Stat 400, section 2.3 & 2.4 Tree Diagrams and Conditional Probability

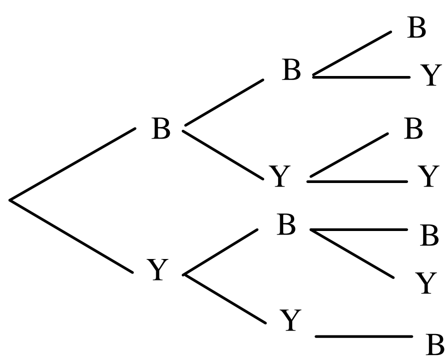
notes by Tim Pilachowski, Spring 2011

One of your classmates asked a very good question after discussion: “How do I know when to multiply and when to add probabilities?” I develop an answer in the notes below.

A tree diagram can be a useful tool for keeping track of the events in a sample space and their probabilities. My Example is relatively simple. For some of the more complicated ones like the textbook homework, it may not be possible to actually draw the tree diagram, but if you can *start* to visualize what it would look like, then the same concepts can be applied.

First approach: A box contains 3 blue blocks and 2 yellow blocks. You pick three blocks without replacement. One way to create this sample space is to designate the block by color and number: B1, B2, B3, Y1, Y2. Then we’d have S = { B1-B2-B3, B1-B2-Y1, B1-B2-Y2, B1-Y1-B2, B1-Y1-B3, B1-Y1-Y2, etc., etc., etc., … } While the outcomes would be equally likely, the process would be extremely cumbersome. Imagine having 333 blue and 222 yellow!

Second approach: A box contains 3 blue blocks and 2 yellow blocks. You pick three blocks without replacement. We’ll consider just blue and yellow colors, and construct a tree diagram.

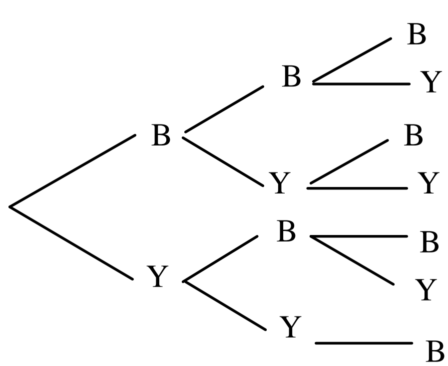


Trial: picking three blocks
Outcomes:

BBB
BBY
BYB
BYY
YBB
YBY
YYB

This is a permutation-type model—the sample space has 7 *mutually exclusive* outcomes: S = { BBB, BBY, BYB, BYY, YBB, YBY, YYB }. Note that for the lowest branch, once both yellow blocks have been picked, the only possibility left is a blue block. Also note that these 7 outcomes are *not* equally likely.

Third approach: A box contains 3 blue blocks and 2 yellow blocks. You pick three blocks without replacement.



Trial: picking three blocks
Events:

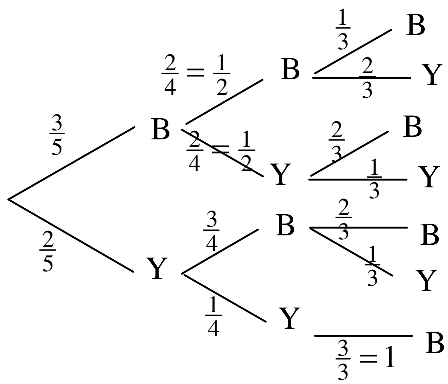
E_1: all three are blue
E_2: 2 blue, 1 yellow
E_2: 2 blue, 1 yellow
E_3: 1 blue 2 yellow
E_2: 2 blue, 1 yellow
E_3: 1 blue, 2 yellow
E_3: 1 blue, 2 yellow

This is a combination-type model—the order of blocks picked does not count. Note that $n(E_1) = 1$, $n(E_2) = 3$, $n(E_3) = 3$. If we were to define E_4 = picking three yellow blocks, then $n(E_4) = 0$.

In a **conditional probability** an outcome or event E is dependent upon another outcome or event F.

