Article

Musicae Scientiae
1–14

DOI: 10.1177/1029864915619000
msx.sagepub.com

Let the music be your master: Power laws and music listening habits

Nicolás Mongiardino Koch
Universidad de Buenos Aires and IEGEBA - CONICET, Argentina

Ignacio M. Soto
Universidad de Buenos Aires and IEGEBA - CONICET, Argentina

Abstract

Music preferences have long been studied owing to their importance in the fields of psychology and sociology. However, previous efforts seldom focused on people's deliberate choices of music in everyday life. In this study, we aimed to analyze music listening behaviors using personal records of music listening activity. We obtained the history of songs listened to by 50 different users of the online database system *Last.fm*, spanning on average five years of activity. With the use of this data set, we are able to confirm that the number of songs reproduced per artist follows a truncated power-law distribution. The scaling parameter of the distribution varies considerably among users, providing a metric that characterizes the way in which different people explore music. We propose that this pattern is consistent with a preferential attachment model, according to which the probability of listening to a given artist at a given time is proportional to the frequency to which the artist was listened to in the past. These results provide new insight regarding the way in which individual music preferences are built.

Keywords

listening behaviors, music, music preferences, power law, preferential attachment, taste

Introduction

Nature is full of regularities, and to many, science is in itself a quest to understand and explain these regularities (Andersen, 2011; Foster, 2004; Wigner, 1964). Among the most pervasive types of regularities are the so-called power laws, which are found to underlie diverse physical, biological and anthropological phenomena, such as the magnitude of earthquakes, wildfires, biological extinctions and wars (Pinto, Lopes, & Tenreiro Machado, 2012). The occurrence of

Corresponding author:
Nicolás Mongiardino Koch, Intendente Güiraldes 2160, Ciudad Universitaria (C1428EGA), Capital Federal, Buenos Aires, Argentina.
Email: nmongiar@ege.fcen.uba.ar

power-law distributions in data sets pertaining to such different disciplines is sometimes called Zipf's law (Gan, Li, & Song, 2006).

Mathematically, a variable x is said to follow a power law if it is drawn from a probability distribution $p(x) \propto x^{-\alpha}$, where α is a constant known as the exponent or scaling parameter (Clauset, Shalizi, & Newman, 2009). Interest in this sort of probability distributions has a long history, dating back to the observation that income (Pareto, 1896), frequency of words in texts (Zipf, 1932) and city sizes (Zipf, 1949) follow such relationships. Power-law behaviors are interesting not only because of their occurrence in extraordinarily diverse phenomena, but also because of their unusual mathematical properties. For instance, power laws are said to be "scale-free" (i.e. invariant under linear rescaling of axes; Deluca & Corral, 2013), meaning that they have no characteristic scale. Power laws also assign sizable probabilities to the occurrence of extremely large events, thus differing from more frequently encountered (e.g. Gaussian) distributions. Furthermore, depending on the value of the scaling parameter, power laws may lack finite moments, meaning that they may have undefined mean and standard deviations (for more information on mathematical properties of such distributions see Sornette, 2006). Beyond the interesting properties of power laws, knowing whether a given quantity follows a power law or a different type of probability distribution is important when searching for generative mechanisms that may underlie the processes under study (Alstott, Bullmore, & Plenz, 2014; Virkar & Clauset, 2014).

Empirical research on the existence of power laws in music was initiated by Zipf himself (Zipf, 1949). Using a limited data set, he concluded that the rank-frequency plots for melodic intervals, as well as for the distance between repetitions of notes, followed a power law. Subsequently, using both musical compositions (such as Bach's Brandenburg Concertos) and continuous 12 hour recordings from radio stations, Voss and Clarke (1975) also showed that several properties of music, such as volume and pitch, fluctuated according to a power law. Many researchers have since proven that several quantities that define the structure of music compositions follow power-law behaviors (Manaris, Romero, & Machado, 2005 and references therein; see also Levitin, Chordia, & Menon, 2012), and have used these statistical characterizations to aid in processes such as computerized music generation (Voss & Clarke, 1978), author attribution and style identification (Machado, Romero, Manaris, Santos, & Cardoso, 2003; Machado, Romero, Santos, Cardoso, & Manaris, 2004).

Nonetheless, the advent of music recommendation systems and other social music services has allowed for a different application of power laws to music research, contributing to a sociological analysis of music. Both artists and listeners form communities, with individuals being connected by collaborations, similarity, taste, etc. (Aucoutier & Pachet, 2007; Cano, Celma, Koppenberger, & Buldu, 2006; de Lima e Silva et al., 2003; Gleiser & Danon, 2003). The properties of such networks can be studied using the tools of graph theory and complex network analysis, an approach that has contributed significantly towards our understanding of music evolution and development, genre structuring and music social dynamics. For example, many sociological aspects of human listening habits, such as the popularity of artists, records and songs, as well as the overall level of activity of users in music service systems, have been shown to follow power-law distributions (Celma & Cano, 2008; Chung & Cox, 1994; Hu & Han, 2008; Lambiotte & Ausloos, 2005).

Music preferences have a long history of research in the fields of sociology and psychology. However, few studies have focused on this topic using a quantitative framework, and even fewer have focused on everyday life choices (Greasley & Lamont, 2006). Here we examine the music listening habits of users of *Last.fm* (www.last.fm), the world's largest social music service (Henning & Reichelt, 2008). Unlike recent efforts towards characterizing listening behaviors

(Celma & Cano 2008; Hu & Han, 2008), we focus for the first time on the artist playcount per user, a quantity directly related to the users' preferences. We are able to show that the distribution of the total amount of listened songs per artist follows a power law. Consequently, we propose that music listening profiles are built following a preferential attachment model, and discuss the ways in which this model may contribute to our understanding of individual music preferences.

