perplexity

The (N+1)/(K+1) Formula: Universal Applications

Excellent question! This formula is **incredibly powerful** and applies to many sampling-without-replacement problems. Let me show you all the variations.[1]

The General Formula

For a collection of **N total items** containing **K special items**:

$$E[\text{draws to first special item}] = \frac{N+1}{K+1}$$

Key requirement: Drawing **without replacement** (no putting items back).

Why the +1 in Numerator and Denominator?

The Intuition

Think of it as **dividing positions among gaps**:

```
Arrangement of N items with K special items:

[regular] [regular] | SPECIAL | [regular] | SPECIAL | ... | SPECIAL | [regular]
    Gap 0               Gap 1       Gap 2          Gap K      Gap K

Total gaps: K + 1
Total positions: N + 1 (including the "end" marker)
Average gap size: (N+1)/(K+1)
```

The "+1" Accounting

Numerator (N+1):

- N items in the collection
- +1 for the "after the last item" position
- This makes the math symmetric for all K+1 gaps

Denominator (K+1):

- K special items create K+1 gaps
- Gap 0: before 1st special

- Gaps 1 to K-1: between special items
- Gap K: after last special

The first gap = draws until first special item (including the special item).

All Problems That Fit This Formula

1. Card Drawing Problems

A. Expected Draws to First Ace

- N = 52 (total cards)
- K = 4 (Aces)
- Answer: (52+1)/(4+1) = **10.6 draws**

B. Expected Draws to First King

- N = 52
- K = 4 (Kings)
- Answer: (52+1)/(4+1) = **10.6 draws**

C. Expected Draws to First Red Card

- N = 52
- K = 26 (red cards)
- Answer: (52+1)/(26+1) = **1.96 draws**

D. Expected Draws to First Spade

- N = 52
- K = 13 (spades)
- Answer: (52+1)/(13+1) = **3.79 draws**

E. Expected Draws to First Face Card

- N = 52
- K = 12 (J, Q, K of each suit)
- Answer: (52+1)/(12+1) = **4.08 draws**

F. Expected Draws to First Picture Card (J, Q, K, A)

- N = 52
- K = 16 (four ranks × four suits)
- Answer: (52+1)/(16+1) = **3.12 draws**

2. Lottery and Raffle Problems

A. Expected Tickets to First Winner

- 1000 tickets sold, 10 are winners
- N = 1000, K = 10
- Answer: (1000+1)/(10+1) = **91 tickets**

B. Expected Tickets to First Grand Prize

- 1000 tickets, 1 grand prize
- N = 1000, K = 1
- Answer: (1000+1)/(1+1) = **500.5 tickets**

C. Expected Raffle Draws to First Employee Winner

- 200 raffle entries, 30 are employees
- N = 200, K = 30
- Answer: (200+1)/(30+1) = **6.48 draws**

3. Manufacturing and Quality Control

A. Expected Items Inspected to Find First Defect

- Batch of 500 items, 20 defective
- N = 500, K = 20
- Answer: (500+1)/(20+1) = **23.9 items**

B. Expected Tests to Find First Positive

- 100 samples, 5 are positive
- N = 100, K = 5
- Answer: (100+1)/(5+1) = **16.83 tests**

C. Expected Widgets Checked to Find First Bad One

- Production run of 10,000, 50 defective
- N = 10,000, K = 50
- Answer: (10,000+1)/(50+1) = **196.1 widgets**

4. Population and Sampling Problems

A. Expected People Surveyed to Find First Vegetarian

- Town of 1000 people, 100 vegetarians
- N = 1000, K = 100
- Answer: (1000+1)/(100+1) = **9.91 people**

B. Expected Students to Find First Left-Handed

- Class of 30 students, 3 left-handed (10%)
- N = 30, K = 3
- Answer: (30+1)/(3+1) = **7.75 students**

C. Expected Voters Polled to Find First Independent

- Electorate of 5000, 500 independents
- N = 5000, K = 500
- Answer: (5000+1)/(500+1) = **9.98 voters**

5. Genetics and Biology

A. Expected DNA Samples to Find First Mutation

- 200 samples, 8 have mutation
- N = 200, K = 8
- Answer: (200+1)/(8+1) = **22.33 samples**

B. Expected Organisms to Find First with Trait

- Population of 1000, 40 have rare trait
- N = 1000, K = 40
- Answer: (1000+1)/(40+1) = **24.41 organisms**

6. Computer Science Problems

A. Expected Array Elements Checked to Find First Target

- Array of 100 elements (random order), 5 target values
- N = 100, K = 5
- Answer: (100+1)/(5+1) = **16.83 checks**

B. Expected Files to Scan for First Virus

- 1000 files, 15 infected
- N = 1000, K = 15
- Answer: (1000+1)/(15+1) = **62.6 files**

C. Expected Nodes Visited to Find First Match

- Graph with 500 nodes, 25 match criteria
- N = 500, K = 25
- Answer: (500+1)/(25+1) = **19.27 nodes**

