Paninimania: sticker rarity and cost-effective strategy

Sylvain Sardy and Yvan Velenik
Section de mathématiques
Université de Genève

Abstract

We consider some issues related to the famous Panini stickers devoted to the football world cup. In particular, we address the following questions: is there a planned shortage of some stickers? What is a good cost-effective strategy to fill in an album?

1 Introduction

The collectors' frenzy over Panini's stickers is now almost a tradition with each new football world cup [1]. In this note we discuss the alleged rarity of certain stickers (famous players, etc.) and propose a cost-effective strategy.

For the purpose of the discussion below, we recall that in Switzerland the stickers can be purchased in three different ways: buying one packet of 5 different stickers for CHF 1; buying one box of 500 different stickers[1] for CHF 100 (actually prices as low as CHF 70 can be found); buying specific individual stickers directly from Panini at a cost of CHF 0.30 apiece with, however, a limitation to at most 50 stickers. The album comprises 660 different stickers.

2 Rare stickers: a myth?

2.1 Testing for uniformity

To test whether all 660 stickers appear with the same frequency, 12 boxes (three in four different Swiss cantons) of 100 packets containing each five different stickers have been collected, which amounts to a total of 6000 stickers. Because of the dependence induced by the fact that stickers bought in a single box are all different, we cannot perform a standard chi-square test. Instead, we perform a Monte-Carlo simulation to estimate the null distribution under the hypothesis that stickers are uniformly distributed overall with the constraints that there are no duplicates neither within a packet nor within a box. Hence we generate 12 independent boxes with such constraints, count the total number of times $N_i,\ i = 1, \ldots, 660$, each sticker occurs, and calculate the statistics

$$S = \sum_{i=1}^{660} \frac{(N_i - e_i)^2}{e_i} \quad \text{with} \quad e_i = \frac{6000}{660}.$$

[1]Although this is true for the analyzed boxes, it seems that some boxes may occasionally contain a few duplicates. We ignore this issue here.

Repeating this experiment $M = 10^4$ times provides an estimate of the null distribution that we plot on the left graph of Figure 1. The statistics $s = 138$ calculated on the collected stickers is also plotted and amounts to a p-value of 0.9974, so that we do not reject the null hypothesis that stickers are uniformly distributed.

Interestingly the p-value is large, which can be explained by the fact that in two of the four cantons, it was possible to complete an entire album by buying three boxes. If boxes were independent of each other the probability of such an event would be very small:

$$\sum_{k=0}^{160} \frac{\binom{k}{160}\binom{500-k}{500}\binom{340+k}{500+k}}{\binom{550}{660}^2} \simeq 3.7\ 10^{-5}. \tag{1}$$

Looking at the serial number of the boxes reveals that successive boxes seem to be dependent, which violates the assumption of the previous test. For a more accurate test, we start by dropping the data from the two cantons where we were able to complete an entire album, and repeat the test. The right graph of Figure 1 now shows the estimated null distribution along with the statistics $s = 168.44$ based on the 3000 measurements. The p-value is now 0.1235, so that we still do not reject the null hypothesis.

