

생존분석을 통한 Car-Sharing 이용자 이용주기 분석

Survival Analysis of Car-Sharing User Utilization Interval

정종우¹⁾, 가동주²⁾, 이재현²⁾, 이청원³⁾

1) 서울대학교 건설환경공학부 석사과정, 2) 서울대학교 건설환경공학부 석박통합과정, 3) 서울대학교 건설환경공학부 교수



1. Introduction

Background

- 카셰어링 양적 증가에 따른 사용자의 이용특성에 대한 이해 필요
- 다양한 이동주기에 따른 카셰어링 이용특성 분석
- 생존분석을 통해 국내 카셰어링 이용특성을 이해함으로써 기존에 선행된 다른 연구 방법들과의 비교
- 대중교통 접근성이 카셰어링 이용주기에 미치는 영향 확인 필요

Purposes

- 생존분석을 통한 카셰어링 이용주기에 영향을 미치는 요소 확인 및 분석

2. Literature Reviews

Domestic Car-Sharing Literature Reviews

- 카셰어링의 양적 성장에 기인한 수요 추정이 주를 이루고 있음
- 양적 성장에 기인한 연구에서 질적 성장을 위한 연구로 발전되고 있는 상황
- 본 연구의 방법론적인 측면인 생존분석을 기반으로 한 국내 선행 연구가 없음

단계	주제	특징
카셰어링 도입 전 : 2009~2010년	카셰어링 도입효과 카셰어링 이용추이	<ul style="list-style-type: none">친환경 교통서비스로써 도입 타당성업무중심형 자동차 공동이용 효과 분석
카셰어링 도입 초기 : 2011~2013년	카셰어링 서비스 모델 편익분석 대여소 입지선정	<ul style="list-style-type: none">카셰어링 서비스의 최적가격 추정차량대여소 입지선정실시간 차량 재배치 전략
카셰어링 양적 성장 : 2014~2017년	교통정책과 수요추정 비즈니스 모델	<ul style="list-style-type: none">카셰어링 이용자 수요 추정지역경제 연계를 위한 로커 거버넌스 구축

3. Methodology

Description of Samples Used in the Analysis

- Null 값 처리 및 데이터 가공
 - 관리자 성격의 사용자 제외 (연 365회 이상 사용)
 - 일회성 카셰어링 이용자 제외 (연 2회 미만 사용)
 - 도심지와 지하철역까지 거리 컬럼 추가
 - 적어도 4달에 1번 이상 사용자 분석
 - 1644obs. 18variables
- 변수 설정
 - ID, AGE, SEX, ADDRESS, NUM, CAR_RATIO, SUB_ROUTE, SUB_DIST, BUS_NUM, CITY_DIST, RES, FIRST, LAST, GAP, GAP_MEAN, RES_GAP, STATUS (RES=재사용 일자, RES_GAP=재사용까지 걸린 일자, STATUS=이용 횟수)

Survival Analysis

- Cox 비례위험모형(Cox's Proportional Hazard Model)
 - $h(t)=h_0(t)\exp(B_1x_1 + \dots + B_kx_k)$
 - $h_0(t)$ =기저 위험함수, x_k =생존시간 T에 영향을 주는 변수, B =회귀계수
- Log Rank Test 검정: 생존함수의 비교
 - 생존함수: 환자가 t시간 이상 생존할 확률
 - 두 그룹의 생존 함수를 총괄적으로 비교하기 위한 가설
 $H_0: S_1(t) = S_2(t)$ for all t
 $H_a: S_1(t) \neq S_2(t)$ for some t

구분	내용
Cox 비례위험모형	<ul style="list-style-type: none">총 19개의 변수들에 대한 다중분석 시행이용특성에 영향을 미치는 변수들의 P-value 값 확인
Kaplan-Meier	<ul style="list-style-type: none">두 그룹에 대한 시각적 표현역세권/비역세권에 따른 카셰어링 이용특성 확인
Log-Rank-Test	<ul style="list-style-type: none">두 그룹에 대한 가설 검정 실시분석결과 도출

