Ehrenberg-Bass Institute Working Paper
Forthcoming in the *International Journal of Market Research*

"The attendance at sporting events: A generalised theory and its implications"

Author:

Giang Trinh - Ehrenberg-Bass Institute

University of South Australia
Ehrenberg-Bass Institute for Marketing Science

The attendance at sporting events: A generalised theory and its implications

Giang Trinh

Ehrenberg Bass Institute for Marketing Science

University of South Australia Business School

Email: giang.trinh@marketingscience.info

Abstract

This paper generalises the well-known NBD theory to attendance behaviour at sporting events. Using data from a large national survey across a range of sporting events in Australia including Australian football, rugby league, soccer (outdoor), horse racing, motor sports, rugby union, cricket (outdoor), netball (indoor and outdoor), basketball (indoor and outdoor), harness racing, and dog racing, we show that the NBD is very robust in describing sporting event attendance behaviour. This result has implications for sporting event marketing activities, such as which attendee segments should be targeted, how to increase ticket sales, as well as predicting future attendance behaviour.

Introduction and background

The negative binomial distribution (NBD) has been widely used in marketing to model purchase frequency counts, particularly in the packaged goods context. It is known as one of the true marketing laws (Ehrenberg, 1996; Sharp, 2010), and was first applied to marketing science by Ehrenberg (1959) to model brand purchasing behaviour. Ehrenberg made two assumptions: (1) Purchases of a given consumer in successive time periods follow a Poisson distribution. This implies that the variance of purchases within individual consumers is "as if" random over time (i.e. Poisson process). (2) The mean rates of purchasing of different consumers in the long run differ and their distribution is a Gamma distribution. Following these assumptions, the frequency of consumers making 0, 1, 2, 3, … *x* purchases in a given time period can be modelled by the NBD. Ehrenberg (1959) shows the earliest published example of the NBD model fit to purchasing data in consumer packaged goods, where the theoretical values closely match the actual data. This means that brand buying behaviour follows a predictable pattern, and this pattern can provide a useful baseline for managers to evaluate their marketing activities as well as how to grow their brands (Sharp, 2010).

Since the original study by Ehrenberg (1959), the NBD model has been applied to numerous different brands, categories, time periods and countries (Chatfield et al. 1966; Ehrenberg 1988; Ehrenberg 1996; Romaniuk & Sharp, 2016). Recently, the NBD model has been applied to different types of behaviour, such as gambling (Mizerski et al. 2004; Lam & Mizerski 2009), consumption of mobile phone services (Lee et al. 2011), radio listening (Lees & Wright, 2009), blood donation (Faulkner et al. 2013), industrial buying (McCabe, 2015), mental brand associations (Stocchi & Romaniuk, 2009; Romaniuk, 2013). Due to its robustness, the

NBD model has been referred to as a fundamental pattern in marketing (Sharp, 2010), and is one of the most useful managerial models for brand and product purchases (Schmittlein et al., 1985).

In this article we apply the NBD model to sporting event attendance. Previous literature suggests that sporting event attendance behaviour may be a boundary for the law-like patterns generally observed in consumer package good purchasing behaviour (Sharp, 2010). Neale and Funk (2008) initially tested this model on Australian football attendance, using survey data from 651 attendees at an Australian football home game. Their results show that overall, the NBD model is unable to predict attendance behaviour at Australian football games. However, if the seasonal ticket holders are removed from the data, Australian football attendance follows the NBD model. As the data was collected at a football home game, it is skewed toward the seasonal ticket holders (about 60% of the total respondents), making the overall misfit of the NBD model not unusual. For the general population, there should be a very small proportion of seasonal ticket holders compared to non-seasonal ticket holders. Therefore, the NBD model is expected to give a good fit to the general population, similarly to that of non-seasonal ticket holders only. In addition, due to only interviewing the attendees at the football match, the study does not include non-attendees, which is a shortcoming for fitting the NBD model. Theoretically, non-attendees should be included as the NBD is a model of non-negative counts. Lastly, as the study only examines attendance behaviour at Australian football games, it is worthwhile to test if the NBD holds for other sporting events.