A box contains 3 blue blocks and 2 yellow blocks. The tree diagram for randomly picking three blocks **without replacement**, with associated probabilities, would look like this:

Event: picking three blocks
Outcomes:



all three are blue
2 blue, 1 yellow
2 blue, 1 yellow
1 blue 2 yellow
2 blue, 1 yellow
1 blue, 2 yellow
1 blue, 2 yellow

Note the changes in probabilities for the second and third picks. This is because the second pick is *dependent* upon the first pick, (and the third pick will be dependent upon both the first and second picks).

More formally, we can define events B_1 = blue on the first pick, B_2 = blue on the second pick, B_3 = blue on the third pick, Y_1 = yellow on the first pick, Y_2 = yellow on the second pick, Y_3 = yellow on the third pick.

For the second pick:

P(picking blue second given a blue was picked first) $= P(B_2 \mid B_1) = \dfrac{\text{2 blue left}}{\text{4 blocks left}} = \dfrac{1}{2}$

$P(Y_2 \mid B_1) = \dfrac{\text{2 yellow left}}{\text{4 blocks left}} = \dfrac{1}{2}$, $P(B_2 \mid Y_1) = \dfrac{\text{3 blue left}}{\text{4 blocks left}} = \dfrac{3}{4}$, $P(Y_2 \mid Y_1) = \dfrac{\text{1 yellow left}}{\text{4 blocks left}} = \dfrac{1}{4}$

For the third pick:

$P(B_3 \mid B_1 \cap B_2) = \dfrac{\text{1 blue left}}{\text{3 blocks left}} = \dfrac{1}{3}$, $P(Y_3 \mid B_1 \cap B_2) = \dfrac{\text{2 yellow left}}{\text{3 blocks left}} = \dfrac{2}{3}$,

$P(B_3 \mid B_1 \cap Y_2) = \dfrac{\text{2 blue left}}{\text{3 blocks left}} = \dfrac{2}{3}$, $P(Y_3 \mid B_1 \cap Y_2) = \dfrac{\text{1 yellow left}}{\text{3 blocks left}} = \dfrac{1}{3}$,

$P(B_3 \mid Y_1 \cap B_2) = \dfrac{\text{2 blue left}}{\text{3 blocks left}} = \dfrac{2}{3}$, $P(Y_3 \mid Y_1 \cap B_2) = \dfrac{\text{1 yellow left}}{\text{3 blocks left}} = \dfrac{1}{3}$,

$P(B_3 \mid Y_1 \cap Y_2) = \dfrac{\text{3 blue left}}{\text{3 blocks left}} = 1$

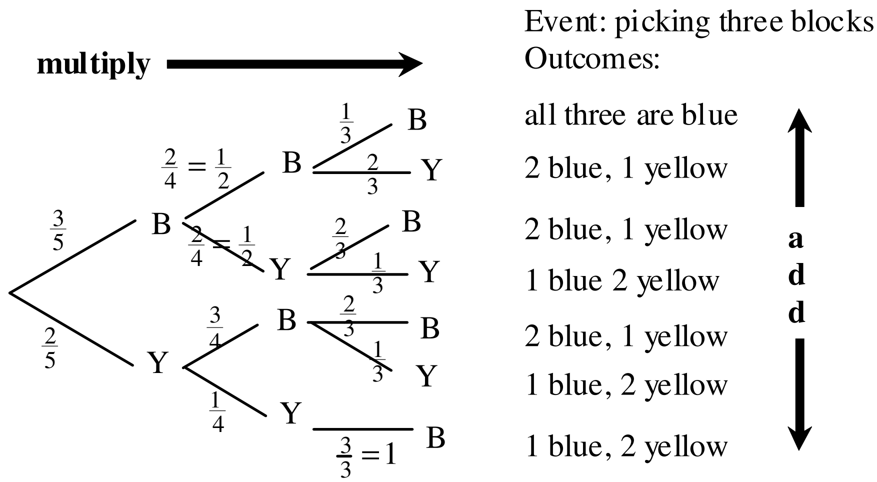
Once both yellow blocks have been picked, the probability of picking a blue block is a certainty., i.e. given a yellow block picked first and second, then the probability of picking a blue block third = 1.

$P(Y_3 \mid Y_1 \cap Y_2) = \dfrac{\text{0 yellow left}}{\text{3 blocks left}} = 0$

Once both yellow blocks have been picked, the probability of picking a yellow block is an impossibility, i.e. given a yellow block picked first and second, P(picking a yellow block third) = 0.

