

## 다운로드/스트리밍이 선호되는 디지털 음원 특성에 대한 고찰

What to Download and What to Stream?:

Investigating Music Characteristics Driving Preferences in Digital Music Consumption Modes

유재원 • Yoo, Jae Won, 형지현 • Hyeong, Ji Hyeon, 이재영 • Lee, Jae Young

디지털 음원 시장은 IT기술의 발달과 함께 빠른 성장을 거듭해왔으며, 특히 가장 대표적 디지털 소비 방식인 다운로드(download)와 스트리밍(streaming)은 음악 산업 전체 수입의 50%를 차지하게 되었다. 디지털 음원 시장을 양분하는 두 디지털 음원 소비 방식에 대해 보다 깊이 있게 이해하고자, 본 연구는 각 소비 방식이 선호되는 음원의 특징들을 규명하고자 한다. 유동(liquid) 및 고정(solid) 소비 이론에 착안하여, 본 연구는 인지도가 높아 애착도가 평균적으로 낮은 음원의 경우 스트리밍이, 인지도가 낮아 애착도가 평균적으로 높은 음원의 경우 다운로드가 보다 선호될 것으로 예상하고, 이를 실증한다. 실증 분석에는 2014년 발표된 주간 가온 뮤직 차트 데이터를 이용하였으며, 분석에는 각 음원이 가지는 이질성(heterogeneity)과 차트 데이터가 가지는 단절적(censored) 특성을 통제하기 위해 베이지언 토빗 모형(Bayesian Tobit model)을 이용하였다. 그 결과, 인지도가 높을 것으로 예상되는 경력이 긴 가수의 음원, 앨범 타이틀 음원, 그리고 음악 프로그램 방송 횟수가 많은 음원 등에 대해서는 스트리밍이 상대적으로 선호되는 것으로 드러났고, 이는 본 연구의 가설을 지지한다. 본 연구의 결과는 디지털 음원 소비 방식 간의 차이를 이해할 이론적 토대를 제공하며, 음원 판매 및 음악 추천 시스템의 효과적인 전략을 위한 실무적 시사점을 남긴다.

핵심주제어: 디지털 음원 시장, 다운로드, 스트리밍, 유동적 소비, 고정적 소비

본 연구는 2016년 대한민국 과학기술정보통신부와 한국연구재단의 지원을 받아 수행된 연구임(NRF- 2016-11-0350).

유재원 | 연세대학교 석사과정(jaewonu1@gmail.com)

형지현 | 연세대학교 석사과정(erecnis@yonsei.ac.kr)

이재영 | 연세대학교 경영대학 마케팅 교수(jaelee@yonsei.ac.kr), 교신저자

## ABSTRACT

With the help of technological advancement, the digital music industry has experienced a drastic growth, leading to download and streaming becoming popular modes of music consumption today. As these two main business portfolios account for 50% of the total revenue of the music industry, the importance of understanding them is growing. This research aims to investigate the song-specific features of digital music that drives the preference of one consumption mode over the other. Based on the theory of liquid and solid consumption, we hypothesize that songs with generally high attachment are consumed preferably through download due to its low awareness and songs with generally low attachment are consumed preferably through streaming due to its high awareness. We collected data from Gaon music chart, which is an aggregation of weekly music charts in Korea, for analysis. Our analysis of the data was run via the Bayesian Tobit model to control for its censored characteristic and song heterogeneity. Results supporting our hypothesis showed that streaming is more preferred when a song is of singers with a long career, a title track from an album or frequently introduced on TV programs. This study provides a theoretical background for understanding the difference between two digital music consumption modes and practical implications giving strategic viewpoints on music sales and the recommendation system.

**Keywords:** Digital Music Market, Downloads, Streaming, Liquid Consumption, Access-based Consumption

This research was funded by National Research Foundation of Korea (NRF) and the Ministry of Science, ICT & Future Planning (NRF- 2016-11-0350).

**Jae Won Yoo** | Graduate Student, School of Business, Yonsei University(jaewonu1@gmail.com)

**Ji Hyeon Hyeong** | Graduate Student, School of Business, Yonsei University(erecnis@yonsei.ac.kr)

**Jae Young Lee** | Assistant Professor, School of Business, Yonsei University(jaelee@yonsei.ac.kr), Corresponding Author

## I. 서론

IT 기술의 발달에 따라 CD, LP 등 전통적인 오프라인 음반 소비를 벗어나 다운로드(download), 스트리밍(streaming) 등 새로운 디지털 음원 소비 방식들이 등장하면서 음악 시장은 빠르게 변화해 왔다. 그 결과, 2016년 전 세계 음악 시장은 총 157억 달러의 시장 규모와 더불어 전년 대비 5.9% 성장률을 보임으로써 규모와 전망 측면에서 압도적인 산업임을 증명하였다. 특히 2016년 국제 음반 산업 협회(International Federation of the Phonographic Industry: IFPI)의 연간 보고서에 따르면, 디지털 음원 소비는 전 세계 음악 시장의 50%를 차지할 만큼 음악 시장의 주 수입원으로 성장하였다. 그 중에서도 다운로드와 스트리밍은 디지털 음원 시장을 양분하는 두 음원 소비 방식이다. 90년대 인터넷의 보급과 함께 널리 이용되어 온 다운로드는 일정 금액을 지불하고 정해진 기간 동안 제한된 개수의 곡을 자신의 전자기기에 저장하여 음원을 소비하는 방식이다(Bhattacharjee et al. 2006; Datta, Knox, and Bronnenberg 2017; Peitz and Waelbroeck 2005). 반면 스트리밍은 최근 무선 네트워크의 보급과 함께 도입된 소비 방식으로, 고정된 금액에 일정 기간 자유롭게 음원에 접근할 수 있는 서비스를 통해 음원을 소비하는 방식이다(Nguyen, Dejean, and Moreau 2012, 2014; Wlömert and Papies 2016). 2016년 기준 전 세계 전체 디지털 음원 수익의 45%를 다운로드 서비스가, 43%를 스트리밍 서비스가 차지하는 만큼(IFPI 2016), 두 디지털 음원 소비 방식은 공존하면서 음악 시장을 이끌어가고 있다.

디지털 음원 시장의 성장에 따라 그와 관련된 다양한 연구들 역시 진행되었으며, 이는 크게 두 가지 흐름으로 정리될 수 있다. 첫째로, 디지털 음원 소비 방식의 도입

이 전통적인 오프라인 음반 소비에 미치는 영향에 대한 연구가 존재한다(Graham et al. 2004; Papies, Eggers, and Wlömert 2011; Peitz and Waelbroeck 2004). 일례로, Peitz and Waelbroeck (2004)는 디지털 음원 소비 방식이 도입됨에 따라 오프라인 음반에 대한 해적 행위(piracy)가 증가하여 CD 판매량이 감소하였음을 보였으며, Graham et al. (2004)는 디지털 음원 소비 방식의 도입이 대형 음반 기획사들의 공급망을 전반적으로 변화시켰음을 밝혔다. 둘째로, 스트리밍이라는 새로운 디지털 음원 소비 방식이 전체 음원 소비 시장에 미친 영향에 대한 연구가 존재한다(Aguiar and Waldfogel 2015; Datta, Knox, and Bronnenberg 2017; Hughes and Lang 2003). 일례로, Hughes and Lang (2003)은 스트리밍이 소비자와 아티스트를 직접 연결시켜 음반 시장 내 대형 기업의 영향력을 감소시켰음을 밝혔으며, Aguiar and Waldfogel (2015)는 스트리밍이 다른 음원 소비 방식들의 사용을 잠식하게 될 것임을 보였다.

물론 상기 선행 연구들은 새로운 음원 소비 방식의 도입에 따른 음악 시장 구조의 변화를 조망한다는 면에서 그 의의를 가진다. 하지만 디지털 음원 소비 방식으로서 다운로드와 스트리밍이 공존하는 현재 시점에서는, 단순한 변화의 양태를 넘어 두 개별적인 음원 소비 방식이 어떻게 활용되고 있는지에 대한 보다 구체적인 이해가 요구된다. 이에 본 연구는 소비자들이 다운로드와 스트리밍을 이용하는 동기에 어떤 차이가 존재하는지 규명하고, 이러한 동기의 차이와 관련하여 다운로드 혹은 스트리밍이 보다 선호되는 음원의 특성을 밝히는 것을 그 중점 연구 문제로 한다. 이는 비단 공존하는 디지털 음원 소비 방식의 차이를 이해하는 이론적 틀을 제안할 뿐만 아니라, 음원 권리자 및 플랫폼의 수익 극대화 전략에도 시사점을 제시한다는 면에서 그 이론적·실무적 중요성을 가진다.

본 연구는 다운로드와 스트리밍의 차이를 밝히는 이론적인 틀로 고정(solid) 및 유동(liquid) 소비의 이론을 이용한다. Bardhi and Arnould (2017)는 물질적 소유 기반의 소비 방식을 고정적 소비로, 제품 및 서비스에 대해 소유하지 않고 일시적인 접근 권한을 얻는 소비 방식을 유동적 소비로 각각 정의하였다. 해당 개념에 따르면, 디지털 음원을 소유하는 가운데 소비하는 다운로드의 보다 고정적 소비로, 인터넷 망 접근을 통해 일시적으로 음원을 소비하는 스트리밍은 보다 유동적 소비로 구분이 가능하다. 일반적으로 소비자들은 본인들이 높은 애착을 가지는 대상에 대해서는 고정적 소비를 보다 선호하고, 그렇지 않은 대상에 대해서는 유동적 소비를 보다 선호하는 것으로 알려져 있다(Bardhi and Arnould 2017; Bardhi, Eckhardt, and Arnould 2012; Rindfleisch, Burroughs, and Wong 2008). 따라서 소비자들이 평균적으로 높은 애착을 보이는 음원에 대해서는 다운로드의 비중이, 낮은 애착을 보이는 음원에 대해서는 스트리밍의 비중이 각각 상대적으로 높을 것으로 예상된다. 본 연구는 인지도가 높은 대상에 대한 평균적인 애착심은 떨어진다는 기존 연구에 착안하여(Beatty, Homer, and Kahle 1988; Beatty and Smith 1987; Eisentraut 2012) 음원 및 해당 가수에 대한 높은 인지도는 애착도와 반비례할 것으로 예상하였고, 따라서 높은 인지도가 예상되는 음원들에 대해서는 스트리밍이 다운로드보다 선호될 것으로 예상하였다.

