Distribution of Adsense Returns
Philomaths Technical Note - TN10

C. R. Drane

Philomaths Pty. Limited
email chris@philomaths.org

May 13, 2021

Abstract

A deriviation of the probability distribution of Adsense payments.

1 Introduction

Adsense is Google's web advertising platform. It is responsible for much of Google's revenue, and literally millions people use it every day. Insights into the statistical nature of this platform should be of interest to many people. By default, Google presents the results to web site owners in terms of the clicks per day and the total payments on day. For low click rates, distribution of payments is definitely not normally distributed and appears to have a surprisingly large variance given the cost per click should not have a particularly large variance. This paper is a derivation of the distribution of the payments and provides an explanation of why this sort of distribution has an unusually large variance.

2 General Analysis

Assume the average number of clicks in a day is λ per unit time period, e.g. 3.2 clicks per day. Assume the cost per click (CPC) is distributed,

$f(\mu, \sigma^2)$, with mean μ and variance σ^2. This analysis holds for arbitrary unit time periods, such as weeks or months, but for ease of exposition, we will maintain our assumption that the unit time period is one day.

The number of clicks in the i^{th} day is n_i and the total payment in the day is p_i, where

$$P_i = \sum_{j=1}^{n_i} p_{ij} \tag{1}$$

where p_{ij} is the j^{th} payment of the i^{th} time period. p_{ij} are distributed $f(x; \mu, \sigma^2)$. The P_i are distributed

$$f_i(x; \mu_i, \sigma_i^2) \tag{2}$$

where $f_i(\cdot)$ is the distribution that results from adding i random variables with distribution $f(\cdot)$, and takes into account the nature of the joint distribution between the variables, i.e. does not assume they are independent. It is assumed that $f_i(\cdot)$ has a mean μ_i and variance σ_i^2

A large sample of days will result in different total clicks on each day i.e. there will be

- m_0 days with 0 clicks
- m_1 days with 1 click
- ...
- m_N days with N clicks

where N is the maximum number of clicks observed in a day. Let w_k be the weight of days with k clicks.

$$w_k = \frac{m_k}{\sum_{i=0}^{N} m_i} \tag{3}$$

The distribution of the payments for the days with k clicks is

$$P_k(x) = f(x; \mu_k, \sigma_k^2). \tag{4}$$

The distribution of the P_k will be a mixture distribution [1, 2]. The probability density function (p.d.f) of the resultant distribution will be

$$h(y) = \sum_{i=1}^{N} w_i f_i(y; \mu_i, \sigma_i) \tag{5}$$

and the cumulative distribution function (c.d.f) is

$$H(Y) = \sum_{i=1}^{N} w_i \int_{-\infty}^{X} f_i(\mu_i, \sigma_i)\, dx \tag{6}$$

From [3], we have that the central moments of the mixture distribution are given by

$$m_P(j) = \sum_{i=1}^{N} \sum_{k=0}^{j} \binom{j}{k} (\mu_i - \mu_P)^{j-k} w_i m_i(k) \tag{7}$$

where $m_i(k)$ is the k^{th} central moment of $f_i(x)$ and $m_P(j)$ is the j^{th} central moment of the the resulting mixture distribution. For this distribution we will be interested in the mean,

$$\sum_{i=1}^{N} w_i \mu_i, \tag{8}$$

the variance,

$$\sigma_P^2 = m_P(2) \tag{9}$$

the coefficient of variation [4]

$$c_v = \frac{\sigma_P}{\mu_P} \tag{10}$$

the Moment Coefficient of Skewness [5]

$$\gamma_1 = \frac{m_P(3)}{\sigma_P^3}, \tag{11}$$

and the Excess Kurtosis [6]

$$\kappa_e = \frac{m_P(4)}{\sigma_P^4} \tag{12}$$

A heavy tailed distribution has tails that are not exponentially bounded. This is equivalent to the moment generation function $M_H(t)$, being infinite for all $t > 0$ [7]. In this case the definition of the moment generating function [8] is

$$M_H(t) = \sum_{i=1}^{N} w_i \int_{-\infty}^{\infty} e^{-yt} f_i(y; \mu_i, \sigma_i)\, dy \tag{13}$$

3 Poisson Distribution

Assume that the number of clicks in a day are Poisson distributed with parameter λ. The probability of having k clicks in a day is

$$w_k = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{14}$$

where $w_k = Pr(N = k)$. When dealing with a Poisson distribution, we can assume $N = \infty$ Accordingly, we can write equation 7 as

$$m_P(j) = \sum_{i=1}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} \sum_{k=0}^{j} \binom{j}{k} (\mu_i - \mu_P)^{j-k} m_i(k) \tag{15}$$

4 Reference distribution

First assume that there is only one bidder, who always makes the same bid, μ. Accordingly $f(x; \mu, \sigma) = \delta(x - \mu)$, where $\delta(\cdot)$ is the Dirac Delta function. In this case σ is zero. When there are i clicks in a day, the distribution of the payments will be

$$f_i(x; \mu_i, \sigma_i) = \delta(x - i\mu) \tag{16}$$

i.e. the $\mu_i = i\mu$ and σ_i is zero.

