Classical Probability examples.

1-9 A red die has face numbers {2, 4, 7, 12, 5, 11}.
A green die has numbers {1, 3, 4, 6, 8, 9, 10}.

1. Write out a 6 by 6 table of all 36 possible outcomes when the two are thrown.

2. Place a "2" in every cell for which $R > G$ and determine $P(R > G)$.

3. Determine $P(R > G \mid G = 3)$ by counting the relative number of cases of $R > G$ among the cases of $G = 3$.

4. Verify that #3 is the same as $P((R > G) \cap (G=3))/P(G=3)$.

5. Events "$R > G$" and "$G = 3$" are said to be INDEPENDENT if your answers to questions #2 and #3 are the SAME. It would mean that the chance of $R > G$ would not be changed if we know that $G = 3$. Are the events independent?

6. Mark with "*" every cell with $R > G$. Mark with "×" every cell with $G = 3$. List the outcomes (among the 36 possible outcomes) comprising event $(R > G) \cap (G = 3)$.

7. Refer to #6. List the outcomes comprising event $(R > G) \cup (G = 3)$. Verify that $P((R > G) \cup (G = 3)) = P(R > G) + P(G = 3) - P((R > G) \cap (G = 3))$ as called for by the addition rule for probabilities.

8. Refer to #2. The multiplication rule for probabilities says $P((R > G) \cap (G = 3))$ is equal to $P(G = 3) \, P(R > G \mid G = 3)$. Evaluate these three probabilities and confirm that is works in this case.

9. Event (R > G) may be broken into two non-overlapping events. One overlays (R > G) onto (G = 3) and is expressed (R > G) ⋂ (G = 3). The other overlays (R > G) onto (G ≠ 3) and is expressed (R> G) ⋂ (G ≠ 3). The LAW OF TOTAL PROBABILITY obtains P(R > G) as the sum P((R > G) ⋂ (G = 3)) + P ((R> G) ⋂ (G ≠ 3)). See that it works in this case by evaluating each of the three probabilities involved.

10-16 A box contains 5 R, 4 G, 9 B balls. Balls will be drawn without-replacement and with equal probability from those then remaining in the box.

10. P(B1) (first ball selected is black).

11. Show what the box looks like if you are told that the first ball is black.

12. From #11 determine P(B2 | B1). Think of drawing out of the box you gave in #9.

13. From #11, #12 determine P(B1 ⋂ B2) = P(B1) P(B2 | B1) (multiplication rule). Verify this works for the three probabilities involved. In future we will simply write B1 B2 for the event (B1 ⋂ B2) (i.e. black on first draw followed by black on second draw). We also say "black on draw one AND black on draw 2."

14. Use the method of #13 (multiplication rule) to determine
     P(R1 B2)
     P(G1 B2)
     P(B1 B2) (done in #11)
The three events
     (R1 B2)
     (G1 B2)
     (B1 B2)
are non-overlapping(i.e. they are **disjoint**). Using the **law of total probability** we may therefore **add the three probabilities above** to obtain P(B2).
     P(B2) = P(R1 B2) + P(G1 B2) + P(B1 B2)

Do so, and reduce your answer for P(B2). **You will find that P(B2) is exactly the same as P(B1) so ORDER OF THE DEAL DOES NOT MATTER even though we are drawing without replacement.** You probably behave as though you accept this anyway. After all, if we are dealing cards (without replacement) around a table do you really think that the cards you may be dealt are better if you sit in a different position?

15. Using the definition of conditional probability, P(B1 | B2) = P(B1 B2) / P(B2). You found both of P(B1 B2) and P(B2) in #14. So calculate P(B1 | B2). You will find that it is exactly the same as P(B2 | B1). Once again, in dealing without replacement and with equal probability, **order of the deal does not matter**.

16. Are the events B1, R2 independent? Why?

17-19 Box I = {5R 4G 9B}, box II = {3R 6G 5B}. Another box has {3I 6II}.

17. From the last box select either a I or II using the equal probability model on the nine possibilities. What are P(I), P(II)?

18. Refer to #17. Having selected a box, think of making TWO selections from that box without replacement and with equal probability for those remaining. Determine P(R1) = P(I R1) + P(II R1).

19. Determine P(I | R1) = P(I R1) / P(R1). This is the probability that we made our ball selection from box I if we are told that the selected ball is red. Using the rules of probability in this way, to deduce the probability of some earlier (prior) event when we have learned of some downstream (a-posteriori) event is attributed to Bayes. Bayes' formula or doing this in general is:

$$P(A1 \mid B) = \frac{P(A1\ B)}{P(B)} = \frac{P(A1)\,P(B \mid A1)}{P(B\ A1) + \ldots\ldots + P(B\ Ak)}$$

$$\frac{P(A1)\,P(B \mid A1)}{P(A1)\,P(B \mid A1) + \ldots\ldots + P(Ak)\,P(B \mid Ak)}$$

$$\frac{P(A1\ B)}{P(B)}$$

$$\frac{P(A1)\,P(B\,|\,A1)}{P(A1)\,P(B\,|\,A1) + ... + P(Ak)\,P(B\,|\,Ak)}$$

where A1, ..., Ak are mutually disjoint (i.e. non-overlapping) events whose union is the entire sample space (set of possible outcomes). As seen above, the formula is usually not needed since everything it does is to break things down using

the definition of conditional probability $P(A1\,|\,B) = \dfrac{P(A1\,B)}{P(B)}$

the law of total probability $P(B) = P(B\,A1) + ... + P(B\,Ak)$

the multiplication rule $P(B\,Ai) = P(Ai)\,P(B\,|\,Ai), i = 1, ..., k.$

The formula does not require any chronology actually. It could just as well be used to determine the conditional probability that the extinct Quagga is a type of Zebra if we learn that there are Zebra who rather resemble the Quagga.