Now, finally, we get the answer to the original question.

When moving left to right on the tree diagram, from root to branch, the probabilities are multiplied.



The probability of picking the permutation $B_1B_2Y_3$ is calculated as follows:

$$P(B_1B_2Y_3) = P(\text{blue on first and blue on second and yellow on third}) = \frac{3}{5}*\frac{1}{2}*\frac{2}{3}=\frac{1}{5}.$$

$$\text{Formally } P(B_1 \cap B_2 \cap B_3) = P(B_1)*P(B_2 \mid B_1)*P(B_3 \mid B_1 \cap B_2)$$

In an “intersection”, i.e. an “and” situation, moving left-to-right on the tree diagram, multiply probabilities.

When moving up and down on the tree diagram, from branch to branch, the probabilities are added.

The probability of ending up with 2 blue blocks and 1 yellow block (represented by branches 2, 3 and 5 from the top) is calculated as follows:

$$\begin{aligned} P(\text{2 blue, 1 yellow}) &= P(\text{blue on first and blue on second and yellow on third}) \\ &\quad \text{or } P(\text{blue on first and yellow on second and blue on third}) \\ &\quad \text{or } P(\text{yellow on first and blue on second and blue on third}) \\ &= \left(\frac{3}{5}*\frac{1}{2}*\frac{2}{3}\right) + \left(\frac{3}{5}*\frac{1}{2}*\frac{2}{3}\right) + \left(\frac{2}{5}*\frac{3}{4}*\frac{2}{3}\right) \\ &= \frac{1}{5} + \frac{1}{5} + \frac{1}{5} = \frac{3}{5} \end{aligned}$$

$$\text{Formally } P(B_1B_2Y_3 \cup B_1Y_2B_3 \cup Y_1B_2B_3) = P(Y_3 \mid B_1 \cap B_2) + P(B_3 \mid B_1 \cap Y_2) + P(B_3 \mid Y_1 \cap B_2)$$

Similarly, the probability of ending up with 1 blue block and 2 yellow blocks (represented by branches 4, 6 and 7 from the top) is calculated as follows:

$$\begin{aligned} P(\text{1 blue, 2 yellow}) &= P(\text{blue on first and yellow on second and yellow on third}) \\ &\quad \text{or } P(\text{yellow on first and blue on second and yellow on third}) \\ &\quad \text{or } P(\text{yellow on first and yellow on second and blue on third}) \\ &= \left(\frac{3}{5}*\frac{1}{2}*\frac{1}{3}\right) + \left(\frac{2}{5}*\frac{3}{4}*\frac{1}{3}\right) + \left(\frac{2}{5}*\frac{1}{4}*1\right) \\ &= \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = \frac{3}{10} \end{aligned}$$

$$\text{Formally } P(B_1Y_2Y_3 \cup Y_1B_2Y_3 \cup Y_1Y_2B_3) = P(Y_3 \mid B_1 \cap Y_2) + P(Y_3 \mid Y_1 \cap B_2) + P(B_3 \mid Y_1 \cap Y_2)$$

In a “union”, i.e. an “or” situation, moving up-and-down on the tree diagram, add probabilities.

Final note and observation: The sum of the conditional probabilities in each "column" of the tree diagram add up to 1, i.e. the two possibilities for the first pick are the sample space for the first pick, the four possibilities for the (first and second) pick are the sample space for the (first and second) pick, and the seven possibilities for the (first and second and third) pick are the sample space for the (first and second and third) pick. Calculate for yourself to show that each of the sums below adds up to 1.

$P(B_1) + P(Y_1) =$

$P(B_1 \cap B_2) + P(B_1 \cap Y_2) + P(Y_1 \cap B_2) + P(Y_1 \cap Y_2) =$

$P(B_1 B_2 B_3) + P(B_1 B_2 Y_3) + P(B_1 Y_2 B_3) + P(B_1 Y_2 Y_3) + P(Y_1 B_2 B_3) + P(Y_1 B_2 Y_3) + P(Y_1 Y_2 B_3) =$

Summary:

multiply	add
intersection	union
"and"	"or"
left-to-right on tree diagram	up-and-down on tree diagram