Method

Data sets

We used the social network and online music database *Last.fm* as our data source. Given that our aim was to characterize music listening habits through the statistical analysis of listening profiles, we only selected users whose profiles exceeded 30,000 listened to songs, aiming at controlling the effects of insufficient sampling. Users were searched at random, and their entire listening profile was downloaded. The procedure was then repeated until obtaining 50 different profiles. The resulting data set represents the music-listening history of 50 people during 5.1 ± 1.3 years (average time-lapse between the creation of each profile and the start of the experiment), and ranges from approximately 35,000 to 194,000 listened songs.

From this database, we compiled the total number of song plays per artist. The lists were manually edited in order to reduce tagging errors, leading to the elimination (e.g. "*[unknown]*") or the merging (e.g. "*Beethoven*" and "*Ludwig van Beethoven*") of certain artist tags. Finally, each user was characterized as a vector, with each position representing the number of songs reproduced per artist, arranged in descending order.

Parameter estimation

Effectively demonstrating that a given quantity follows a power-law distribution is a difficult task (Goldstein, Morris, & Yen, 2004; Newman, 2005). Until fairly recently, whether or not a given set of data was appropriately described by a power law was decided after visual inspection on a double logarithmic plot (on which the data are expected to show a linear behavior), aided with the use of linear least-squares fit. Such methods have been contested on several grounds, including both their validity and accuracy, and are therefore considered to be inefficient at proving and characterizing the power-law behavior of a given data set (for a description of the problems associated with this approach, refer to Bauke, 2007; Clauset et al., 2009; White, Enquist, & Green, 2008).

For our analyses we used the procedure described in Clauset et al. (2009), as implemented in the R package *poweRlaw* (Gillespie, 2015) and the Python package *powerlaw v 1.3.4* (Alstott et al., 2014). Data were handled and plotted after transformation into the complementary cumulative density function (CCDF), defined as the probability that the quantity of interest x is greater than or equal to a given value (i.e. $P(x) = Pr(X \geqslant x)$; in our case, the fraction of bands with a playcount greater than or equal to a given value). This technique largely reduces the errors associated with poor sampling in the tail of the distribution, while at the same time avoiding the need to bin the data (Newman, 2005). The estimation of the scaling parameter (α) for the power-law model was done under maximum likelihood, considered to be the most accurate approach provided sufficient sampling (Bauke, 2007; Clauset et al., 2009). Given that power-law behaviors are often found at the tail of the distributions (hence their classification as "heavy-tailed"), with strong deviations towards lower values of x, a crucial task is the correct identification of the minimum value (or x_{min}) above which the distribution follows a power law (Newman, 2005). This was accomplished using the methodology proposed by Clauset et al. (2009), which aims at finding the value of x_{min} such that the similarity between the probability distribution of the measured data and the best-fit power-law model for $x \geqslant x_{min}$ is maximized. As most recent studies have done, we employed the Kolmogorov-Smirnov (KS) distance to measure this similarity (Deluca & Corral, 2013; p. 1363).

As has been noted, power-law distributions measured in natural data sets will be unlikely to extend indefinitely, given the existence of finite size limitations (Burroughs & Tebbens, 2002; Sornette, 2006). A famous example of this is the Gutenberg-Richter distribution, which describes a power-law relationship between the frequency of earthquakes and their seismic moment (a quantity related to their magnitude). Given that the seismic moment is proportional to the energy dissipated by an earthquake, there has to be an upper bound to the magnitude of earthquakes, given by the fact that a finite Earth cannot release an infinite amount of energy (Burroughs & Tebbens, 2002; Knopoff & Kagan, 1977). Such finite size limitations are modeled using a truncated version of the power law, which shows a power-law behavior over some range, but is truncated by an exponentially bounded tail as it approaches the system size (i.e. it decays faster than a pure power law would; Alstott et al., 2014; Jensen, 1998). In our case, we also expect the data to suffer from such limitations, given that there is an upper limit to the number of song plays an artist can accrue during a finite amount of time. This upper limit will probably be user specific, and will depend on several factors, such as the total amount of time devoted to music listening, the relative distribution of time among different artists, and the length of the songs reproduced. We therefore fitted each user's data to a truncated power law as well.

Comparison to alternative distributions

Finding the parameters of the power law that best describes the data does nothing to prove that the data follows a power-law distribution. Different methods have been proposed to address this question (Clauset et al., 2009), the best of which relies on comparing the power-law model against alternative distributions through some measure of goodness-of-fit (Alstott et al., 2014). The KS statistic provides a first approximation towards comparing the goodness-of-fit of alternative models, given that the model that best characterizes the empirical distribution is also expected to yield the smallest KS distance (Klaus, Yu, & Plenz, 2011). Similarly, log-likelihood ratio (LLR) tests can be used for direct pairwise comparisons of plausible competing distributions (Clauset et al., 2009), and the significance of the test can be analyzed using the method proposed by Vuong (1989). Following Clauset et al. (2009), we used a *p*-value for significance of 0.1. Alternative hypotheses tested (besides the power law and the truncated power law, which were compared using the modified LLR test for nested distributions) were the exponential, stretched exponential, lognormal and Poisson distributions. As noted by Alstott et al. (2014), the exponential distribution is the absolute minimum alternative candidate to be tested, given that the operational definition of a heavy-tailed distribution is that it is not exponentially bounded (Asmussen, 2003). The remaining distributions are all plausible alternative candidates regularly inspected in the literature. However, it should be noted that, finding a significant statistical support for a power-law model over a lognormal or stretched exponential is extremely difficult (Clauset et al. 2009; Malevergne, Pisarenko, & Sornette, 2005; Mitzenmacher, 2004), given the similarity between these distributions over the ranges of values regularly studied. Furthermore, lognormal and stretched exponential distributions have an extra degree of freedom with respect to the power law, making rejections of this last distribution using LLR tests

