JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 93 (1998) 95–105

Asymptotics for the random coupon collector problem

Vassilis G. Papanicolaou [a,*], George E. Kokolakis [b], Shahar Boneh [c]

[a]Department of Mathematics and Statistics, Wichita State University, Wichita, KS 67260-0033, United States
[b]Department of Mathematics, Section of Statistics, National Technical University of Athens, 157 80 Zografou, Athens, Greece
[c]Department of Mathematics, Metropolitan State College of Denver, P. O. Box 173362 Denver, CO 80217-3362, United States

Received 3 June 1997; received in revised form 10 February 1998

Abstract

We develop techniques of computing the asymptotics of the expected number of items that one has to check in order to detect all N existing kinds, as $N \to \infty$. The occurring frequencies of the differend kinds are random variables.

AMS classification: 60C05

Keywords: Random coupon collector problem; Laplace integral

1. Introduction

The coupon-collector problem (CCP) can be briefly described as follows: Suppose there are a total of N objects (coupons), sampled independently with replacement. Let $p_1, p_2, \ldots, p_N$ denote the sampling probabilities, where

$$p_1 + p_2 + \cdots + p_N = 1, \qquad p_n > 0, \quad \text{for } 1 \leqslant n \leqslant N.$$

An object that has been sampled at least once is said to have been detected. The CCP is concerned with various random variables associated with this process. For example:
$D(t)$ is the number of different objects detected after t trials;
T_N is the number of trials it takes until all the N objects are detected.
The CCP as described above was found to be a useful mathematical model for studying probabilistic algorithms, as well as a model for a variety of natural phenomena and engineering applications.

* Corresponding author. E-mail: papanico@cs.twsu.edu.

PII S0377-0427(98)00058-2

We now mention some applications of the CCP. The first three examples introduce probabilistic computational algorithms which can be modeled by the CCP.

1. *Constraint classification in mathematical programming*: In 1983, Karwan et al. [5] described a class of randomized algorithms for classifying all the constraints in a mathematical programming problem as necessary or redundant. The basic algorithm, also known as PREDUCE (Probabilistic REDUCE), can be briefly described as follows: Given an interior feasible point, each iteration consists of generating a ray in a random direction, and recording the first constraint it intersects. Such a constraint is a necessary one. The algorithm generates rays until a stopping rule is satisfied. Then, all the constraints which were not hit at all are classified as redundant – possibly erroneously. Each iteration corresponds to drawing one coupon, with N being the number of necessary constraints. Thus, the CCP model can help to determine an efficient stopping rule.

2. *Multistart methods in global optimization*: In an environment with multiple local optima, this method chooses a random initial point and uses any optimization algorithm to find a local optimum of the objective function. We repeatedly pick different random initial points, which in turn lead to different local optima, according to the partition of the domain to attraction regions. Here, the local optima are the coupons, and the probability of finding any local optimum is proportional to the size of its attraction region. The global optimizer wants to know how long it is expected to take until all the local optima are found, thus yielding the global optimum. For more information we refer to [6].

3. *Determining the convex closure of a point set S in $\mathbb{R}^n$*: To determine the subset of S that spans its convex closure, we generate a random $(n-1)$-dimensional line and measure the orthogonal distance of each point from that line. It can be shown that the point with the greatest distance belongs to the convex closure. This is done repeatedly until a stopping rule indicates that all the spanning points were detected.

The next two examples illustrate applications of the CCP in engineering and natural sciences.

4. *Engineering application*: Fault detection in electronic hardware. Many fault detection procedures in electronic hardware are based on repeated tests of the hardware. In each such test, a certain number of faults are detected, while some other faults are not detected. Here, each detection of a fault corresponds to a drawing of a coupon. It is of interest to determine the expected number of tests necessary until all the faults are detected.

5. *Biological application*: Estimating the number of species. Consider a complex ecosystem inhabited by a great variety of species. Biologists who explore that environment want to detect all the different species of a certain type. For example, all the different species of birds in the Galapagos Islands near Equador. Observing an animal from a given species constitutes a detection of that species. The reader should easily see that this scenario falls into the category of the CCP.

