W02 - Examples

Bayes' Theorem

10 - Bayes' Theorem: COVID tests

Assume that 0.5% of people have COVID. Suppose a COVID test gives a (true) positive on 96% of patients who have COVID, but gives a (false) positive on 2% of patients who do not have COVID. Bob tests positive. What is the probability that Bob has COVID?

Solution

1. Label events.
    - Event A_P: Bob is actually positive for COVID
    - Event A_N: Bob is actually negative; note $A_N = A_P^c$
    - Event T_P: Bob tests positive
    - Event T_N: Bob tests negative; note $T_N = T_P^c$
2. Identify knowns.
    - Know: $P[T_P \mid A_P] = 96\%$
    - Know: $P[T_P \mid A_N] = 2\%$
    - Know: $P[A_P] = 0.5\%$ and therefore $P[A_N] = 99.5\%$
    - We seek: $P[A_P \mid T_P]$
3. Translate Bayes' Theorem.
    - Using $A = T_P$ and $B = A_P$ in the formula:

    $$P[A_P \mid T_P] = P[T_P \mid A_P] \cdot \frac{P[A_P]}{P[T_P]}$$

    - We know all values on the right except $P[T_P]$
4. Use Division into Cases.
    - Observe:

    $$T_P = T_P \cap A_P \bigcup T_P \cap A_N$$

    - Division into Cases yields:

    $$P[T_P] = P[A_P] \cdot P[T_P \mid A_P] + P[A_N] \cdot P[T_P \mid A_N]$$

    - Important to notice this technique!
        - It is a common element of Bayes' Theorem application problems.
        - It is frequently needed *for the denominator*.
    - Plug in data and compute:

    $$\gg\gg \quad P[T_P] = \frac{5}{1000} \cdot \frac{96}{100} + \frac{995}{1000} \cdot \frac{2}{100} \quad \gg\gg \quad \approx 0.0247$$

5. Compute answer.

- Plug in and compute:

$$P[A_P \mid T_P] = P[T_P \mid A_P] \cdot \frac{P[A_P]}{P[T_P]}$$

$$\gg\gg \quad 0.96 \cdot \frac{0.005}{0.0247} \quad \gg\gg \quad \approx 19\%$$

Independence

13 - Independence by hand: red and green marbles

A bin contains 4 red and 7 green marbles. Two marbles are drawn.

Let R_1 be the event that the first marble is red, and let G_2 be the event that the second marble is green.

- (a) Show that R_1 and G_2 are independent if the marbles are drawn *with replacement*.
- (b) Show that R_1 and G_2 are not independent if the marbles are drawn *without replacement*.

Solution

(a) With replacement.

1. Identify knowns.
   - Know: $P[R_1] = \frac{4}{11}$
   - Know: $P[G_2] = \frac{7}{11}$
2. Compute both sides of independence relation.
   - Relation is $P[R_1G_2] = P[R_1] \cdot P[G_2]$
   - Right side is $\frac{4}{11} \cdot \frac{7}{11}$
   - For $P[R_1G_2]$, have $4 \cdot 7$ ways to get R_1G_2, and 11^2 total outcomes.
   - So left side is $\frac{4 \cdot 7}{11^2}$, which equals the right side.

(b) Without replacement.

1. Identify knowns.
   - Know: $P[R_1] = \frac{4}{11}$ and therefore $P[R_1^c] = \frac{7}{11}$
   - We seek: $P[G_2]$ and $P[R_1G_2]$
2. Find $P[G_2]$ using Division into Cases.
   - Division into cases:

   $$G_2 = G_2 \cap R_1 \bigcup G_2 \cap R_1^c$$

   - Therefore:

   $$P[G_2] = P[R_1] \cdot P[G_2 \mid R_1] + P[R_1^c] \cdot P[G_2 \mid R_1^c]$$

   - Find these by counting and compute:

   $$\gg\gg \quad P[G_2] = \frac{4}{11} \cdot \frac{7}{10} + \frac{7}{11} \cdot \frac{6}{10} \quad \gg\gg \quad \frac{70}{110}$$
3. Find $P[R_1G_2]$ using Multiplication rule.

- Multiplication rule (implicitly used above already):

$$P[R_1G_2] = P[R_1] \cdot P[G_2 \mid R_1] \quad \gg\gg \quad \frac{4}{11} \cdot \frac{7}{10} \quad \gg\gg \quad \frac{28}{110}$$

4. Compare both sides.
   - Left side: $P[R_1G_2] = \frac{28}{110}$
   - Whereas, right side:

$$P[R_1] \cdot P[G_2] = \frac{4}{11} \cdot \frac{70}{110} = \frac{28}{121}$$

   - But $\frac{28}{110} \neq \frac{28}{121}$ so $P[R_1G_2] \neq P[R_1] \cdot P[G_2]$ and they are *not independent.*

Tree diagrams

14 - Marble transferred, marble drawn

Setup:

- Bin 1 holds five red and four green marbles.
- Bin 2 holds four red and five green marbles.