To address these shortcomings of the previous study, we use a more comprehensive data set covering a national representative sample across a wide range of sporting events, to see if the NBD model is capable of describing attendance at sporting events. In order to apply the NBD in this context we first assume that attendance frequency x of a given consumer in successive time periods follow a Poisson distribution with parameter λ

$$f(x)_{poisson} = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad (1)$$

with mean

$$E[x] = \lambda$$

and mean frequencies of attendance λ of different consumers in the long run differ and their distribution is a Gamma distribution.

$$f(\lambda; k, a) = \lambda^{k-1} \frac{e^{\frac{-\lambda}{a}}}{a^{k}\Gamma(k)} \qquad (2)$$

where k and a are the shape and scale parameters of the Gamma distribution, respectively

Combining (2) and (1), the probability density function of attendance frequency x is

$$f(x)_{NBD} = (1 + a)^{-k} \frac{\Gamma(x+k)}{x!\Gamma(k)} \left(\frac{a}{1+a}\right)^{x} \quad (3)$$

The maximum likelihood estimation method (Rungie, 2003) then was used to fit the NBD model to the observed data of each sporting even independently.

If attendance at sporting events follows the NBD model, this could have an impact on sports institutions' marketing strategies. For example, a marketing message to persuade people to increase their loyalty to a sporting event would have little effect, since there is strong habitual behaviour associated with this model.

Data

The data was collected from the Australian Bureau of Statistics 2009-10 Multipurpose Household Survey (MPHS). The survey covered all areas except people living in very remote parts of Australia due to operational difficulties, and people living in special dwellings such as prisons, hospitals and boarding schools. The survey was restricted to individuals 15 years old or above. The initial sample was 38,655 private dwellings, but only 32,760 remained in the survey after sample loss. One person in each dwelling was randomly selected based on a computer algorithm. If the randomly selected individual was under 18, permission was obtained from a parent or guardian before conducting the interview. In case permission was not obtained, the parent or guardian answered the survey on behalf of the under 18 individual. Data were collected using a Computer Assisted Interviewing (CAI) technique, where the interviewer recorded responses directly onto an electronic questionnaire in a notebook computer. Of the private dwellings that remained in the survey, 87% fully responded to the MPHS. Approximately half of these respondents were asked questions on attendance at sporting events, including Australian football, rugby league, soccer (outdoor), horse racing, motor sports, rugby union, cricket (outdoor), netball (indoor and outdoor), basketball (indoor and outdoor), harness racing, and dog racing (ABS, 2010).

Results

Table 1 presents the fit of the NBD model to each type of event. As we can see, the NBD model is very robust in describing frequency of attendance at different sporting events. The observed data are closely matched the theoretical estimates. The average mean absolute percentage error (MAPE) is 0.08, ranking from 0.00 (cricket) to 0.18

(rugby league). Consistent across all events, there are many non and light attendees, but very few heavy attendees. This result fits well with the theory of consumer behaviour, where non and light buyers are crucial for brand growth (Ehrenberg, 1988; Sharp, 2010).

Insert Table 1 about here

Implications and future research

Several practical implications can be drawn from the findings. Firstly, in order to increase ticket sales, the marketing of a sporting event should focus on previously non-or light attendees. Although they only buy tickets occasionally, the high number of light attendees (as shown in Table 1) makes them an important focus. While the light attendees are important, previous non-attendees are even more important. This can be demonstrated by the prediction of the NBD model. By using the NBD model to predict attendance at Australian football games, many non-attendees this year are likely to become attendees in future years (see Figure 1). Therefore, sporting event marketers should not ignore these non-attendees when communicating their marketing activity, as in the future they will become important in contributing to ticket sales. One possible promotion strategy to increase the frequency of non and light attendees is offering them a variety of lighter packages than the usual full season ticket holders such as a quarter or a half season ticket holders to satisfy the need of different types of attendees.