These examples show the wide scope of applications of the CCP. It is also apparent from these examples that the particular quantity which bares the most interest is T_N, the number of coupons (or trials) needed to detect all the N different coupons. This random variable will be the focus of this paper.

The CCP has been studied for quite a while. In particular, if $\overline{T_N}$ denotes the average of T_N, it has been established that

$$\overline{T_N} = \sum_{k=1}^{N} (-1)^{k+1} \sum_{1 \leqslant n_1 < \cdots < n_k \leqslant N} \frac{1}{p_{n_1} + \cdots + p_{n_k}} = \int_0^\infty \left[1 - \prod_{n=1}^{N} (1 - \mathrm{e}^{-p_n t}) \right] \mathrm{d}t. \tag{1}$$

The first part of (1) follows from the inclusion–exclusion principle, while the proof of the second part can be found in [4]. General asymptotic estimates of $\overline{T_N}$, as $N \to \infty$, have been obtained in [4, 2], and elsewhere.

In the classic CCP, the occurring probabilities $p_1, p_2, \ldots, p_N$, are taken as fixed parameters. However, in many applications, these probabilities behave more like random variables rather than parameters. For example, in the species detection situation, these probabilities are proportional to the size of the populations of the different species. These populations change randomly, and their exact sizes are typically unknown. Such a version of the CCP has not been dealt with. In this paper, we extend asymptotic results regarding the mean of T_N, from the standard CCP to the case where the occurring probabilities are themselves random variables. We shall refer to it as the random CCP (RCCP).

Let $\alpha = \{a_n\}_{n=1}^{\infty}$ be a sequence of (strictly) positive independent random variables with

$$\mu_n = E[a_n] > 0, \qquad g_n(t) = E[\mathrm{e}^{-ta_n}] \tag{2}$$

($E[\cdot]$ stands, of course, for the expectation associated to the sequence α). For a given N, we can create sampling probabilities $p_1, p_2, \ldots, p_N$ by taking

$$p_n = \frac{a_n}{A_N}, \quad \text{where } A_N = \sum_{n=1}^{N} a_n. \tag{3}$$

Notice that each p_n, $1 \leqslant n \leqslant N$, is a random variable which depends on N and α. This is the set-up of what we call random coupon-collector problem (RCCP). Our main interest here is the study of the quantity

$$q_N = E[\overline{T_N}],$$

namely, the expected number of trials it takes until all the objects are detected.

In Section 2 we examine briefly the general case, while, in Section 3, we focus on the case where the a_n's are (independent and) identically distributed. For this case we obtain the exact asymptotics of q_N, as $N \to \infty$. Finally, the implementation of our results is illustrated by three general examples (in Section 4).

2. Some formulas for the general case

If we denote

$$S_N^{\alpha} = \sum_{k=1}^{N} (-1)^{k+1} \sum_{1 \leqslant n_1 < \cdots < n_k \leqslant N} \frac{1}{a_{n_1} + \cdots + a_{n_k}}$$

and

$$\lambda\alpha = \{\lambda a_1, \lambda a_2, \ldots\}, \quad \text{where } \lambda > 0,$$

then

$$S_N^{\lambda\alpha} = \frac{1}{\lambda} S_N^{\alpha}$$

and therefore by (1) and (3),

$$q_N = E[S_N^{\alpha/A_N}] = E[A_N S_N^{\alpha}]. \tag{4}$$

Furthermore, as in (1), we have

$$S_N^{\alpha} = \int_0^{\infty} \left[1 - \prod_{k=1}^{N} (1 - \mathrm{e}^{-a_k t})\right] \mathrm{d}t, \tag{5}$$

hence (4) can be written as

$$q_N = E\left[\left(\sum_{n=1}^{N} a_n\right) \int_0^{\infty} \left\{1 - \prod_{k=1}^{N} (1 - \mathrm{e}^{-a_k t})\right\} \mathrm{d}t\right].$$

