Sampling from a mixture of different groups of coupons*

Aristides V. Doumas[1] **and Vassilis G. Papanicolaou**[2]

Department of Mathematics
National Technical University of Athens,
Zografou Campus, 157 80, Athens, GREECE

[1]aris.doumas@hotmail.com [2]papanico@math.ntua.gr

November 28, 2018

Abstract

A collector samples coupons with replacement from a pool containing g *uniform* groups of coupons, where "uniform group" means that all coupons in the group are equally likely to occur (while coupons of different groups have different probabilities to occur). For each $j = 1, \dots, g$ let T_j be the number of trials needed to detect Group j, namely to collect all M_j coupons belonging to it at least once. We first derive formulas for the probabilities $P\{T_1 < \cdots < T_g\}$ and $P\{T_1 = \bigwedge_{j=1}^{g} T_j\}$. After that, without severe loss of generality, we restrict ourselves to the case $g = 2$ and compute the asymptotics of $P\{T_1 < T_2\}$ as the number of coupons grows to infinity in a certain manner. Then, we focus on $T := T_1 \vee T_2$, i.e. the number of trials needed to collect all coupons of the pool (at least once), and determine the asymptotics of $E[T]$ and $V[T]$, as well as the limiting distribution of T (appropriately normalized) as the number of coupons becomes large.

Keywords. Coupon collector problems; urn problems; asymptotics, limiting distribution, Gumbel distribution.

2010 AMS Mathematics Classification. 60C05; 60F05; 60F99; 60J10.

1 Introduction of the problem and main results

Coupon collector problems (CCP's) are a popular class of urn problems due to their mathematical elegance, as well as their applications in several areas of

*Available online in Cornell University Library's online repository.

science, from computer science and biology to linguistics and the social sciences. The original problem dates back to De Moivre's treatise *De Mensura Sortis* (1712) and Laplace's *Theorie Analytique des Probabilités* (1812). Nevertheless, new variants of CCP keep arising.

In this paper we study the following CCP version: Suppose we sample coupons independently with replacement from a mixture of g groups of coupons. The first group consists of M_1 coupons each of which having probability p_1 to occur, the secong group of M_2 coupons each of which having probability p_2 to occur, and so on (all numbers M_j, p_j, $j = 1, \dots, g$, are assumed strictly positive). We call "Group j coupons" the coupons of the j-th group. Notice that under our assumptions we must have

$$M_1 p_1 + \cdots + M_g p_g = 1. \tag{1}$$

Thus, for each $j = 1, \dots, g$ the j-th group is a *uniform* family of M_j coupons, where the term "uniform" indicates that all coupons of the group have the same probability p_j to occur. For instance, we can visualize Group 1 as a set of M_1 cards of color 1 (say red), numbered from 1 to M_1, Group 2 as a set of M_2 cards of color 2 (say green), numbered from 1 to M_2, and so on, where each card of color 1 has probability p_1 to occur, each card of color 2 has probability p_2 to occur, and so on.

Suppose we keep drawing coupons one at a time. Naturally, one quantity of interest is the number T of trials (i.e. draws) needed to detect all $M_1 + \cdots + M_g$ coupons (at least once). Some "intermediate" quantities having their own interest are $T_j :=$ the number of trials needed to detect all Group j coupons, $j = 1, \dots, g$. Clearly, T can be expressed as

$$T = \bigvee_{j=1}^{g} T_j, \tag{2}$$

namely the maximum of $T_1, \dots, T_g$.

It is worth mentioning that if we view the coupon sampling process as a sequence $\{C_n\}_{n \geq 1}$ of independent and identically distributed random variables, where each C_n takes values in $\{1, 2, \dots, (M_1 + \cdots + M_g)\}$, namely the set of all existing coupons, with $P\{C_n = i_1\} = p_1$ for $i_1 = 1, 2, \dots, M_1$, $P\{C_n = i_2\} = p_2$ for $i_2 = (M_1 + 1), (M_1 + 2), \dots, (M_1 + M_2)$, and so on (so that, $\{C_n = i\}$ is identified with the event that the type-i coupon is selected at the n-th trial), then T_j, $j = 1, \dots, g$, as well as T are stopping times of the "coupon filtration"

$$\mathcal{F}_n = \sigma(C_1, \dots, C_n), \qquad n \geq 1. \tag{3}$$

Our first quantities of study are the probabilities $P\{T_1 < \cdots < T_g\}$ and $P\{T_1 = T_{\min}\}$, where

$$T_{\min} := \bigwedge_{j=1}^{g} T_j, \tag{4}$$

namely the minimum of $T_1, \dots, T_g$ (thus $P\{T_1 = T_{\min}\}$ is the probability that the Group 1 is the first group to be detected in its entirety). Notice that the equality $T_j = T_k$ is impossible unless, of course, $j = k$.

Theorem 1.

$$P\{T_1 < \cdots < T_g\} =$$
$$K \int_0^{\infty} \cdots \int_0^{t_3} e^{-(p_g t_g + \cdots + p_2 t_2)} \left(1 - e^{-p_g t_g}\right)^{M_g - 1} \cdots \left(1 - e^{-p_2 t_2}\right)^{M_2 - 1} \left(1 - e^{-p_1 t_2}\right)^{M_1} \, dt_2 \cdots dt_g, \tag{5}$$

where

$$K = p_2 p_3 \cdots p_g M_2 M_3 \cdots M_g. \tag{6}$$

Also,

$$P\{T_1 = T_{\min}\} =$$
$$K \int_0^{\infty} \cdots \int_0^{\infty} e^{-(p_g t_g + \cdots + p_2 t_2)} \left(1 - e^{-p_g t_g}\right)^{M_g - 1} \cdots \left(1 - e^{-p_2 t_2}\right)^{M_2 - 1} \left[1 - e^{-p_1 (t_2 \wedge \cdots \wedge t_g)}\right]^{M_1} \, dt_2 \cdots dt_g, \tag{7}$$

Proof. Following a suggestion of Professor Sheldon M. Ross [6] we prove the formulas by applying the powerful technique of "Poissonization."
Let $Z(t)$, $t \geq 0$, be a Poisson process with rate $\lambda = 1$. We imagine that each Poisson event associated to this process is a sampled coupon, so that $Z(t)$ is the number of sampled coupons at time t. Next, for $i = 1, \dots, (M_1 + \cdots + M_g)$, let $Z_i(t)$ be the number of type-i coupons collected at time t. Then, the processes $\{Z_i(t)\}_{t \geq 0}$, $i = 1, \dots, (M_1 + \cdots + M_g)$, are independent Poisson processes with rates p_1 for $i = 1, \dots, M_1$, p_2 for $i = (M_1 + 1), \dots, (M_1 + M_2)$, ..., and, finally, p_g for $i = (M_1 + \cdots + M_{g-1} + 1), \dots, (M_1 + \cdots + M_g)$ [7]. Of course, $Z(t) = Z_1(t) + \cdots + Z_{M_1 + \cdots + M_g}(t)$.
If X_i, $i = 1, \dots, (M_1 + \cdots + M_g)$, denotes the time when the first type-i coupon is collected, i.e. the time of the first Poisson event of the process $Z_i(t)$, then the variables $X_1, \dots, X_{M_1 + \cdots + M_g}$ are clearly independent (being associated to independent processes) and exponentially distributed with parameters p_1 for $i = 1, \dots, M_1$, p_2 for $i = (M_1 + 1), \dots, (M_1 + M_2)$ and so on. We now set

$$\tilde{T}_1 := \bigvee_{i=1}^{M_1} X_i, \qquad \tilde{T}_2 := \bigvee_{i=M_1+1}^{M_1+M_2} X_i, \qquad \dots, \qquad \tilde{T}_g := \bigvee_{i=M_1+\cdots+M_{g-1}+1}^{M_1+\cdots+M_g} X_i. \tag{8}$$

Thus $\tilde{T}_j$, $j = 1, \dots, g$, is the time when all Group j coupons have been detected (at least once) by the process $Z(t)$ and, hence,

$$P\{T_1 < \cdots < T_g\} = P\{\tilde{T}_1 < \cdots < \tilde{T}_g\} \quad \text{and} \quad P\{T_1 = T_{\min}\} = P\{\tilde{T}_1 = \tilde{T}_{\min}\}, \tag{9}$$

where, of course, $\tilde{T}_{\min} := \bigwedge_{j=1}^{g} \tilde{T}_j$.

From the independence of the exponential random variables X_i, $i = 1, \dots, (M_1 + \cdots + M_g)$, it follows that the variables $\tilde{T}_1, \tilde{T}_2, \dots, \tilde{T}_g$ are also independent and, furthermore, (8) implies

$$F_j(t) := P\left\{\tilde{T}_j \le t\right\} = \left(1 - e^{-p_j t}\right)^{M_j}, \qquad t \ge 0, \;\; j = 1, \dots, g, \tag{10}$$

and

$$f_j(t) := F_j'(t) = p_j M_j e^{-p_j t}\left(1 - e^{-p_j t}\right)^{M_j - 1}, \qquad t \ge 0, \;\; j = 1, \dots, g. \tag{11}$$

Therefore,

$$\begin{aligned} P\{\tilde{T}_1 < \cdots < \tilde{T}_g\} &= \int_0^\infty \cdots \int_0^{t_3} \int_0^{t_2} f_g(t_g) \cdots f_2(t_2) f_1(t_1)\, dt_1 dt_2 \cdots dt_g \\ &= \int_0^\infty \cdots \int_0^{t_3} f_g(t_g) \cdots f_2(t_2) F_1(t_2)\, dt_2 \cdots dt_g \end{aligned}$$

and, in view of (9), (10), (11), and (6) the above formula is equivalent to (5). Likewise,

$$\begin{aligned} P\{\tilde{T}_1 = \tilde{T}_{\min}\} &= \int_0^\infty \cdots \int_0^\infty \int_0^{t_2 \wedge \cdots \wedge t_g} f_g(t_g) \cdots f_2(t_2) f_1(t_1)\, dt_1 dt_2 \cdots dt_g \\ &= \int_0^\infty \cdots \int_0^\infty f_g(t_g) \cdots f_2(t_2) F_1(t_2 \wedge \cdots \wedge t_g)\, dt_2 \cdots dt_g, \end{aligned}$$

which establishes (7). ■

Notice that one consequence of formulas (5), (6), and (7) is that the probabilities $P\{T_1 < \cdots < T_g\}$ and $P\{T_1 = T_{\min}\}$ depend only on the ratios $p_2/p_1, \dots, p_g/p_1$.