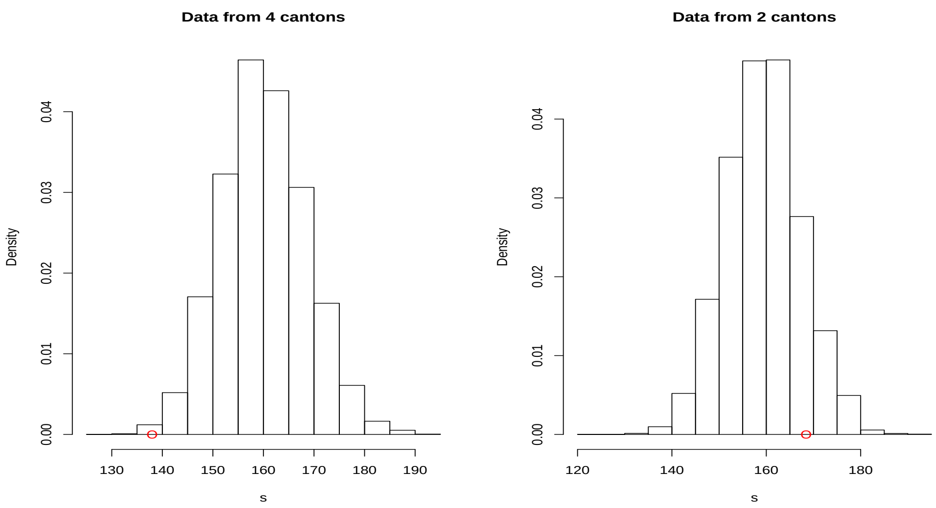

Figure 1: Density estimate of the distribution of the statistics S under the null hypothesis that stickers are uniformly distributed along the value of the statistics calculated from the data: (left) based on the 6000 stickers of 4 cantons, in which case the independence assumption seems violated by boxes with successive serial numbers; (right) based on 3000 stickers of 2 cantons.

2.2 How to explain the myth

Most people who try to fill in the Panini album believe some stickers are indeed rare [1]. To explain the myth, we considered two scenarios and made some calculations.

Number of packets bought	233	+233	+233	+233
Number of different stickers obtained	550	+90	+17	+3

Table 1: Single person scenario. Average number of different stickers obtained as a function of the number of packets bought.

Consider first the situation where a single person buys independent packets of five stickers. The average number of packets he needs to complete the album is 931 as one can read on Figure 2. And on average, with one fourth (i.e., 233 packets), he can obtain 550 different stickers; with another 233 packets, he can obtain 640 different stickers; with another 233 packets, 657 different stickers; finally another 233 packets are needed to complete the album with 660 different stickers. These average calculations are summarized in Table 1. Hence the last three stickers missing seem rare since it takes so long to get them.

Let us turn now to a more realistic scenario which takes into account sticker swapping. To model swapping we consider optimal swapping between k individuals who buy their cards together until they fill k albums. Other swapping procedures could also be considered (e.g., each individual swapping his duplicates for missing cards), but this would lead to less cost effective strategies.

Consider ten friends who buy 100 packets each, one by one (as opposed to buying boxes) and perform optimal swapping. Since a total of 5000 stickers are bought, one has the feeling that it is unlikely that at least one sticker is missing to all the ten friends. But calculations show that this event actually occurs slightly more than 25% of the time. Hence, a group of friends that has at least one sticker missing will incorrectly believe that these stickers are rare.

3 Strategies

3.1 Without swapping

Consider first the situation of a person trying to fill in his album without swapping stickers. This version is very reminiscent of the classical *coupon collector* problem [3, Section IX.3]. In the latter, one of n different coupons is obtained when buying one instance of some product; how many instances are necessary in order to complete the collection? This problem can easily be solved explicitly, yielding in particular an expected number of required purchases equal to $n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) \simeq n(\log n + 0.577)$. In particular for $n = 660$, this leads to approximately 4666.3 stickers, or CHF 933.27.

In our case, the situation is slightly more complicated, since 5 *different* stickers are obtained each time a packet is purchased. Of course, one should not expect the latter constraint to affect significantly the conclusion, since 5 objects sampled (with replacement) from a pool of 660 objects are all different with probability $\binom{660}{5}/660^5 \simeq 0.985$. This situation has also been treated analytically [4, 2] yielding for example an expectation of the number of required packets of five stickers equal to

$$\binom{660}{5}\sum_{j=1}^{660}(-1)^{j+1}\frac{\binom{660}{j}}{\binom{660}{5}-\binom{660-j}{5}} \simeq 930.84.$$

3.2 Effect of swapping

Let us now turn to the effect of swapping on the overall expected cost of filling an album. As above we consider optimal swapping between k individuals. Figure 2 shows the evolution of the expected cost per collector as a function of k.