Using now the independence of the a_n's, Tonelli's theorem (for switching expectation and integral), and the notation of (2) we get

$$q_N = \sum_{n=1}^{N} \int_0^{\infty} \left\{ \mu_n - E[a_n - a_n \mathrm{e}^{-a_n t}] \prod_{\substack{k=1 \\ k \neq n}}^{N} [1 - g_k(t)] \right\} \mathrm{d}t. \tag{6}$$

But, again by (2),

$$E[a_n \mathrm{e}^{-a_n t}] = -g_n'(t)$$

(notice that $g_n'(t)$ is finite for $t > 0$, and $g_n'(0) = -\mu_n$ which may be infinite), thus (6) becomes

$$q_N = \sum_{n=1}^{N} \mu_n \int_0^{\infty} \left\{ 1 - \frac{1 - g_n'(t)/g_n'(0)}{1 - g_n(t)} \prod_{k=1}^{N} [1 - g_k(t)] \right\} \mathrm{d}t. \tag{7}$$

Proposition. *Let $F_n(x)$ be the distribution function of the random variable a_n and ε be any positive number. If for some n we have*

$$\int_0^{\varepsilon} \frac{\mathrm{d}F_n(x)}{x} = \infty \quad \text{or } \mu_n = \infty,$$

then $q_N = \infty$, for all $N \geqslant \max\{n, 2\}$.

Proof. We can assume, without loss of generality, that $n = 1$, and then observe that it is enough to prove the statement only for $N = 2$. From the first equality in (1) we get

$$q_2 = E\left[\frac{1}{p_1} + \frac{1}{p_2}\right] - 1,$$

or, by invoking (3) and the independence of a_1, a_2

$$q_2 = E\left[\frac{a_2}{a_1} + \frac{a_1}{a_2}\right] + 1 = \mu_2 E\left[\frac{1}{a_1}\right] + \mu_1 E\left[\frac{1}{a_2}\right] + 1. \tag{8}$$

Thus, if $\mu_1=\infty$, then $q_2=\infty$. On the other hand,

$$\int_0^{\varepsilon}\frac{\mathrm{d}F_1(x)}{x}=\infty \quad \text{if and only if} \quad E\left[\frac{1}{a_1}\right]=\infty$$

and therefore, if $\int_0^{\varepsilon}\mathrm{d}F_1(x)/x=\infty$, (8) implies that $q_2=\infty$. □

Remarks. (i) Intuitively,

$$\int_0^{\varepsilon}\frac{\mathrm{d}F_n(x)}{x}=\infty \tag{9}$$

means that the values of a_n have a high concentration near 0. If in particular a_n has a density $f_n(x)$, such that f_n is continuous on $[0,\varepsilon)$ and $f_n(0)\neq 0$, then (9) is valid. For example, if a_1 is exponentially distributed, then $q_N=\infty$, for all $N\geqslant 2$.

(ii) Since

$$E\left[\frac{1}{a_n}\right]=E\left[\int_0^{\infty}\mathrm{e}^{-ta_n}\,\mathrm{d}t\right]=\int_0^{\infty}E[\mathrm{e}^{-ta_n}]\,\mathrm{d}t=\int_0^{\infty}g_n(t)\,\mathrm{d}t,$$

we have that

$$\int_0^{\varepsilon}\frac{\mathrm{d}F_n(x)}{x}=\infty \quad \text{if and only if} \quad \int_b^{\infty}g_n(t)\,\mathrm{d}t=\infty.$$

In particular, if $g_1(t)\sim ct^{-\gamma}$, as $t\to\infty$, where $c>0$ and $0<\gamma\leqslant 1$, then $q_N=\infty$, for all $N\geqslant 2$.

(iii) The proposition has a converse: If, for all a_n, $1\leqslant n\leqslant N$,

$$\int_0^{\varepsilon}\frac{\mathrm{d}F_n(x)}{x}<\infty \quad \text{and} \quad \mu_n<\infty,$$

then $q_N<\infty$.