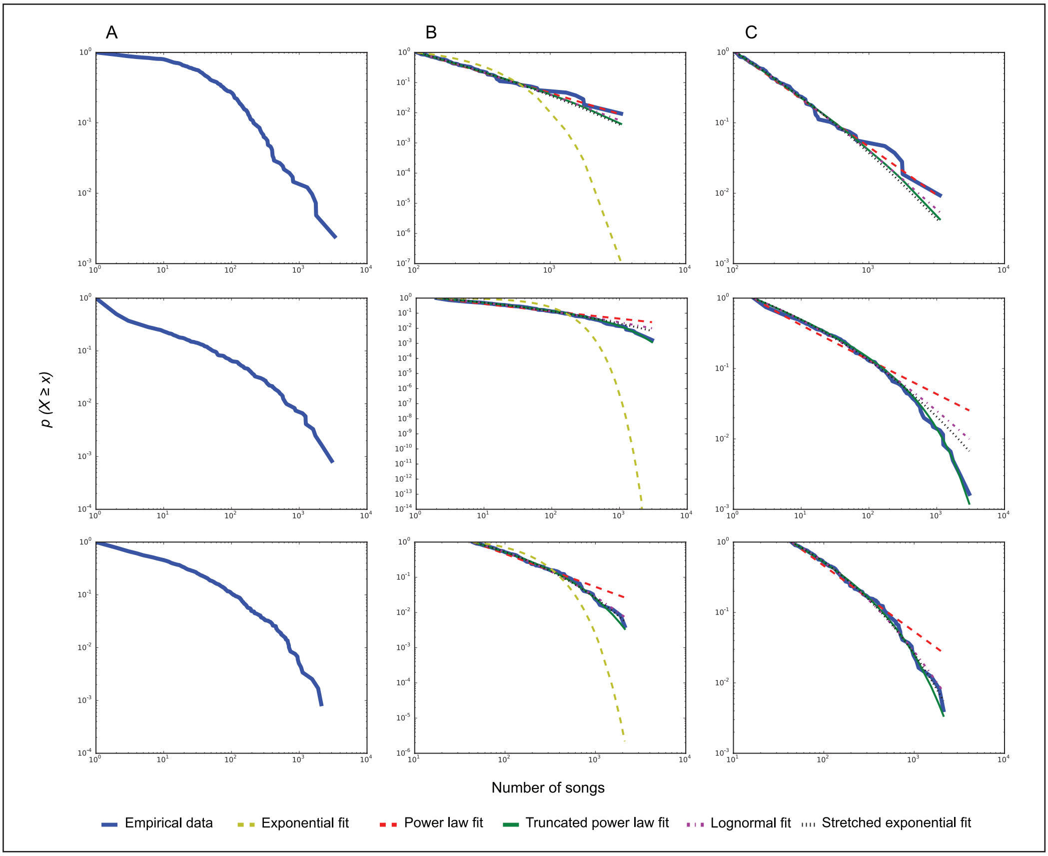

Figure 1. Examples of the CCDF of the number of song plays per artist for different users (rows), along with their best fit to different models (a key for line types is found at the bottom). A. Empirical data. B. Fit to all models, showing that the data is poorly described by an exponential model. C. Fit to "heavy-tailed" models. The tail of the distribution behaved somewhat differently depending on the user (see text). Plots B and C show only values above x_{min}.

difficult to interpret. Given that truncated power laws were generally found to be better descriptors than pure power laws (see Results section), candidate distributions with two parameters were compared to truncated power laws (an approach proposed by Klaus et al., 2011).

Results

Example distributions can be seen in Figure 1. When plotted on a double logarithmic axis, the data was found to follow a straight line only in its central region, with deviations from linearity generally present at both extremes (Figure 1a). After finding the x_{min} value for each user and restricting the analysis to the tail of the distribution, the data was confirmed to be "heavy-tailed". As can be seen from the three examples shown in Figure 1b, the data strongly deviates from an exponential model, with extreme events (i.e. artists with a very large playcount) obtaining sizable probabilities. From the "heavy-tailed" models tested, the power law seemed to be an appropriate descriptor for some of the users (Figure 1c, top row), while for the vast majority, the

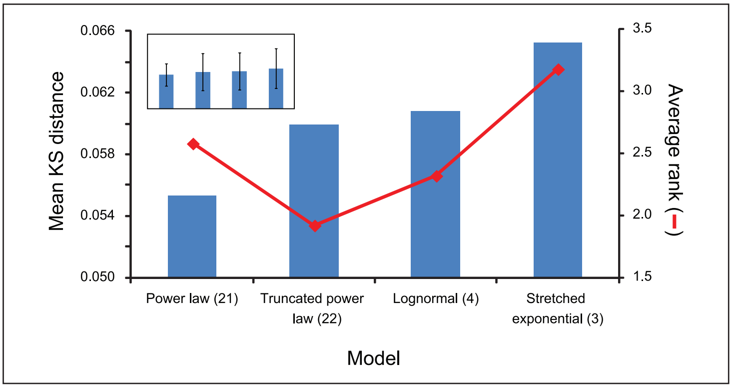

Figure 2. Comparison of alternative distribution models using the Kolmogorov-Smirnov distance metric. Values shown are the mean KS distance (bars) and the average rank (line) derived from the KS metric. Differences in mean KS distance between models are not significant, probably as a consequence of the measure's large variability between users (standard deviations are shown as vertical bars in the figure inset). On the other hand, the truncated power-law model was found to be ranked significantly lower than all other remaining models, while the stretched exponential model was ranked significantly higher than all others (all $p < .05$). Values shown between parentheses are the number of times a given model was ranked first.

deviations toward the largest values of x resulted in a better fit to a truncated power law (Figure 1c, middle row). However, finding the "heavy-tailed" model that best described the data was not always feasible from a graphical standpoint, with many cases in which different models behaved too similarly to allow any distinction (Figure 1c, bottom row).