4. Results

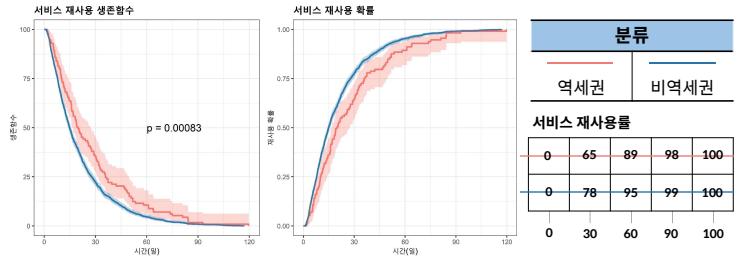
Cox Proportional Hazard Model Estimation Result

구분	P-Value	설명
Sub_dist	0.026	
Age	0.034	
Sex	0.690	
Car_ratio	0.015	
Sub_route	0.750	
Bus_route	0.025	
Bus_num	0.310	
City_dist	0.440	
Res_gap	$3.6e^{-0.8}$	
City_sub	0.410	

구분	Exp(coef)	Exp(-coef)	Lower .95	Upper .95
Sub_dist	1.033	0.9681	1.004	1.063

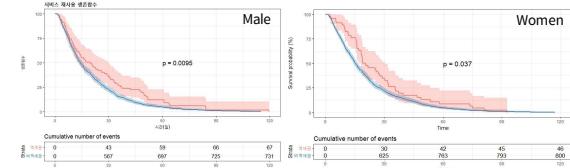
- 단변량 분석 결과 지하철역까지 거리가 1km 증가할 때 이용주기 3.3% 증가

Log Rank Test Result



- 1st: 지하철역까지 거리에 역세권 개념(10분 이내)을 도입, 역세권 유무로 Grouping
- 2nd: H_0 (귀무가설): '역세권 유무가 카셰어링 이용주기에 영향을 미치지 않는다'
- 3rd: P-value가 0.00083으로 귀무가설 기각
- 결론 '두 집단의 생존함수가 다르다는 충분한 근거가 있다'
- 지하철역까지 역세권 유무는 이용주기에 따른 차이가 존재

<성별 변수 추가 Kaplan-Meier>



- 남녀 모두 역세권에 가까울 수록 카셰어링 이용주기가 감소함을 확인

5. Conclusions

요약

- 본 연구에서는 실증 운영사례 데이터를 활용하여 영향관계 분석
- Cox 비례위험모형을 통해 지하철역까지 거리, 성별에 따른 이용주기 특성 이해
- Log Rank Test를 통해 역세권 유무가 카셰어링 이용주기에 미치는 영향 확인

시사점

- 기존 카셰어링 이용특성연구에서 사용된 적 없는 생존분석을 통하여 본 연구를 진행
- 실증 운영사례 데이터를 통한 결과 도출로써 보다 공학적인 분석을 시행

6. References

- Susan A. Shaheen, Mollyanne Meyn(2003). U.S. Shared-use Vehicle Survey Findings. California PATH Research Report
- Susan A. Shaheen, Ph.D.(2002). SHARED-USE VEHICLE SERVICES: A SURVEY OF NORTH AMERICAN MARKET DEVELOPMENTS
- DJ Roe, S Cretin(1997) Estimating effectiveness of cardiac arrest interventions: a logistic regression survival model. (PMID:9396421)
- Kleinbaum,Klein. Survival Analysis a Self Learning Text. Third Edition
- 정준석,노정현.(2017).P2P 카셰어링 도입 및 활성화 방안 연구 18(2),47-60
- 최현수,박준태.(2014)서울시의 카셰어링 이용도에 대한 지역적 요인특성분석,2014년_춘계학술대회
- 김숙희,이규진.(2016)대중교통 접근성이 카셰어링 이용수요에 미치는 영향. ITS학회논문지,15(4),01-11
- 송경일,최종수.(2008).(SPSS 15를 이용한) 생존자료의 분석,401-407

Acknowledgement

이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2017R1D1A1B03028218).