이를 실증하기 위하여, 본 연구는 2014년 가온 뮤직 차트(www.gaonchart.co.kr)의 주간 다운로드 및 스트리밍 상위 100 위에 오른 바 있는 총 1,442 곡의 다운로드 및 스트리밍 횟수를 분석하였다. 개별 음원이 가지는 이질성(heterogeneity)과 순위 데이터가 가진 절단성(censoring)을 고려하여 베이지언 토빗 모형(Bayesian Tobit model)을 추정하였고(Bolton, Kannan, and

Bramlett 2000; Schweidel, Fader, and Bradlow 2008; Wang et al. 2016), 그 결과 인지도가 높을 것으로 예상되는 경력이 긴 가수의 노래, 앨범의 타이틀곡, 그리고 음악 프로그램에 방송된 횟수가 많은 노래의 경우 다운로드의 비중이 상대적으로 낮고 스트리밍의 비중이 상대적으로 높은 것으로 드러났다. 위의 결과는 소비자들이 고정적인 다운로드와 유동적인 스트리밍을 이용함에 있어 그 동기에 차이가 있음을 방증하며, 이는 디지털 음원 시장을 이해하는 토대를 제공한다는 면에서 그 이론적 의의가 있다. 더 나아가, 해당 결과는 음원 권리자 및 음원 플랫폼이 개별 음원을 어떤 소비 방식을 통해 주로 판매해야 할지에 대한 전략적 방향성을 제안한다는 면에서 실무적 시사점을 가진다.

이어지는 장에서 본 연구의 이론적 토대를 살피고, 3장에서 본 연구의 가설을 수립한다. 4장과 5장에서는 실증 분석을 위해 진행한 데이터 가공, 모델 분석 및 결과에 대해 설명한다. 마지막으로 6장에서는 본 연구를 정리하며, 본 연구의 시사점 및 한계에 논의한다.

## II. 이론적 배경

본 연구의 중점 연구 문제에 대한 가설을 도출하기에 앞서, 본 장에서는 다운로드 및 스트리밍과 관련된 기존 연구를 우선적으로 살피고자 한다.

다운로드와 스트리밍은 디지털 음원의 대표적인 소비 방식으로, 크게 두 가지 맥락에서 대별된다. 첫째, 일반적으로 소비자들은 스트리밍을 통한 음원 소비보다 다운로드를 통한 음원 소비에 높은 기회비용을 지불한다(Bockstedt, Kauffman, and Riggins 2005; Datta, Knox, and Bronnenberg 2017; Maftei, Gerogiannis, and Papageorgiou 2016). 이는 두 방식의 구독

(subscription) 방법 및 과금의 차이에서 쉽게 확인할 수 있다. 다운로드 서비스의 경우, 사용자들은 고정된 금액을 지불하면 주어진 기간 동안 제한된 개수 내에서 음악을 다운로드 할 수 있다. 반면 스트리밍 서비스의 경우, 사용자들은 스트리밍 서비스를 구독하여 정해진 기간 동안 음원 플랫폼이 보유하고 있는 전체 음원에 대한 접근 권한을 얻는다(Nguyen, Dejean, and Moreau 2014; Wlömert and Papies 2016). 이와 같은 비즈니스 모델에서의 차이는 사용자가 느끼는 기회비용의 차이로 이어진다. 다운로드 사용자들은 서비스 이용을 위해 스트리밍 사용자보다 높은 금전적 비용을 지불하지만, 다운로드 가능한 곡의 개수가 제한되어 있기 때문에 상대적으로 높은 기회비용을 느낀다. 반면, 스트리밍 사용자들은 서비스 사용에 제약이 덜 하기 때문에 상대적으로 낮은 기회비용을 느낀다.

다음으로, 다운로드를 음원을 소유할 수 있는 권리를 판매하는 반면, 스트리밍은 음원에 접근할 수 있는 권리를 판매한다(Datta, Knox, and Bronnenberg 2017; Nguyen, Dejean, and Moreau 2014; Varian 2000). 다운로드를 디지털 음원 파일을 전자 기기에 저장하는 행위이며, 따라서 사용자들은 다운로드된 음원을 영구적으로 소유할 수 있다(Bhattacharjee et al. 2006). 반면, 스트리밍은 음악 플랫폼 서비스를 이용하는 동안의 일시적인 음원 소비 행위이며, 별도의 저장 서비스를 제공하지 않아 음원의 소유가 이루어지지 않는다.

이처럼 다운로드와 스트리밍은 크게 비용의 크기와 음원 소유 가능 여부에서 대조되며, 이러한 관계를 고정적 소비(Solid consumption)와 유동적 소비(Liquid consumption)의 개념을 통해 보다 체계적으로 설명할 수 있다(Bardhi and Eckhardt 2017; Bauman 2000). 우선 고정적 소비(Solid consumption)는 전통적인 재화의 소비 방식으로, 물질의 소유에 기반하여

지속적으로 이루어지는 소비를 의미한다. 반면, 유동적 소비는 최근 네트워크 연결 기술의 발달과 함께 등장한 새로운 개념으로, 물질적 형태에 의존하지 않고 접근을 기반으로 일시적으로 이루어지는 소비를 의미한다. 위의 정의에 비추어 보았을 때, 높은 기회비용을 지불하여 음원을 소유하며 소비하는 다운로드를 디지털 음원의 고정적 소비 방식으로, 낮은 기회비용을 지불하여 음원에 접근하며 소비하는 스트리밍은 유동적 소비 방식으로 분류될 수 있다.

한편, 선행 연구에 따르면 해당 제품에 대한 애착이 적고 사용의 효용만을 얻고자 할 때 소비자들은 접근 중심적인 유동적 소비를 택하며, 애착이 높을 때 소비자들은 소유 중심적인 고정적 소비를 통해 제품을 소유함으로써 제품과의 지속적인 연결 고리를 형성하거나 소비와 자신의 정체성을 일치시키고자 하는 것으로 알려져 있다(Bardhi and Eckhardt 2017; Dommer and Swaminathan 2012; Lawson et al. 2016; Schaeffers, Lawson, and Kukar-Kinney 2016). 따라서 다운로드를 통한 음원 소비는 보다 애착이 가는 음악에 대하여 선별적으로 이루어지는 반면, 스트리밍을 통한 음원 소비는 애착은 적지만 가볍게 듣고자 하는 음악에 대하여 선별적으로 이루어질 것으로 예상해 볼 수 있다.

### III. 가설 설정

앞선 논의에 따르면, 소비자의 음원 소비 방식은 개개인이 개별 음악에 대해 가지는 애착도에 따라 달라질 것으로 예상된다. 물론, 소비자 개개인의 선호를 파악하여 어떤 음원 방식을 선호하게 될지 예측하는 것은 그 나름의 의의가 있으나, 개개인의 개별 음악에 대한 애착도 데이터의 수집은 실무자 입장에서도 어려운 바, 그 실무



적 의의에도 한계가 있다. 따라서 본 연구는 보다 종합적인 층위(aggregate level)에서 일체감 및 애착에 대한 대리 변수(proxy)를 제안하고, 이에 따라 어떤 음악에 대해 다운로드 혹은 스트리밍을 통한 소비의 비중이 상대적으로 높은지 규명하는 것을 목적으로 한다. 이는 종합적 층위의 음원 소비 데이터를 이용해 소비자들의 음원 소비 양상을 분석한 선행 연구의 전통과도 부합한다(Aguiar and Waldfogel 2015; Bhattacharjee et al. 2007; Dewan and Ramaprasad 2014; Elberse 2010).

일반적으로, 음악에 대한 인지도(awareness)는 해당 음원을 소비하는 소비자들의 평균적인 애착도와 깊은 연관이 있을 것으로 기대된다. 이는 인지도가 높은 유명 브랜드의 고객보다 마이너 브랜드의 고객이 평균적으로 충성도가 높다는 기존 연구와 유사한 논의로 볼 수 있다(Besanko, Gupta, and Dube 2005; Dowling and Uncles 1997). 일반적으로, 인지도가 높은 유명 브랜드의 경우 제품 탐색에 소요되는 노력이 적은 반면 인지도가 낮은 마이너 브랜드의 경우 제품 탐색에 소요되는 노력이 높다(Beerli, Martin, and Quintana 2004; Heilman, Bowman, and Wright 2000). 브랜드에 대한 애착심과 관여도가 높은 소비자일수록 제품 탐색 과정에 더 많은 노력을 기울이며(Beatty, Homer, and Kahle 1988; Beatty and Smith 1987), 따라서 제품 탐색에 더 많은 노력이 요구되는 마이너 브랜드의 경우 제품 탐색에 적은 노력이 요구되는 유명 브랜드에 비해 고객 충성도가 일반적으로 높다.

위와 같은 자기 선택적(self-selective) 과정은 다운로드와 스트리밍에도 유사하게 적용될 것으로 예상된다. 인지도가 낮고 대중들에게 잘 알려지지 않은 음악의 경우, 해당 음악을 접하기 위해 비교적 많은 탐색의 노력이 필요한 만큼 해당 음원을 소비하는 소비자들은 음악

과의 높은 일체감과 애착을 가진 사람에 제한될 것으로 예상된다. 반면 인지도가 높고 대중들에게 잘 알려진 음악의 경우, 상대적으로 소비층의 스펙트럼이 넓어 위와 같은 적극적 선호층 뿐 아니라 해당 음악에 대한 가벼운 소비 의향을 지닌 사람들까지도 소비층으로 포함하고 있을 것으로 예상된다. 그 결과 인지도가 높은 음악은 인지도가 낮은 음악에 비해 일체감과 애착이 상대적으로 떨어지는 가벼운 소비층이 전체 소비층에서 차지하는 비중이 높을 것으로 예상된다. 따라서 전 장의 논의와 종합하여 볼 때 인지도가 상대적으로 높은 음원에 대해서는 스트리밍을 통한 소비의 비중이, 인지도가 상대적으로 낮은 음원에 대해서는 다운로드를 통한 소비의 비중이 높을 것으로 예상된다. 이에 따라 본 연구는 다음과 같은 핵심 가설을 제안한다.