Corollary 1. For $\ell = 1, \dots, g$ we have

$$P\{T_\ell = T_{\min}\} = \\ (-1)^g \sum_{k_g=1}^{M_g} \cdots \sum_{k_1=1}^{M_1} (-1)^{k_1 + \cdots + k_g} \binom{M_1}{k_1} \cdots \binom{M_g}{k_g} \frac{k_\ell p_\ell}{k_1 p_1 + \cdots + k_g p_g}. \tag{12}$$

In particular, for $g = 2$ we have

$$P\{T_1 < T_2\} = \sum_{k=1}^{M_2} \sum_{j=1}^{M_1} (-1)^{j+k} \binom{M_1}{j} \binom{M_2}{k} \frac{p_1 j}{p_1 j + p_2 k} \tag{13}$$

and

$$P\{T_2 < T_1\} = \sum_{k=1}^{M_2} \sum_{j=1}^{M_1} (-1)^{j+k} \binom{M_1}{j} \binom{M_2}{k} \frac{p_2 k}{p_1 j + p_2 k}. \tag{14}$$

Proof. It is enough to prove (12) only for the case $\ell = 1$. For $j = 2, \dots, g$ we have

$$p_j M_j e^{-p_j t_j}\left(1 - e^{-p_j t_j}\right)^{M_j - 1} = -\sum_{k_j=1}^{M_j} (-1)^{k_j} \binom{M_j}{k_j} p_j k_j e^{-k_j p_j t_j}, \tag{15}$$

while

$$\left[1 - e^{-p_1(t_2 \wedge \cdots \wedge t_g)}\right]^{M_1} = \sum_{k_1=0}^{M_1} (-1)^{k_1} \binom{M_1}{k_1} e^{-k_1 p_1 (t_2 \wedge \cdots \wedge t_g)}. \tag{16}$$

Substituting (15), (16), and (6) in (7) yields

$$P\{T_1 = T_{\min}\} =$$
$$(-1)^{g-1} \sum_{k_g=1}^{M_g} \cdots \sum_{k_2=1}^{M_2} \sum_{k_1=0}^{M_1} (-1)^{k_1+\cdots+k_g} \binom{M_g}{k_g} \cdots \binom{M_1}{k_1} I, \tag{17}$$

where

$$I := \int_0^\infty \cdots \int_0^\infty k_g p_g \cdots k_2 p_2 \, e^{-(k_g p_g t_g + \cdots + k_2 p_2 t_2)} e^{-k_1 p_1 (t_2 \wedge \cdots \wedge t_g)} dt_2 \cdots dt_g. \tag{18}$$

A quick look at (18) reveals that

$$I = E\left[e^{-k_1 p_1 Y_{\min}}\right], \tag{19}$$

where $Y_{\min}$ is the minimum of the independent exponential random variables $Y_2, \ldots, Y_g$ with parameters $(k_2 p_2), \ldots, (k_g p_g)$ respectively. Since (as it is well known) $Y_{\min}$ is exponentially distributed with parameter $k_2 p_2 + \cdots + k_g p_g$, it follows from (19) that

$$I = \frac{k_2 p_2 + \cdots + k_g p_g}{k_1 p_1 + (k_2 p_2 + \cdots + k_g p_g)} \tag{20}$$

and the substitution of (20) in (17) gives

$$P\{T_1 = T_{\min}\} =$$
$$(-1)^{g-1} \sum_{k_g=1}^{M_g} \cdots \sum_{k_2=1}^{M_2} \sum_{k_1=0}^{M_1} (-1)^{k_1+\cdots+k_g} \binom{M_g}{k_g} \cdots \binom{M_1}{k_1} \frac{k_2 p_2 + \cdots + k_g p_g}{k_1 p_1 + k_2 p_2 + \cdots + k_g p_g}, \tag{21}$$

or, equivalently,

$$P\{T_1 = T_{\min}\} =$$
$$(-1)^{g} \sum_{k_g=1}^{M_g} \cdots \sum_{k_2=1}^{M_2} \sum_{k_1=0}^{M_1} (-1)^{k_1+\cdots+k_g} \binom{M_g}{k_g} \cdots \binom{M_1}{k_1} \frac{k_1 p_1}{k_1 p_1 + k_2 p_2 + \cdots + k_g p_g}$$
$$- (-1)^{g} \sum_{k_g=1}^{M_g} \cdots \sum_{k_2=1}^{M_2} \sum_{k_1=0}^{M_1} (-1)^{k_1+\cdots+k_g} \binom{M_g}{k_g} \cdots \binom{M_1}{k_1}. \tag{22}$$

In the first multiple sum of the right-hand side of (22), clearly, the value $k_1 = 0$ of the dummy variable k_1 can be omitted since it does not contribute anything to the sum. Hence we may as well take k_1 to vary from 1 to M_1 (instead of 0 to M_1). As for the second multiple sum of the right-hand side of (22), just notice that it can be factored as

$$\left[\sum_{k_g=1}^{M_g}(-1)^{k_g}\binom{M_g}{k_g}\right]\cdots\left[\sum_{k_2=1}^{M_2}(-1)^{k_2}\binom{M_g}{k_g}\right]\left[\sum_{k_1=0}^{M_1}(-1)^{k_1}\binom{M_1}{k_1}\right],$$

where, obiously, the last factor is equal to 0 (being the binomial expansion of $(1-1)^{M_1}$). Hence the whole multiple sum vanishes, and (22) reduces to (12) (for $\ell = 1$).

Formulas (13) and (14) follow immediately from (12). ■

We can number the groups so that $p_1 < p_2 < \cdots < p_g$. Then, our CCP problem is stochastically bounded between two "extreme" cases where we have only two groups of coupons: (i) one group consisting of M_1 coupons each of which having probability p_1 to occur and another group consisting of $M_2 + \cdots + M_g$ coupons each of which having probability p_2 to occur and (ii) one group consisting of M_1 coupons each of which having probability p_1 to occur and another group consisting of $M_2 + \cdots + M_g$ coupons each of which having probability p_g to occur. Hence, the case $g = 2$ is quite important since it can, at least, provide upper and lower estimates for the more general case of an arbitrary number of groups. With this in mind, let us spell out an immediate corollary of Theorem 1.

Corollary 2. We have

$$P\{T_1 < T_2\} = p_2 M_2 \int_0^\infty e^{-p_2 t}\left(1-e^{-p_1 t}\right)^{M_1}\left(1-e^{-p_2 t}\right)^{M_2-1} dt. \tag{23}$$

Also (by substitution $x = e^{-t}$ in the above integral),

$$\begin{aligned} P\{T_1 < T_2\} &= p_2 M_2 \int_0^1 (1-x^{p_1})^{M_1}(1-x^{p_2})^{M_2-1} x^{p_2-1} dx \\ &= -\int_0^1 (1-x^{p_1})^{M_1}\left[(1-x^{p_2})^{M_2}\right]' dx. \end{aligned} \tag{24}$$

The observation that $P\{T_1 < T_2\}$ depends only on the ratio

$$\lambda := \frac{p_2}{p_1} \tag{25}$$

yields two slightly simplified equivalent versions of (24), namely

$$P\{T_1 < T_2\} = \lambda M_2 \int_0^1 x^{\lambda-1}(1-x)^{M_1}\left(1-x^{\lambda}\right)^{M_2-1} dx \tag{26}$$

and

$$P\{T_1 < T_2\} = M_2 \int_0^1 \left(1-x^{1/\lambda}\right)^{M_1}(1-x)^{M_2-1} dx \tag{27}$$

(formula (23) too can be simplified a little by using the fact that $P\{T_1 < T_2\}$ depends only on λ).

In the sequel, we will assume that $g = 2$, namely that we have only two groups of coupons. Our goal is to understand the behavior of certain quantities as $\lambda = p_2/p_1$ stays fixed, while M_1 and M_2 become large in such a way that

$$M_1 = \nu_1 M \qquad \text{and} \qquad M_2 = \nu_2 M, \tag{28}$$

where $\nu_1 \geq 1$ and $\nu_2 \geq 1$ are fixed integers, while the integer M is allowed to grow. Notice that, under these assumptions p_1 and p_2 depend on M. Then, recalling (1), namely that $M_1 p_1 + M_2 p_2 = 1$, the quantities

$$\alpha_1 := M_1 p_1 = \frac{\nu_1}{\nu_1 + \lambda\nu_2} \qquad \text{and} \qquad \alpha_2 := M_2 p_2 = \frac{\lambda\nu_2}{\nu_1 + \lambda\nu_2} = 1 - \alpha_1 \tag{29}$$

are independent of M too.

In the rest of the paper we study the asymptotic behavior of certain quantities related to $T_1 = T_1(M)$, $T_2 = T_2(M)$, and $T = T(M) = T_1 \vee T_2$, as the integer M grows large. It is notable that our results determine the order of magnitude of the corresponding quantities for the case of g groups, for any $g > 2$.

In Section 2 we derive the asymptotic formula (Theorem 2)

$$P\{T_1 < T_2\} \sim \frac{\nu_2 \lambda \Gamma(\lambda)}{\nu_1^{\lambda}} \cdot \frac{1}{M^{\lambda - 1}}, \qquad M \to \infty,$$

under the assumption that λ of (25) is > 1 (as usual, the notation $f(M) \sim g(M)$ means that $f(M)/g(M) \to 1$ as $M \to \infty$).

Section 3 contains some auxiliary topics including a key example discussed in Subsection 3.1. These topics are used in Section 4 in order to determine the asymptotic behavior of the expectation, the variance, as well as the distribution of T_1, T_2, and T as $M \to \infty$. Some indicative results of Section 4 are:

(i) A formula for the asymptotics of the expectation of T

$$E[T] = (\nu_1 + \lambda\nu_2) M \ln M + (\nu_1 + \lambda\nu_2)(\gamma + \ln \nu_1) M + O\left(M^{2-\lambda} \ln M\right), \qquad M \to \infty.$$

This formula follows immediately from Theorem 6.

(ii) A formula for the asymptotics of the variance of T

$$V[T] \sim \frac{\pi^2 (\nu_1 + \lambda\nu_2)^2}{6} M^2, \qquad M \to \infty$$

(this is Corollary 4).

(iii) The limiting distribution of T (appropriately normalized). We have shown that the random variable

$$\frac{T - (\nu_1 + \lambda\nu_2) M \ln M}{(\nu_1 + \lambda\nu_2) M} - \ln \nu_1$$

converges in distribution to the standard Gumbel random variable as $M \to \infty$ (Theorem 8).

The above three results are presented in of Subsection 4.2 and hold under the assumption that $\lambda > 1$.
Finally, for the betterment of the flow of the paper, the derivations of formulas (63) and (105) are placed in the appendix (Section 5).