3. The case of identically distributed sampling probabilities

We now consider the case where the independent random variables a_n are also identically distributed with common distribution function $F(x)$. We set

$$\mu=E[a_n], \quad g(t)=E[\mathrm{e}^{-ta_n}] \tag{10}$$

and, in view of the Proposition of Section 2 and the remarks following it, we assume that

$$0<\mu<\infty \quad \text{and} \quad \int_0^{\infty}g(t)\,\mathrm{d}t=E\left[\frac{1}{a_n}\right]<\infty. \tag{11}$$

In this set-up, (7) becomes

$$q_N=N\mu\int_0^{\infty}\{1-[1-g'(t)/g'(0)][1-g(t)]^{N-1}\}\,\mathrm{d}t. \tag{12}$$

Our main goal is to compute the asymptotics of q_N above, as $N \to \infty$. In fact, it is enough to find the asymptotics of

$$I(N) = \int_0^\infty \{1 - [1 - g'(t)/g'(0)][1 - g(t)]^{N-1}\}\,\mathrm{d}t, \tag{13}$$

since, by (12),

$$q_N = \mu N I(N). \tag{14}$$

We have

$$I(n) - I(n-1) = \int_0^\infty g(t)\left[1 - \frac{g'(t)}{g'(0)}\right][1 - g(t)]^{n-2}\,\mathrm{d}t,$$

or

$$I(n) - I(n-1) = \int_0^\infty g(t)[1 - g(t)]^{n-2}\,\mathrm{d}t - \int_0^\infty \frac{g(t)g'(t)}{g'(0)}[1 - g(t)]^{n-2}\,\mathrm{d}t.$$

The second integral can be computed explicitly (substitute $u = 1 - g(t)$ and remember that $g(t)$ is monotone decreasing with $g(0) = 1$ and $g(\infty) = 0$), and we get

$$I(n) - I(n-1) = \int_0^\infty g(t)[1 - g(t)]^{n-2}\,\mathrm{d}t - \frac{1}{g'(0)}\left[\frac{1}{n} - \frac{1}{n-1}\right]$$

or

$$I(n) - I(n-1) = J(n) - \frac{1}{g'(0)}\left[\frac{1}{n} - \frac{1}{n-1}\right], \tag{15}$$

where

$$J(n) = \int_0^\infty g(t)[1 - g(t)]^{n-2}\,\mathrm{d}t. \tag{16}$$

Lemma 1.

$$\sum_{n=2}^{\infty} J(n) = \infty.$$

Proof. By (16)

$$\sum_{n=2}^{N+1} J(n) = \int_0^\infty \{1 - [1 - g(t)]^N\}\,\mathrm{d}t.$$

Since $0 < g(t) < 1$ for all $t > 0$, monotone convergence implies

$$\sum_{n=2}^{\infty} J(n) = \lim_N \int_0^\infty \{1 - [1 - g(t)]^N\}\,\mathrm{d}t = \int_0^\infty \mathrm{d}t = \infty. \qquad \square$$

If we sum (15) from $n=2$ to N, we obtain

$$I(N)-I(1)=\sum_{n=2}^{N}J(n)-\frac{1}{g'(0)}\left[\frac{1}{N}-\frac{1}{1}\right].$$

But $I(1)=-1/g'(0)=1/\mu$, thus

$$I(N)=\frac{1}{\mu N}+\sum_{n=2}^{N}J(n)$$

and, by (14),

$$q_N=1+\mu N\sum_{n=2}^{N}J(n). \tag{17}$$

Next, we compute the asymptotics of $J(n)$, as $n\to\infty$. By (16)

$$J(n)=\int_0^\infty g(t)\mathrm{e}^{(n-2)\ln[1-g(t)]}\,\mathrm{d}t.$$

Substituting $u=g(t)$ (again, remember that $g(t)$ is monotone decreasing with $g(0)=1$ and $g(\infty)=0$) we get

$$J(n)=-\int_0^1 u\mathrm{e}^{(n-2)\ln(1-u)}\frac{\mathrm{d}u}{g'[g^{-1}(u)]}.$$

This is a so-called *Laplace integral* (see [1, Section 6.4]). Since the maximum of $\ln(1-u)$ on $[0,1]$ occurs at $u=0$, it is well known that, as $n\to\infty$,

$$J(n)\sim-\int_0^\varepsilon u\mathrm{e}^{(n-2)\ln(1-u)}\frac{\mathrm{d}u}{g'[g^{-1}(u)]},\quad\text{for any }\varepsilon\text{ in }(0,1).$$