Experiment:

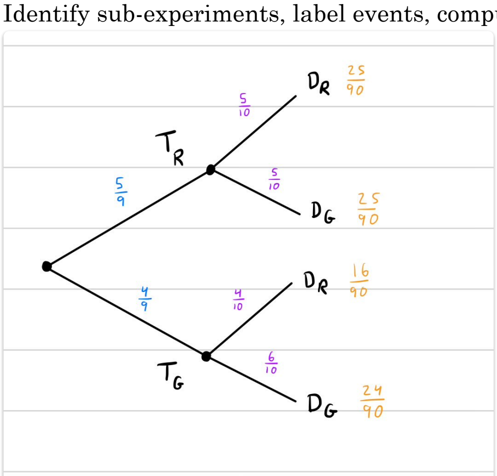
- You take a random marble from Bin 1 and put it in Bin 2 and shake Bin 2.
- Then you draw a random marble from Bin 2 and look at it.

Questions:

- (a) What is the probability you *draw* a red marble?
- (b) Supposing that you drew a red marble, what is the probability that a red marble was *transferred*?

Solution

1. Construct the tree diagram.
   - Identify sub-experiments, label events, compute probabilities:



2. For (a), compute $P[D_R]$.
   - Add up leaf numbers for D_R at leaf:

$$P[D_R] = \frac{25}{90} + \frac{16}{90} = \frac{41}{90}$$

3. For (b), compute $P[T_R \mid D_R]$.
   - Conditional probability:

$$P[T_R \mid D_R] = \frac{P[T_R D_R]}{P[D_R]}$$

   - Plug in data and compute:

$$\gg\gg \quad \frac{25/90}{41/90} \quad \gg\gg \quad \frac{25}{41}$$

   - Interpretation: mass of desired pathway over mass of possible pathways.

Counting

16 - Haley and Hugo from 2 groups of 3

The class has 40 students. Suppose the professor chooses 3 students Wednesday at random, and again 3 on Friday. What is the probability that Haley is chosen today and Hugo on Friday?

Solution

1. Count total outcomes.
   - Have $\binom{40}{3}$ possible groups chosen Wednesday.
   - Have $\binom{40}{3}$ possible groups chosen Friday.
   - Therefore $\binom{40}{3} \times \binom{40}{3}$ possible groups in total.
2. Count desired outcomes.
   - Groups of 3 with Haley are same as groups of 2 taken from others.
   - Therefore have $\binom{39}{2}$ groups that contain Haley.
   - Have $\binom{39}{2}$ groups that contain Hugo.
   - Therefore $\binom{39}{2} \times \binom{39}{2}$ total desired outcomes.
3. Compute probability.
   - Let E label the desired event.
   - Use formula:

$$P[E] = \frac{|E|}{|S|}$$

   - Therefore:

$$P[E] \quad \gg\gg \quad \frac{\binom{39}{2} \times \binom{39}{2}}{\binom{40}{3} \times \binom{40}{3}}$$

$$\gg\gg \quad \left(\frac{\frac{39 \cdot 38}{2!}}{\frac{40 \cdot 39 \cdot 38}{3!}} \right)^2 \quad \gg\gg \quad \left(\frac{3}{40} \right)^2$$

17 - Counting VA license plates

A VA license plate has three letters (with no I, O, or Q) followed by four numerals. A random plate is seen on the road.

- (a) What is the probability that the numerals are in increasing order?
- (b) What is the probability that at least one number is repeated?

Solution

(a)

1. Count ways to have 4 numerals in increasing order.
   - Any four distinct numerals have a single order that's increasing.
   - There are $\binom{10}{4}$ ways to choose 4 numerals from 10 options.
2. Count ways to have 3 letters in order except I, O, Q.
   - 26 total letters, 3 excluded, thus 23 options.
   - Repetition allowed, thus $23 \cdot 23 \cdot 23 = 23^3$ possibilities.
3. Count total plates with increasing numerals.
   - Multiply the options:

$$23^3 \cdot \binom{10}{4}$$

4. Count total plates.
   - Have $23 \cdot 23 \cdot 23$ options for letters.
   - Have $10 \cdot 10 \cdot 10 \cdot 10$ options for numbers.
   - Thus $23^3 \cdot 10^4$ possible plates.
5. Compute probability.
   - Let E label the event that a plate has increasing numerals.
   - Use the formula:

$$P[E] = \frac{|E|}{|S|}$$

   - Therefore:

$$P[E] \quad \gg\gg \quad \frac{23^3 \cdot \binom{10}{4}}{23^3 \cdot 10^4} \quad \gg\gg \quad \frac{\frac{10!}{4!6!}}{10000} \quad \gg\gg \quad \frac{21}{1000}$$

(b)

1. Count plates with at least one number repeated.
   - "At least" is hard! Try *complement*: "no repeats".
   - Let E^c be event that *no* numbers are repeated. All distinct.
   - Count possibilities:

$$|E^c| = 23 \cdot 23 \cdot 23 \cdot 10 \cdot 9 \cdot 8 \cdot 7$$

   - Total license plates is still $23^3 \cdot 10^4$.
   - Therefore, license plates with *at least one number repeated*:

$$|E| = |S| - |E|$$

$$\gg\gg \quad 23^3 \cdot 10^4 - 23^3 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \quad \gg\gg \quad 60348320$$

2. Compute probability.
   - Desired outcomes over total outcomes:

$$\frac{|E|}{|S|} \quad \gg\gg \quad \frac{60348320}{23^3 \cdot 10^4} \quad \gg\gg \quad 0.496$$