Insert Figure 1 about here

Secondly, sporting event managers could use the NBD model to predict changes in attendance behaviour over time using the conditional expectation of the NBD model (Goodhardt and Ehrenberg, 1967; Schmittlein et al., 1985). For example, if a person is a light attendee to a sporting event in one period, what is the probability that he/she is still a light attendee of that sporting event in the next period? Further, what is the probability of that person converting to a heavy attendee in the next period? Consequently, the NBD model could be used as a benchmark tool to evaluate any real change in attendance behaviour (as opposed to stochastic change) caused by marketing activities, such as promotions and advertising.

For future research, one could extend the univariate NBD model to a bivariate NBD model to examine the differences in heterogeneity in attendance behaviour between different consumer segments (e.g. do women attend sporting events less than men?); or between different sporting events (e.g. if an individual has a higher rate of football attendance, does she/he also have a higher attendance rate at tennis?). The bivariate NBD model can be used to answer these questions. The correlation of the bivariate NBD model will indicate how consistent attendance behaviour is across segments or categories, without the results being dominated by differences in the mean rates of attendance. The higher the correlation, the higher the consistency is in attendance behaviour. The second direction of future research is to apply the NBD-Dirichlet model developed by Goodhardt et al. (1984) to competing sporting events within a sport category. This would be useful to identify if other law-like patterns such as double jeopardy and duplication of purchase also hold for sporting event attendance behaviour. The third direction of future research is to compare the fit of the NBD across non-season ticket holders and season ticket holders. Despite our results show that the NBD fits the overall population including the non attendees, it may or may not fit the season ticket holders. Last but not least, although the NBD describes attendance behaviour at sporting events well, it does not have an upper bound to frequency. Whereas, sporting events usually have fixtures. A more theoretical appealing model in this context is the Beta Binomial Distribution (BBD), which has been used for magazine readership (Danaher, 1992), where the number of issues is limited. Future research therefore could apply the BBD to attendance behavior at sporting events and compare with the NBD.

References

ABS (2010). Spectator attendance at sporting events. Australian Bureau of Statistics.

Chatfield, C., Ehrenberg, A.S.C., & Goodhardt, G.J. (1966). Progress on a simplified model of stationary purchasing behaviour. *Journal of the Royal Statistic Society A, 129* (3), 317-367.

Danaher, P. J. (1992). Some statistical modeling problems in the advertising industry: A look at media exposure distributions. *The American Statistician*, *46*(4), 254-260.

Ehrenberg, A.S.C. (1959). The pattern of consumer purchases. *Applied Statistic, 8*, 26-41.

Ehrenberg, A.S.C. (1988). *Repeat buying*. (2nd ed.). London: Griffin. (reprinted in *J. Empirical Generalisations in Marketing Science, 5*, 392–770, available from http://www.empgens.com/).

Ehrenberg, A.S.C. (1996). Towards an integrated theory of consumer behaviour. *International Journal of Market Research, 38* (4), 395-427.

Ehrenberg, A.S.C., Uncles, M.D., & Goodhardt, G.J. (2004). Understanding brand performance measures: using Dirichlet benchmarks. *Journal of Business Research*, *57*, 1307-1325.

Goodhardt, G.J, Ehrenberg, A.S.C., & Chatfield, C. (1984). The Dirichlet: a comprehensive model of buying behaviour. *Journal of Royal Statistical Society Series A*., *147*, 621– 55.

Faulkner, M, Stern. P., & Romaniuk, J. (2013). Do charities comprise a Dirichlet Market?, *Australian New Zealand Marketing Academy Conference*, Auckland.

Goodhardt, G. J., & Ehrenberg, A. S. C. (1967) Conditional trend analysis: A breakdown by initial purchasing level. *Journal of Marketing Research*, *4* (2), 155-161.

Lam, D., & Mizerski, R. (2009). An investigation into gambling purchases using the NBD and NBD–Dirichlet models. *Marketing Letters, 20*, 263-276.

Lee, R., Rungie, C., & Wright, M. (2011). Regularities in the consumption of a subscription service. *Journal of Product & Brand Management, 20* (3), 182-189.

Lees, G., & Wright, M. (2009). Do the NBD-Dirichlet's empirical generalisations hold true for radio listening?, *23rd ANZAM Conference: Sustainable Management and Marketing Proceedings*

McCabe, J. D. J. (2015). Capital Equipment Purchasing: A Stochastic Model of Industrial Buying Behaviour. In *Revolution in Marketing: Market Driving Changes* (pp. 155-159). Springer International Publishing.