7. Game Theory and Puzzles

A. Expected Draws to Get First Prize Ball

- Lottery drum with 50 balls, 5 are prizes
- N = 50, K = 5
- Answer: (50+1)/(5+1) = **8.5 balls**

B. Expected Cards to Draw First Joker

- Deck with 54 cards (52 + 2 jokers)
- N = 54, K = 2
- Answer: (54+1)/(2+1) = **18.33 cards**

C. Expected Tiles to Draw First Blank (Scrabble)

- 100 Scrabble tiles, 2 blanks
- N = 100, K = 2
- Answer: (100+1)/(2+1) = **33.67 tiles**

8. Financial and Trading

A. Expected Stocks Analyzed to Find First Undervalued

- Portfolio of 200 stocks, 20 undervalued
- N = 200, K = 20
- Answer: (200+1)/(20+1) = **9.57 stocks**

B. Expected Trades to Find First Profitable

- Historical data: 365 trading days, 200 profitable
- N = 365, K = 200
- Answer: (365+1)/(200+1) = **1.82 days**

9. Extreme Cases (Edge Cases)

A. Only One Special Item

- N items, K = 1
- Answer: (N+1)/(1+1) = **(N+1)/2**
- Example: 52 cards, 1 Ace → 53/2 = **26.5 draws**

B. Half Are Special

- N items, K = N/2
- Answer: (N+1)/(N/2+1)
- Example: 100 items, 50 special → 101/51 ≈ **1.98 draws**

C. All Are Special

- N items, K = N
- Answer: (N+1)/(N+1) = **1 draw** (always get special immediately!)

D. None Are Special

- N items, K = 0
- Answer: (N+1)/(0+1) = **N+1 draws** (must draw all items + never find one)

Comparison Table: Different Values of K

For **N = 52** (deck of cards):

K (special items)	Description	Formula	Answer
1	One specific card	53/2	26.5
2	Two jokers	53/3	17.67
4	Four Aces	53/5	**10.6**
12	Face cards	53/13	4.08
13	One suit	53/14	3.79
26	Red cards	53/27	1.96
52	All cards	53/53	1.0

Pattern: More special items → Fewer draws needed (on average)

Why N/K is Wrong (The Common Mistake)

Wrong Intuition:

```
"52 cards, 4 Aces → every 13 cards on average"
E[X] = N/K = 52/4 = 13 ✖
```

Why It Fails:

1. Doesn't Account for Position Symmetry

```
The formula N/K treats the deck as if you're sampling WITH replacement.
Without replacement, early positions are more likely to contain the first Ace.
```

2. Missing the "Gap" Structure

```
N/K = 52/4 = 13 is the average spacing between Aces.
But the first Ace appears BEFORE the first spacing ends!

Positions: [gap0] ACE [gap1] ACE [gap2] ACE [gap3] ACE [gap4]
           ^^^^^^
           This is what we want!

Average gap size = (N+1)/(K+1) = 53/5 = 10.6
```

3. Mathematical Proof

```
Consider extreme case: N = 2, K = 1
- Two cards, one is special
- By symmetry, special card is at position 1 or 2 equally likely
- E[X] = (1 + 2)/2 = 1.5

Wrong formula: N/K = 2/1 = 2 ✖
Correct formula: (N+1)/(K+1) = 3/2 = 1.5 ✓
```

Visual Proof: Why (N+1)/(K+1)?

Arranging N=12 items with K=3 special ones:

```
Possible arrangement:
o o o o ★ o o ★ o ★ o o
       ↑     ↑   ↑
     Gap0  Gap1 Gap2 Gap3

Gap sizes: 4, 2, 1, 2 (total = 9 regular items)
First special at position: 4+1 = 5
```

By **symmetry**, each of the 4 gaps gets an equal share of:

- The N regular items
- Plus "virtual positions" that make the math work

Average gap: (N + K virtual items)/(K+1 gaps) = (12 + 1)/(3 + 1) = **3.25**

First gap = First special position = 3.25 ✓

This is slightly different from naive N/K = 12/3 = 4.

Interview Pro Tip: Recognize the Pattern

When you see:

- "Expected draws/picks until first..."
- "Sampling without replacement"
- "From a finite population"

Immediately think: $(N+1)/(K+1)$

Then verify:

1. Is it without replacement? ✓
2. Is there a fixed population? ✓
3. Are we looking for the FIRST occurrence? ✓

If all three are yes → Apply the formula!

Summary: Master Formula

Problem Type	N	K	Answer
Cards: First Ace	52	4	10.6 [1]
Cards: First Red	52	26	1.96
Lottery: First Winner	1000	10	91
QC: First Defect	500	20	23.9
Survey: First Vegetarian	1000	100	9.91
DNA: First Mutation	200	8	22.33
Files: First Virus	1000	15	62.6
Scrabble: First Blank	100	2	33.67

The formula $(N+1)/(K+1)$ is universal for all "first occurrence in sampling without replacement" problems!

Does this clarify why the +1 is crucial and show you all the problem types this applies to?

⁂

1. Coding-Questions-Report-QUANT.pdf