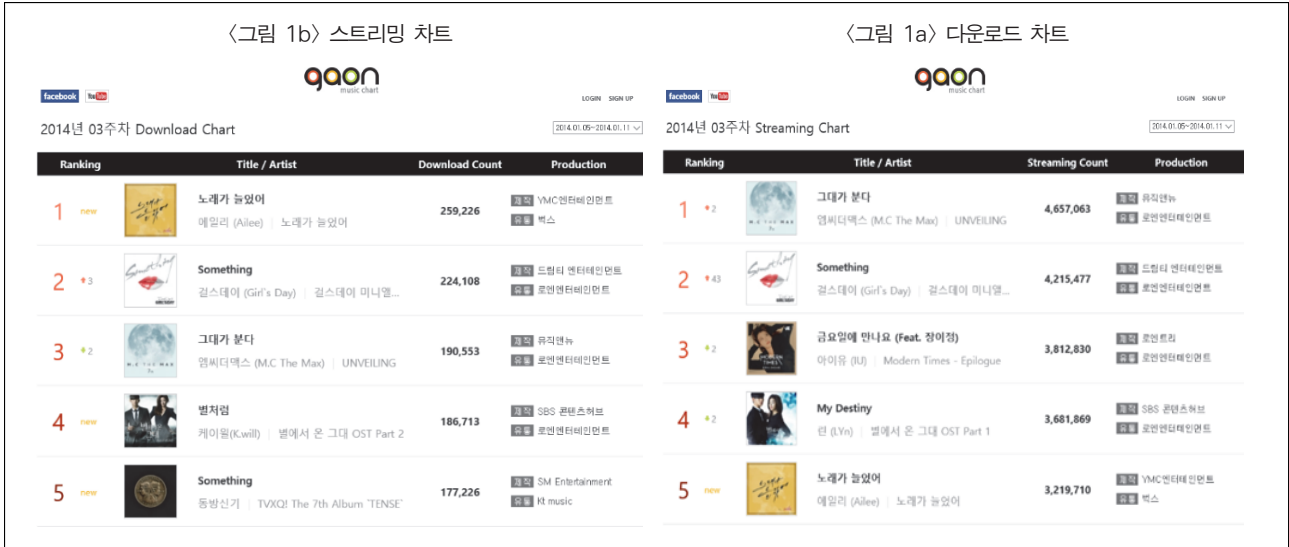
가설: 인지도가 높은 음악일수록 다운로드를 통한 소비의 비중이 상대적으로 낮다.

#### IV. 데이터 및 변수 소개

본 연구의 가설 검정을 위해 가온 뮤직 차트(www.gaonchart.co.kr)로부터 개별 곡들의 다운로드 수와 스트리밍 수에 대한 데이터를 주 단위로 수집하였다. 가온 차트는 한국 음악 콘텐츠 산업 협회(KMCIA)에서 운영하는 대한민국 공인 음악 차트로, 다음과 같은 두 가지 특징을 지닌다.

첫째로, 일반적인 음악 차트들이 iTunes나 Spotify와 같은 특정 음악 플랫폼에만 제한된 것과 달리 가온 차트는 국내 주요 디지털 음악 플랫폼의 음원 소비량을 망라하여 집계하므로 보다 포괄적이다. 둘째로, <그림 1>에서 확인할 수 있듯, 가온 차트는 다운로드와 스트리밍

〈그림 1〉 가온 차트 - 다운로드 vs. 스트리밍



차트를 분리하여 한 주 단위로 발표한다. 각 차트는 Melon, Bugs 등을 포함한 총 9개의 국내 주요 디지털 음악 플랫폼의 음원 소비량을 집계하여 상위 100곡의 가수-곡명을 총 소비량과 함께 발표한다. 이는 실제 소비량이 발표된다는 점에서 종합 순위만을 공개하는 빌보드 등의 타 종합 차트들과 대조된다. 소비자들의 다운로드-스트리밍 행동을 종합적 층위에서 분석하고자 하는 본 연구의 목적을 고려해 보았을 때, 가온 차트는 본 연구를 위한 데이터를 얻는 출처로서 적합하다고 볼 수 있다.

본 연구는 2014년에 발매된 곡들의 다운로드 수와 스트리밍 수를 주 별로 합산하여 2014년의 마지막 주까지 총 52주에 걸쳐 수집하였다. 이는 다운로드 차트에서 얻은 546명의 아티스트의 1,438 곡과 스트리밍 차트에서 얻은 377명의 아티스트의 777곡을 포함한다. 다운로드 차트의 곡은 평균적으로 3.3주 동안 차트 안에 머물렀고, 스트리밍 차트의 곡은 보다 길게 5.9주 동안 차트 안에 머물렀다. 두 차트를 종합하여 살펴보면, 52주 동안 총 1,442 곡이 다운로드 혹은 스트리밍 차트의 상위 100위 안에 들었다. 이 중 773곡은 모든 차트에 들었

고, 665곡은 다운로드 차트에만 들었으며, 4곡은 스트리밍 차트에만 들었다. 본 연구는 모형 분석을 위해 두 차트 중 적어도 하나의 차트이라도 등장했던 1,442곡의 음악에 집중하였다.

## 1. 다운로드 & 스트리밍 횟수

가온 차트의 데이터는 상위 100위 안에 들었던 곡에 대해서만 실제 음원 소비량을 제공한다. 따라서 음악이 상위 100위에서 벗어난 시점의 실제 음원 소비량은 관측이 불가능하다. 순위 데이터의 이러한 절단성(censoring)을 고려하여,  $t$  시점에 아티스트  $i$ 의 곡  $j$ 에 대하여 관측된(observed) 다운로드 수( $D_{ijt}$ )와 관측된 스트리밍 수( $S_{ijt}$ )를 다음과 같이 정의하였다.

$$D_{ijt} = D_{ijt}^* \text{ when } \bar{D}_t \leq D_{ijt}^*, \text{ \& } 0 \text{ otherwise. (1)}$$

$$S_{ijt} = S_{ijt}^* \text{ when } \bar{S}_t \leq S_{ijt}^*, \text{ \& } 0 \text{ otherwise.}$$

위의 식에서,  $D_{ijt}^*$  ( $S_{ijt}^*$ )는  $t$  시점에 아티스트  $i$ 의 곡  $j$

가 실제 다운로드(스트리밍) 된 횟수를 나타내며, 이를 실질(true) 다운로드(스트리밍) 수로 지칭한다. 앞서 설명한 바와 같이, 실질 소비량은 주간 순위에서 상위 100위 안에 들었던 곡에 대해서만 관측 가능하며, 이에 해당하지 않는 곡의 경우 관측이 불가능하다.  $\bar{D}_t(\bar{S}_t)$ 는 다운로드(스트리밍) 차트 상위 100위 안에 진입하기 위해  $t$ 시점에 요구되는 최소 다운로드(스트리밍) 수, 즉 해당 주간 차트에서 100위에 오른 곡의 관측 다운로드(스트리밍) 수로 정의하였다. 마지막으로, 관측 다운로드( $D_{ijt}$ ) 및 스트리밍( $S_{ijt}$ ) 횟수는 데이터의 특성을 고려하여 해당 곡이 상위 100위 안에 든 경우에만 실질 음원 소비량과 동일하며, 그렇지 않을 경우에는 0의 값을 갖도록 정의되었다.

결과적으로, 총 1,442곡에 대하여 관측 다운로드 및 스트리밍 횟수가 정의되었으며, 이 중 해당 곡이 최소 하나의 차트의 Top 100위에 진입해 있는 유의미한 관측치는 총 5,278쌍에 이른다. 이는 다시 세 가지의 범주로 나뉘는데, 1) 관측 다운로드 수가 0이 아니고 관측 스트리밍 수가 0이 아닌 데이터 3,629개, 2) 관측 다운로드 수가 0이 아니고 관측 스트리밍 수가 0인 데이터 930개, 3) 관측 다운로드 수가 0이고 관측 스트리밍 수가 0이 아닌 데이터 719개로 구성된다. 이를 시점  $t$ 에 아티스트  $i$ 의 곡  $j$ 가 속한 범주를 나타내는 지지 변수(indicator)를 통하여 다음과 같이 나타내었다.

$$y_{1ijt} = 1 \text{ when } D_{ijt} > 0 \ \& \ S_{ijt} > 0, \ \& \ 0 \text{ otherwise,} \quad (2)$$

$$y_{2ijt} = 1 \text{ when } D_{ijt} > 0 \ \& \ S_{ijt} = 0, \ \& \ 0 \text{ otherwise,}$$

$$y_{3ijt} = 1 \text{ when } D_{ijt} = 0 \ \& \ S_{ijt} > 0, \ \& \ 0 \text{ otherwise,}$$

where  $y_{1ijt} + y_{2ijt} + y_{3ijt} = 1$  for all the observations.

총 5,278 관측치에 대한 관측된 음원 소비량의 분포는 모두 여타 순위 관련 데이터와 유사하게 우편향(right-skewed)의 특성을 갖는다(Hendricks and Sorensen 2009). 또한, 다운로드 소비량은 평균 44,927, 표준편차 45,574, 스트리밍 소비량은 평균 1,418,142, 표준편차 980,704로 스트리밍의 소비량이 다운로드의 소비량을 상회하고 있는 것으로 나타났다.

## 2. 다운로드 / 스트리밍 비율

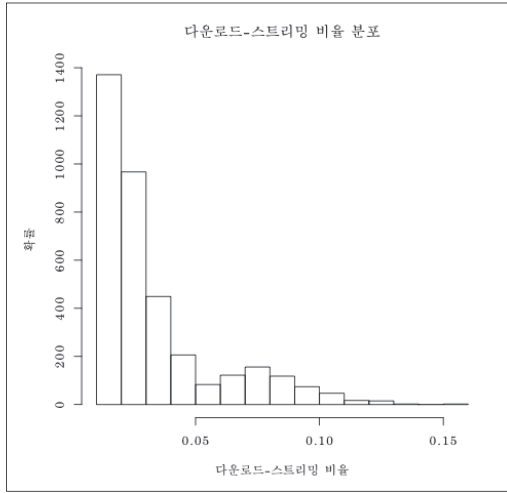
본 연구는 다운로드와 스트리밍 간 상대적 선호의 차이를 이해하기 위해 다운로드-스트리밍 간의 소비 비율을 분석하고자 한다. 다운로드-스트리밍 간의 소비 비율은 다운로드 횟수를 스트리밍 횟수로 나누어 준 값으로, 해당 비율은 그 값이 클수록 다운로드가 상대적으로 선호됨을 의미하고, 작을수록 스트리밍이 상대적으로 선호됨을 의미한다. 단, 앞서 설명한 데이터의 절단성을 반영하여, 다운로드-스트리밍 간의 소비 비율을 관측 소비 비율( $R_{ijt} = D_{ijt} / S_{ijt}$ )과 실질 소비 비율( $R_{ijt}^* = D_{ijt}^* / S_{ijt}^*$ )로 나누어 정의하고자 한다. 여기에서 관측 소비 비율은 '관측된 다운로드, 스트리밍 횟수를 바탕으로 계산된 다운로드-스트리밍 간의 소비 비율'을, 실질 소비 비율은 '관측될 수도, 관측되지 않을 수도 있는 실제 다운로드, 스트리밍 횟수를 바탕으로 한 실제 다운로드-스트리밍 간의 소비 비율'을 의미한다.

