Math-UA.233: Theory of Probability
Lecture 7

Tim Austin

tim@cims.nyu.edu
cims.nyu.edu/~tim

Independence (Ross section 3.4)

Suppose we roll two dice. Let E and F be the events of getting a six from the first and second die, respectively.

If we assume equally likely outcomes, then

$$P(E \cap F) = P(\{(6,6)\}) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(E)P(F).$$

Another way to describe this modeling assumption:

- First: all faces of the first die are equally likely.
- Second: all faces of the second die remain equally likely, no matter what the first die gave, and therefore

$$\underbrace{P(E \cap F) = P(E)P(F \mid E)}_{\text{multiplication rule}} = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

Put another way, we are assuming

(i) that each die by itself gives equally likely outcomes, and
(ii) that the first die has no ‘influence’ on the second die.

Mathematical expression of assumption (ii):

> *The conditional probabilities for the second die given the value of the first die are the same as the unconditioned probabilities.*

This idea is very important in general.

Definition

Given any sample space S and probability distribution P, two events E and F are ***independent*** *if*

$$\boxed{P(E \cap F) = P(E)P(F).}$$

We may also say E ***is independent of*** *F (or that F is* ***independent of*** *E).*

If this does <u>not</u> hold, then E and F are ***dependent****.*

OBSERVE: By the multiplication rule, this is equivalent to

$$P(E \mid F) = P(E) \quad \text{(provided } P(F) > 0\text{)}$$

and also to

$$P(F \mid E) = P(F) \quad \text{(provided } P(E) > 0\text{)}.$$

Example

An urn contains three red and three blue balls. If two balls are drawn out at random...

(a) *... with replacement, then their colours are independent.*
(b) *... with<u>out</u> replacement, then their colours are <u>not</u> independent.*

Very often, an assumption of independence is part of choosing a probability model.

We typically assume that two events are independent if we believe that "what causes E to happen is disconnected from what causes F to happen" — as in the example of the two dice.

But independence in probability theory is not really an assertion about causes. It's just the equality between the numbers $P(E \cap F)$ *and* $P(E)P(F)$*. We might assume it because of some story that we tell about causes, but it might just happen anyway.*

Example (Similar to Ross E.g. 4c)

In the rolling of two fair dice, let E be the event that the first die shows 6, let G be the event that the sum of the values is 12, and let H be the event that the sum of the values is 7. Are E and G independent? Or E and H? Or G and H?

ANS = NO; YES; NO

Note that, even though E and H are independent, in this case it isn't because we can 'separate their underlying causes', or anything like that. It just happens to be true.

Example (Ross E.g. 3.4a)

A card is selected at random from an ordinary deck of 52 cards. If E is the event that the selected card is an ace and F is the event that it is a spade, then E and F are independent.

Worth knowing:

Proposition (Ross Prop 3.4.1)

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

STORY: So if E and F are independent, then the probability of E occurring is unaffected by the knowledge of whether *or not* F occurs.

This is a good time for a GENERAL WARNING:

DISJOINTNESS AND ***INDEPENDENCE*** are ***DIFFERENT***.

Disjoint: E and F cannot both occur.

Independent: E does not tell us anything about whether F occurs.

Digression: When should we assume that two events are independent?

This is another modeling question, rather than a math one.

Sometimes we assume independence as an approximation, rather than an exact statement about causes. This is very important in opinion polling. The reliability of a poll depends on selecting a sample from the population in a very 'unbiased' way.

What 'unbiased' means here: we must select people to ask in a random way so that the events

$$E = \{\text{the person responds to the opinion poll}\}$$

and

$$F = \{\text{the person holds opinion X}\}$$

are *independent*.

Example

Before the 1936 presidential election, Literary Digest magazine mailed a questionnaire to 10 million people, and got 2.4 million responses. Based on those, they predicted that Alf Landon would win the election with 57% of the vote.

What's odd about this?

Independence for several events

Next we need to generalize independence to more than two events. Before doing that, we start with a cautionary example.

Example (Ross E.g. 4e)

Two fair dice are rolled. Let E be the event that the sum is 7, F the event that the first die shows 4, and G the event that the second die shows 3.

Then E and F are independent, and E and G are independent, and F and G are independent.

BUT $F \cap G \subset E$, so $P(E \mid F \cap G) = 1$, so E is <u>not</u> independent of $F \cap G$.

So if we have three events E, F and G, and we know that any two of them are independent, there may still be other combinations of them which are *de*pendent. To extend the notion of independence to a triple of events, we explicitly rule this out.