The use of the KS distance metric to compare the goodness-of-fit of the different models to each user's data produced a similar result. Both the exponential model and the Poisson model were found to be extremely poor descriptors of the empirical distributions, with KS distances generally an order of magnitude above those of the other models (across-user mean KS values of 0.29 and 0.71, respectively). The remaining four models consistently showed very low KS values, which were furthermore quite similar to each other, as can be seen in Figure 2. The power law and truncated power-law models showed the smallest KS values overall, although differences among models were found to be non-significant (Kruskal-Wallis test, $p = .74$). This lack of significance was seemingly a result of the huge standard deviations within each model (see Figure 2 inset). Therefore, we used the Friedman test to study whether the models significantly differed in their ranks. By comparing ranks instead of absolute KS values, this test allows us to compare the performance of each model without being affected by the huge variance between users. Significant differences were in fact detected in the ranking of the four contending models ($p = 10^{-6}$). With the use of the post-hoc test of Conover (1999, p. 371), we were able to conclude that the truncated power-law model was significantly better ranked than all of the remaining models (see Figure 2 and caption).

Distribution models were subsequently compared using the LLR test. First, the power-law and truncated power-law models were compared using the nested version of the LLR test. For the 50 profiles studied, 76% were found to be significantly better described by the truncated power-law model. This value increased to 98% when considering all instances in which the truncated power law was found to be better, irrespective of whether such differences were significant or not. This means that for only one profile the pure power law was a better model than the truncated power law (and even in this case, the difference in fit was not significant). Given that truncated power laws were found to be statistically better than pure power laws for most cases using both the LLR test and the average ranking derived from KS distances, stretched exponential and lognormal models were compared against truncated power laws, avoiding the problems of comparing distributions with different numbers of parameters (Klaus et al., 2011). For the same reason, exponential and Poisson distributions were compared with pure power laws.

Overall results of the LLR tests supported the power-law hypothesis. None of the 200 LLR tests performed comparing power laws to alternative models resulted in a statistically significant result favoring the alternative model. A case by case description of the results is synthesized in Table 1. Data corresponding to the LLR tests against Poisson distributions are not shown, given that power laws were found to be significantly better models in all cases. Overall, the proportion of cases in which the power-law model resulted in a significantly superior fit to the data was 76% against exponential models, 56% against stretched exponentials and 62% against lognormals. Such values increase to 92, 94 and 92%, respectively, when considering all results favoring power laws, not discriminating between significant and non-significant results.

Finally, given that accurately discriminating among power laws, stretched exponentials and lognormal distributions generally requires an extremely large dataset (Clauset et al., 2009; Malevergne et al., 2005), we tested whether profiles for which power laws were not the best alternative were those with the least amount of data. This could justify the occasional lack of efficacy of the described methods to tell the different distributions apart. The results confirmed this: the number of data points above x_{min} (i.e. the number of data points used to fit the different models) was significantly smaller in the eight profiles for which at least one alternative distribution had a higher likelihood (distributions with at least one light gray cell in Table 1; one-way ANOVA: $F = 6.32$, $p = .015$). On average, profiles for which power laws were the model with the overall highest likelihood had 1.65 times the number of data points to fit than profiles for which the model with the highest likelihood was an alternative one.

Conclusions

Music is one of the most ubiquitous cultural expressions of mankind, and has been considered by some to be one of the most biologically significant activities in human life (Cross, 1999). Given its relevance to the understanding of human psychological and sociological dynamics, there has been a long tradition of inquiry into music preferences (Martin, 1995; Sloboda, 1985). Sloboda, Lamont, and Greasley (2011) identified four basic functions music can play in our daily life: music as a distracting, energizing, entertaining and meaning-enhancing activity. Music can also be used to define our social identity and guide our relationships with others (Laiho, 2004; Trepte, 2006), or may simply be enjoyed because of its aesthetic appeal. Given this breadth in uses and purposes, music preferences are expected to be an amalgam of many complex underlying factors, including personality traits, familiarity and repetition, social context and musical training (see Lamont & Greasley, 2009 and references therein). Our knowledge on how these variables interact with the characteristics of music, resulting in differences in preferences, have been mostly derived from studies focusing on people's verbal or behavioral responses to music. However, these experiments are generally performed with the use of music either artificially contrived or chosen by the experimenter (North & Hargreaves, 1997), hardly ever focusing on people's unassisted tastes and choices (but see Greasley & Lamont, 2006). Furthermore, most studies focus on obtaining responses either to individual songs or to genres/styles, without analyzing the intermediate category level of the artist, even though many people spontaneously describe their music preferences referring to this level (Greasley, Lamont, & Sloboda, 2013).

Table 1. Data sets employed, along with the parameters of the best fit truncated power-law models (α and λ) and the results of the LLR tests.