2 Asymptotics of $P\{T_1 < T_2\}$

Equation (26) can be written as

$$P\{T_1 < T_2\} = \lambda \nu_2 M \, I_M, \quad \text{where } I_M = \int_0^1 x^{\lambda-1} (1-x)^{\nu_1 M} \left(1 - x^{\lambda}\right)^{\nu_2 M - 1} dx. \tag{30}$$

Thus, the asymptotic behavior of $P\{T_1 < T_2\}$ as $M \to \infty$ reduces to the asymptotic behavior of I_M.
For convenience we will assume from now on, without loss of generality, that

$$\lambda = \frac{p_2}{p_1} > 1. \tag{31}$$

Formula (30) yields immediately the following upper bound for I_M:

$$I_M < \int_0^1 x^{\lambda-1} (1-x)^{\nu_1 M} dx = B(\lambda, \nu_1 M + 1) = \frac{\Gamma(\lambda)\,\Gamma(\nu_1 M + 1)}{\Gamma(\lambda + \nu_1 M + 1)}, \tag{32}$$

where $B(\,\cdot\,,\cdot)$ and $\Gamma(\,\cdot)$ denote the Beta and Gamma funtion respectively, while an immediate consequence of Stirling's formula is that

$$\int_0^1 x^{\lambda-1} (1-x)^{\nu_1 M} dx = \frac{\Gamma(\lambda)\,\Gamma(\nu_1 M + 1)}{\Gamma(\lambda + \nu_1 M + 1)} \sim \frac{\Gamma(\lambda)}{\nu_1^{\lambda}} \cdot \frac{1}{M^{\lambda}}, \qquad M \to \infty. \tag{33}$$

Next, we need to find a satisfactory lower bound for I_M. Let $0 < \varepsilon < 1 - (1/\lambda)$, so that $(1/\lambda) + \varepsilon < 1$. Then, (30) implies

$$I_M > I_M^{\flat} := \int_0^{M^{-(1/\lambda)-\varepsilon}} x^{\lambda-1} (1-x)^{\nu_1 M} \left(1 - x^{\lambda}\right)^{\nu_2 M - 1} dx. \tag{34}$$

For $x \in [0, M^{-(1/\lambda)-\varepsilon}]$, we have

$$0 \geq (\nu_2 M - 1) \ln\left(1 - x^{\lambda}\right) \geq (\nu_2 M - 1) \ln\left(1 - \frac{1}{M^{1+\lambda\varepsilon}}\right) = -\frac{\nu_2}{M^{\lambda\varepsilon}} + O\left(\frac{1}{M^{1+\lambda\varepsilon}}\right) \tag{35}$$

as $M \to \infty$ (uniformly in x). Hence

$$1 \geq \left(1 - x^{\lambda}\right)^{\nu_2 M - 1} \geq \exp\left(-\frac{\nu_2}{M^{\lambda\varepsilon}}\right) \left[1 + O\left(\frac{1}{M^{1+\lambda\varepsilon}}\right)\right] = 1 + O\left(\frac{1}{M^{\lambda\varepsilon}}\right). \tag{36}$$

Using (36) in (34) yields

$$\int_0^{M^{-(1/\lambda)-\varepsilon}} x^{\lambda-1}(1-x)^{\nu_1 M}\,dx \geq I_M^\flat \geq \left[1+O\left(\frac{1}{M^{\lambda\varepsilon}}\right)\right]\int_0^{M^{-(1/\lambda)-\varepsilon}} x^{\lambda-1}(1-x)^{\nu_1 M}\,dx \tag{37}$$

or

$$I_M^\flat \sim \int_0^{M^{-(1/\lambda)-\varepsilon}} x^{\lambda-1}(1-x)^{\nu_1 M}\,dx, \qquad M\to\infty. \tag{38}$$

Finally, we notice that the fact that $(1/\lambda)+\varepsilon<1$ implies

$$\int_{M^{-(1/\lambda)-\varepsilon}}^1 x^{\lambda-1}(1-x)^{\nu_1 M}\,dx < \left(1-\frac{1}{M^{(1/\lambda)+\varepsilon}}\right)^{\nu_1 M} = O\left(\frac{1}{M^r}\right) \quad \text{for any } r>0, \tag{39}$$

thus, in view of (33), formulas (38) and (39) give

$$I_M^\flat \sim \int_0^1 x^{\lambda-1}(1-x)^{\nu_1 M}\,dx \sim \frac{\Gamma(\lambda)}{\nu_1^\lambda}\cdot\frac{1}{M^\lambda}, \qquad M\to\infty. \tag{40}$$

Hence, the combination of (32), (33), (34), and (40) yields

$$I_M \sim \frac{\Gamma(\lambda)}{\nu_1^\lambda}\cdot\frac{1}{M^\lambda}, \qquad M\to\infty. \tag{41}$$

Therefore, by applying (41) in (30) we obtain the following result.

Theorem 2. If $\lambda = p_2/p_1 > 1$, then

$$P\{T_1<T_2\} \sim \frac{\nu_2\lambda\Gamma(\lambda)}{\nu_1^\lambda}\cdot\frac{1}{M^{\lambda-1}} = \frac{\nu_2\Gamma(\lambda+1)}{\nu_1^\lambda}\cdot\frac{1}{M^{\lambda-1}} = \frac{\nu_2}{\nu_1}\cdot\frac{\Gamma(\lambda+1)}{M_1^{\lambda-1}} \tag{42}$$

as $M\to\infty$.

Notice that, no matter how big the ratio $M_2/M_1=\nu_2/\nu_1$ is, the probability $P\{T_1<T_2\}$ approaches 0 as $M\to\infty$, as long as λ is bigger than 1 (even slightly).

3 Auxiliary material

Suppose we sample independently with replacement from a pool of N coupons, where the probability of the j-th coupon to occur is q_j, $j=1,\dots,N$ (the q_n's are usually referred as the "coupon probabilities"). Let $S=S_N$ denote the number of trials needed in order to detect all N coupons. Obviously, the possible values of S_N are $N, N+1, \dots$ (it is easy to see that $P\{S_N<\infty\}=1$ as long as $q_j>0$ for all j; actually, from the generating function $E\left[z^{-S_N}\right]$, as computed in [4], one can easily see that $P\{S_N=k\}$ deays exponentially as $k\to\infty$).

For the purposes of this paper we will need a formula for the expectation $E[S_N^{(r)}]$ for any real $r>0$, where

$$s^{(r)} := \frac{\Gamma(s+r)}{\Gamma(s)} \tag{43}$$

is the “natural” extension of the so-called Pochhammer function.
If we denote by W_j the number of trials needed in order to detect the j-th coupon, then, it is clear that W_j is a geometric random variable with parameter q_j and

$$S_N = \bigvee_{j=1}^{N} W_j.$$

However, the above formula for S_N is not very useful, since the W_j's are not independent. Instead, we can employ again the “Poissonization technique” (see, e.g., [7]) in order to get an explicit formulas for $E[S_N^{(r)}]$.
As in the proof of Thorem 1, we take $Z(t)$, $t \geq 0$, to be a Poisson process with rate $\lambda = 1$. We imagine that each Poisson event associated to Z is a collected coupon, so that $Z(t)$ is the number of detected coupons at time t. Next, for $j = 1, \ldots, N$, let $Z_j(t)$ be the number of times that the j-th coupon has been detected up to time t. Then, the processes $\{Z_j(t)\}_{t\geq 0}$, $j = 1, \ldots, N$, are independent Poisson processes with rates q_j respectively [7] and, of course, $Z(t) = Z_1(t) + \cdots + Z_N(t)$. If X_j, $j = 1, \ldots, N$, denotes the time of the first event of the process Z_j, then $X_1, \ldots, X_N$ are obviously independent (being associated to independent processes), while their maximum

$$X = \bigvee_{j=1}^{N} X_j \tag{44}$$

is the time when all different coupons have been detected at least once.
Now, for each $j = 1, \ldots, N$ the random variable X_j is exponentially distributed with parameter q_j, i.e.

$$P\{X_j \leq t\} = 1 - e^{-q_j t}, \qquad t \geq 0. \tag{45}$$

It follows from (44) and the independence of the X_j's that

$$P\{X \leq t\} = \prod_{j=1}^{N} \left(1 - e^{-q_j t}\right), \qquad t \geq 0. \tag{46}$$

Next, we observe that S_N and X are related as

$$X = \sum_{k=1}^{S_N} U_k, \tag{47}$$

where $U_1, U_2, \ldots$ are the interarrival times of the process Z. It is common knowledge that the U_j's are independent and exponentially distributed random variables with parameter 1. Hence for any integer $m \geq 1$ the sum $U_1 + \cdots + U_m$ follows the Erlang distribution with parameters m and 1. Therefore,

$$E\left[\phi\left(\sum_{k=1}^{m} U_k\right)\right] = \int_0^{\infty} \phi(\xi)\, \frac{\xi^{m-1}}{(m-1)!}\, e^{-\xi} d\xi, \tag{48}$$

where $\phi(x)$ is any (Lebesgue) measurable function on $(0, \infty)$ for which the integral in (48) makes sense (i.e. converges absolutely). Noticing that S_N is independent of the U_j's, formulas (47) and (48) imply

$$E\left[\phi(X) \mid S_N\right] = \int_0^\infty \phi(\xi) \, \frac{\xi^{S_N - 1}}{(S_N - 1)!} \, e^{-\xi} d\xi \tag{49}$$

and, consequently (by taking expectations)

$$E\left[\phi(X)\right] = E\left[\int_0^\infty \phi(\xi) \, \frac{\xi^{S_N - 1}}{(S_N - 1)!} \, e^{-\xi} d\xi\right]. \tag{50}$$

If we take $\phi(x) = x^r$ for a fixed real number $r > 0$, then (50) becomes

$$E\left[X^r\right] = E\left[\int_0^\infty \frac{\xi^{S_N + r - 1}}{(S_N - 1)!} \, e^{-\xi} d\xi\right] = E\left[\frac{\Gamma(S_N + r)}{(S_N - 1)!}\right] = E\left[S_N^{(r)}\right]. \tag{51}$$

Finally, by using (46) in (51) we obtain the following result.

Lemma 1. For any real number $r > 0$ we have

$$E\left[S_N^{(r)}\right] = E\left[\frac{\Gamma(S_N + r)}{\Gamma(S_N)}\right] = r \int_0^\infty t^{r-1} \left[1 - \prod_{j=1}^{N} \left(1 - e^{-q_j t}\right)\right] dt. \tag{52}$$

In particular for $r = 1$ we have

$$E\left[S_N\right] = \int_0^\infty \left[1 - \prod_{j=1}^{N} \left(1 - e^{-q_j t}\right)\right] dt, \tag{53}$$

while for $r = 2$ we have

$$E\left[S_N^{(2)}\right] = E\left[S_N(S_N + 1)\right] = 2 \int_0^\infty t \left[1 - \prod_{j=1}^{N} \left(1 - e^{-q_j t}\right)\right] dt. \tag{54}$$

Remark 1. Let the random variables $\Xi_1, \dots, \Xi_N$ be independent and exponentially distributed with parameters $q_1, \dots, q_N$ respectively. If

$$\Xi_{\max} := \bigvee_{j=1}^{N} \Xi_j,$$

then formula (52) tells us that for any real number $r > 0$ we have

$$E\left[S_N^{(r)}\right] = E\left[\frac{\Gamma(S_N + r)}{\Gamma(S_N)}\right] = E\left[\Xi_{\max}^r\right].$$

Let us also notice that by expanding the product inside the integral in (52) and integrate the resulting sum term by term we obtain the expression

$$E\left[S_N^{(r)}\right] = \Gamma(r+1) \sum_{\substack{J \subset \{1,\dots,N\} \\ J \neq \emptyset}} \frac{(-1)^{|J|-1}}{\left(\sum_{j \in J} q_j\right)^r}$$
$$= \Gamma(r+1) \sum_{m=1}^{N} (-1)^{m-1} \sum_{1 \leq j_1 < \cdots < j_m \leq N} \frac{1}{(q_{j_1} + \cdots + q_{j_m})^r}, \tag{55}$$

where $|J|$ denotes the cardinality of J.

Remark 2. Let us first observe that since S_N is always a positive integer, the quantity

$$\frac{\Gamma(r+S_N)}{\Gamma(r)\,\Gamma(S_N)} \tag{56}$$

makes sense for every $r \in \mathbb{C}$; actually, it is entire in r (the poles of $\Gamma(r+S_N)$ are cancelled by the zeros of $\Gamma(r)^{-1}$). Now, let us look at the function

$$H(r) := \frac{1}{\Gamma(r)}\, E\left[S_N^{(r)}\right] = E\left[\frac{\Gamma(r+S_N)}{\Gamma(r)\,\Gamma(S_N)}\right] = \sum_{k=N}^{\infty} \frac{\Gamma(k+r)}{\Gamma(r)\,(k-1)!}\, P\{S_N = k\}. \tag{57}$$

Since (i) $\Gamma(k+r)/(k-1)! \sim k^r$ as $k \to \infty$ (see, e.g., formula (69) below) and (ii) $P\{S_N = k\}$ decays exponentially in k, it follows that the series in (57) converges uniformly (and absolutely) in r on any compact subset of the complex plane $\mathbb{C}$. Therefore, $H(r)$ is an entire function and consequently formula (57) implies that $E[S_N^{(r)}]$ is meromorphic in r whose poles are located at $-N, -(N+1), \dots$. Although the fact that $E[S_N^{(r)}]$ is meromorphic also follows from (55), it is not obvious from this formula that there are no poles at $-1, -2, \dots, -(N-1)$.