Definition

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

$$\begin{aligned} P(E \cap F) &= P(E)P(F) \\ P(E \cap G) &= P(E)P(G) \\ P(F \cap G) &= P(F)P(G) \\ \underline{\text{and}} \quad P(E \cap F \cap G) &= P(E)P(F)P(G). \end{aligned}$$

Why these four assumptions?

It turns out that, once we have these four, we can deduce that

> *E is independent from any other event that can be defined just in terms of F and G.*

For example,

Proposition

If E, F and G are independent, then E is independent of $F \cap G^c$.

(More challenging example for you to think about: E is independent of $F \cup G$. Still more challenging: give a general statement and proof.)

Here's the generalization to an arbitrary sequence of events:

Definition

A (finite or infinite) sequence of events $E_1, E_2, \ldots$ is ***independent*** *if we have*

$$P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m}) = P(E_{i_1})P(E_{i_1}) \cdots P(E_{i_m}).$$

for any choice of finitely many indices $i_1 < i_2 < \cdots < i_m$.

That is, "for any intersection that we can form from our sequence, its probability factorizes".

(In the case of an infinite sequence, we require this only for choices of finitely many indices. In fact this implies a version for infinitely many, but we leave that aside.)

As with pairs of events, we often assume independence for several events because they are all caused by different things.

But independence is not *equivalent* to that assumption: probability theory knows nothing about causes, only about whether numbers are equal or not.

Example (Ross E.g. 4g)

A system composed of n separate components is called 'parallel' if it functions whenever at least one of those components is working. Suppose that component i is working with probability p_i, $i = 1, 2, \ldots, n$, and that these events are independent. What is the probability that the system functions?

The previous example used a fairly useful general formula, which it's worth explaining.

Proposition

If $E_1, \ldots, E_k$ are independent, then

$$P(E_1 \cup \cdots \cup E_k) = 1 - (1 - P(E_1))(1 - P(E_2)) \cdots (1 - P(E_k)).$$

(See Pishro-Nik p49.)

NOTICE: This is nothing like the rule for *disjoint* events.

(... because ***DISJOINTNESS*** AND ***INDEPENDENCE*** are ***DIFFERENT***.)

‘Trials’

Many experiments consist of a finite or infinite sequence of smaller ‘subexperiments’, and we assume that the outcomes of those subexperiments do not influence each other: for example, repeatedly tossing a coin.

In this setting, if $E_1, E_2, \ldots$ are events for which E_i depends only on the outcome of the i^{th} subexperiment, then it would be natural to assume that this sequence is independent.

If the subexperiments are repeatedly doing the same thing, so they all have the same possible outcomes, then these subexperiments are often called **trials**.

Example (Ross E.g. 4f)

An infinite sequence of independent trials is performed. Each trial results in a success with probability p and a failure with probability $1 - p$. What is the probability that

(a) *at least 1 success occurs in the first n trials;*

(b) *exactly k successes occur in the first n trials;*

(c) *all (infinitely many) trials result in successes?*

Example (Ross E.g. 4j: The problem of the points)

Independent trials are performed, each resulting in success with probability p and failure with probability $1 - p$. Given positive whole numbers n and m, what is the probability that n successes occur before m failures?

$$\text{ANS} = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}$$

Interesting variant: service protocol in a serve and rally game (Ross E.g. 3.4k)

Example (Ross E.g. 4m: The gambler's ruin)

A and B, bet on the outcomes of successive flips of a p-biased coin. A takes $1 *from B if the coin lands heads, and vice-versa if it lands tails. This continues until one of them runs out of money. What is the probability* P_i *that A wins if s/he starts with* $\$i$ *and B starts with* $\$(n-i)$*?*

IDEA: Condition on the outcome of the first flip. This gives

$$P_i = pP_{i+1} + (1-p)P_{i-1} \quad \text{for } 1 \leq i \leq n-1.$$

Solving this gives

$$P_i = \begin{cases} \frac{1-((1-p)/p)^i}{1-((1-p)/p)^n} & \text{if } p \neq 1/2 \\ \\ i/N & \text{if } p = 1/2 \end{cases}$$

Another use of this idea: finding the prob of *n consecutive* successes before *m consecutive* failures (Ross E.g. 3.5c).

Another nice practice example:

Example (Pishro-Nik E.g. 1.23)

Two players take turns to shoot a basketball. The first one to make a basket wins. On each shot, Player 1 (who goes first) has probability p_1 of success, and Player 2 has probability p_2. Shots are assumed independent.

(a) *Find $P(\text{Player 1 wins})$.*

(b) *For what values of p_1 and p_2 do they have equal probability of winning?*