Multinomial Skittles

MathsJam 2025

Ben Handley

Prior work

- “No two packs are the same” is clearly false. Thanks Clare!
- 16-17 skittles per pack, 5 colours
- Only about 10,000 possible packs
- So 120 packs should give us roughly 50/50 odds

Remaining mysteries

- We reliably get duplicates even with fewer than 100 packs
- 100 possible packs have 15+ of the same colour but we never see them

Possible explanations

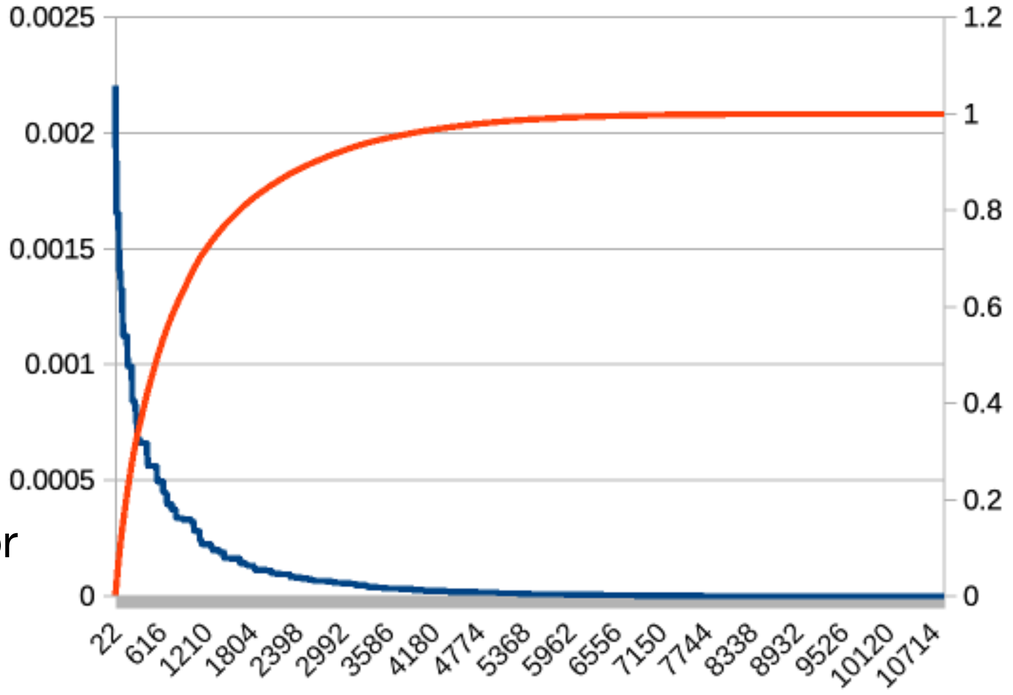
- Skittles aren’t drawn independently. Maybe they try to avoid some bag types?
- The maths is more complicated

Are they independent?

Metric	Expected	Actual
Num missing colours	0.14	0.16
8+ one colour	0.035	0.032
All max 4	0.18	0.169

Non uniform probabilities

- Only one way to make the bag with 16 reds
- Lots of ways to make 3,3,3,3,4
- $p = \frac{n!}{5^n \prod c_i!}$
- Statistical breadth: $\frac{1}{\sum p^2}$
- Works out to about 1580
- This suggests about 47 bags for 50/50
- But this is an approximation



Random sampling

- Either sample randomly from the theoretical distribution
- Or from Clare’s dataset: without replacement!
- Theoretical: About 48 for 50%
- Clare’s dataset: also 48!

Conclusion

- There is no evidence that skittles are anything other than independently uniformly random
- Roughly 50 bags for 50%,
  100 is almost guaranteed