Now let us consider the "uniform" case, namely the case where all N coupons are equally likely to occur, i.e.

$$q_j = \frac{1}{N} \qquad \text{for } \; j = 1, \dots, N. \tag{58}$$

In this case formula (52) becomes

$$E\left[S_N^{(r)}\right] = r \int_0^\infty t^{r-1} \left[1 - \left(1 - e^{-t/N}\right)^N\right] dt. \tag{59}$$

Substituting $t = Ns$ in the above integral gives

$$E\left[S_N^{(r)}\right] = N^r \int_0^\infty r s^{r-1} \left[1 - (1 - e^{-s})^N\right] ds. \tag{60}$$

Next, we integrate by parts and get

$$E\left[S_N^{(r)}\right] = N^{r+1} \int_0^\infty s^r (1 - e^{-s})^{N-1} e^{-s} ds. \tag{61}$$

Then, we make the substitution $s = \ln N - \ln x$ (so that $x = Ne^{-s}$) in the integral of (61) and obtain

$$E\left[S_N^{(r)}\right] = N^r \ln^r N \int_0^N \left(1 - \frac{x}{N}\right)^{N-1} \left(1 - \frac{\ln x}{\ln N}\right)^r dx. \tag{62}$$

Starting from (62), it can be shown that, for any given $r > 0$, the asymptotic behavior of $E[S_N^{(r)}]$ as $N \to \infty$ is

$$E\left[S_N^{(r)}\right] = N^r (\ln N)^r \sum_{k=0}^{n} \binom{r}{k} \frac{(-1)^k}{\ln^k N} \int_0^\infty e^{-x} (\ln x)^k dx + o\left(\frac{1}{\ln^n N}\right) \tag{63}$$

for every $n = 1, 2, \ldots$. Here, $\binom{r}{k}$ stands for the generalized binomial coefficient (in the sense that r is not necessarily a positive integer—see formulas (172) and (173) in Subsection 5.1 of the Appendix).
Intuitively, it is not hard to see why (62) implies (63). However the complete proof is quite long and for this reason is given in the Subsection 5.1 of the Appendix.
To further simplify (63), let us first notice that if we differentiate k times the Gamma function $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ and then set $z = 1$, we get

$$\Gamma^{(k)}(1) = \int_0^\infty e^{-x} (\ln x)^k dx. \tag{64}$$

Of course, $\Gamma^{(0)}(1) = \Gamma(1) = 1$. As for the derivatives $\Gamma^{(k)}(1)$, $k = 1, 2, \ldots$, there are some known expressions and recursions (see, e.g., [2] and the references therein). For instance,

$$\Gamma^{(1)}(1) = -\gamma, \quad \Gamma^{(2)}(1) = \frac{\pi^2}{6} + \gamma^2, \quad \Gamma^{(3)}(1) = -\left[2\zeta(3) + \frac{\pi^2}{2}\gamma + \gamma^3\right], \quad \text{etc.}, \tag{65}$$

where $\gamma = 0.5772...$ is the Euler (or Euler-Macheroni) constant and $\zeta(\cdot)$ is the Riemann Zeta function.
Using (64) we can write (63) as

$$E\left[S_N^{(r)}\right] = N^r (\ln N)^r \left[\sum_{k=0}^{n} \binom{r}{k} \frac{(-1)^k \, \Gamma^{(k)}(1)}{\ln^k N} + o\left(\frac{1}{\ln^n N}\right)\right], \quad N \to \infty, \tag{66}$$

for every $n = 1, 2, \ldots$. Formula (66) can be written equivalently as an asymptotic series (for the definition of the asymptotic series and the associated usage of the symbol $\sim$ see, e.g., [1])

$$E\left[S_N^{(r)}\right] \sim N^r (\ln N)^r \sum_{k=0}^{\infty} \binom{r}{k} \frac{(-1)^k \, \Gamma^{(k)}(1)}{\ln^k N}, \qquad N \to \infty, \tag{67}$$

for any $r > 0$ (of course, if r is an integer, $\binom{r}{k} = 0$ for $k > r$ and the series becomes a finite sum). In particular, the leading behavior of $E[S_N^{(r)}]$ is

$$E\left[S_N^{(r)}\right] \sim N^r (\ln N)^r, \qquad N \to \infty. \tag{68}$$

Let us also mention that in the case where r is a positive integer there are more detailed expressions for $E[S_N^{(r)}]$ (see, e.g., [4] and the references therein).
Finally, from (68) we can easily obtain an asymptotic formula for $E[S_N^r]$ as $N \to \infty$. For a fixed $r > 0$ Stirling's formula yields

$$s^{(r)} = \frac{\Gamma(s+r)}{\Gamma(s)} \sim s^r, \qquad s \to \infty. \tag{69}$$

Since $S_N \geq N$, formula (70) implies that, for any $\varepsilon > 0$ there is a $N_0 = N_0(\varepsilon)$ such that

$$(1-\varepsilon)\, S_N^{(r)} \leq S_N^r \leq (1+\varepsilon)\, S_N^{(r)} \qquad \text{for any} \;\; N \geq N_0, \tag{70}$$

and consequently,

$$(1-\varepsilon)\, E\left[S_N^{(r)}\right] \leq E\left[S_N^r\right] \leq (1+\varepsilon)\, E\left[S_N^{(r)}\right] \qquad \text{for any} \;\; N \geq N_0, \tag{71}$$

i.e., in view of (68),

$$E\left[S_N^r\right] \sim E\left[S_N^{(r)}\right] \sim N^r (\ln N)^r, \qquad N \to \infty, \tag{72}$$

for any $r > 0$.

3.1 A preliminary example

Suppose our set of coupons is $\{0, 1, \ldots, N\}$ with corresponding probabilities

$$q_0 = \theta \qquad \text{and} \qquad q_j = \frac{1-\theta}{N}, \quad j = 1, \ldots N, \tag{73}$$

where $\theta \in (0,1)$ is a given number. Let $S(\theta) = S(\theta; N)$ be the number of trials needed until all $N+1$ coupons are detected (thus $S(\theta; N) = S_{N+1}$ under the previous notation). Then, (53) gives (in the sequel, the dependence of $S(\theta)$ on N will be suppressed for typographical convenience)

$$E\left[S(\theta)\right] = \int_0^\infty \left[1 - \left(1 - e^{-\theta t}\right)\left(1 - e^{-(1-\theta)t/N}\right)^N\right] dt$$

or

$$E\left[S(\theta)\right] = J_1(N;\theta) + J_2(N;\theta), \tag{74}$$

where

$$J_1(N;\theta) := \int_0^\infty \left[1 - \left(1 - e^{-(1-\theta)t/N}\right)^N\right] dt \tag{75}$$

and

$$J_2(N;\theta) := \int_0^\infty e^{-\theta t} \left(1 - e^{-(1-\theta)t/N}\right)^N dt. \tag{76}$$

The integral $J_1(N;\theta)$ of (75) reminds the expectation of S_N in the case where all N coupons are equally likely to occur. This is very easy to see via the substitution $y = 1 - e^{-(1-\theta)t/N}$ which yields

$$J_1(N;\theta) = \frac{N}{1-\theta}\int_0^1 \frac{1-y^N}{1-y}\,dy = \frac{N}{1-\theta}\int_0^1 \left(\sum_{j=0}^{N-1} y^j\right) dy = \frac{N}{1-\theta}\sum_{j=1}^{N}\frac{1}{j}, \quad (77)$$

or

$$J_1(N;\theta) = \frac{NH_N}{1-\theta}, \qquad \text{where} \qquad H_N := \sum_{j=1}^{N}\frac{1}{j}. \quad (78)$$

The quantity H_N is called the N-th harmonic number and its full asymptotic expansion, as $N \to \infty$, is well known (see, e.g., [1]):

$$H_N \sim \ln N + \gamma + \frac{1}{2N} + \sum_{k=1}^{\infty}\frac{B_{2k}}{2k}\cdot\frac{1}{N^{2k}}, \quad (79)$$

where B_m is the m-th Bernoulli number defined by the formula

$$\frac{z}{e^z - 1} = \sum_{m=1}^{\infty}\frac{B_m}{m!}z^m. \quad (80)$$

Since $z(e^z-1)^{-1} + z/2$ is an even function of z, we have that $B_{2k+1} = 0$ for all $k \geq 1$.
Next, let us bound the integral $J_2(N;\theta)$ of (76). For any fixed $\rho > 0$ formula (76) implies

$$\begin{aligned} J_2(N;\theta) &\leq \int_0^{\rho N}\left(1 - e^{-(1-\theta)t/N}\right)^N dt + \int_{\rho N}^{\infty} e^{-\theta t}dt \\ &\leq \rho N\left(1 - e^{-(1-\theta)\rho}\right)^N + \frac{1}{\theta\rho N}\,e^{-\theta\rho N}, \end{aligned} \quad (81)$$

Hence, there is an $\varepsilon_1 > 0$ such that for any fixed $\varepsilon \in (0, \varepsilon_1)$ we have

$$J_2(N;\theta) = O\left(e^{-\varepsilon N}\right), \qquad N \to \infty \quad (82)$$

(ε is a symbol we recycle).
Using (78) and (82) in (74) yields

$$E\left[S(\theta)\right] = \frac{NH_N}{1-\theta} + O\left(e^{-\varepsilon N}\right), \qquad N \to \infty \quad (83)$$

(let us recall that in the case where all N coupons are equally likely to occur we have $E[S_N] = NH_N$).
The full asymptotic expansion of $E\left[S(\theta)\right]$ can be obtained immediately by applying (79) in (83). In particular,

$$E\left[S(\theta)\right] = \frac{N\ln N}{1-\theta} + \frac{\gamma N}{1-\theta} + O(1), \qquad N \to \infty. \quad (84)$$

In the same way we can get the asymptotics of the second rising moment $E\left[S(\theta)^{(2)}\right]$ of $S(\theta)$. By (54) and (73) we get

$$E\left[S(\theta)^{(2)}\right] = \tilde{J}_1(N;\theta) + \tilde{J}_2(N;\theta), \tag{85}$$

where

$$\tilde{J}_1(N;\theta) := 2\int_0^\infty t\left[1 - \left(1 - e^{-(1-\theta)t/N}\right)^N\right] dt \tag{86}$$

and

$$\tilde{J}_2(N;\theta) := 2\int_0^\infty te^{-\theta t}\left(1 - e^{-(1-\theta)t/N}\right)^N dt. \tag{87}$$

The approach we used to bound $J_2(N;\theta)$ applies to $\tilde{J}_2(N;\theta)$ as well and it implies that there is an $\varepsilon_2 > 0$ such that for any fixed $\varepsilon \in (0, \varepsilon_2)$ we have

$$\tilde{J}_2(N;\theta) = O\left(e^{-\varepsilon N}\right), \qquad N \to \infty. \tag{88}$$