User	No. of songs	No. of artists	x_{min}	% in tail	α	λ	vs. EX	vs. SE	vs. LN
1	49141	1440	32	81	2.107	5.32^{-04}			
2	72868	356	177	49	1.718	2.37^{-04}			
3	58508	971	1	100	1.358	4.74^{-04}			
4	46274	1284	99	52	1.000	6.23^{-03}			
5	58313	1831	79	59	2.360	5.17^{-04}			
6	193579	502	107	65	1.697	2.19^{-05}			
7	64085	2776	34	84	1.989	4.82^{-04}			
8	35861	2955	49	60	2.353	3.98^{-03}			
9	51323	1590	3	99	1.455	8.75^{-04}			
10	97609	1519	10	96	1.547	2.84^{-04}			
11	167622	2610	101	73	1.495	1.44^{-03}			
12	35299	1476	1	100	1.449	1.23^{-03}			
13	112503	2119	1	100	1.292	8.98^{-04}			
14	43613	1397	9	95	1.374	2.64^{-03}			
15	52473	300	301	31	1.000	1.46^{-03}			
16	56018	2444	1	100	1.389	1.40^{-03}			
17	58568	707	25	84	1.631	1.34^{-04}			
18	46058	503	130	45	1.814	7.68^{-04}			
19	57443	1178	44	79	1.535	1.01^{-03}			
20	86489	373	509	27	1.656	3.64^{-04}			
21	72262	2715	15	95	1.751	1.01^{-03}			
22	98604	821	131	55	1.631	3.22^{-04}			
23	64478	629	174	37	2.253	2.60^{-05}			
24	37376	352	25	83	1.483	6.85^{-04}			
25	53570	430	5	97	1.201	4.60^{-04}			
26	109018	921	4	99	1.341	2.89^{-04}			
27	120666	4059	5	100	1.561	4.13^{-04}			
28	73884	959	149	47	2.224	1.29^{-04}			
29	41509	412	104	52	2.240	1.90^{-04}			
30	83405	740	253	40	2.467	–			
31	68626	1198	97	62	1.287	8.26^{-04}			
32	94483	620	335	32	1.000	2.17^{-03}			
33	59460	486	95	53	1.658	1.34^{-04}			
34	155107	2344	327	32	1.946	3.81^{-04}			
35	65135	836	196	37	1.828	2.71^{-04}			
36	49994	385	148	51	1.638	4.77^{-04}			
37	50594	729	53	71	1.536	5.50^{-04}			
38	46807	430	142	43	1.614	3.44^{-04}			
39	55166	3155	3	99	1.617	1.10^{-03}			
40	102503	1702	335	23	1.853	2.89^{-03}			
41	92270	1003	62	74	1.465	5.33^{-04}			
42	40616	710	28	82	1.522	6.32^{-04}			
43	121384	1030	1	100	1.259	3.45^{-04}			

Table 1. (Continued)

User	No. of songs	No. of artists	x_{min}	% in tail	α	λ	vs. EX	vs. SE	vs. LN
44	42723	1213	2	100	1.363	6.79^{-04}			
45	123537	2005	81	71	1.603	4.09^{-04}			
46	118438	514	296	46	1.514	2.42^{-04}			
47	71662	657	217	36	2.902	4.98^{-06}			
48	42916	2043	44	73	2.340	2.21^{-04}			
49	100761	4000	2	100	1.586	5.61^{-04}			
50	50876	1661	8	96	1.461	2.57^{-03}			

Note. Only in the case of user 30 (the only one for which the power law showed a superior fit than the truncated power law), the value of α shown is the exponent of the power law. For the LLR tests, black represents that the power-law model was a significantly better fit ($p < .1$) to the data, dark gray that it was better, yet not significantly, and light gray that it was worse, yet not significantly. EX = exponential; SE = stretched exponential; LN = lognormal.

The present study made use of the great wealth of information regarding listening behaviors that is available through online music databases. Such an approach has already significantly enriched our understanding of the social (Lambiotte & Ausloos, 2005) and temporal (Mauch, MacCallum, Levy, & Leroi, 2015) dynamics of music. By analyzing the number of song plays per artist by each user, we were able to determine that these quantities follow a truncated power-law distribution. This result was corroborated by comparing this and other alternative models through measures of goodness-of-fit, derived from both maximum likelihood and the Kolmogorov-Smirnov distance. Overall, this procedure provided a strong statistical support favoring the truncated power-law model over all other alternatives. Even in the few instances in which results were not conclusive, we were able to find evidence that this was likely an effect of insufficient sample size rather than deviations from the proposed model.

Our results have direct ties to what has been called the Long Tailed model (Anderson, 2006), which relies on the popularity of products following a "heavy-tailed" distribution. However, such a pattern reflects the aggregate behavior of consumer communities, and as such, is better suited towards understanding economic dynamics than the psychological bearing of individual choices. Our analysis, on the other hand, can provide several insights for the study of music preferences. By directly employing people's record of musical activity, we were able to derive a metric of preference different from the ones commonly employed in the field, which are usually based on some kind of preference scale (Litle & Zuckerman, 1986; Rentfrow & Gosling, 2003) or derived from intensive interviewing sessions (DeNora, 2000). The playcount of different artists may be superior to all of these as a measure of preference, defined by Price (1986: 154) as the "act of choosing, esteeming, or giving advantage to one thing over another". Given that there exists a limit to the amount of time people can devote to music listening, choosing to listen to a given artist means also choosing not to listen to any other.