우선, 관측된 다운로드 수와 관측된 스트리밍 수 모두 0이 아닌 3,629개의 관측치에 대해서는( $y_{1ijt} = 1$ ) 실질 다운로드-스트리밍 비율( $R_{ijt}^*$ )의 정확한 값을 계산할 수 있으며, 이 때의 실질 다운로드-스트리밍 비율은 관측된 다운로드 수를 관측된 스트리밍 수로 나눈 값과 동일하다( $R_{ijt}^* = R_{ijt}$ ). 해당 3,629개의 관측치( $y_{1ijt} = 1$ )의 다운로드-스트리밍 비율 분포는 <그림 2>와 같으며, <그림

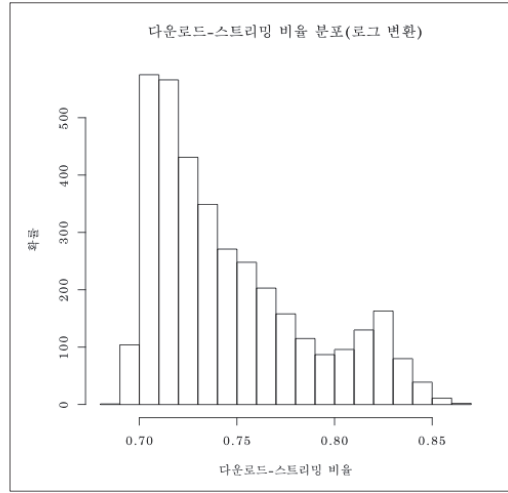


〈그림 2〉 다운로드-스트리밍 비율의 분포

(a) 다운로드-스트리밍 비율 분포



(b) 다운로드-스트리밍 비율 분포 (로그 변환)



2a)는 로그 변환을 취하지 않은 분포이고 〈그림 2b)는 편향성을 완화하여 보다 안정된 값을 얻기 위해 로그 변환을 취한 분포이다. 〈그림 2a)의 분포는 강한 우편향의 분포적 특성을 보이며, 전자는 평균 0.03, 표준편차 0.02로, 후자는 평균 0.75, 표준편차 0.04로 요약된다.

반면, 데이터의 절단성으로 인해 둘 중 하나의 소비량이라도 0의 값을 갖는 관측치에 대해서는 실질 다운로드-스트리밍 비율( $R_{ijt}^*$ )의 정확한 값을 구할 수 없다.<sup>1)</sup> 단, 이와 같은 경우에는 차트 진입에 요구되는 최소 소비량에 대한 정보를 이용하여 실질 다운로드-스트리밍 비율의 범위(range)를 도출할 수 있다. 관측된 다운로드 수가 0이 아니고 관측된 스트리밍 수가 0인 930개의 관측치에 대한( $y_{2ijt} = 1$ ) 실질 스트리밍 수는 최소 스트리밍 수보다 작은 값을 가진다( $S_{ijt}^* < \bar{S}_t$ ). 따라서 실질 다운로드-스트리밍 비율은 관측된 다운로드 수를 최소 스트리밍 수로 나눈 값보다 크다( $R_{ijt}^* > D_{ijt} / \bar{S}_t$ ). 마찬가지로,

관측된 다운로드 수가 0이고 관측된 스트리밍 수가 0이 아닌 719개의 관측치에 대한( $y_{3ijt} = 1$ ) 실질 다운로드 수는 최소 다운로드 수보다 작은 값을 가진다( $D_{ijt}^* < \bar{D}_t$ ). 따라서 이 경우 실질 다운로드-스트리밍 비율은 최소 다운로드 수를 관측된 스트리밍 수로 나눈 값보다 작다( $R_{ijt}^* < \bar{D}_t / S_{ijt}$ ).<sup>2)</sup> 본 연구는 실질 다운로드-스트리밍 비율이 정확하게 관측되는 3,629 개의 관측치 뿐 아니라 실질 다운로드-스트리밍 비율의 범위 정보가 포함된 1,649개의 관측치를 모두 활용한 모형 분석으로 연구 가설을 검증하고자 하며, 해당 모형은 모형 분석에서 보다 자세하게 소개하고자 한다.

### 3. 독립 변수

본 연구는 인지도가 높은 음악과 낮은 음악 간에 음원 소비 방식의 차이가 있을 것이라 예상하였으며, 음악 인

1. 같은 이유에서 해당 관측치에 대한 분포는 도식할 수 없다.

2. 다운로드 수와 스트리밍 수 모두 관측할 수 없었던 관측치에 대해서는 실질 다운로드-스트리밍 비율에 대한 어떠한 정보도 추론할 수 없다. 따라서 실증 분석에 사용된 데이터는 두 소비 방식 중 적어도 하나의 소비량이 0이 아닌 5,278개의 관측치만을 포함하고 있다.

지도에 대한 관측 가능한 세 가지 대리 변수(proxy)로 크게 아티스트의 경력( $YSD_{ij}$ ), 타이틀 곡( $title_{ij}$ ), 음악 프로그램 방송 횟수( $show_{ijt}$ )를 제안한다.

우선, 아티스트 경력은 음원의 인지도와 밀접하게 연관된다. 일반적으로 아티스트가 인지도가 떨어지고 대중으로부터 호응을 이끌어내지 못하는 경우 커리어가 단명하게 된다(Blume, and Easley 2002). 이와 같은 선별 과정(selection process)을 거치며 장기간 커리어를 이어 온 아티스트와 그의 음악은 대중들 사이에서 높은 인지도를 보유한 경우가 많을 것이다. 또한 커리어가 안정된 아티스트의 경우, 그 흥행성으로 인해 보다 집중적인 마케팅 활동이 이루어지며(Kohli and Jaworski 1990), 이는 역시 음악에 대한 인지도의 제고로 이어질 것으로 보인다. 종합하면, 경력이 오래된 가수의 음원은 인지도가 높을 것으로 기대되며, 따라서 다운로드를 통한 소비의 비중이 상대적으로 낮을 것으로 추측할 수 있다. 아티스트의 경력은 가수가 데뷔 이후 활동하고 있는 연차( $years-since-debut$ )로 정의하였다. 보다 구체적으로, 아티스트의 경력을 측정할  $YSD_{ij}$ 는 곡  $j$ 를 녹음한 아티스트  $i$ 가 데뷔 후 몇 년이 지났는지를 나타내며, 2015에서 아티스트의 데뷔 연도를 뺀 값으로 정의되었다. 즉, 2014년에 데뷔한 아티스트는 1의 값을 가지며, 보유 데이터 중 가장 활동 기간이 긴 1973년에 데뷔한 아티스트는 42의 값을 가진다. 다수의 아티스트가 곡을 함께 녹음하였을 경우, 메인 아티스트의 데뷔 연도에 기반하여 계산하였다. 단, 본 데이터는 2014년에 발매된 노래의 해당 연도 소비량을 연구하므로,  $YSD_{ij}$ 는 관측 기간 동안 불변한다.

다음으로, 타이틀곡은 한 앨범을 대표하여 동일 앨범 내의 다른 곡들에 비해 보다 집중적으로 홍보가 이루어지는 곡이다. 실제로 음반 홍보비의 80%가 타이틀곡을 홍보하는 데 지출되며(오정일, 박민수 2006), 따라서

타이틀곡에 대한 인지도는 비 타이틀곡보다 높을 것으로 예상된다. 종합하면, 타이틀곡은 상대적으로 인지도가 높아 다운로드를 통한 소비의 비중이 상대적으로 낮을 것으로 예상된다. 타이틀곡을 나타내는  $title_{ij}$ 은 아티스트  $i$ 의 곡  $j$ 의 타이틀곡 해당 여부에 대한 지시 변수로, 타이틀곡일 경우 1의 값을 갖고 그렇지 않을 경우 0의 값을 갖는다. 본 연구는 해당 곡이 수록된 앨범에 타이틀곡이라고 명시되었을 경우 이를 타이틀곡으로 간주하였으며, 하나의 곡만 수록된 미니 싱글의 경우에도 동일한 기준을 적용하였다.

마지막으로, 음악 프로그램 출연은 아티스트들이 신곡을 홍보하는 가장 일반적인 경로이며, 매스 미디어를 통한 노출은 자연스럽게 해당 음원의 인지도를 상승시킬 것으로 예상된다. 실제 선행 연구에 따르면 음악 프로그램에서 방송되는 것은 음악의 흥행 결정요인이자(김양석, 이상혁 2015) 음원 소비자의 스펙트럼을 넓히는 요인으로(Strobl and Tucker 2000) 드러났다. 따라서 음악 프로그램에 방송되는 것은 해당 음악의 인지도를 높여 상대적으로 낮은 다운로드 소비의 비중과 관련될 것으로 예상된다. 이에 대한 측정 변수인  $show_{ijt}$ 는 시점  $t$ 에 아티스트  $i$ 의 곡  $j$ 가 방송된 음악 프로그램의 개수를 나타낸다. 해당 변수는 세 개의 국내 지상파 방송사의 음악 프로그램과(i.e., KBS 뮤직뱅크, SBS 인기가요, MBC 음악캠프) 한 개의 국내 종합편성 방송사의 음악 프로그램을(i.e., Mnet M! Countdown) 대상으로 수집된 2014년 출연진 데이터에 기반하였다. 각 프로그램은 매주 한 번 방송되며, 방송되는 시간이 서로 겹치지 않는다. 또한 특정 스타일의 음악에만 국한되지 않고 다양한 출연진을 섭외한다. 만약 시점  $t$ 에 아티스트  $i$ 의 곡  $j$ 가 어떠한 프로그램에도 방송되지 않았을 경우 해당 변수는 0의 값을 갖고, 모든 프로그램에 방송되었을 경우 4의 값을 갖는다.

#### 4. 통제 변수

이 외에도 기존의 연구들은 거대 음반 기획사와 같은 아티스트 혹은 곡에 관련된 속성들이 음원 판매에 미치는 영향을 입증해왔으며(Hendricks and Sorensen 2009; Dewan and Ramaprasad 2014), 본 연구는 이와 같은 혼동 요인들을 통제할 목적으로 음원 발매 후 기간, 거대 음반 기획사 여부, 드라마 OST 여부 등의 변수들을 추가적으로 수집하였다.