Mizerski, D., Miller, R., Mizerski, K., & Lam, D. (2004). The stochastic nature of purchasing a state's lottery products. *Australasian Marketing Journal, 12* (3), 56-69.

Neale, L., & Funk, D.C. (2008). Using the Negative Binomial Distribution to Investigate Sports Attendance. *Proceedings from the Australian and New Zealand Marketing Academy (ANZMAC) Conference*. Sydney, AUS.

Romaniuk, J. (2013). Modeling mental market share. *Journal of Business Research*, *66*(2), 188-195.

Romaniuk, J & Sharp, B. (2016). *How brands grow Part 2*, Oxford University Press, Australia.

Rungie, C. (2003). How to estimate the parameters of the Dirichlet model using likelihood theory in Excel. *Marketing Bulletin*, *20*, Technical Note 2, 1-9.

Sharp, B. (2010). *How brands grow*, Oxford University Press, Australia & New Zealand.

Sharp, B., Wright, M., Dawes, J., Driesener, C., Meyer-Waarden, L., Stocchi, L., & Stern, P. (2012). "It's a Dirichlet World: Modelling individuals' Loyalties reveals How Brands Compete, Grow, and Decline", *Journal of Advertising Research*, *52* (2), 203-213.

Schmittlein, D. C., Bemmaor, A., & Morrison, D. G. (1985). Why does the NBD model work? Robustness in representing product purchases, brand purchases and imperfectly recorded purchases. *Marketing Science*, *4* (3), 255-266.

Stocchi, L & Romaniuk, J. (2009). Understanding and Predicting Brand Attribute Frequency Distributions, *Australian New Zealand Marketing Academy Conference*

Table 1. The fit of the NBD model to the frequency of attendance to each sporting event

Frequency of attendance	Australian Football		Basketball		Cricket		Dog Racing	
	O	*T*	O	*T*	O	*T*	O	*T*
Non	83.80	*83.76*	98.74	*98.74*	96.12	*96.12*	98.39	*98.39*
Light (1-2)	6.88	*7.45*	0.51	*0.62*	2.42	*2.42*	1.01	*1.04*
Medium (3-5)	4.27	*3.64*	0.38	*0.28*	0.89	*0.89*	0.41	*0.36*
Heavy (6+)	5.05	*5.15*	0.38	*0.37*	0.58	*0.58*	0.20	*0.21*
MAPE	*0.07*		*0.14*		*0.00*		*0.06*	
Frequency of attendance	**Harness Racing**		**Horse Racing**		**Motor sports**		**Netball**	
	O	T	O	T	O	T	O	T
Non	97.64	97.64	88.90	*88.91*	91.86	*91.87*	98.99	*98.99*
Light (1-2)	1.70	*1.70*	8.20	*8.01*	5.53	*5.26*	0.39	*0.42*
Medium (3-5)	0.49	*0.48*	2.02	*2.30*	1.47	*1.85*	0.22	*0.20*
Heavy (6+)	0.17	*0.18*	0.88	*0.79*	1.14	*1.03*	0.40	*0.38*
MAPE	*0.01*		*0.07*		*0.09*		*0.05*	
Frequency of attendance	**Rugby League**		**Rugby Union**		**Soccer**		**Tennis**	
	O	T	O	T	O	T	O	T
Non	91.05	*91.02*	96.71	*96.71*	94.63	*94.62*	98.32	*98.32*
Light (1-2)	3.54	*4.35*	1.82	*1.86*	2.14	*2.43*	1.38	*1.36*
Medium (3-5)	2.98	*2.02*	0.81	*0.76*	1.48	*1.15*	0.24	*0.27*
Heavy (6+)	2.43	*2.62*	0.66	*0.67*	1.75	*1.79*	0.07	*0.05*
MAPE	*0.18*		*0.03*		*0.11*		*0.09*	

O: Observed, T: Theoretical, MAPE: Mean Absolute Percentage Error

Figure 1. Event penetration built over time in Australian football.