But, near $u=0$, $\ln(1-u)=-u+\mathrm{O}(u^2)$ and $\mathrm{e}^{2u}=1+\mathrm{O}(u)$, hence

$$J(n)\sim-\int_0^\varepsilon u\mathrm{e}^{-nu}\frac{\mathrm{d}u}{g'[g^{-1}(u)]},\quad\text{for any }\varepsilon\text{ in }(0,1) \tag{18}$$

(notice also that $1/g'[g^{-1}(u)]=\mathrm{d}[g^{-1}(u)]/\mathrm{d}u$).

Formula (18) is a big improvement since it enables us to compute the asymptotics of $J(n)$, as $n\to\infty$. In particular, it implies that the asymptotics of $J(n)$ are completely determined by the behavior of $g(t)$ near $t=\infty$, or equivalently (Tauberian theorem – see [3, Section XIII.5]) by the behavior of $F(x)$, as $x\to0^+$. If $g(t)$ is given, one can apply, for example, Watson's Lemma to (18) (see [1, Section 6.4]) and get the asymptotics of $J(n)$ explicitly. Since $J(n)>0$, for all n, and (previous lemma)

$$\sum_{n=2}^{\infty}J(n)=\infty,$$

Lemma 2 of the appendix allows us to substitute (18) into (17) and thus establish the following:

Theorem 2. *If q_N is the expected number of trials it takes until all the objects are detected, as given in* (12), *and μ, $g(t)$ are as in* (10), *then, as $N \to \infty$,*

$$q_N \sim -\mu N \sum_{n=2}^{N} \int_0^{\varepsilon} u\mathrm{e}^{-nu} \frac{\mathrm{d}u}{g'[g^{-1}(u)]}, \quad \textit{for any } \varepsilon \textit{ in } (0,1). \tag{19}$$

In practice (see the examples of the next section), there is often going to be a simple expression $\phi(n) > 0$ such that

$$\phi(n) \sim -\int_0^{\varepsilon} u\mathrm{e}^{-nu} \frac{\mathrm{d}u}{g'[g^{-1}(u)]}, \quad \text{as } n \to \infty$$

(of course, (18) and Lemma 1 imply that $\sum_{n=2}^{\infty} \phi(n) = \infty$). In this case (see Lemma 2 of the appendix), (19) can be written as

$$q_N \sim \mu N \sum_{n=2}^{N} \phi(n), \quad \text{as } N \to \infty,$$

where the asymptotics of the sum $\sum_{n=2}^{N} \phi(n)$ can be obtained by various known methods, such as the powerful Euler–Maclaurin sum formula (see [1, Section 6.7]).

4. Examples

In this section we give some examples that illustrate the theorem of Section 3. The notation is consistent with the previous section (and the conditions of (11) are satisfied in all the examples below).

Example 1. $g(t) \sim (\ln t)^{-\beta} t^{-1}$ as $t \to \infty$, where $\beta > 1$. It follows that

$$g^{-1}(u) \sim \frac{1}{u(-\ln u)^{\beta}} \quad \text{as } u \to 0^+.$$

Next, we observe that the given asymptotics of $g(t)$ as $t \to \infty$, determine the asymptotics of the distribution function $F(x)$ as $x \to 0^+$ (see [3, Section XIII.5]). Thus, the asymptotics of $xF(x)$ as $x \to 0^+$, are known, which, in turn, determine the asymptotics of $g'(t)$ as $t \to \infty$. For this example we have

$$g'(t) \sim \frac{-1}{(\ln t)^{\beta} t^2}, \quad \text{as } t \to \infty.$$

Hence,

$$\frac{u}{g'[g^{-1}(u)]} \sim \frac{-1}{u(-\ln u)^{\beta}}, \quad \text{as } u \to 0^+.$$