우선 음원 발매 후 기간은 음원의 소비 및 소비 방식의 편향성과 밀접한 연관이 있다. Beck (2007)는 음반이나 영화와 같은 창작물들의 판매량은 대체로 발매 직후 정점을 찍고 그 후 빠르게 감소한다고 하였다. 뿐만 아니라 본 연구의 가온 차트 데이터가 다운로드 횟수와 스트리밍 횟수를 집계한 방식을 살펴보면, 다운로드 횟수는 다운로드 이후 소비자가 해당 곡을 반복 재생한 횟수에 대한 정보를 포함하고 있지 않지만, 스트리밍 횟수는 소비자가 해당 곡을 재생할 때마다 이를 집계한다. 따라서 시간이 지남에 따라 스트리밍 횟수가 다운로드 횟수에 비해 필연적으로 많아지므로, 가설 검정을 위해 해당 효과에 대한 통제는 필수적이다. 따라서 본 연구는 아티스트  $i$ 의 곡  $j$ 의 발매 후 경과 주( $WSR_{jit}$ , Weeks Since

Release) 변수를 수집하여 통제하였으며, 해당 변수는 곡이 처음 발매된 주에는 1로 표시되고 그 이후 매주 1씩 증가한다.

다음으로 주요 음반 기획사 역시 음원의 소비 및 그 편향성에 잠재적 영향을 미칠 수 있다. 국내의 대략 160개의 음반 기획사 중 코스닥(KOSDAQ, Korea Securities Dealers Automated Quotation)에 상장된 12개의 주요 기획사들이 전체 국내 음악 시장의 3분의 2 가 넘는 부분을 차지하고 있다(와이지 엔터테인먼트, 2015). 이러한 거대 기획사들은 보다 강력한 시장 지배력을 행사하는 경향이 있으므로 음원 판매에 영향을 주게 된다. 실제로 Asai (2011)에 따르면, 주요 레코드사는 큰 규모의 홍보 활동을 진행할 수 있기 때문에 주요 레코드사와 음반 판매 사이에 유의미한 관계가 존재한다. 따라서 본 연구는 아티스트  $i$ 의 곡  $j$ 가 상장 기획사에서 제작되었는지 여부( $kosdaq_{ij}$ )를 수집하여 통제하였다.

마지막으로 드라마 OST 여부 역시 음원 판매에 유의미한 영향을 미칠 수 있다. 드라마에 배경음악으로 삽입되는 음악은 대중으로부터 크게 선호되므로(Mussulman 1974), 드라마 OST로 활용된 음악은 그렇지 않은 음악들과 구분되는 특성을 지니므로써 음원 판매에 영향을 미칠 수 있다. 따라서 아티스트  $i$ 의 곡  $j$ 가 드라마 OST

〈표 1〉 변수의 기술 통계량

	Mean	SD	Correlation						
			1	2	3	4	5	6	7
Download <sup>a</sup>	9.01	3.63							
Streaming <sup>a</sup>	11.52	5.35	-.08						
Weeks Since Release <sup>a</sup>	1.27	1.01	-.45	.50					
Years Since Debut <sup>a</sup>	1.77	.89	.02	-.02	-.05				
Show	.75	.43	-.06	.20	.22	-.11			
Title Song	.52	1.10	.16	-.11	-.26	-.06	-.22		
KOSDAQ Company	.39	.49	.02	.14	.10	-.29	-.17	.21	
OST	.05	.22	-.01	.01	.02	-.25	.03	-.11	.10

<sup>a</sup> 분포 편향성으로 인해 로그 변환함.

로 활용되었는지의 여부( $ost_{ij}$ )를 수집하여 통제하였다.

〈표 1〉은 5,278개의 관측치 가운데 앞서 소개한 주요 변수들과 통제 변수들의 기술 통계량을 요약한다. 강한 분포 편향성을 보이는 다운로드-스트리밍 비율( $R_{ijt}^*$ ), 아티스트의 경력( $YSD_{ij}$ ), 발매 후 경과 주( $WSR_{ijt}$ ) 등의 변수는 실제 분석에서 로그 변환하여 분석하였으며, 모든 연속(continuous) 변수들은 평균 중심화(mean-centered)하여 추정치의 해석력을 높였다.

## V. 모형 분석

### 1. 분석 모형

앞서 설명한 바와 같이, 본 연구는 실질 다운로드-스트리밍 비율의 관측치 뿐 아니라, 실질 다운로드-스트리밍 비율의 범위에 대한 정보까지 활용하여 모형을 추정하고자 가설을 검정하고자 한다. 이를 위하여, 우선 실질 다운로드-스트리밍 비율을 아래와 같이 정의한다.

$$\ln(R_{ijt}^*) = \beta_{0j} + \beta_1 \ln(YSD_{ij}) + \beta_2 title_{ij} + \beta_3 show_{ijt} + X_{ijt} \beta_4 + \varepsilon_{ijt}. \quad (4)$$

기본적인 다운로드-스트리밍 비율은  $\beta_{0j}$ 에 반영된다. 단, 음원 간에 상존할 수 있는 소비율의 이질성을 반영하기 위해 기본 다운로드-스트리밍 비율을 음원 별로 다르게 정의하였고, 특히  $\beta_{0j}$ 는  $N(\beta_0, \sigma_0^2)$ 의 정규분포를 따른다고 가정한다. 다음으로  $\beta_1$ 은 아티스트 경력의 퍼센티지 변화가 다운로드-스트리밍 비율에 미치는 퍼센티

지 효과를 나타낸다.<sup>3)</sup>  $\beta_1 < 0$ 일 때 아티스트의 경력이 늘어날수록 다운로드-스트리밍 비율이 감소하며, 이는 경력이 긴 아티스트의 음악보다 경력이 짧은 아티스트의 음악에 다운로드가 상대적으로 선호됨을 의미한다.  $\beta_2$ 는 타이틀곡 여부가 다운로드-스트리밍 비율에 미치는 효과를 보여주며,  $\beta_2 < 0$ 인 경우 타이틀곡보다 비 타이틀곡에서 다운로드가 상대적으로 선호되는 것으로 볼 수 있다. 마지막  $\beta_3$ 는 음악 프로그램 방송 횟수가 다운로드-스트리밍 비율에 미치는 효과를 나타내며,  $\beta_3 < 0$ 인 경우 음악 프로그램 방송 횟수가 많은 음악보다 적은 음악에 다운로드가 상대적으로 선호되는 것으로 볼 수 있다.

앞서 언급된 주요 변수 외에도 〈식 4〉는 통제 변수들에 대한 벡터인  $X_{ijt}$ 를 포함하며,  $\beta_4$ 는 통제 변수들에 대한 추정치를 갖는 벡터이다.  $X_{ijt}$ 는  $t$  시점에 측정된 아티스트  $i$ 의 곡  $j$ 의 발매 후 경과 주를 의미하는 변수인  $WSR_{ijt}$ , 아티스트  $i$ 의 곡  $j$ 가 상장 기획사에서 제작되었는지 여부를 의미하는 지시 변수인  $kosdaq_{ij}$ , 아티스트  $i$ 의 곡  $j$ 가 드라마의 OST로 이용되었는지 여부를 의미하는 지시 변수인  $ost_{ij}$ , 그리고 시간에 따른 다운로드-스트리밍 비율의 체계적인 변화를 통제하기 위한 고정 효과(fixed effect)를 포함한다.  $\varepsilon_{ijt}$ 는 에러항으로  $N(0, \sigma_\varepsilon^2)$ 의 정규분포를 따른다고 가정한다.

단, 데이터의 절단성으로 인해 실질 다운로드-스트리밍 비율은  $y_{1ijt} = 1$ 인 관측치에 대해서만 정확하게 관측되며,  $y_{2ijt} = 1$ 인 관측치에 대해서는  $R_{ijt}^*$ 의 하한선( $D_{ijt}/\bar{S}_t$ )을,  $y_{3ijt} = 1$ 인 관측치에 대해서는  $R_{ijt}^*$ 의 상한선( $\bar{D}_t/S_{ijt}$ )을 구하게 된다. 위의 절단성을 아티스트  $i$ 의 곡  $j$ 의  $t$ 시점의 실질 다운로드-스트리밍 비율에 대한 우도(likelihood)로 표현하면 아래와 같다. 단, 이 때

3. 이는 양 변수가 모두 로그 변환되었기 때문이며, 앞서 설명한 바와 같이 관측 다운로드-스트리밍 비율( $R_{ijt}$ )과 아티스트의 경력( $ysd_{ij}$ )의 강한 우편향성으로 인해 로그 변환 후 분석을 진행하였다.

$\Pr(\cdot)$ 는 어떤 사건이 발생할 확률을 의미한다.

$$L_{ijt} = \Pr(D_{ijt}/S_{ijt} = R_{ijt}^*)^{y_{1ijt}} \times \Pr(D_{ijt}/\bar{S}_t < R_{ijt}^*)^{y_{2ijt}} \\ \times \Pr(\bar{D}_t/S_{ijt} > R_{ijt}^*)^{y_{3ijt}}, \quad (5)$$

마지막으로 <식 5>를 구성하는 각 확률의 양 변에 로 그 변환을 취해줌으로써 <식 4>를 <식 5>에 대입할 수 있고, 그 결과 쌍방향으로 절단된 Tobit 모형을 도출할 수 있다(Bolton, Kannan, and Bramlett 2000; Schweidel, Fader, and Bradlow 2008; Wang et al. 2016). 보다 구체적으로,  $y_{1ijt} = 1$ 인 관측치는 일반적인 선형 회귀 모형이,  $y_{2ijt} = 1$ 인 관측치는  $R_{ijt}^*$ 가 하한선  $D_{ijt}/\bar{S}_t$  보다 클 확률을 구하는 probit 모형이,  $y_{3ijt} = 1$ 인 관측치는  $R_{ijt}^*$ 가 상한선  $\bar{D}_t/S_{ijt}$  보다 작을 확률을 구하는 probit 모형이 추정한다.

음원 별로 존재하는 이질성(heterogeneity)과 관측치의 절단성을 보다 용이하게 추정하기 위해 추정에는 Hierarchical Bayesian Markov Chain Monte Carlo(MCMC)방식을 사용한다(Jedidi, Jagpal, and Manchanda 2003; Manchanda, Ansari, and Gupta 1999; Rossi, McCulloch, and Allenby 1996). 추

정을 위해 15,000번 반복되는 깁스(Gibbs) 샘플링을 시행하였고, 매 시행에 대한 시계열 자료를 추적하여 수렴값을 얻었다. 본 분석의 결과는 초기 5,000번의 시행을 제외한 10,000번의 시행에 기반하며, 추정과 관련된 세부사항은 부록(Appendix)을 참고하도록 한다.