We have that $\mu_i = i\mu_i$, so

$$\mu_P = \sum_{i=1}^{N} w_i i\mu \tag{17}$$

For a poisson distribution

$$\sum_{i=1}^{N} iw_i = \lambda \tag{18}$$

so

$$\mu_P = \mu\lambda \tag{19}$$

Now let us consider the central moments of f_i in this case, i.e. assuming constant bids.

$$m_i(k) = \int_{-\infty}^{\infty} \delta(x - i\mu)(x - i\mu)^k \, dx \tag{20}$$

so that we have $m_i(k) = 0$, except for $k = 0$, where $m_i(0) = 1$. Substituting these valued for $m_i(k)$ into equation 15 we have

$$m_P(j) = \sum_{i=1}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} \binom{j}{0} (\mu_i - \mu_P)^j \tag{21}$$

This equation can easily be evaluated with Mathematica with the result;

$$m_P(2) = \lambda\mu^2, \tag{22}$$

$$m_P(3) = \lambda\mu^3, \tag{23}$$

$$m_P(4) = \lambda\mu^4 + 3\lambda^2\mu^4. \tag{24}$$

Now using equations 10. . . 12, we can write that

$$c_v = \frac{\mu\lambda}{\sqrt{\lambda\mu^2}} = \sqrt{\lambda}, \tag{25}$$

$$\gamma_1 = \frac{1}{\sqrt{\lambda}}, \tag{26}$$

$$\kappa_e = \frac{1}{\lambda}. \tag{27}$$

It can be seen the coefficient of variation grows with large λ, but the skewness and the excess kurtosis both approach zero. This confirms that for large λ, the distribution becomes normally distributed. However the story is very different for small lambda. For instance if $\lambda = \frac{1}{9}$, then $\gamma_1 = 3$ and $\kappa_e = 9$, so the distribution is highly asymmetric and has a high propensity to produce outliers. The observation of this behaviour is what motivated this study.

The Moment generating function in this case will be

$$M_H(t) = \sum_{i=1}^{\infty} w_i \int_{-\infty}^{\infty} e^{yt}\delta(y - i\mu)\, dy, \tag{28}$$

so

$$M_H(t) = \sum_{i=1}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} \int_{-\infty}^{\infty} e^{i\mu t}\delta(y - i\mu)\, dy, \tag{29}$$

so, using Mathematica

$$M_H(t) = e^{\lambda(e^{t\mu}-1)}. \tag{30}$$

This is finite for finite t, so although the distribution can have large excess kurtosis, it is not heavy tailed in the formal sense.

5 General Discrete Distribution

Suppose that the bids for advertisements are distributed as

$$f(x;\vec{a}) = \sum_{i=1}^{M} a_k \delta(x - a_k g_k). \tag{31}$$

where $\vec{a} = (a_1, a_2, ..., a_M)^T$ and,

$$\sum_{i=1}^{M} a_k = 1 \tag{32}$$

The a_k represent the probability of a bid of g_k winning. This might represent a situation where the highest bidder wins with a price g_1, then price g_2 wins and so on. Although during a particular day, this will not be a stationary process, looking from day to day, it is reasonable to assume it stationary on a day to day basis.

6 Conclusion

The distribution of Adsense returns and has been found to be a mixture distribution. An explicit solution was found for the case of a single bidder and Poisson distributed clicks. This showed that for a low click frequency, the distribution will be non-normal, with a high degree of skewness and propensity for outliers. However, the distribution is not heavy tailed in the formal sense.

7 Resources

This technical note are available for access at `https://github.com/philomaths-org`. You can access earlier versions of this note on Github by clicking on the tag corresponding to the earlier technical note's version number. This paper has not been refereed, so informed comments are especially welcome.

8 Acknowledgments

This work was supported by the University of Technology, Sydney, and in particular Professor Hung Nguyen. All errors are due to the author.

9 Appendix

The translation property of the Dirac delta function [9], states that for a function $f(x)$

$$\int_{-\infty}^{\infty} \delta(x-\mu) f(x)\, dx = f(\mu) \tag{33}$$

Accordingly

$$\int_{-\infty}^{\infty} \delta(x-\mu)(x-\mu)^k\, dx = (\mu-\mu)^k = 0 \tag{34}$$

for $k > 0$. If $k = 0$, the value of the integral is 1.

References

[1] B. S. Everitt and D. Hand, *Finite Mixture Distributions.* 11 New Fetter Lane, London: Chapman and Hall, 1981.

[2] U. E. M. D. M. Titterington, A. F. M. Smith, *Statistical analysis of finite mixture distributions.* Chichester: John Wiley and Sons, 1985.

[3] Wikipedia, "Mixture distribution." https://en.wikipedia.org/wiki/Mixture_distribution, 2015. [Online; accessed 14-July-2015].

[4] Wikipedia, "Coefficient of variation." https://en.wikipedia.org/wiki/Coefficient_of_variation, 2016. [Online; accessed 10-May-2016].

[5] Wikipedia, "Binomial coefficient." https://en.wikipedia.org/wiki/Skewness, 2016. [Online; accessed 1-Mar-2016].

[6] Wikipedia, "Binomial coefficient." https://en.wikipedia.org/wiki/Kurtosis, 2016. [Online; accessed 1-Mar-2016].

[7] Wikipedia, "Heavy tailed distribution." https://en.wikipedia.org/wiki/Heavy_tailed_distribution, 2016. [Online; accessed 10-May-2016].

[8] Wikipedia, "Moment generating function." https://en.wikipedia.org/wiki/Moment-generating_function, 2016. [Online; accessed 17-May-2016].

[9] Wikipedia, "Dirac delta function." https://en.wikipedia.org/wiki/Dirac_delta_function, 2016. [Online; accessed 19-April-2016].