Thus (18) becomes

$$J(n) \sim \int_0^{\varepsilon} \mathrm{e}^{-nu} \frac{\mathrm{d}u}{u(-\ln u)^{\beta}}, \quad \text{for any } \varepsilon \text{ in } (0,1),$$

as $n \to \infty$. Therefore (see [3, Section XIII.5])

$$J(n) \sim \frac{1}{\beta - 1}(\ln n)^{1-\beta}.$$

It follows that

$$\sum_{n=2}^{N} J(n) \sim \frac{1}{\beta - 1} N(\ln N)^{1-\beta}, \quad \text{as } N \to \infty,$$

hence, (19) gives

$$q_N \sim \frac{\mu}{\beta - 1} N^2 (\ln N)^{1-\beta}.$$

Example 2. $g(t) \sim (\ln t)^{-\beta} t^{-\gamma}$ as $t \to \infty$, where β is any real number and $\gamma > 1$ (the gamma distribution falls in this category). In a similar manner, we obtain

$$\frac{u}{g'[g^{-1}(u)]} \sim \frac{-\gamma^{(\beta/\gamma)-1}}{u^{1/\gamma}(-\ln u)^{\beta/\gamma}} \quad \text{as } u \to 0^+,$$

and by using this in (18) we get

$$J(n) \sim \gamma^{(\beta/\gamma)-1} \int_0^{\varepsilon} \mathrm{e}^{-nu} \frac{\mathrm{d}u}{u^{1/\gamma}(-\ln u)^{\beta/\gamma}}, \quad \text{for any } \varepsilon \text{ in } (0,1).$$

Hence (see [3, Section XIII.5])

$$J(n) \sim \gamma^{(\beta/\gamma)-1} \Gamma\left(\frac{\gamma - 1}{\gamma}\right) \frac{1}{n^{1-(1/\gamma)}(\ln n)^{\beta/\gamma}} \quad \text{as } n \to \infty,$$

where $\Gamma(\cdot)$ is the gamma function. It follows that

$$\sum_{n=2}^{N} J(n) \sim \gamma^{\beta/\gamma} \Gamma\left(\frac{\gamma - 1}{\gamma}\right) \frac{N^{1/\gamma}}{(\ln N)^{\beta/\gamma}}, \quad \text{as } N \to \infty,$$

and, finally, (19) gives

$$q_N \sim \mu \gamma^{\beta/\gamma} \Gamma\left(\frac{\gamma - 1}{\gamma}\right) \frac{N^{1+(1/\gamma)}}{(\ln N)^{\beta/\gamma}}.$$

Example 3. $g(t) \sim \exp(-ct^{\delta})$ as $t \to \infty$, where $c > 0$ and $0 < \delta \leqslant 1$ ($g(t)$ cannot decay faster than an exponential). We assume that

$$\frac{\mathrm{d}}{\mathrm{d}u}[g^{-1}(u)] \sim -\frac{1}{\delta c^{1/\delta}} \frac{(-\ln u)^{(1/\delta)-1}}{u}, \quad \text{as } u \to 0^+$$

(this is in agreement with the given asymptotics of $g(t)$, as $t \to \infty$). Then (18) becomes

$$J(n) \sim \frac{1}{\delta c^{1/\delta}} \int_0^{\varepsilon} \mathrm{e}^{-nu}(-\ln u)^{(1/\delta)-1}\, \mathrm{d}u, \quad \text{for any } \varepsilon \text{ in } (0,1).$$

Thus (see [3, Section XIII.5])

$$J(n) \sim \frac{1}{\delta c^{1/\delta}} \frac{(\ln n)^{(1/\delta)-1}}{n}, \quad \text{as } n \to \infty.$$