## 2. 분석 결과

추정 결과는 아래의 <표 2>에 정리되어 있다. 모델의 적합도와 복잡도를 평가하기 위해 deviance information criterion(DIC)을 제시하였으며(Montoya, Netzer, and Jedidi 2010; Sonnier, Ainslie, and Otter 2007; Spiegelhalter et al. 2002), 이는 더 나은 모델일수록 작은 값을 가진다. 또한 일반적인 Bayesian 분석과 같이 사후 평균과 95% 신뢰 구간을 제시하여 추정치의 사후 분포를 설명하였다(Manchanda, Ansari, and Gupta 1999; Rossi, McCulloch, and Allenby 1996).

세 가지 핵심 발견은 다음과 같다. 첫째로, 경력이 긴 아티스트의 곡일수록 다운로드-스트리밍 소비 비율이 낮다. 다시 말해, 디지털 음원 소비는 경력이 짧아 아직 자

<표 2> 모델 추정 결과

모수	추정치	95% 신뢰도
절편: $\beta_0$	-0.02	(-.22, .23)
In(YSD): $\beta_1$	-.12**	(-.19, -.05)
Title: $\beta_2$	-.46**	(-.57, -.35)
Show: $\beta_3$	-.02**	(-.02, -.01)
In(WSR)	-.94**	(-.95, -.92)
Kosdaq	-.39**	(-.53, -.28)
OST	-.33**	(-.62, -.10)
절편의 표준편차: $\sigma_0$	1.05**	(1.00, 1.10)
에러항의 표준편차: $\sigma_e$	.16**	(.15, .16)
DIC	-3917.062	

참고: 본 모델은 51 주의 고정 효과를 포함한다. 더미 변수에 대한 추정치는 설명의 편의 상 보고되지 않았으나, 요청에 따라 제공이 가능하다.  
\*는 95% 신뢰수준에서의 유의성을 나타내며, 이에 대응하는 신뢰구간을 괄호 안에 표현하였다. \*\*는 99% 수준에서의 유의성을 나타낸다.



리를 잡지 못한 아티스트의 곡보다 경력이 길어 보다 안정적인 위치에 있는 아티스트의 곡이 상대적으로 다운로드의 비중이 낮은 구조를 갖는다. 본 연구의 모델 추정치에 따르면, 데뷔 횟수가 10% 길어질 때 다운로드-스트리밍 비율은  $1.1\% (= 1 - (1.1)^{-0.12})$  감소한다. 본 연구는 아티스트의 경력을 음악 인지도의 대리 변수로 보고 안정적인 위치에 있는 아티스트가 지닌 더 높은 인지도로 인해 곡에 대한 애착심이 적은 소비가 촉발될 것이라 예상하였으며, 따라서 해당 결과는 본 연구의 가설을 지지한다.

둘째로, 타이틀곡일수록 다운로드-스트리밍 소비 비율이 낮다. 모델 분석으로부터 얻은 추정치에 따르면, 타이틀곡일 때 다운로드-스트리밍 비율은  $36.9\% (= 1 - e^{-0.46})$  감소한다. 타이틀곡은 비 타이틀곡에 비해 집중적인 마케팅 활동으로 형성된 높은 인지도를 특징으로 하기 때문에, 이 결과는 인지도가 높을수록 다운로드 비중이 낮아질 것이라는 본 연구의 주장과 일치한다.

셋째로, 음악 프로그램에서 방송된 횟수가 많을수록 다운로드-스트리밍 간의 소비 비율이 낮다. 본 연구의 모델이 추정한 바에 따르면, 음악 프로그램에서 방송된 횟수가 한 번 늘어날 때 다운로드-스트리밍 비율은  $1.7\% (= 1 - e^{-0.02})$  감소한다. 본 연구는 음악 프로그램 방송 횟수를 음악 인지도에 대한 대리 변수로 상정하여, 이는 해당 음원에 대한 인지도를 높여 소비층을 넓히기 때문에 곡에 대한 애착심이 높지 않은 소비자들을 유입시켜 고정적 소비인 다운로드의 비중을 감소시킬 것이라 예상하였으며, 따라서 해당 결과는 본 연구의 주장과 일치한다.

이 외에도 통제 변수들에 대한 분석 결과는 본 연구가 제안한 메커니즘의 타당성을 제고한다. 보다 구체적으로, 발매 후 오랜 기간이 지난 곡, 주요 음반 기획사의 곡, 그리고 드라마의 OST였던 곡은 상대적으로 인지도

가 높다는 공통점을 지니며, <표 2>에 따르면 해당 노래들은 그렇지 않은 노래들에 비해 다운로드를 통한 소비가 덜 선호되는 것으로 드러났다. 본 연구는 인지도가 높은 음악일수록 높은 애착심을 수반하는 소비가 전체 소비에서 차지하는 비중이 감소하기 때문에 고정적 소비인 다운로드의 비중이 감소한다는 메커니즘을 제안하였으며, 통제 변수들에 대한 이와 같은 분석 결과는 본 연구가 제안한 메커니즘과 상통하는 내용으로, 이에 대한 타당성을 추가적으로 입증한다.

## VI. 결론

### 1. 연구 요약

본 연구는 다운로드와 스트리밍이 음원 소비 방식으로서 가지는 본질적인 차이를 유동·고정 소비의 개념을 통해 이해하는 것에서 출발한다(Bardhi and Arnould 2017; Bardhi, Eckhardt, and Arnould 2012). 소유를 수반하는 음원 소비 방식인 다운로드는 보다 고정적인 음원의 소비로, 소유를 수반하지 않는 음원 소비 방식인 스트리밍은 보다 유동적인 음원의 소비로 정리되었다. 일반적으로 소비자는 높은 애착을 가지는 대상에 대해 고정적 소비를 보다 선호하고, 반대의 경우 유동적 소비를 보다 선호하는 것으로 알려져 있으며(Bardhi and Arnould 2017; Bardhi, Eckhardt, and Arnould 2012; Rindfleisch, Burroughs, and Wong 2008), 본 연구는 이를 디지털 음원 소비 방식의 맥락에 적용하였다. 단, 인지도가 높은 대상에 대한 평균적인 애착심은 떨어진다는 기존 연구에 착안하여(Beatty, Homer, and Kahle 1988; Beatty and Smith 1987; Eisentraut 2012), 인지도가 높을 것으로 예상되는 경

력이 긴 아티스트의 곡, 앨범의 타이틀곡, 그리고 음악 프로그램에 방송된 횟수가 많은 곡의 경우 다운로드의 비중이 상대적으로 낮을 것으로 예상하였고, 실증 분석 결과 해당 가설은 지지되는 것으로 드러났다. 해당 분석은 연구 목표에 부합하는 데이터와 절단성 및 이질성을 고려한 엄밀한 모형 분석을 통해 이루어졌고, 이는 본 연구 결과의 타당성을 제고한다.

## 2. 학문적 시사점

본 연구가 가지는 학문적 시사점은 크게 두 가지 층위에서 논의될 수 있다. 첫째, 본 연구는 새로운 음원 소비 방식이 가져온 변화를 넘어, 각 소비 방식 간의 차이를 총체적으로 이해하는 이론적 틀을 제안하였다. 현재 디지털 음원 시장에는 다양한 음원 소비 방식이 공존하고 있으며, 따라서 각 소비 방식들 간의 차이로부터 비롯된 소비 동기의 차이와 특정 소비 방식에 대한 선호를 이끄는 음원의 특성 등을 보다 유기적이고 포괄적으로 이해할 필요가 있다. 본 연구는 이러한 시대적 필요에 부합하는 큰 이해의 틀을 제공하였다는 점에서 그 이론적 의의를 찾을 수 있다. 이는 두 소비 방식의 차이에서 더 나아가, 향후 두 소비 방식 간 존재하는 상호 작용과 같은 총체적인 디지털 음원 시장의 역학 관계에 대한 이해로 발전될 여지가 있어 관련 연구자에게 시사하는 바가 크다. 예컨대, 음원 발매 직후 인지도가 낮은 초반에는 애착도가 높은 소비자들 위주의 다운로드가 스트리밍 소비를 견인하는 한편, 어느 정도 음악이 알려진 이후에는 스트리밍이 다운로드를 견인하는 상황도 예상되며, 이를 체계적으로 분석한다면 그 이론적 의의는 상당할 것으로 기대된다.

둘째, 본 연구는 디지털 음원 시장을 이해하기 위해 고정적·유동적 소비의 개념을 도입하였다. 다운로드와 스

트리밍이 고정적·유동적 소비 개념과 정확히 부합하는 만큼, 해당 개념을 이용하여 본 연구가 제안한 애착심이나 인지도 외에도 다양한 소비 동기 및 소비 양상의 차이를 밝힐 수 있을 것으로 보인다. 예컨대, 유동적 소비가 고정적 소비에 비해 사회 관계망(social network)과 보다 밀접한 연관을 가진다는 선행연구가 있으며(Bardhi and Arnould 2017), 이를 바탕으로 SNS 상에서 많이 언급되는 노래의 경우 다운로드보다 스트리밍의 비율이 높을 것으로 추정해볼 수도 있을 것이다. 또, 유동적 소비는 사용상의 가치가 보다 강조된다는 선행연구에 따라(Bardhi, Eckhardt, and Arnould 2012), 크리스마스 캐롤과 같이 특정 이벤트와 관련되어 관습적으로 사용되는 노래의 경우 다운로드보다 스트리밍의 비율이 높을 것으로 예상할 수 있다. 이와 같이 본 연구는 새로운 개념을 도입하여 향후 디지털 음원 소비 연구가 나아가야 할 방향을 제시한다는 점에서 그 이론적 시사점을 가질 수 있다.

## 3. 실무적 시사점

본 연구는 특정 음원 소비 방식을 선호되게 만드는 음원의 특징을 규명하였고, 본 연구 결과가 가지는 실무적 시사점은 크게 다음의 두 가지 층위에서 논의될 수 있다.

첫째, 개별 음원의 주 판매 경로에 대한 전략적 방향성을 제시한다. 현재 한국 법규를 기준으로 보았을 때, 다운로드의 수익은 곡마다 300원의 수익을, 한 번의 스트리밍은 곡마다 6원의 수익을 발생시킨다. 곡 당 수익 면에서 50배 가까이 차이나는 만큼 디지털 음원 관련 실무자들은 스트리밍 보다는 다운로드를 통한 수익 극대화를 추구할 것으로 보인다. 하지만 본 연구 결과에 따르면 인지도가 상대적으로 높을 것으로 보이는 경력이 긴 아티스트의 노래, 앨범의 타이틀곡, 그리고 음악 프로그램 방송 횟

수가 많은 노래 등의 경우, 다운로드가 아닌 스트리밍을 통한 수익 극대화가 보다 효율적일 수 있다. 따라서 음원의 인지도에 따라 음원 판매 전략을 탄력적으로 조정할 필요가 있을 것으로 보인다. 예컨대, 인지도가 높은 음악의 경우, 스트리밍 서비스를 경쟁 우위로 삼고 이에 대한 집중적 투자가 이루어지는 음악 플랫폼 위주의 음원 유통을 고려할 수 있다. 반면 인지도가 낮은 음악의 경우, 고음질 다운로드 음원을 제공하는 등 다운로드 서비스에 대한 다양한 프로모션을 제공하는 음악 플랫폼 위주의 음원 유통을 고려할 수 있다. 실제로 음악 플랫폼 별 추구하는 포지셔닝이 다르기 때문에<sup>4)</sup> 유통 파트너에 대한 전략적 선택이 필요하며, 이는 한정된 예산을 가진 인지도가 낮은 음원에 있어 특히 효과적일 것으로 기대된다.