It is significant that all user profiles conform to the same type of distribution. This means that, despite their differences in music preferences, and even despite the fact that they may listen to music in different ways and with different purposes, all of the users studied shared the same overall listening habits, distributing listening time among artists following similar underlying rules. Nonetheless, beyond the similarities found, it should be noted that the exponent of the power law strongly varied between users, adopting absolute values $1 < \alpha < 3$. Lower values show that listening time is distributed more homogenously among artists, while higher values represent a stronger asymmetry in the time devoted to each artist. Given that people who prefer a wider range of styles are expected to listen to a wider breadth of artists in a repeated and consistent way, rather than anecdotally, the scaling coefficient of the power law seems to be a quantity somewhat related to the "omnivorousness" of each person's music taste (Coulangeon & Lemel, 2007; Peterson & Kern, 1996; Peterson & Simkus, 1992). People whose listening habits result in high exponent values are probably specialists (or "univores"), listening over and over to a limited number of artists (and therefore, a limited number of styles), while people with lower exponent values are likely to be much more eclectic in their tastes, appreciating the aesthetics of many distinctive forms of music. Although Greasley et al. (2013) pointed out a possible relationship between the level of engagement with music and the "omnivorousness" of musical taste, we were not able to find any correlation between the scaling exponent and the absolute number of songs or artists listened to by the users. It should however be pointed out that all of the users in this study would probably be regarded as having a strong level of engagement with music. Further research on this and other ways of quantifying the asymmetry of time distribution among artists must be undertaken before drawing any further conclusions.

Understanding the way in which a quantity is distributed is extremely useful when trying to propose generative mechanisms, and those leading to power-law distributions have been extensively reviewed (Mitzenmacher, 2004; Newman, 2005). Of these, the preferential attachment mechanism (also called Yule process) is regularly cited as a possible mechanism underlying certain human dynamics (Abbasi, Hossain, & Leydesdorff, 2012; Barabási & Albert, 1999; Jones & Hancock, 2003), and is also a plausible scenario for the type of data being handled here. This model was originally proposed by Yule (1925) to explain the power-law distribution of the number of species per genus. It was later generalized by Simon (1955) for its application to a variety of situations, and gained much attention as a possible model behind the growth of the World Wide Web (Barabási & Albert, 1999), as well as other types of evolving networks (Vázquez, 2003). The mechanism describes the dynamics of a system composed of a group of objects, each one of which also possesses an associated quantity. At each time step, new objects are added to the system, and the quantity associated with the existing objects is increased by a magnitude proportional to the value already attained by them. In our case, the user's profile represents the system, with artists as the objects and the playcount as the measured quantity. As time progresses, new artists are discovered and incorporated into the profile, while more songs are listened to from already known artists, increasing their playcount. In case the probability of listening to a given artist is proportional to the number of times the artist was listened to in the past, the result of the process will be a power-law distribution. Adding a finite size limitation to the number of played songs an artist can accrue results in the truncated power law found in this study.

The way in which the dynamics proposed by the preferential attachment model relates to what we already know regarding patterns of music listening habits still needs to be explored. However, repetition has been unanimously considered a key aspect defining aesthetic preferences (Berlyne, 1971; Greasley & Lamont, 2006; Hargreaves, 1984; Johnston, 2015; Russell, 1986; Walker, 1980; Zajonc, 1968). It should be noted that most past research, and the models derived from it (such as the inverted U-shaped relationship between familiarity and liking; Russell, 1986) are based on experiments involving non-voluntary exposure to music, and may be for that same reason more suitable for explaining the changes in preferences towards "radio hits" than towards personal favorites. In this sense, research focusing on people's own music collections and everyday life experiences has revealed listening habits and preferences that are more complex than the ones suggested by the inverted U-shape model. When variables such as the context of the listening experience, its effect on personal mood and its emotional associations are taken into account, the richness and complexity of listening behaviors become evident (Greasley et al., 2013; Lamont & Webb, 2010; North & Hargreaves, 1996). Such is the case, for example, of the possible decoupling between short- and long-term musical preferences (Lamont & Webb, 2010), of the conscious regulation of exposure to avoid over-familiarization, and of the resulting cyclical pattern of choices in order to refresh personal favorites (Greasley & Lamont, 2006). It is in this context that the present study should be interpreted.

Finally, recommender systems have historically aimed at providing accurate suggestions based on predicting and matching user information needs (Cremonesi, Koren, & Turrin, 2010). However, the fact that recommender systems also need to provide novel and diverse suggestions has been recently recognized by many (McNee, Riedl, & Konstan, 2006; Vargas & Castells, 2011). These three dimensions are involved in a trade-off (Ribeiro, Lacerda, Veloso, & Ziviani, 2012): independently maximizing one of them generally leads to poor results in the others. The power-law behavior of each user's profile is a rendition of the way in which music is explored, and as such, it combines aspects relevant to the calculation of the relative importance of all three dimensions. Therefore, this information could provide user-specific weights for the dimensions involved – a critical step towards finding optimal algorithms (Ribeiro et al., 2012).

Acknowledgements

We would like to thank Jeff Alstott for providing help with the use of software, and Aaron Clauset for insightful comments at early stages of the research. Furthermore, we are grateful to Bhart-Anjan S. Bhullar for language correction, as well as two anonymous reviewers whose contributions improved the present manuscript. Nicolás Mongiardino Koch was supported by a postgraduate fellowship from CONICET. Ignacio M. Soto is a member of Carrera del Investigador Científico (CONICET).

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Abbasi, A., Hossain, L., & Leydesdorff, L. (2012). Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics*, *6*, 403–412.

Alstott, J., Bullmore, E., & Plenz, D. (2014). powerlaw: A Python package for analysis of heavy-tailed distributions. *PLoS ONE*, *9*, e85777.

Andersen, H. K. (2011). Mechanisms, laws, and regularities. *Philosophy of Science*, *78*, 325–331.