Bayes was unsure of the merits of his discovery, which has become extremely important in signal processing, image restoration, character/voice/face recognition, fixing the Hubble telescope, strategies for gambling/investment, setting risk, competitive engagements, procurement, and countless other pressing applications. Not bad for a simple one line formula Bayes may have found startling, yet hauntingly superficial in some sense. How, might he have wondered, could it possibly be important and yet how could it have escaped notice? Anyone might see the formula as only a tautological play on arithmetical rules, a trifling result of pushing around a few symbols. Perhaps true, and yet in our time, when coupled with high speed computing and ideas borrowed from physics and statistics, the formula is a key element in cracking secrets of the genome, economic forecasting, weather forecasting, code breaking, ...

20. **Three doors problem.** This is a stripped-down version of the TV show version. It is an old problem in probability. There are three doors. Behind one door is a prize of great value. There is nothing of value behind the other two doors. The "host of the TV program" invites us to select a door. We choose a door at random with equal probability 1/3 for each door. The host then opens one of the two other doors to show that there is nothing behind it. We are then given the option of sticking with our previ-

ously selected door or switching to the one we did not select and which was not opened by the host. Should we keep our originally selected door or switch? Here is a classical model assuming that door 1 has the prize. The cases in which door 2 or door 3 hold the prize are the same. Our door choice is indicated by *.

prize	empty	empty	keeper wins	switcher wins
*			yes	no
	*		no	yes
		*	no	yes

Confirm that the switcher or keeper wins are correct in each contingency. It is seen that P(keeper wins) = 1/3 whereas P(switcher wins) = 2/3.

Even simpler. We agree that P(keeper wins) = 1/3 since the keeper chooses a door at random and stays with it. The key is that when the keeper wins the switcher loses and visa-versa. So the probability that the switcher wins is, by the rule of complements, 1 - P(keeper wins) = 1 - 1/3 = 2/3.

21-22. **Second guessing.** *Asked to guess the hostess' age I say 34. Dr. Steele says "Oh no, she is younger than that!" I don't know what age HE has in mind, maybe he thinks her age is something very low such as 21, but the damage is done. We find out that he is right, she is younger than 34.*

21. Having learned my lesson (never go on record with a guess of the hostess' age unless it is an underestimate) I will now say my guess was not actually 34. The question posed and answered by Dr. Steele was how often he should be right when he second guesses according to this simple rule:

     a. Get the other person to announce their estimate of some numerical quantity neither of you should be expected to know exactly (e.g. number of pages in a book, weight of a person, tons of steel in a bridge, etc.). Perhaps enlist a third party to pick the question.

     b. Having formed your own estimate, when you hear the first guesser (a) announce theirs simply say "you are too high" or "you are too low" as you judge it.

Here is a stipped-down version of Dr. Steele's solution. This version is set in a classical model which his was not. The model is intended to describe the possibilities for two equally good guessers. Asterisk means that person is closer to the truth. This model allows no ties.

first guesser (targeted for humiliation)	second guesser	2nd is correct
low	low*	yes
low*	low	no
low	high*	yes
low*	high	yes
high	low*	
high*	low	
high	high*	
high*	high	

By "2nd is correct" we mean that the second guesser is proven correct, in my case I really did overestimate the age of the hostess and Dr. Steele claimed. It was never learned whether he believed she was as young as (say) 21. All I know is that he must have judged her younger than 34.

Fill out the rest of the table and using it determine P(second guesser is correct).

22. Why is the classical model obtained by removing the rows 3 and 5 not a reasonable embodiment of equally good guessers? In that model what is P(second guesser is correct)?

low	low*	yes
low*	low	no
low*	high	yes
high*	low	
high	high*	
high*	high	

23-26. Suppose a field may contain OIL. There is a seismic test which is useful in helping to judge whether oil is present. By "+" or "-" we mean that the test comes back positive or negative respectively. So P(+ | OIL) is the conditional probability for OIL if the test comes back positive. A reasonably good test should have P(OIL | +) > P(OIL), meaning that a positive test increases the chance above what it was before the test came back positive. Suppose

> **P(OIL) = 0.2 (probability of no oil is 0.8)**
> **P(+ | OIL) = 0.9 (false negative probability of 0.1)**
> **P(+ | no OIL) = 0.3 (false positive probability of 0.3)**

23. Fill out the tree diagram.

$\qquad\qquad$ P(+ | OIL) = $\qquad\qquad\qquad\qquad$ P(OIL +) =

$\quad$ P(OIL) =

$\qquad\qquad$ P(- | OIL) = $\qquad\qquad\qquad\qquad$ P(OIL -) =

$\qquad$ conditional total = 1 $\qquad$ (law of total probability) $\Rightarrow$ total = P(OIL)

$\qquad\qquad$ P(+ | no OIL) = $\qquad\qquad\qquad$ P(no OIL +) =

$\quad$ P(no OIL) =

$\qquad\qquad$ P(- | no OIL) = $\qquad\qquad\qquad$ P(no OIL -) =

$\qquad$ conditional total = 1 $\qquad$ (law of total probability) $\Rightarrow$ total = P(no OIL)

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ grand total = 1

24. Determine P(+) = P(OIL+) + P(OIL-) by the law of total probability.

25. Determine P(OIL | +) = $\dfrac{P(\text{OIL}+)}{P(+)}$. This is our revised probability for OIL if the test comes back positive.

26. Determine P(OIL | -). This is our revised probability for OIL if the test comes back negative.