둘째, 고객 만족도를 보다 제고할 수 있는 추천 시스템의 개선 방향을 제시한다. Melon, Bugs, Genie 등 다운로드와 스트리밍 서비스를 동시에 제공하는 음악 플랫폼들은 대부분 소비자의 소비 이력에 따라 그 구미에 맞는 음악들을 추천하는 추천 시스템을 운영하고 있다. 하지만 본 연구 결과는 음원 뿐 아니라 그 소비 방식까지 고려하여 추천하는, 보다 정교한 추천 시스템의 도입을 가능케 한다. 예컨대 사용자들이 기존에 들던 음악적 취향을 고려하여 그와 비슷한 음악을 추천하되, 그 중 보다 인지도가 낮은 음악은 다운로드 구독자에게, 보다 인지도가 높은 음악은 스트리밍 서비스 구독자에게 추천함으로써 추천 성공률을 더욱 높일 수 있다. 이는 소비자들에게 더욱 편리한 추천 시스템을 제공함으로써 그들의 만족도를 높이고 장기적으로는 서비스 이용을 증가시킬 것으로 예상된다.

#### 4. 한계점 및 향후 연구 방향

본 연구는 다운로드와 스트리밍 간의 관계를 조명하고 이를 실증적으로 밝혀냄으로써 기존 연구들의 공백을 채우고자 하였다. 그러나 향후 연구를 위해 본 연구가 갖는 한계점에 대한 언급이 필요하다.

우선, 본 연구는 절단된 소비 데이터를 이용하였다. 이와 같은 절단성을 다루기 위해 Tobit 모형을 사용하여 세심하게 모형을 계획하였으나, 그럼에도 상위 100위에 들지 못한 곡들에 대해 본 연구 결과가 동일하게 적용될 수 있을지에 대해서는 조심스럽게 접근할 필요가 있을 것으로 판단된다.

다음으로, 본 연구는 곡 별 음원 소비량에 대한 데이터를 통해 진행되었다. 만약 개인 별 소비 데이터를 수집할 수 있었다면 애착심에 대한 음원 소비의 동기적인 측면을 보다 깊게 살펴볼 수 있었을 것이다. 추가적으로, 나이나 음악 관련 전문성과 같은 소비자 개개인에 대한 인구통계학적 및 행동적 변수에 대한 이해가 향후 연구에서 이루어질 수 있을 것이다. 예컨대, 젊은 소비자는 다운로드보다 스트리밍을 더 선호하고, 이들이 스트리밍 유저의 대부분을 차지한다는 연구 결과도 존재한다(IFPI 2015; Weijters, Goedertier, and Verstrecken 2014). 따라서 소비자들의 나이나 전문성이 곡, 아티스트 및 장르에 따른 음원 소비에 어떻게 영향을 미치는지를 추가적으로 살펴볼 수 있을 것이다.

〈최초투고일: 2017년 10월 30일〉

〈수정일: 1차: 2018년 1월 2일, 2차: 2018년 2월 12일〉

〈게재확정일: 2018년 2월 14일〉

4. <http://www.kinews.net/news/articleView.html?idxno=69317>

## 참고문헌

- Aguiar, Luis, and Joel Waldfogel(2015), “Streaming Reaches Flood Stage: Does Spotify Stimulate or Depress Music Sales?,” Working Paper, *National Bureau of Economic Research*.
- Asai, Sumiko(2011), “Demand analysis of hit music in Japan,” *Journal of Cultural Economics*, 35(2), 101-117.
- Bardhi, Fleura, Giana M. Eckhardt, and Eric J. Arnould(2012), “Liquid relationship to possessions,” *Journal of Consumer Research*, 39(3), 510-529.
- Bardhi, Fleura, and Giana M. Eckhardt(2017), “Liquid Consumption,” *Journal of Consumer Research*, 44(3), 582-597.
- Bauman, Zygmunt(2000), *Liquid modernity*, Polity, Cambridge.
- Beatty, Sharon E., Pamela Homer, and Lynn R. Kahle(1988), “The involvement—commitment model: Theory and implications,” *Journal of Business research*, 16(2), 149-167.
- Beatty, Sharon E. and Scott M. Smith(1987), “External search effort: An investigation across several product categories,” *Journal of consumer research*, 14(1), 83-95.
- Beck, Jonathan(2007), “The sales effect of word of mouth: a model for creative goods and estimates for novels,” *Journal of Cultural Economics*, 31(1), 5-23.
- Berli, Asunción, Josefa D. Martin, and Agustín Quintana (2004), “A model of customer loyalty in the retail banking market,” *European Journal of Marketing*, 38(1/2), 253-275.
- Besanko, David, Jean-Pierre Dubé, and Sachin Gupta(2005), “Own-brand and cross-brand retail pass-through,” *Marketing Science*, 24(1), 123-137.
- Bhattacharjee, Sudip, Ram D. Gopal, Kaveepan Lertwachara, and James R. Marsden(2006), “Impact of legal threats on online music sharing activity: An analysis of music industry legal actions,” *The Journal of Law and Economics*, 49(1), 91-114.
- Blume, Lawrence E., and David Easley(2002), “Optimality and natural selection in markets,” *Journal of Economic Theory*, 107(1), 95-135.
- Bockstedt, Jesse, Robert J. Kauffman, and Frederick J. Riggins (2005, January), “The move to artist-led online music distribution: Explaining structural changes in the digital music market,” *In System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on (pp. 180a-180a)*, IEEE.
- Bockstedt, Jesse C., Robert J. Kauffman, and Frederick J. Riggins(2006), “The move to artist-led on-line music distribution: a theory-based assessment and prospects for structural changes in the digital music market,” *International Journal of Electronic Commerce*, 10(3), 7-38.
- Bolton, Ruth N., P. K. Kannan, and Matthew D. Bramlett (2000), “Implications of loyalty program membership and service experiences for customer retention and value,” *Journal of the academy of marketing science*, 28(1), 95-108.
- Datta, Hannes, George Knox, and Bart J. Bronnenberg(2017), “Changing their tune: How consumers’ adoption of online streaming affects music consumption and discovery,” Working paper, *Marketing Science*.
- Dewan, Sanjeev, and Jui Ramaprasad(2014), “Social media, traditional media, and music sales,” *Mis Quarterly*, 38(1).
- Dommer, Sara L., and Vanitha Swaminathan(2012), “Explaining the endowment effect through ownership: The role of identity, gender, and self-threat,” *Journal of Consumer Research*, 39 (5), 1034-1050.
- Dowling, Grahame R., and Mark Uncles(1997), “Do customer loyalty programs really work?,” *Sloan management review*, 38(4), 71.
- Eisentraut, Jochen(2012), *The accessibility of music: participation, reception, and contact*, Cambridge University Press.
- Elberse, Anita(2010), “Bye-bye bundles: The unbundling of music in digital channels,” *Journal of Marketing*, 74.3, 107-123.
- Graham, Gary, Bernard Burnes, Gerard J. Lewis, and Janet Langer(2004), “The transformation of the music industry supply chain: A major label perspective,” *International Journal of Operations and Production Management*, 24(11),

- 1087-1103.
- Heilman, Carrie M., Douglas Bowman, and Gordon P. Wright (2000), "The evolution of brand preferences and choice behaviors of consumers new to a market," *Journal of Marketing Research*, 37(2), 139-155.
- Hendricks, Ken, and Alan Sorensen(2009), "Information and the skewness of music sales," *Journal of political Economy*, 117(2), 324-369.
- Hughes, Jerald, and Karl R. Lang(2003), "If I had a song: The culture of digital community networks and its impact on the music industry," *International Journal on Media Management*, 5(3), 180-189.
- IFPI(2015), "Digital Music Report," (accessed June 19, 2016), [available at <http://www.ifpi.org/downloads/Digital-Music-Report-2015.pdf>].
- IFPI(2016), "Digital Music Report," (accessed October 13, 2017), [available at <http://www.ifpi.org/downloads/GMR2016.pdf>].
- Jedidi, Kamel, Sharan Jagpal, and Puneet Manchanda(2003), "Measuring heterogeneous reservation prices for product bundles," *Marketing Science*, 22(1), 107-130.
- Jeung-Il Oh and Min-Soo Park(2006), "Korean Online-Music Industry : Current Issues and Policy Implications," *Journal of Industrial Economics and Business*, 19(1), 157-178.
- Kohli, Ajay K., and Bernard J. Jaworski(1990), "Market orientation: the construct, research propositions, and managerial implications," *The Journal of Marketing*, 1-18.
- Lawson, Stephanie J., Mark R. Gleim, Rebeca Perren, and Jiyoung Hwang(2016), "Freedom from ownership: An exploration of access-based consumption," *Journal of Business Research*, 69(8), 2615-2623.
- Maftai, Vlad A., Vassilis C. Gerogiannis, and Elpiniki I. Papageorgiou(2016), "Critical success factors of online music streaming services-a case study of applying the fuzzy cognitive maps method," *International Journal of Technology Marketing*, 11(3), 276-300.
- Manchanda, Puneet, Asim Ansari, and Sunil Gupta(1999), "The "shopping basket": A model for multicategory purchase incidence decisions," *Marketing science*, 18(2), 95-114.
- Montoya, Ricardo, Oded Netzer and Kamel Jedidi(2010), "Dynamic allocation of pharmaceutical detailing and sampling for long-term profitability," *Marketing Science*, 29(5), 909-924.
- Mussulman, Joseph A.(1974), *The uses of music: An introduction to music in contemporary American life*, Prentice Hall.
- Nguyen, Godefroy D., Sylvain Dejean, and François Moreau (2012), "Are streaming and other music consumption modes substitutes or complements?," Available at [ssrn.com/abstract=2025071](http://ssrn.com/abstract=2025071). (March 16)
- Nguyen, Godefroy D., Sylvain Dejean, and François Moreau (2014), "On the complementarity between online and offline music consumption: the case of free streaming," *Journal of Cultural Economics*, 38(4), 315-330.
- Papies, Dominik, Felix Eggers, and Nils Wlömert(2011), "Music for free? How free ad-funded downloads affect consumer choice," *Journal of the Academy of Marketing Science*, 39 (5), 777-794.
- Peitz, Martin, and Patrick Waelbroeck(2004), "The effect of internet piracy on CD sales: Cross-section evidence," Working Paper, *CESifo*, Available at SSRN: <https://ssrn.com/abstract=511763>
- Peitz, Martin, and Patrick Waelbroeck(2006), "Why the music industry may gain from free downloading—the role of sampling," *International Journal of Industrial Organization*, 24(5),907-913.
- Gopal, Ram D., Sudip Bhattacharjee, and G. L. Sanders(2006), "Do artists benefit from online music sharing?," *The Journal of business*, 79(3), 1503-1533.
- Rindfleisch, Aric, James E. Burroughs, and Nancy Wong(2008), "The safety of objects: Materialism, existential insecurity, and brand connection," *Journal of Consumer Research*, 36 (1), 1-16.
- Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby (1996), "The value of purchase history data in target marketing," *Marketing Science*, 15(4), 321-340.