It follows that

$$\sum_{n=2}^{N} J(n) \sim \frac{1}{c^{1/\delta}} (\ln N)^{1/\delta}, \quad \text{as } N \to \infty.$$

Thus, (19) gives

$$q_N \sim \frac{\mu}{c^{1/\delta}} N (\ln N)^{1/\delta}. \tag{20}$$

Remark. If we take $\delta = 1$ in the last example, then $g(t) \sim \exp(-ct)$, as $t \to \infty$, and this forces a_n to be deterministic, i.e. $a_n = c = \mu$ a.s. In this case (20) becomes

$$q_N \sim N \ln N,$$

which is the well-known formula for the (deterministic case of) uniform probabilities. Notice that, among all the cases we consider, this case gives the asymptotically smallest q_N. We believe that a variational analysis of (12) will establish that q_N becomes minimum, if $g(t) = \exp(-ct)$, i.e. when $a_n = c$ a.s., for all n.

Appendix A

Here we prove the following lemma:

Lemma A.1. *Let $\psi(n) > 0$, for $n = 1, 2, 3, \ldots$, and*

$$\sum_{n=2}^{\infty} \psi(n) = \infty.$$

If

$$\phi(n) \sim \psi(n) \quad \textit{as } n \to \infty, \quad \textit{i.e. if} \quad \lim_n \frac{\phi(n)}{\psi(n)} = 1,$$

then, for any fixed n_0,

$$\sum_{n=n_0}^{N} \phi(n) \sim \sum_{n=n_0}^{N} \psi(n), \quad \textit{as } N \to \infty.$$

Proof. Fix a $\delta \in (0, 1)$ and choose $n_* = n_*(\delta) > n_0$ such that

$$(1 - \delta)\phi(n) < \psi(n) < (1 + \delta)\phi(n), \quad \text{for all } n \geqslant n_*.$$

Then

$$(1-\delta)\sum_{n=n_*}^{N}\phi(n)<\sum_{n=n_*}^{N}\psi(n)<(1+\delta)\sum_{n=n_*}^{N}\phi(n),$$

or

$$(1-\delta)\left[\sum_{n=n_0}^{N}\phi(n)-\sum_{n=n_0}^{n_*-1}\phi(n)\right]<\sum_{n=n_0}^{N}\psi(n)-\sum_{n=n_0}^{n_*-1}\psi(n)<(1+\delta)\left[\sum_{n=n_0}^{N}\phi(n)-\sum_{n=n_0}^{n_*-1}\phi(n)\right].$$

Next, we divide all sides of the above inequality by $\sum_{n=n_0}^{N}\phi(n)$. Now $\sum_{n=n_0}^{\infty}\phi(n)=\infty$, thus we can make

$$(1-\delta)(1-\delta)<\frac{\sum_{n=n_0}^{N}\psi(n)}{\sum_{n=n_0}^{N}\phi(n)}<1+2\delta$$

by taking N sufficiently large. Since δ is arbitrary, the proof is finished. □

References

[1] C.M. Bender, S.A. Orszag, Advanced Mathematical Methods for Scientists and Engineers, McGraw-Hill, New York, 1978.

[2] S. Boneh, V.G. Papanicolaou, General asymptotic estimates for the coupon collector problem, J. Comput. Appl. Math. 67 (2) (1996) 277–289.

[3] W. Feller, An Introduction to Probability Theory and Its Applications, vol. II, Wiley, New York, 1966.

[4] P. Flajolet, D. Gardy, L. Thimonier, Birthday paradox, coupon collectors, caching algorithms and self-organizing search, Discrete Appl. Math. 39 (1992) 207–229.

[5] M. Karwan, V. Lotfi, J. Telgen, S. Zionts, Redundancy in Mathematical Programming, Springer, Berlin, 1983.

[6] A. Torn, A. Zallinskas, Global Optimization, Lecture Notes in Computer Science, No. 350, Springer, New York, 1989.