Anderson, C. (2006). *The long tail: Why the future of business is selling less of more*. New York, NY: Hyperion.

Asmussen, S. (2003). *Applied probability and queues*. Berlin, Germany: Springer.

Aucoutier, J. J., & Pachet, F. (2007). A scale-free distribution of false positives for a large class of audio similarity measures. *Pattern Recognition*, *41*, 272–284.

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*, 509–512.

Bauke, H. (2007). Parameter estimation for power-law distributions by maximum likelihood methods. *The European Physical Journal B - Condensed Matter and Complex Systems*, *58*, 167–173.

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts.

Burroughs, S. M., & Tebbens, S. F. (2002). The upper-truncated power law applied to earthquake cumulative frequency-magnitude distributions: Evidence for a time-independent scaling paramenter. *Bulletin of the Seismological Society of America*, *92*, 2983–2993.

Cano, P., Celma, O., Koppenberger, M., & Buldu, J. M. (2006). Topology of music recommendation networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *16*, 013107.

Celma, O., & Cano, P. (2008). From hits to niches? Or how popular artists can bias music recommendation and discovery. In *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition* (p. 5). New York, NY: ACM.

Chung, K. H., & Cox, R. A. K. (1994). A stochastic model of superstardom: An application of the Yule distribution. *The Review of Economics and Statistics*, *76*, 771–7775.

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*, 661–703.

Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). New York, NY: Wiley.

Coulangeon, P., & Lemel, Y. (2007). Is distinction really outdated? Questioning the meaning of the omnivorization of musical taste in contemporary France. *Poetics*, *35*, 93–111.

Cremonesi, P., Koren, Y., & Turrin, R. (2010). Performance of recommender algorithms on top-n recommendation tasks. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (pp. 39–46). New York, NY: ACM.

Cross, I. (1999). Is music the most important thing we ever did? Music, development and evolution. In S. W. Yi (Ed.), *Music, mind, science* (pp. 10–39). Seoul, South Korea: Seoul University Press.

de Lima e Silva, D., Soares, M. M., Henriques, M. V. C., Alves, M. S., de Aguiar, S. G., de Carvalho, T. P., & Lucena, L. S. (2004). The complex network of the Brazilian Popular Music. *Physica A: Statistical Mechanics and its Applications*, *332*, 559–565.

DeNora, T. (2000). *Music in everyday life*. Cambridge, UK: Cambridge University Press.

Deluca, A., & Corral, A. (2013). Fitness and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophysica*, *61*, 1351–1394.

Foster, J. (2004). *The divine lawmaker: Lectures on induction, laws of nature, and the existence of God*. Oxford, UK: Clarendon.

Gan, L., Li, D., & Song, S. (2006). Is the Zipf law spurious in explaining city-size distributions? *Economics Letters*, *92*, 256–262.

Gillespie, C. S. (2015). Fitting heavy tailed distributions: The poweRlaw package. *Journal of Statistical Software*, *64*(2), 1–16.

Gleiser, P. M., & Danon, L. (2003). Community structure in jazz. *Advances in Complex Systems*, *6*, 565–573.

Goldstein, M. L., Morris, S. A., & Yen, G. G. (2004). Problems with fitting to the power-law distribution. *The European Physical Journal B - Condensed Matter and Complex Systems*, *41*, 255–258.

Greasley, A. E., & Lamont, A. M. (2006). Music preference in adulthood: Why do we like the music we do? In M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition* (pp. 960–966). Bologna, Italy: University of Bologna.

Greasley, A. E., Lamont, A. M., & Sloboda, J. (2013). Exploring musical preferences: An in-depth qualitative study of adults' liking for music in their personal collections. *Qualitative Research in Psychology*, *10*, 402–427.

Hargreaves, D. J. (1984). The effects of repetition on liking for music. *Journal of Research in Music Education*, *32*, 35–47.

Henning, V., & Reichelt, J. (2008). Mendeley – a last.fm for research? In *eScience'08. IEEE Fourth International Conference on eScience* (pp. 327–328). New York, NY: IEEE.

Hu, H. B., & Han, D. Y. (2008). Empirical analysis of individual popularity and activity on an online music service system. *Physica A: Statistical Mechanics and its Applications*, *387*, 5916–5921.

Jensen, J. J. (1998). *Self-organized criticality: Emergent complex behaviour in physical and biological systems*. Cambridge, UK: Cambridge University Press.

Johnston, R. R. (2015). The effect of repetition on preference ratings for select unfamiliar musical examples: Does preference transfer? *Psychology of Music*. Advance online publication. doi: 10.1177/0305735615577248

Jones, J. H., & Handcock, M. S. (2003). An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society of London B*, 270, 1123–1128.

Klaus, A., Yu, S., & Plenz, D. (2011). Statistical analyses support power law distributions found in neuronal avalanches. *PLoS ONE*, *6*, e19779.

Knopoff, L., & Kagan, Y. (1977). Analysis of the theory of extremes as applied to earthquake problems. *Journal of Geophysical Research*, *82*, 5647–5657.

Laiho, S. (2004). The psychological functions of music in adolescence. *Nordic Journal of Music Therapy*, *13*, 47–63.

Lambiotte, R., & Ausloos, M. (2005). Uncovering collective listening habits and music genres in bipartite networks. *Physical Review E*, *72*, 066107.

Lamont, A., & Greasley, A. (2009). Musical preferences. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 160–168). New York, NY: Oxford University Press.

Lamont, A., & Webb, R. (2010). Short- and long-term musical preferences: What makes a favourite piece of music. *Psychology of Music*, *38*, 222–241.

