赠券收集问题的简单探讨

学院：　环境学院

班级：　1625102 班

姓名：　徐斯琪

学号：　1162510207

赠券收集问题的简单探讨

哈尔滨工业大学环境学院 2016 级 1625102 班 徐斯琪

摘要：近年来，随着经济不断增长，市场化程度越来越高，商家常常会通过收集赠券的方法吸引顾客以达到营销目的. 赠券收集问题也成为了一类经典的数学问题. 赠券收集问题是一类计算有 n 种等可能情况存在时，做伯努利试验，若正好每种情况都经历过一次，做伯努利试验的次数的数学期望的问题. 本文通过探讨给定情况下集齐赠券需要购买商品的数学期望，得到赠券收集问题近似的数学计算方法，抽象出赠券收集问题的数学模型并阐述其变型，从而为生活中可以近似成赠券收集问题的问题提供思考方法. 本文共分为四章：第一章阐述了赠券收集问题的背景与研究意义，第二章以最基础条件下的赠券收集问题为例推算出这类问题的数学公式，第三章阐述了赠券收集问题的应用及其他变型问题，第四章作为本文的总结.

关键词：赠券收集问题、数学期望、数学模型

Abstract: In recent years, with the development of economy, the degree of marketization is higher than before. Merchant often attract customers by collecting coupons so that they can achieve their turnover goals. In this background, Coupon Collecting becomes a classical math problem. Coupon Collecting Problem is a kind of problem calculating when there are n same-possibility situation exist, how many times the person should do Bernoulli experiments to ensure that he had experienced all the situations. The dissertation calculate the expectation customs need when they want to collect all kinds of coupons in order to get the approximate mathematical method to solve coupon collecting problems, abstract the model of couple collecting problems and expound its variants in order to provide a thinking method for similar problems in lives. The dissertation contains four chapters: chapter one analyze the background of Coupon Collecting problem and its value for search, chapter two takes the coupon collecting problem in the most basic situation as an example to calculate the method to these kinds of problems, chapter three expounds the usage and variants of coupon collecting problems and chapter four has a conclusion of the dissertation.

Key words: Coupon Collecting problem, mathematical expectation, mathematical model

一、背景分析

随着经济社会的不断发展，市场化程度越来越高，在市场竞争中商家常常会使用各种办法以达到吸引顾客的目的，而赠券收集就是其中很常见的一种．赠券收集问题是一类顾客在购买商品的过程中获得商品中随机附加的赠券，通过购买一定数量的商品集齐赠券的问题．商家将多种不同的赠券随机或按一定比例地放入商品中，让顾客在购买商品的同时收集赠券，集满所有赠券便可获得礼物．过去常用的方式是在商品里随机或按一定比例地放入某一种赠券，但是现在随着网络日渐发达，支付平台上收集赠券也十分普遍，如今年过年时支付宝推出的“集福卡”活动．对于消费者来说，可以预期集齐赠券需要的成本有选择地收集赠券是最理想的，而对商家来讲，从收集赠券活动中获取尽量多的利润是商家最希望做到的．不仅如此，生活中的许多问题也可以近似抽象为赠券收集问题，比如“球与箱子问题”：球独立且随机地放入箱子中，直到所有箱子里至少有一个球，那么需要多少个球?[1]由此可见，赠券收集问题是一类在生活中存在重要应用的问题，具有深入探讨的价值.

二、基础条件下赠券收集问题的数学推导

假设有一种赠券收集活动共有 n 种赠券，每件商品里有其中的一种，赠券放入商品中是独立且随机的．顾客每买一件商品便可收集到商品中含有的赠券，收集到所有 n 种赠券即可获得奖品．假设商品总数足够大（即顾客购买的商品数对商品总体无影响），n 足够大，顾客购买商品是随机的并且在收集赠券的过程中不与其他人合作，则顾客想要收集齐 n 种赠券，需要购买商品的数学期望是多少?

令随机变量X表示收集到 n 种赠券时需要购买的商品总数，则该问题求的是E(X)．令X_i表示已收集到 i-1 种不同赠券时，收集到第 i 张不同的赠券时需要购买的商品数，显然：$\mathrm{X}=\sum_{i=1}^{n} X_i$.

因为赠券放入商品中是独立且随机的，顾客购买每次商品也是独立且随机的，则顾客每次购买都是一次伯努利试验，当顾客已经得到 i-1 张不同赠券时，再得到一张不同赠券所需要购买商品的数量X_i服从几何分布.

$$\mathrm{P}(X_i=\mathrm{k})=q^{k-1}\mathrm{p}_i,(\mathrm{q}=1-\mathrm{p}_i;\ \mathrm{k}=1,2,\dots,\mathrm{n};\ \ \mathrm{i}=1,2,\dots,\mathrm{n})$$ [2]

顾客每购买一次商品，得到每种赠券的概率为：$\mathrm{p}=\frac{1}{n}$，因此当已经收集到 i-1 种赠券时，每一次购买得到新赠券的概率为：

$$p_i=1-\frac{i-1}{n}$$

由几何分布的数学期望公式，$\mathrm{E}(X_i)=\frac{1}{p_i}$[3]，$p_i=1-\frac{i-1}{n}$，可得：

$$\mathrm{E}(X_i)=\frac{n}{n-i+1}$$

由数学期望的性质：$\mathrm{E}(\sum_{i=1}^{n} X_i)=\sum_{i=1}^{n} E(X_i)$

$$\begin{aligned}\mathrm{E(X)}&=\sum_{i=1}^{n} E(X_i)\\&=\sum_{i=1}^{n}\frac{n}{n-i+1}\\&=\frac{n}{n}+\frac{n}{n-1}+\frac{n}{n-2}+\cdots+\frac{n}{1}\\&=\mathrm{n}\sum_{i=1}^{n}\frac{1}{i}\end{aligned}$$