To calculate $\tilde{J}_1(N;\theta)$ we substitute $s = (1-\theta)t$ in the integral of (86) and obtain

$$\tilde{J}_1(N;\theta) = \frac{2}{(1-\theta)^2}\int_0^\infty s\left[1 - \left(1 - e^{-s/N}\right)^N\right] ds \tag{89}$$

The integral in the right-hand side of (89) equals $E[S_N^{(2)}]$, where S_N is the number of trials needed to collect all coupons in the uniform case where all N coupons are equally likely to occur. Since $r = 2$ is a very special value, we can get more precise results than the ones coming directly from formula (67). Indeed, it is not hard to show (see, e.g., [3]) that

$$E\left[S_N^{(2)}\right] = N^2\left(H_N^2 + \sum_{j=1}^N \frac{1}{j^2}\right). \tag{90}$$

Therefore, (89) becomes

$$\tilde{J}_1(N;\theta) = \frac{N^2}{(1-\theta)^2}\left(H_N^2 + \sum_{j=1}^N \frac{1}{j^2}\right). \tag{91}$$

Using (88) and (91) in (85) we finally get

$$E\left[S(\theta)^{(2)}\right] = \frac{N^2}{(1-\theta)^2}\left(H_N^2 + \sum_{j=1}^N \frac{1}{j^2}\right) + O\left(e^{-\varepsilon N}\right), \qquad N \to \infty, \tag{92}$$

for sufficiently small $\varepsilon > 0$. The full asymptotic behavior of $\sum_{j=1}^N j^{-2}$ is well known (see, e.g., [1])

$$\sum_{j=1}^N \frac{1}{j^2} \sim \frac{\pi^2}{6} - \frac{1}{N} + \frac{1}{2N^2} - \sum_{k=1}^\infty \frac{B_{2k}}{N^{2k+1}}, \tag{93}$$

hence we can obtain easily the full asymptotic expansion of $E\left[S(\theta)^{(2)}\right]$ by using (79) and (93) in (92).
For the variance of $S(\theta)$ we have

$$V\left[S(\theta)\right] = E\left[S(\theta)^{(2)}\right] - E\left[S(\theta)\right] - E\left[S(\theta)\right]^2, \tag{94}$$

hence applying (83) and (92) in (94) yields

$$V\left[S(\theta)\right] = \frac{N^2}{(1-\theta)^2}\sum_{j=1}^{N}\frac{1}{j^2} - \frac{NH_N}{1-\theta} + O\left(e^{-\varepsilon N}\right), \qquad N \to \infty, \tag{95}$$

for $\varepsilon > 0$ sufficiently small. Again, the full asymptotic expansion of $V[S(\theta)]$ can be obtained immediately with the help of (79) and (93). In particular,

$$V\left[S(\theta)\right] = \frac{\pi^2 N^2}{6(1-\theta)^2}\left[1 + O\left(\frac{\ln N}{N}\right)\right], \qquad N \to \infty. \tag{96}$$

In a similar fashion, we can compute the asymptotics of the fractional rising moments of $S(\theta)$. For $r > 0$, in view of (73), formula (52) becomes

$$\begin{aligned} E\left[S(\theta)^{(r)}\right] &= r\int_0^\infty t^{r-1}\left[1 - \left(1 - e^{-\theta t}\right)\left(1 - e^{-(1-\theta)t/N}\right)^N\right] dt \\ &= r\int_0^\infty t^{r-1}\left[1 - \left(1 - e^{-(1-\theta)t/N}\right)^N\right] dt + r\int_0^\infty t^{r-1} e^{-\theta t}\left(1 - e^{-(1-\theta)t/N}\right)^N dt, \end{aligned} \tag{97}$$

thus, in the same way we got (82), we can now get

$$E\left[S(\theta)^{(r)}\right] = r\int_0^\infty t^{r-1}\left[1 - \left(1 - e^{-(1-\theta)t/N}\right)^N\right] dt + O\left(e^{-\varepsilon N}\right), \qquad N \to \infty, \tag{98}$$

for $\varepsilon > 0$ sufficiently small. Next, as usual, we substitute $s = (1-\theta)t$ in the integral of (98) and obtain

$$E\left[S(\theta)^{(r)}\right] = \frac{1}{(1-\theta)^r}\int_0^\infty r s^{r-1}\left[1 - \left(1 - e^{-s/N}\right)^N\right] ds + O\left(e^{-\varepsilon N}\right), \quad N \to \infty, \tag{99}$$

In view of (59), the integral in the right-hand side of (99) equals $E[S_N^{(r)}]$, where S_N is the number of trials needed to collect all coupons in the uniform case. Hence, we can use formula (67) in (99) and conclude that

$$E\left[S(\theta)^{(r)}\right] \sim \frac{N^r(\ln N)^r}{(1-\theta)^r}\sum_{k=0}^{\infty}\binom{r}{k}\frac{(-1)^k\,\Gamma^{(k)}(1)}{\ln^k N}, \qquad N \to \infty, \tag{100}$$

for any $r > 0$. In particular,

$$E\left[S(\theta)^{(r)}\right] \sim \frac{N^r(\ln N)^r}{(1-\theta)^r}, \qquad N \to \infty. \tag{101}$$

Furthermore, since $S(\theta) \geq N+1$, in the same way we obtained (72), we can now get

$$E\left[S(\theta)^r\right] \sim \frac{N^r (\ln N)^r}{(1-\theta)^r}, \qquad N \to \infty, \tag{102}$$

for any $r > 0$.
Finally, we will give the limiting distribution of $S(\theta)$ as $N \to \infty$. The formulas for the moments and the variance of $S(\theta)$ suggest that the law of the random variable $(1-\theta)S(\theta)$ must be very close to the law of $S_N =$ the number of trials needed to detect all N coupons in the uniform case where all coupons are equally likely to occur.
The limiting distribution of S_N as $N \to \infty$ has been found in 1961 by Erdős and Rényi [5]:

$$\frac{S_N - N \ln N}{N} \xrightarrow{D} Y \qquad \text{as } \; N \to \infty \tag{103}$$

(the symbol $\xrightarrow{D}$ denotes convergence in distribution) where

$$F(y) = P\{Y \leq y\} = \exp\left(-e^{-y}\right), \qquad y \in \mathbb{R}, \tag{104}$$

namely Y is a standard Gumbel random variable. Therefore, it is not surprising that

$$\frac{(1-\theta)S(\theta) - N \ln N}{N} \xrightarrow{D} Y \qquad \text{as } \; N \to \infty, \tag{105}$$

where, again, Y is a standard Gumbel random variable.
Our proof of formula (105) is based on characteristic functions. The details are given in the Subsection 5.2 of the Appendix.

4 The asymptotic behavior of T_1, T_2, and T

4.1 The random variables T_1 and T_2

If we are only interested in the variable $T_1 = T_1(M)$ alone, namely the number of trials needed to collect all $M_1 = \nu_1 M$ coupons of Group 1, then all the coupons of Group 2 feel the same to us, and consequently we can assume that the Group 2 consists of only one coupon having probability $M_2 p_2 = \alpha_2$ to occur (recall (29)). Under this point of view, the number of trials $S = S_{M_1+1}$ needed to detect the totality of the $M_1 + 1$ existing coupons (i.e. the M_1 coupons of Group 1 plus the single coupon of Group 2) can be identified with the variable $S(\theta) = S(\theta; N)$ studied in Subsection 3.1, where $\theta = \alpha_2$ and $N = M_1 = \nu_1 M$. Although in our notation we will usually suppress the dependence on M for typographical convenience, we should always keep in mind that both T_1 and S below depend on the integer M.
Obviously, $T_1 \leq S$ and the event $\{T_1 < S\}$ happens if and only if the Group 2 coupon occurs last, namely after detecting all $\nu_1 M$ Group 1 coupons. Therefore,

$$P\{T_1 < S\} \leq P\{\nu_1 M < S\} = (1-\alpha_2)^{\nu_1 M} = \alpha_1^{\nu_1 M} \tag{106}$$

(the last equality follows from the fact that, in view of (29), $\alpha_1+\alpha_2=1$). This is a rather crude estimate of the probability of $\{T_1<S\}$, but it will be sufficient for our purpose.
Next, we will estimate the difference $S-T_1$ in the L_1 sense. Let us first notice that,

$$S-T_1=(S-T_1)\,\mathbf{1}_{\{T_1<S\}}. \tag{107}$$

Then, taking expectations in (107) yields

$$E\left[S-T_1\right]=E\left[(S-T_1)\,\mathbf{1}_{\{T_1<S\}}\right]=E\left[S-T_1\,|\,T_1<S\right]P\{T_1<S\}. \tag{108}$$

Now, notice that for $k=1,2,\dots$, we have $P\{S-T_1=k\,|\,T_1<S\}=\alpha_1^{k-1}\alpha_2$. Thus, the conditional distribution of $S-T_1$, given $\{T_1<S\}$, is geometric with parameter α_2. Therefore, $E\left[S-T_1\,|\,T_1<S\right]=1/\alpha_2$ and (108) becomes

$$E\left[S\right]-E\left[T_1\right]=E\left[S-T_1\right]=\frac{1}{\alpha_2}\,P\{T_1<S\}. \tag{109}$$

which, in view of (106), implies that S and T_1 get very close in the L_1 sense as $M\to\infty$. As for the asymptotics of $E[T_1]$, we can use (83) (with $\theta=\alpha_2$ and $N=\nu_1 M$) and (106) in (109) and obtain immediately the following result:

Theorem 3. For every sufficiently small $\varepsilon>0$ we have

$$E\left[T_1\right]=\frac{\nu_1}{\alpha_1}\,MH_{\nu_1 M}+O\left(e^{-\varepsilon M}\right)=(\nu_1+\lambda\nu_2)MH_{\nu_1 M}+O\left(e^{-\varepsilon M}\right),\quad M\to\infty, \tag{110}$$

where H_N is the N-th harmonic number (see (78)).
Likewise,

$$E\left[T_2\right]=\frac{\nu_2}{\alpha_2}\,MH_{\nu_2 M}+O\left(e^{-\varepsilon M}\right)=\left(\lambda^{-1}\nu_1+\nu_2\right)MH_{\nu_2 M}+O\left(e^{-\varepsilon M}\right),\quad M\to\infty. \tag{111}$$

For example, in view of (79), formula (110) implies

$$E\left[T_1\right]=(\nu_1+\lambda\nu_2)M\ln M+(\nu_1+\lambda\nu_2)(\gamma+\ln\nu_1)M+\frac{\alpha_1}{2}+O\left(\frac{1}{M}\right),\quad M\to\infty, \tag{112}$$

where, recalling (29), we have that $\alpha_1=\nu_1/(\nu_1+\lambda\nu_2)$.
We continue by noticing that in a similar way we can also get easily the asymptotics of the second rising moment of T_1. With the help of Schwarz's inequality (and the fact that $S\geq T_1$) we have

$$\begin{aligned}E\left[S^2\right]-E\left[T_1^2\right]&=E\left[S^2-T_1^2\right]=E\left[(S+T_1)\,(S-T_1)\right]\\&\leq E\left[(S+T_1)^2\right]^{\frac{1}{2}}E\left[(S-T_1)^2\right]^{\frac{1}{2}}\leq 2E\left[S^2\right]^{\frac{1}{2}}E\left[(S-T_1)^2\right]^{\frac{1}{2}}.\end{aligned} \tag{113}$$

