CONDITIONAL PROBABILITY, INDEPENDENCE, AND THE BIRTHDAY PARADOX

Lecture 19
Conditional Probability

**Definition:** Let $E$ and $F$ be events with $p(F) > 0$. The conditional probability of $E$ given $F$, denoted by $P(E|F)$, is defined as:

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

**Example:** A bit string of length four is generated at random so that each of the 16 bit strings of length 4 is equally likely.

What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0?
Conditional Probability

Example: A bit string of length four is generated at random so that each of the 16 bit strings of length 4 is equally likely. What is the probability that it contains at least two consecutive 0’s, given that its first bit is a 0?

Solution: Let $E$ be the event that the bit string contains at least two consecutive 0s, and $F$ be the event that the first bit is a 0.

- Since $E \cap F = \{0000, 0001, 0010, 0011, 0100\}$, $p(E \cap F) = 5/16$.
- Because 8 bit strings of length 4 start with a 0, $p(F) = 8/16 = \frac{1}{2}$.

Hence, $p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{5/16}{1/2} = \frac{5}{8}$. 
Conditional Probability

Example: What is the conditional probability that a family with two children has two boys, given that they have at least one boy.
Assume that each of the possibilities BB, BG, GB, and GG is equally likely where B represents a boy and G represents a girl.

Solution: Let $E$ be the event that the family has two boys and let $F$ be the event that the family has at least one boy. Then $E = \{BB\}$, $F = \{BB, BG, GB\}$, and $E \cap F = \{BB\}$.

- It follows that $p(F) = 3/4$ and $p(E \cap F) = 1/4$.

Hence,

$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{1/4}{3/4} = \frac{1}{3}.$$
Independence

Definition:
Events $E$ and $F$ are independent if and only if

$$p(E \cap F) = p(E)p(F).$$

Which events are independent?
1. H H on two consecutive coin tosses.
2. 4 2 on two consecutive rolls of a 6-sided die.
3. Choosing a card from a 52-card deck. Put back and choose a second card. Choosing a jack and then an 8.
4. Choosing a card from a 52-card deck. Not put back and choose a second card. Choosing a jack and then an 8.
3. Choosing a card from a 52-card deck. Put back and choose a second card. Choosing a jack and then an 8.

\[
P(\text{jack}) = \frac{4}{52} \\
P(8) = \frac{4}{52} \\
P(\text{jack and } 8) = \frac{4}{52} \cdot \frac{4}{52}
\]

4. Choosing a card from a 52-card deck. \textbf{Not} put back and choose a second card. Choosing a jack and then an 8.

\[
P(\text{jack and } 8) = \frac{4}{52} \cdot \frac{4}{51}
\]
Independence

Example: $E$ is the event that a randomly generated bit string of length four begins with a 1. $F$ is the event that the bit string contains an even number of 1s. Are $E$ and $F$ independent if the 16 bit strings of length four are equally likely?

Solution: There are 8 bit strings of length 4 that begin with a 1, and 8 bit strings of length 4 that contain an even number of 1s. Since the number of bit strings of length 4 is 16,

$$p(E) = p(F) = \frac{8}{16} = \frac{1}{2}.$$ 

Since $E \cap F = \{1111, 1100, 1010, 1001\}$, $p(E \cap F) = \frac{4}{16} = \frac{1}{4}$. Events $E$ and $F$ are independent, because

$$p(E \cap F) = \frac{1}{4} = \left(\frac{1}{2}\right) \left(\frac{1}{2}\right) = p(E) p(F)$$
Independence

**Example:** Assume that each of the four ways a family can have two children (BB, GG, BG, GB) is equally likely. Are the events $E$, that a family with two children has two boys, and $F$, that a family with two children has at least one boy, independent?

**Solution:** Because $E = \{BB\}$, $p(E) = 1/4$. We saw previously that $p(F) = 3/4$ and $p(E \cap F) = 1/4$. The events $E$ and $F$ are not independent since

$$p(E) p(F) = 3/16 \neq 1/4 = p(E \cap F).$$
Bernoulli Trials

Definition: Suppose an experiment can have only two possible outcomes, e.g., the flipping of a coin or the random generation of a bit.

- Each performance of the experiment is called a Bernoulli trial.
- One outcome is called a success and the other a failure.
- If $p$ is the probability of success and $q$ the probability of failure, then $p + q = 1$.

Many problems involve determining the probability of $k$ successes when an experiment consists of $n$ mutually independent Bernoulli trials.
Bernoulli Trials

Example: A coin is biased so that the probability of heads is $2/3$.
What is the probability that exactly four heads occur when the coin is flipped seven times?

Solution: There are $2^7 = 128$ possible outcomes.
The number of ways four of the seven flips can be heads is $C(7,4)$.
The probability of each of the outcomes is $(2/3)^4(1/3)^3$ since the seven flips are independent.
Hence, the probability that exactly four heads occur is

$$C(7,4) \cdot (2/3)^4 \cdot (1/3)^3 = (35 \cdot 16)/27 = 560/2187.$$
Probability of \( k \) Successes in \( n \) Independent Bernoulli Trials.

**Theorem 2**: The probability of exactly \( k \) successes in \( n \) independent Bernoulli trials, with probability of success \( p \) and probability of failure \( q = 1 - p \), is \( C(n,k)p^kq^{n-k} \).

**Proof**: The outcome of \( n \) Bernoulli trials is an \( n \)-tuple \((t_1,t_2,\ldots,t_n)\), where each is \( t_i \) either \( S \) (success) or \( F \) (failure).

The probability of each outcome of \( n \) trials consisting of \( k \) successes and \( n-k \) failures (in any order) is \( p^kq^{n-k} \).

There are \( C(n,k) \) \( n \)-tuples of \( S \)’s and \( F \)’s that contain exactly \( k \) \( S \)’s.

Hence, the probability of \( k \) successes is \( C(n,k)p^kq^{n-k} \).
The Famous Birthday Problem

How many people are needed in a room to ensure that the probability of at least two of them having the same birthday is more than ½?

Solution: Assume all birthdays are equally likely and a year has 366 days.

Imagine people entering the room one by one. The probability that at least two have the same birthday is $1 - p_n$.

- The probability that the birthday of the second person is different from that of the first is $365/366$.
- The probability that the birthday of the third person is different from the other two, when these have two different birthdays, is $364/366$.

In general, the probability that the $j$th person has a birthday different from those already in the room, assuming that these people all have different birthdays, is $(366 - (j - 1))/366 = (367 - j)/366$. 

The Famous Birthday Problem (2)

How many people are needed in a room to ensure that the probability of at least two of them having the same birthday is more than $\frac{1}{2}$?

Hence, $p_n = \frac{365}{366} \cdot \frac{364}{366} \cdots \frac{367 - n}{366}$.

Therefore, $1 - p_n = 1 - \frac{365}{366} \cdot \frac{364}{366} \cdots \frac{367 - n}{366}$.

Checking various values for $n$ with computation help tells us that for $n = 22$, $1 - p_n \approx 0.457$, and for $n = 23$, $1 - p_n \approx 0.506$.

Consequently, a minimum number of 23 people are needed so that the probability that at least two of them have the same birthday is greater than $1/2$. 