$$A = \{abcd \in S \mid abcd \text{ contains two consecutive 0's}\}$$

Then

1. $|B| = 8$
2. $A \cap B = \{00cd \in S \mid c, d \in \{0,1\}\} \cup \{0100\}$

It is clear that $|\{00cd \in S \mid c,d \in \{0,1\}\}| = 4$
Thus $|A \cap B| = 4+1 = 5$.
This means
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{5/16}{8/16} = \frac{5}{8}. \quad //$$

Example: What is the probability a family with two children has two boys, given that they have at least one boy.
Solution:
The sample space S can be written as
$$S = \{GG, GB, BG, BB\}.$$
Let $E = \{BB\}$ and $F = \{GB, BG, BB\}$.
We have been asked to compute $P(E|F)$.
$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{1/4}{3/4} = \frac{1}{3}. \quad //$$

The Monty-Hall Three-Door Puzzle

The following problem seemingly involves conditional probability, but in actual fact doesn't.

Suppose you are a game show contestant. You have a chance to win a large prize. You are asked to select one of three doors to open; the large prize is behind one of the doors and the other two doors have nothing behind them. The prize is yours if you (finally) choose the door which has the prize behind it. The game is played as follows: Once you choose a door, the game show host opens one of the two other doors and this is a "losing door", i.e. one of the doors behind which there is nothing. Then he offers you a chance to switch your choice. What is a better strategy — switching doors or not switching doors?

SOLUTION ON NEXT PAGE ⟶

Solution: This is not really a conditional probability problem though it is seemingly so.

The probability that you picked a door behind which there is nothing is $2/3$.

If you picked a losing door then on switching you will get the prize. So the probability of winning if you switch is $2/3$.

The only way you can win if you don't switch is if you picked the winning door. The probability of that is $1/3$.

Thus it is better to switch. //

New probability spaces from old

Let (S, P) be a probability space and

$$f : S \longrightarrow T$$

a function from S to another finite set T.

Define

$$P^* : \mathcal{P}(T) \longrightarrow [0,1]$$

by the rule

$$P^*(D) = P(f^{-1}(D)), \quad \forall D \subset T.$$

It is then easy to see that P^* is a probability measure on T, and hence (T, P^*) is a probability space.

One can do something very similar when T is not finite. In that case, we know that $f(S)$ is finite. Let T' be any finite subset of T containing $f(S)$ (e.g. take $T' = f(S)$). Then f can be regarded as a map from S to T' and hence as above we can construct a probability space (T', P^*) with $P^*(D) = P(f^{-1}(D))$ $\forall$ subsets D of T'.

Random Variables

Let (S,P) be a probability space. A random variable (on (S,P)) is a function

$$X: S \longrightarrow \mathbb{R}.$$

From our discussion above, if S' is any finite set containing $X(S)$, $P_X : \mathcal{P}(S') \longrightarrow [0,1]$ the map $P_X(D) = P(X^{-1}(D))$ for all subsets D of S', then (S', P_X) is also a probability space. The nice thing is that now S' is a subset of the real line.

The Expectation of a random variable

Let (S,P) be a probability. Let

$$S = \{s_1, s_2, \dots, s_n\}$$

and let

Recall: This is shorthand for $P(\{s_1\})$.

$$p_1 = P(s_1),\ p_2 = P(s_2),\ \dots,\ p_n = P(s_n)$$

Let

$$X: S \longrightarrow \mathbb{R}$$

be a random variable. The expectation of X, denoted $E(X)$, is

$$E(X) = p_1 X(s_1) + p_2 X(s_2) + \dots + p_n X(s_n).$$

The intuition is: Suppose the experiment that produces the outcomes in S according to the probability law P is repeated N times, where N is a large number. Since the probability that s_i is the outcome is p_i, the proportion of s_i occuring in the N experiments is approximately p_i, and hence:

The number of times s_i occurs in the N experiments $\approx p_i \cdot N$.

Suppose the outcome of the j^{th} experiment is $\sigma_j \in S$. Then

$$X(\sigma_1) + \dots + X(\sigma_N) \approx N \cdot \sum_{i=1}^{N} p_i X(s_i) \text{ (why?)}$$

This implies that the mean of the $X(\sigma_i)$, $i = 1, \dots, N$ is

$$\frac{X(\sigma_1) + \dots + X(\sigma_N)}{N} \approx \sum_{i=1}^{N} p_i X(s_i) = E(X)$$