Levitin, D. J., Chordia, P., & Menon, V. (2012). Musical rhythm spectra from Bach to Joplin obey 1/*f* power law. *Proceedings of the National Academy of Sciences*, *109*, 3716–3720.

Litle, P., & Zuckerman, M. (1986). Sensation seeking and music preferences. *Personality and Individual Differences*, *7*, 575–578.

Machado, P., Romero, J., Manaris, B., Santos, A., & Cardoso, A. (2003). Power to the critics: A framework for the development of artificial critics. In C. Bento, A. Cardoso, & J. Gero (Eds.), *Proceedings of the IJCAI 2003 Workshop on Creative Systems* (pp. 55–64). Coimbra, Portugal: University of Coimbra.

Machado, P., Romero, J., Santos, M. L., Cardoso, A., & Manaris, B. (2004). Adaptive critics for evolutionary artists. In F. Rothlauf, J. Branke, S. Cagnoni, D. W. Corne, R. Drechsler, Y. Jin, & G. Squillero (Eds.), *Applications of Evolutionary Computing* (pp. 437–446). Berlin, Germany: Springer.

Malevergne, Y., Pisarenko, V., & Sornette, D. (2005). Empirical distributions of stock returns: Between the stretched exponential and the power law? *Quantitative Finance*, *5*, 379–401.

Manaris, B., Romero, J., & Machado, P. (2005). Zipf's law, music classification, and aesthetics. *Computer Music Journal*, *29*, 55–69.

Martin, P. (1995). *Sounds and society: Themes in the sociology of music*. Manchester, UK: Manchester University Press.

Mauch, M., MacCallum, R. M., Levy, M., & Leroi, A. M. (2015). The evolution of popular music: USA 1960–2010. *Royal Society Open Science*, *2*, 150081.

McNee, S., Riedl, J., & Konstan, J. (2006). Being accurate is not enough: How accuracy metrics have hurt recommender systems. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems* (pp. 1097–1101). New York, NY: ACM.

Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, *1*, 226–251.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*, 323–351.

North, A. C., & Hargreaves, D. J. (1996). The effects of music on responses to a dining area. *Journal of Environmental Psychology*, *16*, 55–64.

North, A. C., & Hargreaves, D. J. (1997). Experimental aesthetics and everyday music listening. In D. J. Hargreaves & A. C. North (Eds.), *The social psychology of music* (pp. 84–103). Oxford, UK: Oxford University Press.

Pareto, V. (1896). *Cours d'economie politique*. Geneva, Switzerland: Droz.

Peterson, R. A., & Kern, R. M. (1996). Changing highbrow taste: From snob to omnivore. *American Sociological Review*, *61*, 900–907.

Peterson, R. A., & Simkus, A. (1992). How musical taste groups mark occupational status groups. In M. Lamont & M. Fournier (Eds.), *Cultivating differences: Symbolic boundaries and the making of inequality* (pp. 152–186). Chicago, IL: University of Chicago Press.

Pinto, C. M. A., Lopes, A. M., & Tenreiro Machado, J. A. (2012). A review of power laws in real life phenomena. *Communications in Nonlinear Science and Numerical Simulation*, *17*, 3558–3578.

Price, H. E. (1986). A proposed glossary for use in affective response literature in music. *Journal of Research in Music Education*, *34*, 151–159.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: Examining the structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*, 1236–1256.

Ribeiro, M. T., Lacerda, A., Veloso, A., & Ziviani, N. (2012). Pareto-efficient hybridization for multi-objective recommender systems. In *Proceedings of the Sixth ACM Conference on Recommender Systems* (pp. 19–26). New York, NY: ACM.

Russell, P. A. (1986). Experimental aesthetics of popular music recordings: Pleasingness, familiarity and chart performance. *Psychology of Music*, *14*, 33–43.

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika*, *42*, 425–440.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*. New York, NY: Oxford University Press.

Sloboda, J., Lamont, A., & Greasley, A. (2011). Choosing to hear music: Motivation, process and effect. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 431–440). New York, NY: Oxford University Press.

Sornette, D. (2006). *Critical phenomena in natural sciences* (2nd ed.). Berlin, Germany: Springer.

Trepte, S. (2006). Social identity theory. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (pp. 255–271). Mahwah, NJ: Lawrence Erlbaum Associates.

Vargas, S., & Castells, P. (2011). Rank and relevance in novelty and diversity metrics for recommender systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems* (pp. 109–116). New York, NY: ACM.

Vázquez, A. (2003). Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Physical Review E*, *67*, 056104.

Virkar, Y., & Alstott, J. (2014). Power-law distributions in binned empirical data. *The Annals of Applied Statistics*, *8*, 89–119.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, *57*, 307–333.

Voss, R. F., & Clarke, J. (1975). 1/f noise in music and speech. *Nature*, *258*, 317–318.

Voss, R. F., & Clarke, J. (1978). 1/f noise in music: Music from 1/f noise. *Journal of the Acoustical Society of America*, *63*, 258–263.

Walker, E. L. (1980). *Psychological complexity and preference: A hedgehog theory of behavior*. Monterey, CA: Brooks/Cole.

White, E. P., Enquist, B. J., & Green, J. L. (2008). On estimating the exponent of power-law frequency distributions. *Ecology*, *89*, 905–912.

Wigner, E. P. (1964). Events, laws of nature, and invariance principles. *Science*, *145*, 995–999.

Yule, G. (1925). A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London (Series B)*, *213*, 21–87.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology: Monograph Supplement*, *9*, 1–27.

Zipf, G. (1932). *Selective studies and the principle of relative frequency in language*. Cambridge, MA: Harvard University Press.

Zipf, G. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.