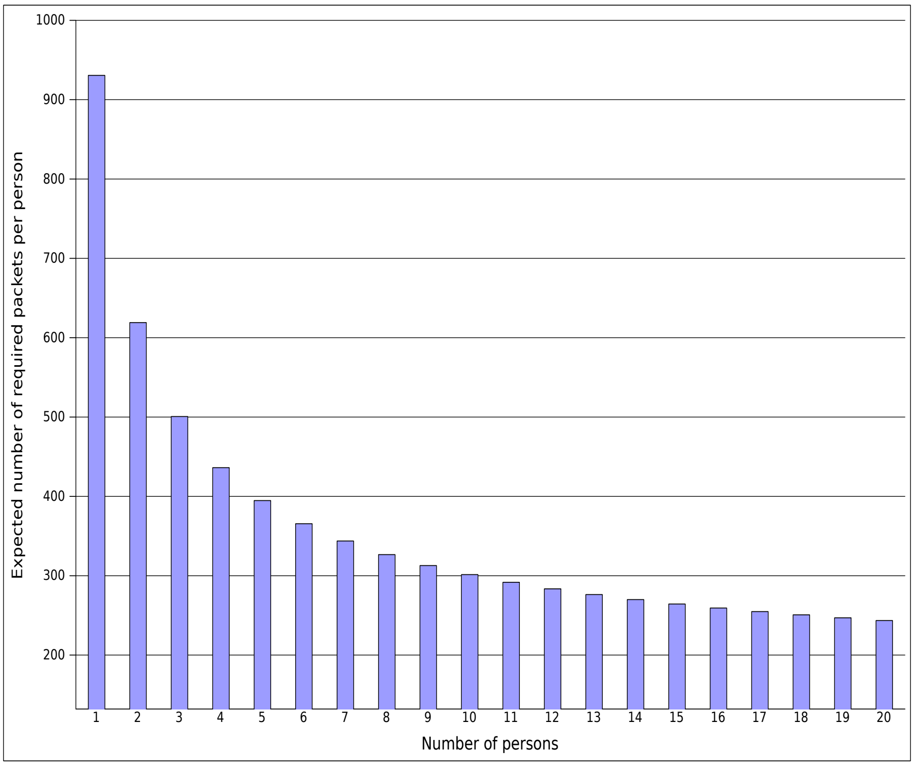

Figure 2: Average number of packets each of k persons have to buy in order for all of them to complete their album, under an assumption of optimal swapping. Notice that in the limiting situation of an infinite number of friends, this minimal number is actually equal to $132 = 660/5$; this value corresponds to the horizontal axis in the picture.

3.3 A good strategy

In order to devise a reasonable strategy, we use the facts that it is possible to buy boxes containing 500 *different* stickers for a price of CHF $70-100$, depending on the vendor, and that it is possible to buy from Panini up to 50 specified stickers for a unit price of CHF .30 each.

Numerical computations suggest the following strategy, when swapping with 9 other persons: 1) buy a box, 2) buy 40 additional packets and swap the duplicates until at most 50 stickers are missing from your collection, 3) order the missing stickers from Panini. The cost of this strategy is between CHF 125 and CHF 155, depending on the price of the box.

If one ignores swapping, it is possible to derive analytically the optimal strategy [4].

References

[1] *À bon entendeur*, Télévision Suisse Romande, 18 mai 2010, `http://www.tsr.ch/emissions/abe/1968926-foot-abe-met-le-maillot-et-sort-son-album-panini.html#1968926`

[2] Adler, Ilan and Ross, Sheldon M.: *The coupon subset collection problem*, J. Appl. Probab. 38 (2001), no. 3, 737–746.

[3] Feller, William: *An introduction to probability theory and its applications. Vol. I.*, Third edition John Wiley & Sons, Inc., New York-London-Sydney 1968.

[4] Stadje, Wolfgang: *The collector's problem with group drawings.*, Adv. in Appl. Probab. 22 (1990), no. 4, 866–882.