其中$\sum 1/i$为调和级数，由于调和级数发散，所以无法求出调和级数的和，这里用到的是调和级数的部分和$\mathrm{H(n)}=\sum_{i=1}^{n} 1/i$．调和级数的第 n 个部分和与$\ln n$的差值收敛于欧拉-马歇罗尼常数$(\gamma)$[4]，即：

$$\lim_{n\to\infty}(\mathrm{H(n)}-\ln n)=\gamma$$

所以$\mathrm{E(X)}=\mathrm{E}(X_i)=n\sum_{i=1}^{n} 1/i=nH(n)=n\ln n+n\gamma$．即顾客若想集齐 n 种赠券，需要购买商品的数量的数学期望是$\mathrm{E(X)}=n\ln n+n\gamma$．

设每件商品的价格为 Q，奖品价值为 T，则为了集齐赠券，顾客所需要花费的成本 W 的数学期望为：

$$\begin{aligned}\mathrm{E(W)}&=\mathrm{QE(X)}\\&=\mathrm{Q}(n\ln n+n\gamma)\end{aligned}$$

商家兑换一件奖品所获利润 G 的数学期望为：

$$\begin{aligned}\mathrm{E(G)}&=\mathrm{E(W)}-\mathrm{T}\\&=\mathrm{Q}(n\ln n+n\gamma)-\mathrm{T}\end{aligned}$$

（注：上述求解过程建立在商品总数和 n 都足够大的条件下，若 n 不够大，则$\gamma=\Theta(\mathrm{n})$．）

三、赠券收集问题的应用

赠券收集问题可以抽象成模型：依次地做独立试验，其中每次实验有 n 种等概率的结果，问要使得每种结果都至少出现一次平均需要做多少次试验?

这个模型可以用于很多实际问题，比如盒子里有 n 种颜色的小球，每种颜色小球有a_i个，每次抽出后放回，则若想知道盒子里小球的所有颜色，需要抽取小球次数的数学期望为多少. 或者一个教室里为使十二个星座里每一个星座都对应至少一人，这个教室里人数的数学期望为多少. [5]这类问题的求解过程与第二部分所述完全一致.

赠券收集模型的经典之处在于它还可以衍生出许多的变型，它的主要变型有三种. [6]

第一种变型：依次地做独立试验，其中每次实验有 n 种等概率的结果，则要使得每种结果都至少出现 m 次平均需要做多少次试验?

第二种变型：依次地做独立试验，其中每次实验有 n 种等概率的结果，实验分批进行，每批进行 k 个实验，则要使得每种结果都至少出现一次平均需要做多少批试验?

第三种变型：依次地做独立试验，其中每次实验有 n 种结果，其概率依次为$p_1, p_2, p_3, \dots, p_n$，则要使得每种结果都至少出现 1 次平均需要做多少次试验? [6]

这三种变型每一种都有其在生活中的广泛应用. 第一种变型可以讨论集齐多套赠券需要购买商品的期望. 第二种变型可以讨论从样本空间中每次随机抽取一定数目的样本之后放回，则抽取多少次后所有样本都至少被抽取过一遍. 例如从题库中每次随机抽取一定数目的题出一套试卷，则多少套试卷后，题库中所有的题都被抽取过至少一次? 第三种变型在生活中更加常见，因为实际生活中商家发行赠券时常常会将某一种赠券的出现概率降低从而使赠券的收集更加困难.

以上三种变型的求解过程显然比第二部分复杂得多，前人已经给出了三种变型的公式，具体求解过程本文不加以赘述.

四、结语

本文主要对赠券收集问题进行了简单的探讨，在第一部分分析了赠券收集问题的产生背景与现实意义，在第二部分推导了最基础条件下的赠券收集问题的求解过程及公式，得出在 n 种赠券等可能且随即均匀分布时，为了收集到所有赠券需要购买商品数量的数学期望为 $\mathrm{E(X)} = \mathrm{n}\sum_{i=1}^{n} 1/i = nH(n)$，在第三部分阐述了由赠券收集问题变型而来的三种变型及其分别对应的现实问题.

赠券收集问题是一类经典问题，在互联网、通信、计算机等等领域都有较为广泛的应用，值得更进一步探讨与分析.

五、参考文献

[1]Michael Mitzenmacher, Eli Upfal. Probability and Computing: Randomized Algorithms and Probabilistic Analysis［M］. Beijing: China Machine Press, 2007. 4

[2]-[3]王勇. 概率论与数理统计（第二版）[M]. 北京：高等教育出版社, 2014. 12

[4]朱永生, 龚晓岚. 欧拉常数γ的性质及在解题中的应用[J]. 高师理科学刊, 2005（3）

[5]J. E. Kobza, A Survey of Coupon Collector’s Problem with Random Sample Sizes, Method. Comput. Appl. Prob 9:573-584, 2007

[6]知之. 歌曲随机播放，每首最少播放一次的期望是多少？[DB/OL]. (2015.9.4) [2017.12.24]. https://www.zhihu.com/question/35334439/answer/62280714