Exercise 3: adopting pets.

Suppose you know a family adopted two pets at different times, and you know each pet is either a cat (C) or a dog (D).

1. What is the probability that the family has two cats, if you're told at least one of the pets is a cat?

$S = \{CD, CC, DC, DD\}$, $|S| = 4$

A: two cats

B: at least 1 is a cat

$P(A) = \frac{1}{4}$

$P(B) = \frac{3}{4}$

$P(A \cap B) = \frac{1}{4}$

$P(A|B) = \frac{1/4}{3/4} = \frac{1}{3}$

2. What is the probability that the family has two dogs, if you are told that the first pet is a dog?

A: two dogs

B: first is a dog

$P(A) = 1/4$

$P(B) = \frac{2}{4} = \frac{1}{2}$

$P(A \cap B) = 1/4$

$P(A|B) = \frac{1/4}{1/2} = \frac{1}{2}$

Main problem for today: how probable is it that Eric Reid was drug tested 7 times within the first 11 weeks of the 2018 NFL season?

"That has to be statistically impossible. I'm not a statistician, but there's no way that's random"

Form groups of 3-4 and start brainstorming how you would analyze this:

- **Assume:** (1) 11 week period, (2) 76 players on the team, (3) 10 people chosen from the team to be tested each week.
- How likely is it that Eric Reid is tested at least seven times?

Here's an idea, let's run some simulations!

```python
import random

NP = 76  # number of players on team
NW = 11  # number of weeks to consider
NT = 10  # number of players tested each week
NS = 50000  # number of simulations

tested = [0] * (NW + 1)
players = list(range(NP))
PLAYER = 25  # any player should work
for s in range(NS):
    N = 0
    for w in range(NW):
        random.shuffle(players)
        for i in range(NT):
            if players[i] == PLAYER:
                N += 1
                break
    tested[N] += 1

for i, m in enumerate(tested):
    print(f"tested {i} times: avg = {round(100 * m / NS, 4)} %")
```

Main problem for today: how probable is it that Eric Reid was drug tested 7 times within the first 11 weeks of the 2018 NFL season?

Assumptions:

- 11 week period, 76 players on the team, 10 people chosen from the team to be tested each week.

Let's break the problem down:

intuition $\frac{10}{76}$ using probability: $\frac{\binom{75}{9}}{\binom{76}{10}}$ ← pick 9 others of 75 ← pick 10 of 76

$\frac{75!}{9!\,66!} \cdot \frac{10!\,66!}{76!} = \frac{10}{76}$ ✓

- In how many ways can a single player be tested in a 1-week period?
- What is the size of the sample space for an 11-week period? $\binom{76}{10}^{11}$ pick 10 of 76 for 11 weeks
- In how many ways can the NFL pick a player k times for testing over an 11-week period? pick k weeks of 11 $\binom{11}{k}$
- In how many ways can the NFL pick other 10 players for drug testing in the other $11 - k$ weeks? 11-k weeks left, pick 10 from 75 (76−1 since not Reid) $\binom{75}{10}^{11-k}$
- In how many ways can a single player be tested k times over an 11-week period, i.e. $|E_k|$? $|E_k| = \binom{11}{k}\binom{75}{9}^{k}\binom{75}{10}^{11-k}$
- Probability of a single player being tested exactly k times, $p(E_k)$, for $0 \leq k \leq 7$? $p_k = \frac{|E_k|}{|S|} = \frac{\binom{11}{k}\binom{75}{9}^{k}\binom{75}{10}^{11-k}}{\binom{76}{10}^{11}}$
- Probability of a single player being tested *at least* 7 times? $p_7 + p_8 + p_9 + p_{10} + p_{11} = 1 - (p_0 + p_1 + p_2 + p_3 + p_4 + p_5 + p_6) \approx 0.0138\%$

Possible code to calculate the probablity of being tested *at least* 7 times.

```python
NP = 76  # number of players on team
NW = 11  # number of weeks to consider
NT = 10  # number of players tested each week
s = math.comb(NP, NT) ** NW  # size of the sample space

def p_tested(k: int) -> float:
    ek = math.comb(NW, k) * math.comb(NP - 1, NT - 1) ** k \
        * math.comb(NP - 1, NT) ** (NW - k)
    return ek / s

# calculate the probability of getting tested
# at least 7 times using 1 - \sum_{k=0}^6 p_k
p_total = 0.
for k in range(7):
    pk = p_tested(k)
    print(f"prob. tested {k} times  = {round(100 * pi, 3)} %")
    p_total += pk
p_at_least_7 = 1. - p_total
print(f"prob. tested >= 7 times = {round(100 * p_at_least_7, 5)} %
```