Now, (107) implies that

$$(S-T_1)^2=(S-T_1)^2\,\mathbf{1}_{\{T_1<S\}} \tag{114}$$

and hence, in the spirit of (108) and (109) we can get

$$E\left[(S-T_1)^2\right] = E\left[(S-T_1)^2 \mid T_1 < S\right] P\{T_1 < S\} = \frac{1+\alpha_1}{\alpha_2^2}\, P\{T_1 < S\}. \tag{115}$$

Using (115) in (113) yields

$$E\left[S^2\right] - E\left[T_1^2\right] \le \frac{2\sqrt{1+\alpha_1}}{\alpha_2}\, E\left[S^2\right]^{\frac{1}{2}} P\{T_1 < S\}^{\frac{1}{2}}. \tag{116}$$

Thus, by (92) (with $\theta = \alpha_2$ and $N = \nu_1 M$) and (106) we get that the quantity in the left-hand side of (116) satisfies

$$E\left[S^2\right] - E\left[T_1^2\right] = O\left(e^{-\varepsilon M}\right), \qquad M \to \infty, \tag{117}$$

for $\varepsilon > 0$ sufficiently small.
Therefore, by applying (92) (with $\theta = \alpha_2$ and $N = \nu_1 M$) in (117) together with Theorem 3 and (83) (with $\theta = \alpha_2$ and $N = \nu_1 M$) we obtain the following result:

Theorem 4. For every sufficiently small $\varepsilon > 0$ we have

$$\begin{aligned} E\left[T_1(T_1+1)\right] &= \left(\frac{\nu_1}{\alpha_1}\right)^2 M^2 \left(H_{\nu_1 M}^2 + \sum_{j=1}^{\nu_1 M} \frac{1}{j^2}\right) + O\left(e^{-\varepsilon M}\right) \\ &= (\nu_1 + \lambda\nu_2)^2 M^2 \left(H_{\nu_1 M}^2 + \sum_{j=1}^{\nu_1 M} \frac{1}{j^2}\right) + O\left(e^{-\varepsilon M}\right) \end{aligned} \tag{118}$$

as $M \to \infty$. Likewise,

$$\begin{aligned} E\left[T_2(T_2+1)\right] &= \left(\frac{\nu_2}{\alpha_2}\right)^2 M^2 \left(H_{\nu_2 M}^2 + \sum_{j=1}^{\nu_2 M} \frac{1}{j^2}\right) + O\left(e^{-\varepsilon M}\right) \\ &= \left(\lambda^{-1}\nu_1 + \nu_2\right)^2 M^2 \left(H_{\nu_2 M}^2 + \sum_{j=1}^{\nu_2 M} \frac{1}{j^2}\right) + O\left(e^{-\varepsilon M}\right) \end{aligned} \tag{119}$$

as $M \to \infty$.
From Theorems 3 and 4 we get immediately the following

Corollary 3. For every sufficiently small $\varepsilon > 0$ we have

$$\begin{aligned} V\left[T_1\right] &= \left(\frac{\nu_1}{\alpha_1}\right)^2 \left(\sum_{j=1}^{\nu_1 M} \frac{1}{j^2}\right) M^2 - \frac{\nu_1}{\alpha_1} M H_{\nu_1 M} + O\left(e^{-\varepsilon M}\right) \\ &= (\nu_1 + \lambda\nu_2)^2 \left(\sum_{j=1}^{\nu_1 M} \frac{1}{j^2}\right) M^2 - (\nu_1 + \lambda\nu_2) M H_{\nu_1 M} + O\left(e^{-\varepsilon M}\right) \end{aligned} \tag{120}$$

as $M \to \infty$. Likewise,

$$V\left[T_2\right] = \left(\frac{\nu_2}{\alpha_2}\right)^2 \left(\sum_{j=1}^{\nu_2 M} \frac{1}{j^2}\right) M^2 - \frac{\nu_2}{\alpha_2} M H_{\nu_2 M} + O\left(e^{-\varepsilon M}\right)$$

$$= \left(\lambda^{-1}\nu_1 + \nu_2\right)^2 \left(\sum_{j=1}^{\nu_2 M} \frac{1}{j^2}\right) M^2 - \left(\lambda^{-1}\nu_1 + \nu_2\right) M H_{\nu_2 M} + O\left(e^{-\varepsilon M}\right) \tag{121}$$

as $M \to \infty$.

In particular,

$$V\left[T_1\right] = \frac{\pi^2(\nu_1 + \lambda\nu_2)^2}{6} M^2 \left[1 + O\left(\frac{\ln M}{M}\right)\right], \qquad M \to \infty. \tag{122}$$

Let us, also, mention that a similar approach can be use to determine the asymptotics of $E[T_1^{(r)}]$ and $E[T_1^r]$. Indeed, formulas (101) and (102) (for $\theta = \alpha_2$ and $N = \nu_1 M$, as usual) imply

$$E\left[T_1^r\right] \sim E\left[T_1^{(r)}\right] \sim E\left[S^{(r)}\right] \sim (\nu_1 + \lambda\nu_2)^r M^r (\ln M)^r, \quad M \to \infty, \qquad r > 0. \tag{123}$$

Likewise,

$$E\left[T_2^r\right] \sim E\left[T_2^{(r)}\right] \sim \left(\lambda^{-1}\nu_1 + \nu_2\right)^r M^r (\ln M)^r, \quad M \to \infty, \qquad r > 0. \tag{124}$$

Finally, for $\theta = \alpha_2$ and $N = \nu_1 M$ formula (105) becomes

$$\frac{S - (\nu_1 + \lambda\nu_2) M \left(\ln M + \ln \nu_1\right)}{(\nu_1 + \lambda\nu_2) M} \xrightarrow{D} Y \qquad \text{as } M \to \infty \tag{125}$$

where Y follows the standard Gumbel distribution displayed in (104). We can rewrite (125) as

$$\frac{T_1 - (\nu_1 + \lambda\nu_2) M \left(\ln M + \ln \nu_1\right)}{(\nu_1 + \lambda\nu_2) M} + \frac{S - T_1}{(\nu_1 + \lambda\nu_2) M} \xrightarrow{D} Y \qquad \text{as } M \to \infty. \tag{126}$$

However, from (106) and (109) we have that $S - T_1 \to 0$ in L_1 and, therefore in probability (actually it is easy to see by using (106) and (109) and Chebyshev's inequality that, for any $\delta > 0$ we have $\sum_{M=1}^{\infty} P\{S - T_1 > \delta\} < \infty$, hence $P\{S - T_1 > \delta \text{ i.o.}\} = 0$ and the convergence is almost surely). It follows that $S - T_1 \to 0$ in distribution as $M \to \infty$. Therefore,

$$\frac{S - T_1}{(\nu_1 + \lambda\nu_2) M} \xrightarrow{D} 0 \qquad \text{as } M \to \infty, \tag{127}$$

hence by combining (126) and (127) we obtain the following theorem regarding the limiting distribution of T_1 (and T_2):

Theorem 5.

$$\frac{T_1 - (\nu_1 + \lambda\nu_2)M\ \ln M}{(\nu_1 + \lambda\nu_2)M} - \ln \nu_1 \xrightarrow{D} Y \qquad \text{as} \;\; M \to \infty \tag{128}$$

where

$$F(y) = P\{Y \le y\} = \exp\left(-e^{-y}\right), \qquad y \in \mathbb{R}, \tag{129}$$

namely Y is a standard Gumbel random variable.
Likewise,

$$\frac{T_2 - \left(\lambda^{-1}\nu_1 + \nu_2\right)M\ \ln M}{\left(\lambda^{-1}\nu_1 + \nu_2\right)M} - \ln \nu_2 \xrightarrow{D} Y \qquad \text{as} \;\; M \to \infty. \tag{130}$$

4.2 The random variable T

We are now ready to determine the asymptotic behavior of the variable $T = T_1 \vee T_2$ as $M \to \infty$. Without loss of generality, as in Section 2, we will assume for convenience that

$$\lambda = \frac{p_2}{p_1} > 1. \tag{131}$$

Let us first observe that we can write

$$T - T_1 = T_1 \vee T_2 - T_1 = (T_2 - T_1)\, \mathbf{1}_{\{T_1 < T_2\}}. \tag{132}$$

Taking expectations in (132) yields

$$E\left[T - T_1\right] = E\left[(T_2 - T_1)\, \mathbf{1}_{\{T_1 < T_2\}}\right] = E\left[T_2 - T_1 \,|\, T_1 < T_2\right] P\{T_1 < T_2\}. \tag{133}$$

From the fact that T_1 and T_2 are stopping times of the coupon filtration (recall (3)) we get

$$E\left[T_2 - T_1 \,|\, T_1 < T_2\right] \le E\left[T_2\right], \tag{134}$$

thus, using (134) in (133) gives

$$E\left[T\right] - E\left[T_1\right] = E\left[T - T_1\right] \le E\left[T_2\right] P\{T_1 < T_2\}. \tag{135}$$

Therefore, by invoking Theorems 2 and 3 we obtain
Theorem 6.

$$E[T] = (\nu_1 + \lambda\nu_2)M H_{\nu_1 M} \;+\; O\left(M^{2-\lambda} \ln M\right), \qquad M \to \infty, \tag{136}$$

where, as usual, H_N denotes the N-th harmonic number.
Since $\lambda > 1$, formula (136) together with (79) imply

$$E[T] = (\nu_1{+}\lambda\nu_2)M \ln M{+}(\nu_1{+}\lambda\nu_2)(\gamma{+}\ln \nu_1)M{+}O\left(M^{2-\lambda} \ln M\right), \qquad M \to \infty. \tag{137}$$

From Theorem 6 we see that the larger the λ, the more accurate the asymptotic formula for $E[T]$ becomes. The value $\lambda = 2$ is somehow critical, since if $\lambda > 2$, then (136) yields

$$E[T] = (\nu_1 + \lambda\nu_2)M \ln M + (\nu_1 + \lambda\nu_2)(\gamma + \ln \nu_1)M + \frac{\nu_1 + \lambda\nu_2}{2\nu_1} + o(1) \quad (138)$$

as $M \to \infty$.

