Coupon collecting with coupons that periodically change

George Kesidis  
School of EECS  
Pennsylvania State University  
University Park, PA, USA  
gik2@psu.edu

CSE Dept Technical Report No. CSE-18-001  
Jan. 29, 2018  
revised Feb. 4, 2018

I. INDEPENDENT POISSON COLLECTION AND COUPONS-RESET PROCESSES

We consider a coupon collector that selects from a set of $M$ different coupons uniformly at random and with replacement. Coupon selection is a stationary Poisson process with rate $\beta$. All $M$ coupons are periodically changed/reset according to an independent Poisson process with rate $\delta$. When the coupons change, the collector starts over with zero currently valid coupons. Let $z = \beta/\delta$.

Proposition 1: The stationary (viewed at a typical time) mean number of different coupons obtained is

$$\frac{Mz}{M + z}.$$  \(1\)

Proof: Suppose that, starting from zero, the collector selects coupons over a period of time $T$ during which the coupons do not change. If $K$ coupons are selected during this time, then the mean number of different coupons obtained is

$$M(1 - (1 - 1/M)^K).$$

Since $K \sim \text{Poisson}(\beta T)$, we can condition on $K$ to get that the mean number of different coupons obtained is simply

$$\sum_{K=0}^{\infty} M(1 - (1 - 1/M)^K) \frac{(\beta T)^K e^{-\beta T}}{K!} = M(1 - e^{-\beta T/M}).$$  \(2\)

If $T \sim \text{exp}(\delta)$ then the expected number of coupons found just before they change (i.e., the “average peak”) is, by simply conditioning on $T$,

$$\int_0^\infty M(1 - e^{-\beta T/M}) \delta e^{-\delta T} dT = \frac{Mz}{M + z}.$$  

Finally, since Poisson Arrivals See Time Averages (PASTA) [2], [1], the stationary view is the same as that just before the coupons change.

II. APPROXIMATE MEAN NUMBER OF DIFFERENT COUPONS JUST BEFORE THEY CHANGE FOR A NON-POISSON COUPON-RESET PROCESS

Again let $T$ be distributed as the time between consecutive coupon resets and $K$ be distributed as the aggregate number of coupon selections over $[0, T]$.

Suppose that $\delta \ll \beta (z \gg 1)$ so that the total number of selections over a period length $t$ on the order $1/\delta$ ($\gg 1/\beta$) is $\approx \beta t$ by the law of large numbers. So,

$$\mathbb{E}(M(1 - (1 - 1/M)^K)|T) \approx M(1 - (1 - 1/M)^{\beta T}).$$
Let \( v = -\log(1 - 1/M) > 0 \) if we approximate the distribution of times between consecutive coupon resets as Gaussian, \( T \sim N(1/\delta, (\kappa/\delta)^2) \), then we can approximate the average peak number of different coupons collected as
\[
EM(1 - (1 - 1/M)^{\beta T}) \approx M(1 - e^{-zv + \frac{1}{2}(\kappa zv)^2}).
\] (3)

So, with parameter \( \kappa \) of this approximation, we can explore smaller variances than that of the exponential distribution.

Alternatively, we can use (2) for Poisson selection (even when \( z \gg 1 \) does not hold) and then apply the Gaussian moment generating function to get,
\[
EM(1 - e^{-\beta T/M}) \approx M(1 - e^{-z/M + \frac{1}{2}(\kappa z/M)^2})
\] (4)
as an approximation of the average peak number of different coupons collected. Note that (3) and (4) are approximately equal since \( v = -\log(1 - 1/M) \approx 1/M \) for large \( M \).

Acknowledgments: Our research is supported by Defense Advanced Research Projects Agency (DARPA) Extreme DDoS Defense (XD3) contract no. HR0011-16-C-0055.

REFERENCES