- Schaefers, Tobias, Stephanie J. Lawson, and Monika Kukar-Kinney(2016), “How the burdens of ownership promote consumer usage of access-based services,” *Marketing Letters*, 27(3), 569-577.
- Schweidel, David A., Peter S. Fader, and Eric T. Bradlow (2008), “A bivariate timing model of customer acquisition and retention,” *Marketing Science*, 27(5), 829-843.
- Sonnier, Garrett, Andrew Ainslie, and Thomas Otter(2007), “Heterogeneity distributions of willingness-to-pay in choice models,” *Quantitative Marketing and Economics*, 5(3), 313-331.
- Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde(2002), “Bayesian measures of model complexity and fit,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.
- Strobl, Eric A., and Clive Tucker(2000), “The dynamics of chart success in the UK pre-recorded popular music industry,” *Journal of Cultural Economics*, 24(2), 113-134.
- Varian, Hal R.(2000), “Buying, sharing and renting information goods,” *The Journal of Industrial Economics*, 48(4), 473-488.
- Wang, Yanwen, Michael Lewis, Cynthia Cryder, and Jim Sprigg (2016), “Enduring effects of goal achievement and failure within customer loyalty programs: A large-scale field experiment,” *Marketing Science*, 35(4), 565-575.
- Weijters, Bert, Frank Goedertier, and Sofie Verstreken(2014), “Online music consumption in today’s technological context: Putting the influence of ethics in perspective,” *Journal of Business Ethics*, 124(4), 537-550.
- Wlömert, Nils, and Dominik Papies(2016), “On-demand streaming services and music industry revenues—Insights from Spotify’s market entry,” *International Journal of Research in Marketing*, 33(2), 314-327.
- Yang Sug Kim and Sang Houck Lee(2015), “A study to Find Factors of the Box Office Record in Korean Popular Music : by Control Effects of Management Companies and Distribution Companies,” *Asia-Pacific Journal of Business Venturing and Entrepreneurship*, 10(1), 23-31.
- YG Entertainment(2015), “Business Report,” (accessed October 29, 2017), [available at <https://dart.fss.or.kr/dsaf001/main.do?rcpNo=20150331002022>].

## 부록 (Appendix)

### 1. 사전 분포(prior distribution)

논의의 편의를 위하여 본문에서 정의된 변수 및 상수들을 아래와 같이 재정의한다.

우선, 종속 변수의 벡터인  $\underline{R}^*$ 는 길이  $n$ (i.e, 전체 데이터의 관측치 수)을 가지며, 각 성분(entry)은 실질 다운로드-스트리밍 비율( $R_{ijt}^*$ )의 각 값에 해당한다.  $\underline{R}_j^*$ 는  $\underline{R}^*$ 의 부분 벡터로, 길이는  $n_j$  (i.e, 노래  $j$ 에 대한 데이터의 관측치 수)이며, 각 성분은 해당 곡의 실질 다운로드-스트리밍 비율의 각 값에 해당한다. 한편, 독립변수의 행렬인  $Z$ 는  $(1, \ln(YSD), title, show, X)$ 의  $n \times k$  (i.e., 절편을 포함한 전체 설명 변수의 개수), 각 성분은 해당 관측치에 대한 독립 변수의 값으로 정의된다.  $Z_j$ 는  $Z$ 의 부분 행렬로,  $n_j \times k$ 의 행렬이며, 노래  $j$ 에 대한 독립 변수 정보만을 포함하고 있다.

상수 벡터의 재정의에 앞서, 본문에서 소개된 노래 별 절편(song-specific intercept)은 다음과 같이 표현 가능하다.

$$\beta_{0j} = \beta_0 + \eta_j, \text{ where } \eta_j \sim \mathcal{N}(0, \sigma_0^2) \quad (A1)$$

위의 표현에서  $\eta_j$ 는 평균 절편( $\beta_0$ )으로부터 노래  $j$ 의 절편( $\beta_{0j}$ )이 떨어진 차이(deviation)를 의미한다. 한편,  $\eta$ 는 길이  $n$ 의 벡터로, 각 성분은 노래  $j$ 에 상응하는  $\eta_j$ 에 해당한다.  $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)$ 는 길이  $k$ 의 벡터로 각 성분은 그에 상응하는 모수에 해당한다. 본 모형에서 추정코자 하는 상수(parameter)들의 사전분포는 아래와 같이 정의한다.

$$\begin{aligned} \beta &\sim \mathcal{N}(0, \Sigma) \text{ where } \Sigma = 10^3 \times I_K & (A2) \\ \sigma_0 &\sim \text{InverseGamma}(\nu_0, \nu_0) \text{ where } \nu_0 = 10^{-3}, \nu_0 = 10^{-3}, \\ \text{and } \sigma_\varepsilon &\sim \text{InverseGamma}(\nu_\varepsilon, \nu_\varepsilon) \text{ where } \nu_\varepsilon = 10^{-3}, \nu_\varepsilon = 10^{-3}, \end{aligned}$$

위의 사전 분포는 선행 문헌에서 제안한 diffuse conjugate prior에 부합한다(Gelman et al. 2014; Rossi, Allenby, and McCulloch 2012).

### 2. 샘플링 알고리즘 (sampling algorithm)

앞서 설명하였듯이 모수의 사전분포는 모두 conjugate prior로, 모수 추정은 깁스(Gibbs) 샘플링으로 시행하였다. 보다 구체적으로, 조건부 사후 분포(conditional posterior distribution)로부터의 샘플링은 아래와 같은 알고리즘을 따라 이루어졌다.

우선 알고리즘은 실질 다운로드-스트리밍 비율( $\ln(R_{ijt}^*)$ )의 샘플링에서 출발한다. 본문에서 언급하였듯,  $\ln(R_{ijt}^*)$ 의 관측은  $y_{1ijt} = 1$ 인 3,629개의 관측치에 제한되어 있다. 따라서 모형의 추정을 위해서는  $y_{2ijt} + y_{3ijt} = 1$ 인 1,649개의 관측치에 대한  $\ln(R_{ijt}^*)$ 가 샘플링 되어야 하며, 관측치의 절단성을 고려하여 매 시행 마다 아래와 같은 절단 정규분포(truncated normal distribution)에서의 샘플링이 시행된다.

$$\begin{aligned} \ln(R_{ijt}^*) \mid \underline{\beta}, \underline{\eta}_j, \sigma_\varepsilon &= \ln(R_{ijt}), & \text{when } y_{1ijt} = 1, & (A3) \\ &\sim N(Z_{ijt}\beta + \underline{\eta}_j, \sigma_\varepsilon) \text{ conditional on } R_{ijt}^* > D_{ijt}/\bar{S}_t, & \text{when } y_{2ijt} = 1, \\ &\sim N(Z_{ijt}\beta + \underline{\eta}_j, \sigma_\varepsilon) \text{ conditional on } R_{ijt}^* < \bar{D}_t/S_{ijt}, & \text{when } y_{3ijt} = 1. \end{aligned}$$

다음으로, 본 연구의 핵심 모수인  $\beta$ 를 샘플링한다. 앞서 언급하였듯,  $\ln(R_{ijt}^*)$ 의 샘플링 이후의 과정은 random effect regression model의 추정과 동일하며, 모수 벡터  $\beta$ 는 아래와 같은 조건부 사후 분포에서 샘플링 된다.

$$\beta \mid \ln(R^*), \underline{\eta}, \sigma_\varepsilon \sim N((Z'Z)\sigma_\varepsilon^{-2} + \Sigma^{-1})^{-1}(\sigma_\varepsilon^{-2}Z'(\ln(R^*) - \underline{\eta})), ((Z'Z)\sigma_\varepsilon^{-2} + \Sigma^{-1})^{-1} \quad (A4)$$

각 노래 별 이질성을 통제하는 이질적 절편의 편차( $\eta_j$ )를 매 곡  $j$ 마다 샘플링하며, 이의 조건부 사후 분포는 아래와 같다.

$$\eta_j \mid \ln(R_j^*), \beta, \sigma_0, \sigma_\varepsilon \sim N((n_j\sigma_\varepsilon^{-2} + \sigma_0^{-2})^{-1}\Sigma_j(\ln(R_j^*) - Z_j\beta), (n_j\sigma_\varepsilon^{-2} + \sigma_0^{-2})^{-1}) \quad (A5)$$

노래 별 이질성의 표준편차는 아래와 같은 조건부 사후 분포로부터 샘플링한다.

$$\sigma_0 | \underline{\eta} \sim \text{InverseGamma}(v_0 + n_j/2, v_0 + \underline{\eta}'\underline{\eta}/2) \quad (\text{A6})$$

에러항의 표준편차는 아래와 같은 조건부 사후 분포로부터 샘플링한다.

$$\sigma_\varepsilon | \ln(\underline{R}_j^*), \beta, \underline{\eta} \sim \text{InverseGamma}(v_\varepsilon + n/2, v_\varepsilon + (\ln(\underline{R}_j^*) - Z_j\beta - \underline{\eta})' (\ln(\underline{R}_j^*) - Z_j\beta - \underline{\eta})/2) \quad (\text{A7})$$

모수의 추정은 위의 알고리즘을 10,000번 반복 시행하여 수렴값을 얻음으로써 이루어졌다.

#### 참고 문헌 - 부록

Gelman, A. John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin(2014), *Bayesian data analysis* (Vol. 2), Boca Raton, FL: CRC press.

Rossi, Peter E., Greg M. Allenby, and Rob McCulloch(2012), *Bayesian statistics and marketing*, John Wiley & Sons.