We continue with the asymptotics of the second rising moment of T. We will follow the approach used in the previous subsection for $E[T_1^{(2)}]$. For better estimates, instead of the Schwarz's inequality we use here the more general Hölder inequality (and the fact that $T \leq T_1 + T_2$) to get

$$\begin{aligned} E\left[T^2\right] - E\left[T_1^2\right] &= E\left[T^2 - T_1^2\right] = E\left[(T + T_1)(T - T_1)\right] \\ &\leq E\left[(2T_1 + T_2)^r\right]^{\frac{1}{r}} E\left[(T - T_1)^s\right]^{\frac{1}{s}}, \end{aligned} \quad (139)$$

where

$$r > 1 \qquad \text{and} \qquad s = \frac{r}{r-1}. \quad (140)$$

An immediate upper bound of the first factor of the right-hand side of the inequality in (139) is given by the Minkowski inequality:

$$E\left[(2T_1 + T_2)^r\right]^{\frac{1}{r}} \leq 2E\left[T_1^r\right]^{\frac{1}{r}} + E\left[T_2^r\right]^{\frac{1}{r}}. \quad (141)$$

Now, (132) implies

$$(T - T_1)^s = (T - T_1)^s \, \mathbf{1}_{\{T_1 < T_2\}} \quad (142)$$

and hence, in the spirit of (134)

$$E\left[(T - T_1)^s\right] = E\left[(T_2 - T_1)^s \mid T_1 < T_2\right] P\{T_1 < T_2\} \leq E\left[T_2^s\right] \, P\{T_1 < T_2\}. \quad (143)$$

Using (141) and (143) in (139) yields

$$E\left[T^2\right] - E\left[T_1^2\right] \leq \left(2E\left[T_1^r\right]^{\frac{1}{r}} + E\left[T_2^r\right]^{\frac{1}{r}}\right) E\left[T_2^s\right]^{\frac{1}{s}} \, P\{T_1 < T_2\}^{\frac{1}{s}}. \quad (144)$$

Thus, by using (123), (124), and the result of Theorem 2 in (144) we obtain

$$E\left[T^2\right] - E\left[T_1^2\right] = O\left(M^{2-(\lambda-1)/s} \ln^2 M\right), \qquad M \to \infty. \quad (145)$$

If we had used the Schwarz's inequality, then we would have been forced to take $r = s = 2$. By using Hölder inequality, we are free to choose r as large as we wish and, consequently, in view of (140), we can take s arbitrarily close to 1. Thus, formula (145) is valid for any $s > 1$ and we can write it as

$$E\left[T^2\right] = E\left[T_1^2\right] + O\left(M^{3-\lambda+\varepsilon}\right), \qquad M \to \infty, \quad (146)$$

for any $\varepsilon > 0$. Hence, by Theorems 3 and 6 formula (146) becomes

$$E\left[T^{(2)}\right] = E\left[T_1^{(2)}\right] + O\left(M^{3-\lambda+\varepsilon}\right), \qquad M \to \infty. \quad (147)$$

Therefore, by using Theorem 4 in (147) we obtain the following result:

Theorem 7. For every $\varepsilon > 0$ we have

$$E\left[T^{(2)}\right] = (\nu_1 + \lambda\nu_2)^2 M^2 \left(H_{\nu_1 M}^2 + \sum_{j=1}^{\nu_1 M} \frac{1}{j^2}\right) + O\left(M^{3-\lambda+\varepsilon}\right), \qquad M \to \infty. \tag{148}$$

Notice that, since $\lambda > 1$ and ε can be taken arbitrarily close to 0, the exponent $3 - \lambda + \varepsilon$ in the error term can be always assumed to be less than 2 (hence formula (147) is meaningful for any $\lambda > 1$). In particular, from (148) we can immediately deduce that

$$E\left[T^2\right] \sim E\left[T^{(2)}\right] \sim (\nu_1 + \lambda\nu_2)^2 M^2 \ln^2 M, \qquad M \to \infty \tag{149}$$

and, furthermore, in a similar manner we can show that for any $r > 0$ we have

$$E\left[T^r\right] \sim E\left[T^{(r)}\right] \sim (\nu_1 + \lambda\nu_2)^r M^r \ln^r M, \qquad M \to \infty. \tag{150}$$

From Theorems 6 and 7 we get the following corollary.

Corollary 4.

$$V\left[T\right] \sim \frac{\pi^2(\nu_1 + \lambda\nu_2)^2}{6} M^2, \qquad M \to \infty. \tag{151}$$

Finally, let us determine the limiting distribution of T as $M \to \infty$. Formula (128) can be written as

$$\left[\frac{T - (\nu_1 + \lambda\nu_2)M\,\ln M}{(\nu_1 + \lambda\nu_2)M} - \ln\nu_1\right] - \frac{T - T_1}{(\nu_1 + \lambda\nu_2)M} \xrightarrow{D} Y, \qquad M \to \infty, \tag{152}$$

where Y is a standard Gumbel random variable. Moreover, by using (42) and (110) in (135) we get

$$E\left[\frac{T - T_1}{(\nu_1 + \lambda\nu_2)M}\right] = O\left(\frac{\ln M}{M^{\lambda-1}}\right), \qquad M \to \infty. \tag{153}$$

Since $\lambda > 1$, formula (153) implies that, as $M \to \infty$,

$$\frac{T - T_1}{(\nu_1 + \lambda\nu_2)M} \to 0 \qquad \text{in the } L_1 \text{ sense.} \tag{154}$$

Hence the above convergence is also in probability and, consequently, in distribution. Therefore, by using (154) in (152) we obtain immediately the limiting distribution of the random variable T as $M \to \infty$:

Theorem 8. Let T be the number of trials required to detect all Group 1 and Group 2 coupons. Then,

$$\frac{T - (\nu_1 + \lambda\nu_2)M\,\ln M}{(\nu_1 + \lambda\nu_2)M} - \ln\nu_1 \xrightarrow{D} Y \qquad \text{as } M \to \infty,$$

where Y is a standard Gumbel random variable.

5 APPENDIX

5.1 Proof of formula (63)

Let $\alpha \in (0,1)$. For typographical convenience we set

$$U(N;\alpha) := e^{\ln^\alpha N} \tag{155}$$

(so that for any constant $\beta > 0$ we have $\ln^\beta N << U(N;\alpha) << N^\beta$ as $N \to \infty$) and then we write (62) as

$$E\left[S_N^{(r)}\right] = N^r \ln^r N \left[I_1(N) + I_2(N)\right], \tag{156}$$

where

$$I_1(N) := \int_0^{U(N;\alpha)} \left(1 - \frac{x}{N}\right)^{N-1} \left(1 - \frac{\ln x}{\ln N}\right)^r dx \tag{157}$$

and

$$I_2(N) := \int_{U(N;\alpha)}^{N} \left(1 - \frac{x}{N}\right)^{N-1} \left(1 - \frac{\ln x}{\ln N}\right)^r dx. \tag{158}$$

We will first estimate $I_2(N)$ as $N \to \infty$.

$$0 < I_2(N) < \left(1 - \frac{U(N;\alpha)}{N}\right)^{N-1} \int_{U(N;\alpha)}^{N} \left(1 - \frac{\ln x}{\ln N}\right)^r dx$$
$$< N\left[\left(1 - \frac{U(N;\alpha)}{N}\right)^{\frac{N-1}{U(N;\alpha)}}\right]^{U(N;\alpha)} \sim Ne^{-U(N;\alpha)}. \tag{159}$$

In particular (159) implies

$$I_2(N) = o\left(\frac{1}{N^\kappa}\right) \qquad \text{for any } \ \kappa > 0. \tag{160}$$

To estimate $I_1(N)$ we first notice that for $0 \le x \le U(N;\alpha) = e^{\ln^\alpha N}$ we have

$$(N-1)\ln\left(1 - \frac{x}{N}\right) = -(N-1)\left[\frac{x}{N} + O\left(\frac{x^2}{N^2}\right)\right] = -x + o\left(\frac{1}{N^\theta}\right) \quad \text{for any } \ \theta \in (0,1). \tag{161}$$

Exponentiating (161) yields

$$\left(1 - \frac{x}{N}\right)^{N-1} = e^{-x}\left[1 + o\left(\frac{1}{N^\theta}\right)\right] \quad \text{for any } \ \theta \in (0,1). \tag{162}$$

We, then, substitute (162) in (157) and obtain

$$I_1(N) = \int_0^{U(N;\alpha)} e^{-x}\left(1 - \frac{\ln x}{\ln N}\right)^r dx + o\left(\frac{1}{N^\theta}\right)\int_0^{U(N;\alpha)} e^{-x}\left(1 - \frac{\ln x}{\ln N}\right)^r dx. \tag{163}$$

Now,

$$\int_0^{U(N;\alpha)} e^{-x}\left(1-\frac{\ln x}{\ln N}\right)^r dx \le \int_0^1 e^{-x}\left(1-\frac{\ln x}{\ln N}\right)^r dx + \int_1^{U(N;\alpha)} e^{-x}dx = O(1) \tag{164}$$

as $N \to \infty$. Thus, by using (164) in (163) we arrive at

$$I_1(N) = \int_0^{U(N;\alpha)} e^{-x}\left(1-\frac{\ln x}{\ln N}\right)^r dx \, + \, o\left(\frac{1}{N^\theta}\right) \qquad \text{for any } \; \theta \in (0,1). \tag{165}$$

Since $\ln x \to -\infty$ as $x \to 0^+$, we need to estimate the "bottom tail" of the integral in (165). We have

$$0 < \int_0^{U(N;\alpha)^{-1}} e^{-x}\left(1-\frac{\ln x}{\ln N}\right)^r dx < \left(\frac{2^r}{\ln^r N} + \frac{2^r}{e^{\ln^\alpha N}}\right)\int_0^{U(N;\alpha)^{-1}} (-\ln x)^r dx \tag{166}$$

(recall that $U(N;\alpha)^{-1} = e^{-\ln^\alpha N}$, where $0 < \alpha < 1$). Integration by parts gives

$$\int_0^{U(N;\alpha)^{-1}} (-\ln x)^r dx = e^{-\ln^\alpha N}(\ln N)^{r\alpha}\left[1+o(1)\right], \tag{167}$$

hence, by using (167) in (166) we obtain

$$\int_0^{U(N;\alpha)^{-1}} e^{-x}\left(1-\frac{\ln x}{\ln N}\right)^r dx = o\left(e^{-\ln^\alpha N}\right). \tag{168}$$

In view of (168), formula (165) implies

$$I_1(N) = \int_{U(N;\alpha)^{-1}}^{U(N;\alpha)} e^{-x}\left(1-\frac{\ln x}{\ln N}\right)^r dx \, + \, o\left(e^{-\ln^\alpha N}\right). \tag{169}$$

Finally, by substituting (160) and (169) in (156) we get

$$E\left[S_N^{(r)}\right] = N^r(\ln N)^r J(N;\alpha), \tag{170}$$

where

$$J(N;\alpha) := \int_{U(N;\alpha)^{-1}}^{U(N;\alpha)} e^{-x}\left(1-\frac{\ln x}{\ln N}\right)^r dx \, + \, o\left(e^{-\ln^\alpha N}\right) \tag{171}$$

as $N \to \infty$.

Before we continue let us recall that, for any given $r > 0$, the function $h(y) := (1-y)^r$ has the Taylor series expansion

$$h(y) = (1-y)^r = \sum_{k=0}^{\infty}(-1)^k\binom{r}{k}y^k, \qquad y \in [-1,1], \tag{172}$$

where

$$\binom{r}{0} = 1 \qquad \text{and} \qquad \binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k!} \quad \text{for } \; k = 1,2,\dots. \tag{173}$$

Thus, if $n \geq 1$ is a fixed integer, then

$$h(y) = (1-y)^r = \sum_{k=0}^{n} (-1)^k \binom{r}{k} y^k + \frac{h^{(n+1)}(\xi)}{(n+1)!} y^{n+1}, \tag{174}$$

where ξ lies between 0 and y.
Going back to formula (171) we see that due to the limits of integration the dummy variable x satisfies

$$-\frac{\ln^\alpha N}{\ln N} \leq \frac{\ln x}{\ln N} \leq \frac{\ln^\alpha N}{\ln N}. \tag{175}$$

Hence, if we set $y = \frac{\ln x}{\ln N}$ in (174), then the quantity $\frac{h^{(n+1)}(\xi)}{(n+1)!}$ is bounded and we get

$$\left(1 - \frac{\ln x}{\ln N}\right)^r = \sum_{k=0}^{n} (-1)^k \binom{r}{k} \left(\frac{\ln x}{\ln N}\right)^k + O\left(\frac{\ln^{\alpha(n+1)} N}{\ln^{n+1} N}\right) \tag{176}$$

uniformly in x, as long as the range of values of x is given by (175).
It follows from (176) that if (for our given n) we choose an α so that

$$0 < \alpha < \frac{1}{n+1}, \tag{177}$$

then

$$\left(1 - \frac{\ln x}{\ln N}\right)^r = \sum_{k=0}^{n} (-1)^k \binom{r}{k} \left(\frac{\ln x}{\ln N}\right)^k + o\left(\frac{1}{\ln^n N}\right) \tag{178}$$

again uniformly in x, within the range of values given by (175). Thus, we can substitute (178) in (171) and get

$$J(N;\alpha) = \sum_{k=0}^{n} \binom{r}{k} \frac{(-1)^k}{\ln^k N} \int_{U(N;\alpha)^{-1}}^{U(N;\alpha)} e^{-x} (\ln x)^k dx + o\left(\frac{1}{\ln^n N}\right) \tag{179}$$

as $N \to \infty$.
Next, we observe that in the same way we derived (168) we can also get

$$\int_0^{U(N;\alpha)^{-1}} e^{-x} (\ln x)^k dx = o\left(e^{-\ln^\alpha N} \ln^k N\right). \tag{180}$$

Also, it is easy to see that

$$\int_{U(N;\alpha)}^{\infty} e^{-x} (\ln x)^k dx = O\left(e^{-U(N;\alpha)} (\ln N)^{\alpha k}\right) \tag{181}$$

(recall that $U(N;\alpha) = e^{\ln^\alpha N}$). Therefore, by using (180) and (181) in (179) we obtain

$$J(N;\alpha) = \sum_{k=0}^{n} \binom{r}{k} \frac{(-1)^k}{\ln^k N} \int_0^{\infty} e^{-x} (\ln x)^k dx + o\left(\frac{1}{\ln^n N}\right), \qquad N \to \infty, \tag{182}$$

and, finally, by substituting (182) in (170) we arrive at (63). ■

5.2 Proof of formula (105)

We start by introducing the generating functions

$$G(z) := E\left[z^{-S(\theta)}\right] = 1-(z-1)\int_0^\infty e^{-(z-1)t}\left[1-\left(1-e^{-\theta t}\right)\left(1-e^{-(1-\theta)t/N}\right)^N\right]dt. \tag{183}$$

Notice that if

$$\Re\{z\} > 1-\frac{1-\theta}{N}, \tag{184}$$

the integral appearing in (183) is absolutely convergent.
We will derive formula (105) via characteristic functions. Let us fix a $\xi \in \mathbb{R}$ and set

$$\zeta := e^{-i\xi}. \tag{185}$$

Then, in view of (183), the characteristic function of $[(1-\theta)S(\theta) - N\ln N]/N$ is

$$\phi_N(\xi) = E\left[\zeta^{-\frac{(1-\theta)S(\theta)-N\ln N}{N}}\right] = \zeta^{\ln N} E\left[\left(\zeta^{\frac{(1-\theta)}{N}}\right)^{-S(\theta)}\right] = \zeta^{\ln N} G\left(\zeta^{\frac{(1-\theta)}{N}}\right). \tag{186}$$

Now,

$$\zeta^{\frac{(1-\theta)}{N}} = e^{-\frac{i(1-\theta)\xi}{N}} = 1-\frac{i(1-\theta)\xi}{N}+O\left(\frac{1}{N^2}\right), \qquad N\to\infty. \tag{187}$$

In particular $z = \zeta^{(1-\theta)/N}$ satisfies (184) for all N sufficiently large.
Next, by using (183) and (187) in (186) we get

$$\zeta^{-\ln N}\phi_N(\xi) = 1+\left[\frac{i(1-\theta)\xi}{N}+O\left(\frac{1}{N^2}\right)\right]\left[\chi_1(N)+\chi_2(N)\right], \tag{188}$$

where

$$\chi_1(N) := \int_0^\infty e^{-\left(\zeta^{\frac{(1-\theta)}{N}}-1\right)t}\left[1-\left(1-e^{-(1-\theta)t/N}\right)^N\right]dt \tag{189}$$

and

$$\chi_2(N) := \int_0^\infty e^{-\left(\zeta^{\frac{(1-\theta)}{N}}-1\right)t} e^{-\theta t}\left(1-e^{-(1-\theta)t/N}\right)^N dt. \tag{190}$$

Regarding $\chi_2(N)$, in the same way we got formula (82) from (81) we can obtain

$$\chi_2(N) = O\left(e^{-\varepsilon N}\right), \qquad N\to\infty, \tag{191}$$

for any sufficiently small $\varepsilon > 0$.
Now, using (187) in (189) yields

$$\chi_1(N) = \int_0^\infty e^{[i(1-\theta)\xi + O(N^{-1})]\,t/N}\left[1-\left(1-e^{-(1-\theta)t/N}\right)^N\right]dt \tag{192}$$

or, after the substitution $s = (1-\theta)t/N$ in the above integral

$$\chi_1(N) = \frac{N}{1-\theta}\int_0^\infty e^{[i\xi + O(N^{-1})]\,s}\left[1-\left(1-e^{-s}\right)^N\right]ds. \tag{193}$$

Therefore, by substituting (191) and (193) in (188) we obtain

$$\begin{aligned}\zeta^{-\ln N}\phi_N(\xi) &= 1 + \left[i\xi + O\left(\frac{1}{N}\right)\right]\int_0^\infty e^{[i\xi + O(N^{-1})]\,s}\left[1-\left(1-e^{-s}\right)^N\right]ds \\ &= 1 + i\xi\int_0^\infty e^{i\xi s}\left[1-\left(1-e^{-s}\right)^N\right]ds + O\left(\frac{1}{N}\right)\end{aligned} \tag{194}$$

as $N \to \infty$.
Let A_N be a (real) quantity which grows to ∞ with N so that

$$\frac{N}{e^{A_N}} = o\left(\frac{1}{N}\right) \qquad \text{as } N \to \infty \tag{195}$$

(we do not need to be more specific about A_N). Noticing that by (185) we have $\zeta^{-\ln N} = N^{i\xi}$, we rewrite (194) as

$$N^{i\xi}\phi_N(\xi) = 1 + K_1(N) + K_2(N) + O\left(\frac{1}{N}\right), \tag{196}$$

where

$$K_1(N) := \int_0^{A_N} i\xi\, e^{i\xi s}\left[1-\left(1-e^{-s}\right)^N\right]ds \tag{197}$$

and

$$K_2(N) := \int_{A_N}^\infty i\xi\, e^{i\xi s}\left[1-\left(1-e^{-s}\right)^N\right]ds. \tag{198}$$

Applying integration by parts in (197) yields

$$K_1(N) = e^{i\xi A_N}\left[1+\left(1-e^{-A_N}\right)^N\right] - 1 + N\int_0^{A_N} e^{i\xi s}\left(1-e^{-s}\right)^{N-1}e^{-s}ds, \tag{199}$$

which, in view of (195), implies

$$K_1(N) = e^{i\xi A_N} - 1 + N\int_0^{A_N} e^{i\xi s}\left(1-e^{-s}\right)^{N-1}e^{-s}ds + o\left(\frac{1}{N}\right). \tag{200}$$

Next, by substituting $s = \ln N - \ln x$ in the integral appearing in the right-hand side of (200) we obtain

$$K_1(N) = e^{i\xi A_N} - 1 + N^{i\xi}\int_{Ne^{-A_N}}^N x^{-i\xi}\left(1-\frac{x}{N}\right)^{N-1}dx + o\left(\frac{1}{N}\right). \tag{201}$$

We, then, use (201) in (196) and get

$$N^{i\xi}\phi_N(\xi) = e^{iA_N\xi} + N^{i\xi}\int_{Ne^{-A_N}}^N x^{-i\xi}\left(1-\frac{x}{N}\right)^{N-1}dx + K_2(N) + O\left(\frac{1}{N}\right). \tag{202}$$

Let us, now, turn our attention to the integral $K_2(N)$ of formula (198).
Assume first that $\xi > 0$. We complexify the dummy variable s by setting $z = s + i\tau$ and for N (temporarily) fixed we choose $R > A_N$ and consider the close contour C_R formed by (i) the interval $[A_N, R]$ of the (real) s-axis, (ii) the circular arc $Re^{i\theta}$, $0 \le \theta \le \arccos(A_N/R)$, and (iii) the line segment $A_N + i\tau$, $0 \le \tau \le \sqrt{R^2 - A_N^2}$. Then, Cauchy's Theorem implies

$$\oint_{C_R} i\xi\, e^{i\xi z} \left[1 - \left(1 - e^{-z}\right)^N\right] dz = 0 \qquad \text{for every} \;\; R > A_N. \tag{203}$$

Next (keeping N fixed), we take limits in (203) as $R \to \infty$. It is a standard exercise in contour integration to show that the integral on the circular piece of C_R, namely on the arc $Re^{i\theta}$, $0 \le \theta \le \arccos(A_N/R)$, vanishes. Hence, in view of (198), formula (203) implies

$$K_2(N) = -e^{i\xi A_N} \int_0^\infty \xi\, e^{-\xi\tau} \left[1 - \left(1 - e^{-A_N} e^{-i\tau}\right)^N\right] d\tau. \tag{204}$$

Now, we can allow N to grow large. Thus, in view of (195), formula (204) yields

$$K_2(N) = -e^{i\xi A_N} + o\left(\frac{1}{N}\right) \tag{205}$$

and, hence, by substituting (205) in (202) we obtain

$$\phi_N(\xi) = \int_{Ne^{-A_N}}^N x^{-i\xi} \left(1 - \frac{x}{N}\right)^{N-1} dx + O\left(\frac{1}{N}\right). \tag{206}$$

Formula (206) was obtained under the assumption that $\xi > 0$. However, if $\xi < 0$, then the same approach works if we choose the contour C_R to be the symmetric of the previous one with respect to the (real) s-axis. Therefore, formula (206) is valid for all $\xi \in \mathbb{R} \setminus \{0\}$, while for $\xi = 0$ formulas (185) and (186) imply immediately that

$$\phi_N(0) = 1 \qquad \text{for all} \;\; N. \tag{207}$$

Finally, as in the previous subsection (see, e.g., (157) and (162)), formulas (206) and (207) imply

$$\lim_{N\to\infty} \phi_N(\xi) = \int_0^\infty x^{-i\xi} e^{-x} dx = \Gamma(1 - i\xi) \qquad \text{pointwise for} \;\; \xi \in \mathbb{R}, \tag{208}$$

where $\Gamma(1 - i\xi)$ is recognized as the characteristic function of the standard Gumbel distribution. ■

References

[1] C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*, Springer-Verlag, New York, 1999.

[2] G. Boros and V.H. Moll, *Irresistible Integrals: Symbolics, Analysis and Experiments in the Evaluation of Integrals*, Cambridge University Press, Edinburgh, 2004.

[3] A.V. Doumas and V.G. Papanicolaou, The Coupon Collector's Problem Revisited: Asymptotics of the Variance, *Adv. Appl. Prob.* **44** (1) (2012) 166–195.

[4] A.V. Doumas and V.G. Papanicolaou, Asymptotics of the rising moments for the Coupon Collector's Problem, *Electron. J. Probab.* **Vol. 18** (Article no. 41) (2012) 1–15.

[5] P. Erdős and A. Rényi, On a classical problem of probability theory, *Magyar. Tud. Akad. Mat. Kutató Int. Közl.*, **6** (1961), 215–220.

[6] S.M. Ross, private communication, 2017.

[7] S.M. Ross, *Introduction to Probability Models*, Tenth Edition, Elsevier Inc., Burlington, MA, 2010.